Mobile Access Bandwidth in Practice: Measurement, Analysis, and Implications

Xinlei Yang[†*], Hao Lin[†*], Zhenhua Li[†], Feng Qian[§], Xingyao Li[†], Zhiming He[†]
Xudong Wu[†], Xianlong Wang[†], Yunhao Liu[†], Zhi Liao[‡], Daqiang Hu[‡], Tianyin Xu[¶]

[†]Tsinghua University [‡]UUSense Technology Inc.
[§]University of Minnesota, Twin Cities [¶]University of Illinois at Urbana-Champaign

ABSTRACT

Recent advances in mobile technologies such as 5G and WiFi 6E do not seem to deliver the promised mobile access bandwidth. To effectively characterize mobile access bandwidth in the wild, we work with a major commercial mobile bandwidth testing app to analyze mobile access bandwidths of 3.54M end users in China, based on fine-grained measurement and diagnostic information. Our analysis presents a surprising and frustrating fact—in the past two years, the average WiFi bandwidth remains largely unchanged, while the average 4G/5G bandwidth decreases remarkably. Our analysis further reveals the root causes—the bottlenecks in the underlying infrastructure (*e.g.,* devices and wired Internet access) and side effects of aggressively migrating radio resources from 4G to 5G—with implications on closing the technology gaps. Additionally, our analysis provides insights on building ultra-fast, ultra-light bandwidth testing services (BTSes) at scale. Our new design dramatically reduces the test time of the commercial BTS from 10 seconds to 1 second on average, with a 15× reduction on the backend cost.

CCS CONCEPTS

• **Networks** → **Mobile networks**; **Network measurement**; **Network performance analysis**;

KEYWORDS

Mobile Network; LTE/5G Network; WiFi Network; Access Bandwidth; Bandwidth Testing

ACM Reference Format:
Xinlei Yang, Hao Lin, Zhenhua Li, Feng Qian, Xingyao Li, Zhiming He, Xudong Wu, Xianlong Wang, Yunhao Liu, Zhi Liao, Daqiang Hu, Tianyin Xu. 2022. Mobile Access Bandwidth in Practice: Measurement, Analysis, and Implications. In *ACM SIGCOMM 2022 Conference (SIGCOMM '22), August 22–26, 2022, Amsterdam, Netherlands.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3544216.3544237

*Co-primary authors. Zhenhua Li is the corresponding author.

SIGCOMM '22, August 22–26, 2022, Amsterdam, Netherlands

ACM ISBN 978-1-4503-9420-8/22/08.
https://doi.org/10.1145/3544216.3544237

1 INTRODUCTION

Mobile access technologies have made significant progress in recent years. For example, 5G and WiFi 6E, the latest cellular and WiFi technologies, can support up to 20 Gbps and 9.6 Gbps bandwidth respectively. Those exciting new mobile technologies are the key enabler for a wide range of emerging applications such as Metaverse, autonomous vehicles, and 3D Ultra-HD videos. However, despite the aggressive deployment of 5G and WiFi 6E, reports from large-scale bandwidth testing services (BTSes) reveal that as of late 2021, the median 5G bandwidth merely reaches 135 Mbps in the US and 304 Mbps in China, while the median WiFi bandwidth is only 137 Mbps in the US and 153 Mbps in China [26]. Apparently, the promises of new wireless technologies are significantly under delivered in real-world deployments.

Understanding the root causes of undesirable wireless performance in the wild is a first step towards improving the state of the art. However, it is hampered by the complexity of wireless protocol stacks, the wide spectrum of the mobile ecosystem, and a lack of large-scale measurements. For example, existing studies on commercial 5G performance are based on controlled experiments at limited scales [54, 55, 74]. While some major BTSes do report the landscape of mobile Internet performance, their data are coarse-grained and are limited by (mostly) web-based tools which are incapable of capturing rich, cross-layer diagnostic data.

Cross-Layer and Cross-Technology Measurement. To fill the critical gap, we take a unique opportunity to work with a major Android BTS app named UUSpeedTest [38] (abbreviated as BTS-APP), which has 17M users (mostly located in China) and serves ~0.2M bandwidth test requests per day on average. Its bandwidth testing uses the standard "probing by flooding" approach [77] which is also used by almost all the commercial BTSes today (*e.g.,* Speedtest [27] and SpeedOf [22])—upon a test request, BTS-APP first downloads large files from a nearby server for ten seconds, and then samples the throughput statistics over time to estimate the access bandwidth.

BTS-APP faces two fundamental challenges. First, the coarse-grained data prevent it from pinpointing the root causes of undesirable mobile access bandwidth, which customers are eager to learn. Second, the BTS infrastructure is not scalable—as the wireless access bandwidth keeps increasing, the monetary cost for operating the BTS infrastructure is growing considerably. Note that these two challenges are faced by all the major BTS providers, as they are in essence using the same testing approach described above.

Our collaboration with BTS-APP addresses both challenges. To gain deep insights into the undesirable access bandwidth, we enhance the client of BTS-APP by continuously collecting important PHY- and MAC-layer data through standard Android APIs during a bandwidth test. Our enhancement is implemented as a lightweight plugin for BTS-APP without requiring any additional privileges, making it easy to deploy. Under informed user consent and a proper IRB, over four months (Aug. to Nov. 2021), 3.54M customers used the enhanced BTS-APP to perform 23.6M bandwidth tests, which cover all four major ISPs in China, 2.04M cellular base stations (4G/5G), and 4.47M WiFi APs (WiFi 4/5/6). To the best of our knowledge, this constitutes one of the largest mobile Internet performance datasets reported in the literature.

Data Analysis. Our analysis of the above dataset yields several major findings. A surprising and frustrating finding is that, over the past two years (2020 and 2021), despite the increasing deployment of WiFi 6 and 5G, the average WiFi bandwidth remains largely unchanged: 132 Mbps in 2020 *vs.* 137 Mbps in 2021, and the average 4G/5G bandwidth even decreases: 68/343 Mbps in 2020 *vs.* 53/305 Mbps in 2021. In this paper, we reveal the root causes of such counter-intuitive results from cross-technology perspectives.

For 4G, we observe that three 4G LTE bands (Bands 1, 28, and 41), which occupy 58.2% of the entire high-bandwidth LTE spectrum, were "refarmed" for 5G use in early 2021 [56]. This leads to a sharp decrease in the LTE performance observed in our data: from 2020 to 2021, the average 4G bandwidth has dropped by 22% to 53 Mbps, achieving only 18% of the ISPs' claimed 300 Mbps bandwidth. On the other hand, we find that the top 6.8% of tests where the bandwidth exceeds 300 Mbps were mostly conducted alongside major urban roads where eNodeBs are equipped with LTE-Advanced (with features such as carrier aggregation and enhanced MIMO) to cope with large traffic volumes.

With regard to 5G, we still observe an 11% decrease in the average bandwidth from 2020 to 2021. This is attributed to several factors. First, spectrum refarming may play a negative role. For example, the average bandwidth of the refarmed Band 1 and Band 28 is as low as 103 Mbps and 113 Mbps respectively, because of the thin ($\leq$60 MHz) spectrum refarmed within each band. Second, our data suggest that a strong received signal strength (RSS) level does not necessarily translate into high 5G bandwidth: the average bandwidth under excellent (level 5) RSS is even lower than that under weaker (level 3 and 4) RSS. We find that excellent-RSS tests are more likely to be performed in crowded urban areas where complex multipath interference incurred by buildings [58], load balancing issues caused by heavy population [76], and poor handover problems due to the dense 5G gNodeBs [17] all become prominent. In contrast, in our 4G data, RSS and measured bandwidth are more positively correlated, given the much more mature, well-provisioned 4G infrastructure deployed for more than 10 years.

As to WiFi access, we find that while WiFi 5 is superior to WiFi 4, their average bandwidths are close (195 Mbps *vs.* 208 Mbps) over the 5 GHz band.[1] Regarding the WiFi 6 access, its average bandwidth goes up to 345 Mbps, which is still far below its advertised capability. Our data indicate the reason to be the slow wired Internet, as ~64% of the WiFi customers are still using $\leq$200-Mbps fixed "broadband" Internet access that offsets the advantages of WiFi 5/6.

[1]Note that WiFi 4 and WiFi 6 use both the 2.4 GHz and 5 GHz bands, while WiFi 5 uses the 5 GHz band only.

Implications. Combining the above analysis of 4G, 5G, and WiFi, our analysis depicts a holistic, complete picture of today's mobile access bandwidth. In particular, it quantitatively reveals the side effect of aggressively (and perhaps imprudently) migrating radio resources from 4G, which still owns the vast majority of today's cellular users, to 5G. The side effect is further aggravated by the bottlenecks of 5G and other infrastructures (*e.g.,* devices and the wired Internet access). Although spectrum refarming is inevitable as cellular technology evolves, the current LTE spectrum resources are severely fragmented. This makes contiguous high-bandwidth spectrum available for refarming rather scarce, leading to low 5G bandwidth as exhibited in our data.

Our results advocate more effective band defragmentation and refarming strategies. Meanwhile, since 4G and 5G will coexist for a very long time, our findings also call for strengthening existing LTE infrastructure in a cost-effective manner, such as widening the LTE-Advanced deployment. Our findings also bring implications for other stakeholders. For example, wired ISPs and content providers should take into account the emerging wireless technologies when budgeting their network infrastructure. In addition, customers should be better informed and educated to understand the performance and bottlenecks of new technologies.

Rearchitecting BTS-APP. The study also provides deep insights on rearchitecting the BTS-APP system to make it ultra-fast and ultra-light, addressing the infrastructure scalability challenge. We make two key observations from our data. First, for high-speed mobile networks, TCP slow start accounts for a significant fraction of time taken by a bandwidth test, but it does not contribute useful bandwidth samples. This drives us to reduce BTS-APP's bandwidth probing duration. Second, we find that mobile access bandwidth can be well modeled by a multi-modal Gaussian distribution for a given access technology. This dictates a novel data-driven approach for selecting the initial probing bandwidth without lengthy, expensive calibration: using the bandwidth that best matches the client's access technology with high probability and then fine-tuning the probing data rate with the client feedback for fast convergence.

We mechanize the above ideas by re-implementing BTS-APP's bandwidth testing logic, as well as several system-level improvements such as a UDP-based protocol allowing customized bandwidth probing and an informed server deployment strategy for accommodating the workloads. We deploy and evaluate the new system (dubbed Swiftest) for a whole month, serving 0.2M users and 0.31M test requests. The large-scale, real-world evaluation indicates that Swiftest is highly effective on every dimension. For test duration, Swiftest takes merely 1.19 seconds on average to accomplish a bandwidth test, and 55% of tests are finished within one second (including the initial PING latency). In contrast, BTS-APP takes 10 seconds to complete a test. For infrastructure cost, Swiftest cuts the server expense by $\sim$15×: it only requires 20 100-Mbps servers to support a realistic workload of $\sim$10K tests per day (with margins) while BTS-APP needs 50 1-Gbps servers. For accuracy, the average difference between Swiftest and BTS-APP is as small as 5%.

Code/Data Release. We have released the code and data at https://MobileBandwidth.github.io/ to help the community understand large-scale mobile access bandwidth and to develop Swiftest for customized mobile measurements.

2 STUDY METHODOLOGY

This section first presents BTS-APP's bandwidth testing architecture, and then describes the lightweight plugin we build for collecting in-depth network information in the wild.

BTS-APP's System Architecture. The bandwidth testing logic and system deployment of BTS-APP are quite similar to those of Speedtest, a state-of-the-art BTS system that owns the largest user scale (around 15M user requests served per day) and server pool (16,190 test servers deployed across the globe as of January 2022) [27]. Some additional adaptations are made by the development team to fit the specific workload of BTS-APP.

Upon a test request, BTS-APP first measures the PING latency from the user client to a subset of its deployed test servers, so as to find a nearby server with the lowest latency. Then, during the actual bandwidth testing process, it continuously downloads large files from the selected server via HTTP connections to probe the access bandwidth for 10 seconds, and acquires a bandwidth sample every 50 milliseconds in the meantime (therefore generating a total of 200 samples). Here the probing duration (10 seconds) is shorter than that of Speedtest (15 seconds) because almost all of BTS-APP's user requests come from Mainland China (meaning shorter RTTs for data transmission). In order to ensure that the user-side access bandwidth is fully saturated, BTS-APP progressively sets up new HTTP connections to other nearby test servers, if the latest bandwidth sample reaches a predefined threshold (*i.e.,* 25 Mbps, 35 Mbps, and so on, following Speedtest's design).

To produce the test result, BTS-APP first partitions the collected bandwidth samples into 20 groups, each containing 10 samples. Then, to address the noises caused by TCP slow start and network dynamics, it discards 5 groups (of samples) with the lowest average bandwidth and 2 groups with the highest. The remaining groups' average bandwidth is used as the final result. All the empirical parameters used in this stage conform to those of Speedtest [27], whose robustness has been extensively evaluated in the real world.

BTS-APP's current infrastructure consists of 352 test servers distributed across Mainland China, whose bandwidths range from 1 Gbps to 10 Gbps. In particular, 62 of the test servers are directly provided by ISPs through commercial negotiations, which are close to the Internet backbone networks (IXPs) and thus are especially high-speed. In each test, 5 (out of the 352) geographically nearby (determined by the IP addresses) servers are PINGed to find the nearest server; in contrast, 10 out of the 16,190 servers are PINGed in Speedtest. This seemingly "degraded" configuration is acceptable in practice, as it can well handle the present workload of BTS-APP (serving ~0.2M user requests per day generally issued from Mainland China) without harming the test accuracy.

Fine-Grained Data Collection. Despite being able to provide reliable and accurate bandwidth testing service in the past 7 years or so, BTS-APP cannot give an in-depth analysis of its results, making it hard to understand the root causes of undesirable access bandwidths. This is because although BTS-APP is an Android app, the implementation of its bandwidth testing logic is mostly web-based (similar to other BTSes introduced in §1). Thereby, BTS-APP cannot capture critical underlying network information (such as frequency band and signal strength) at the client side for in-depth performance analysis and troubleshooting. To overcome the shortcoming, we build a data collection module for BTS-APP to capture fine-grained data.

In order to run on heterogeneous mobile devices at scale, such a module needs to be both *lightweight* and *privacy-preserving*. For the former, we resort to passive monitoring of critical PHY- and MAC-layer information during a bandwidth test using generic Android APIs. For the latter, we carefully avoid any data collection that would require additional privileges the original app does not possess, so as to minimize users' privacy concerns. Concretely, with respect to cellular networks, we focus on user device-side signal conditions (*e.g.,* signal strength and signal-to-noise ratio), as well as base station (BS)-side connection information (*e.g.,* BS ID, frequency band, and channel number). Regarding WiFi networks, we are interested in the connected access points' capabilities (*e.g.,* WiFi standard, radio frequency, and MAC-layer transmission speed), as well as the local network status (*e.g.,* states of the other WiFi APs that are detected).

The above module is implemented as a small-size (1K lines of code and 110 KB binary) plugin that BTS-APP can dynamically load at runtime. During a bandwidth test, the plugin carries out data collection every second, which incurs negligible (≤2%) CPU and (≤1 MB) memory overhead on even a low-end phone. After the test, the result and the collected data are uploaded via WiFi (whenever possible) to our data server for subsequent detailed analysis.

Crowdsourcing and Ethical Considerations. Thanks to BTS-APP's development team, we manage to deploy the plugin on almost all of BTS-APP's users (the remaining minority of users choose to opt out). From Aug. 1st to Nov. 30th in 2021, a total of 3.54M users perform 23.6M access bandwidth tests. None of our measurements violate BTS-APP's user agreements. The users involved in this study opted in with informed user consent, and the analysis is conducted under a well-established IRB. During the study, no personally identifiable information was collected, and we have no way of linking the data with users' actual identities.

3 MEASUREMENT FINDINGS

In this section, we first present the general statistics from our measurement (§3.1), and then zoom in on the respective bandwidth characteristics of 4G (§3.2), 5G (§3.3), and WiFi (§3.4) in terms of both technical and non-technical factors.

3.1 General Statistics

During our four-month measurement, 3,542,179 user devices conducted 23,636,352 access bandwidth tests, 99.97% of which are located in China, involving four mobile ISPs, 2,041,586 BSes, and 4,473,362 WiFi APs. More specifically, we record the results of 21,051 3G tests, 1,632,616 4G tests, 905,471 5G tests, and 21,077,214 WiFi tests, along with the cross-layer, in-situ network information as discussed in §2. To enable the longitudinal analysis, we also refer to BTS-APP's measurement reports in recent years when necessary.

Bandwidth Variation over Time. As the commercial prosperity of 5G and WiFi 6 commenced in late 2019, mobile access bandwidth was expected to grow constantly in the subsequent years (2020 and 2021) in response to the increasing deployment of 5G BSes and WiFi 6 APs. Nevertheless, we find that the average WiFi bandwidth remains mostly unchanged: 132 Mbps in 2020 *vs.* 137 Mbps in 2021. More surprisingly, the average 4G bandwidth decreases from

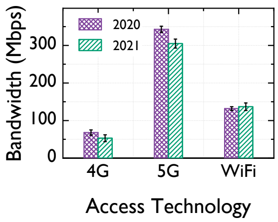

Figure 1: Avg. 4G/5G/WiFi bandwidth over time.

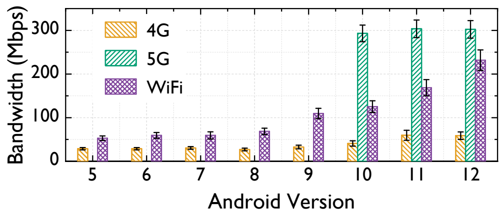

Figure 2: Average 4G, 5G and WiFi bandwidth for different OSes.

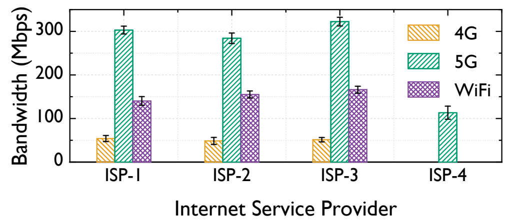

Figure 3: Average 4G, 5G and WiFi bandwidth for different ISPs.

68 Mbps in 2020 to 53 Mbps in 2021, and the average 5G bandwidth drops from 343 Mbps in 2020 to 305 Mbps in 2021. If it is any consolation, the average overall cellular bandwidth (taking 2G/3G/4G/5G all into consideration) increases from 117 Mbps in 2020 to 135 Mbps in 2021, which is expected because the user percentage of 5G almost doubled in 2021 (33%) as compared to that in 2020 (17%).

Furthermore, we closely examine the bandwidth variation with regard to the same user group (that belong to the same ISP in the same city), including those China Unicom, China Mobile, and China Telecom users in Beijing, Shanghai, Guangzhou, and Shenzhen. We also observe declines in average 4G and 5G bandwidths for the same user group, which are 12%–31% and 5%–23%, respectively.

The above findings reveal that in real-world deployment, the advance in wireless technologies is far from being fully exploited. In particular for cellular access, the QoS for the majority of users (*i.e.,* 4G users) is in fact damaged despite the well expected improvement of the "average overall" QoS. This, in our opinion, is unknown and hardly acceptable to 4G users, and thus may hurt users' confidence and do harm to the mobile ecosystem. Worse still, even 5G users who are prioritized are experiencing deteriorated QoS. We will investigate the undesirable situations in the remainder of this section from cross-technology and cross-layer perspectives.

Spatial Disparity. We further examine the bandwidth variation across different cities in China during our measurement period (Aug. to Nov. 2021), including 21 mega cities, 51 medium cities, and 254 small cities. In general, there is noticeable difference among the access bandwidths of 4G (28–119 Mbps), 5G (113–428 Mbps), and WiFi (83–256 Mbps) with regard to these cities. A mega city (such as Guangzhou) does not necessarily possess high 4G, 5G, and WiFi bandwidths (55 Mbps, 301 Mbps, and 136 Mbps, respectively) even with dense infrastructure deployment, probably due to the severe network resource contention among plenty of users. Besides, 41% cities are subject to unbalanced development of 4G and 5G networks; for example, Shanghai has higher 5G bandwidth (337 Mbps) as compared to the national average (305 Mbps), while its 4G bandwidth (48 Mbps) is 9% lower than the national average. On average, the 4G and 5G access bandwidth in urban areas is 24% and 33% higher than that in the rural areas of the same cities, respectively, mostly owing to their distinct densities of infrastructure deployment.

User-side Hardware and Software. We also study the impact of user-side hardware and software (*i.e.,* the Android system that actually manages the wireless data connectivity) on the access bandwidth. In our dataset, there are 191 mobile phone vendors and 2,381 device models whose hardware configurations vary from rather low-end to very high-end. At first glance, it appears that mobile access bandwidth is in general positively correlated with the superiority of hardware. Closer examination, however, indicates that this is merely a common illusion caused by missing a key factor at play—software that bridges the hardware and mobile access networks.

Figure 2 lists the average 4G, 5G and WiFi bandwidth for different Android versions, illustrating that it might well be the Android version that essentially determines the access bandwidth (in a statistical sense). This is quite understandable in principle, given the considerable improvements made in the cellular/WiFi management modules by higher-version Android systems. In contrast, when a low-end device model and a high-end device model are equipped with the same Android version, usually we do not observe obvious difference in mobile access bandwidth between them—the standard deviation for the same access technology is $\leq$23 Mbps. Consequently, the fact that higher-end mobile phones often (but do not necessarily) possess higher access bandwidths is only because they have more up-to-date hardware that is more often used for running higher-version OSes.

ISP-side Infrastructure Investment. Our study involves all the four major ISPs in China: China Mobile, China Unicom, China Telecom, and China Broadcast Network, who provide both cellular and fixed broadband services for Internet users. They are referred to as ISP-1, ISP-2, ISP-3, and ISP-4 henceforth. Figure 3 presents their average 4G, 5G and WiFi bandwidths. As shown, while their average 4G bandwidths are quite similar (probably owing to their wide deployment of mature and similar 4G infrastructure), there is noticeable difference among their average 5G bandwidths. In particular, as a newly-founded ISP that focuses on 5G, ISP-4 bears obviously lower 5G bandwidth, since its 5G service is based on a special low-bandwidth 700 MHz band originally designated for 4G and radio broadcast services. In other words, ISP-4 is trading bandwidth for low-cost deployment.

We also note that ISP-3 outperforms ISP-1 and ISP-2 in both 5G and WiFi bandwidths. The former is because ISP-3 deploys 5G mostly on an advantageous frequency range of a dedicated 3 GHz band (detailed in §3.3). The latter is due to ISP-3's heavier investment in its fixed broadband infrastructure (detailed in §3.4).

3.2 4G (LTE) Access Bandwidth

As illustrated in Figure 4, the average 4G access bandwidth in our measurement is only 53 Mbps, which is far below the ISPs' claimed bandwidth limit (*i.e.,* 300 Mbps). According to BTS-APP's measurement reports, this is even 22% lower than that in 2020. While in the

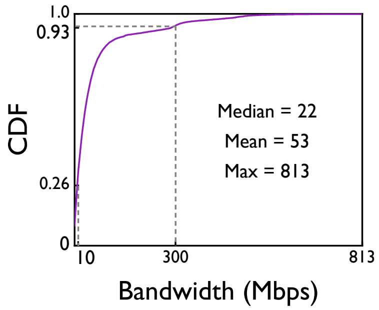

Figure 4: Bandwidth distribution for 4G access in our measurement.

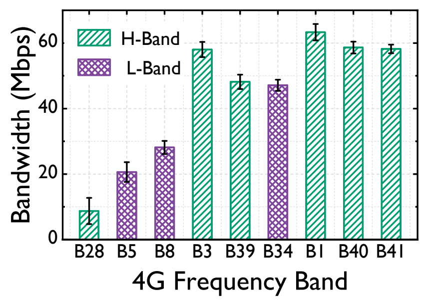

Figure 5: Average access bandwidth of each LTE band involved in our study.

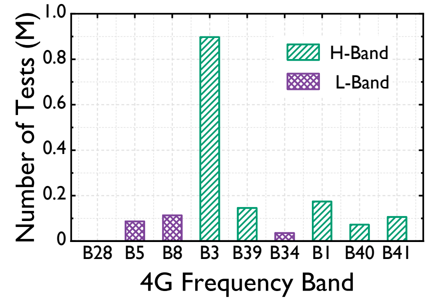

Figure 6: Number of access bandwidth tests conducted on each LTE band.

Table 1: The nine LTE bands involved in our study, ordered by their downlink (DL) spectrum.

Band	DL Spectrum	Max Channel Bandwidth	ISPs
Band 28	758 – 803 MHz	20 MHz	ISP-4
Band 5	869 – 894 MHz	10 MHz	ISP-3
Band 8	925 – 960 MHz	10 MHz	ISP-1, 2
Band 3	1805 – 1880 MHz	20 MHz	ISP-1, 2, 3
Band 39	1880 – 1920 MHz	20 MHz	ISP-1
Band 34	2010 – 2025 MHz	15 MHz	ISP-1
Band 1	2110 – 2170 MHz	20 MHz	ISP-2, 3
Band 40	2300 – 2400 MHz	20 MHz	ISP-1
Band 41	2496 – 2690 MHz	20 MHz	ISP-1

top 6.8% of tests the bandwidth exceeds 300 Mbps, in over a quarter (26.3%) of tests the result is below 10 Mbps. In this part we explain the above phenomena by delving into the radio characteristics of LTE, the migration of radio resources from LTE to 5G, and the deployment of the novel LTE-Advanced technology.

Radio Characteristics. Frequency range (*a.k.a,* spectrum) and channel bandwidth are among the key radio characteristics that determine the performance of cellular access. Each LTE band is unique in the two characteristics. In theory, lower-frequency bands have less signal propagation loss, and thus can bring better radio coverage and signal-to-noise ratio (SNR). On the other hand, channel bandwidth has a more direct impact on the access bandwidth—the limit of access bandwidth linearly grows as the maximum channel bandwidth increases, as dictated by the Shannon-Hartley theorem [64]. Given the above theoretical radio features of different bands, we are particularly interested in their actual impact on the access bandwidth.

We have captured all the nine LTE bands used in China, referred to as Band 1, 3, 5, 8, 28, 34, 39, 40 and 41 following 3GPP's definition [1]. Table 1 lists each band's downlink spectrum (recall that our study concentrates on the download bandwidths of mobile devices), maximum supported channel bandwidth, and corresponding ISP(s)—note that one band can be multiplexed by multiple ISPs. According to 3GPP's LTE specifications [1], the channel bandwidth should reach 20 MHz to realize the theoretical bandwidth limit of 4G access, so we denote the bands that support the 20 MHz channel bandwidth as *high-bandwidth* bands (H-Bands for short), and the others as *low-bandwidth* bands (L-Bands for short).

Figure 5 lists the average access bandwidths of the nine LTE bands. Note that Band 28, which is assigned to the 5G-first ISP-4, was only used in two LTE bandwidth tests (see Figure 6) so its result is highly biased here. Not surprisingly, H-Bands (except Band 28) yield higher access bandwidths than L-Bands. However, the average bandwidth of Band 39 is as low as 48.2 Mbps, even close to that (47.1 Mbps) of Band 34 which is an L-Band. This is because Band 39 is dedicated to serving rural areas where LTE BSes are sparsely deployed [32]. In comparison, Band 40 is used for penetrating indoor environments where LTE BSes are usually densely deployed, and thus offers better signal strength—an average of -88 dBm for Band 40 *vs.* -94 dBm for Band 39. These special purposes explain the low correlation between spectrum and access bandwidth for certain bands as shown in Figure 5.

Radio Resource Migration. Since H-Bands are superior to L-Bands in terms of access bandwidth, most mobile users should be served by H-Bands, which are reflected on Figure 6, where the majority (85.6%) of LTE bandwidth tests are conducted on H-Bands. In particular, Band 3 alone serves 55% tests. More specifically, for all the three ISPs (ISP-1, ISP-2 and ISP-3) that deploy LTE on Band 3, the percentage of Band-3 LTE bandwidth tests is the highest among their used bands, *i.e.,* 31%, 63% and 76% respectively.

We attribute this skewed workload distribution to the recent migration of radio resources from other LTE H-Bands to 5G. In early 2021, a large portion of LTE H-Band spectrum was "refarmed" for 5G usage [56]; the affected bands include Band 1, Band 28, and Band 41, which in together occupy 58.2% of the entire H-Band spectrum. Such an aggressive migration constitutes an important cause of the sharp decrease in LTE access bandwidth from 2020 to 2021 (as mentioned in §3.1). In detail, the average bandwidths of the refarmed Band 1 (63 Mbps) and Band 41 (58 Mbps) have fallen below the average LTE bandwidth in 2020 (68 Mbps).

LTE-Advanced Deployment. As mentioned in the beginning of this part, although the average LTE bandwidth is rather low (53 Mbps), we do observe that in 6.8% LTE bandwidth tests the result is higher than 300 Mbps, averaging at 403 Mbps and peaking at 813 Mbps. A closer examination reveals that the majority of these tests are performed alongside urban main roads, where ISPs deploy the LTE-Advanced [18] technology for the nearby LTE BSes (termed eNodeBs) to deal with the large traffic volume. LTE-Advanced makes significant improvements on conventional LTE

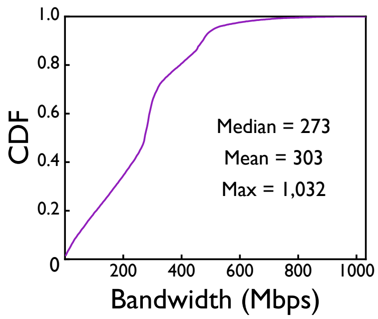

Figure 7: Bandwidth distribution for 5G access in our measurement.

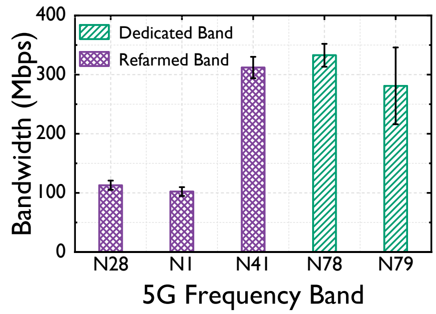

Figure 8: Average access bandwidth of each 5G band involved in our study.

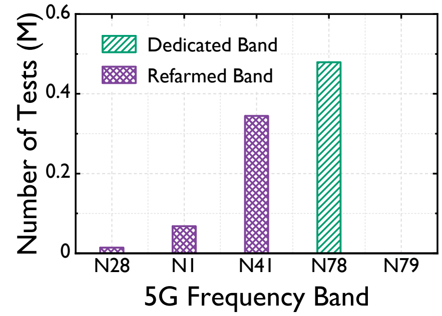

Figure 9: Number of access bandwidth tests conducted on each 5G band.

Table 2: The five 5G bands involved in our study, ordered by their downlink (DL) spectrum.

Band	DL Spectrum	Max Channel Bandwidth	ISPs
N28	758 – 803 MHz	20 MHz	ISP-4
N1	2110 – 2170 MHz	20 MHz	ISP-2, 3
N41	2496 – 2690 MHz	100 MHz	ISP-1
N78	3300 – 3800 MHz	100 MHz	ISP-2, 3
N79	4400 – 5000 MHz	100 MHz	ISP-1, 4

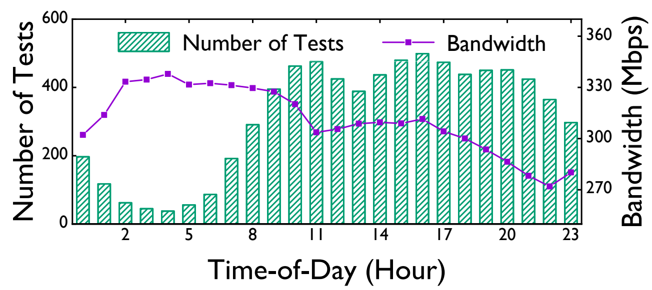

Figure 10: Number of 5G tests and average 5G bandwidth in different times of a typical day.

bandwidth (which can only reach 150 Mbps) through a suite of innovations such as carrier aggregation, multi-antenna technology, enhanced MIMO and mobility. As a result, LTE-Advanced can achieve up to 2 Gbps bandwidth, comparable to the bandwidth of today's commercial 5G. More importantly, LTE-Advanced is technically mature, easy-to-deploy, and cost-effective.

3.3 5G Access Bandwidth

As the state-of-the-art cellular technology, 5G can offer up to 20 Gbps access bandwidth along with ultra-low latency (*e.g.,* 5 ms) and ultra-high service capacity (*e.g.,* 1M devices per square kilometer). Over the past two years, ISPs have made enormous investments on 5G's infrastructure and commercial promotion. Particularly, as revealed in §3.2, even 4G's infrastructure (radio spectrum) has been refarmed to this end. Nevertheless, Figure 7 shows an undesired outcome of the above efforts: the average 5G bandwidth is 303 Mbps, which has decreased by 11% as compared to that in 2020 (according to BTS-APP's measurement reports). To demystify this "lose-lose" situation of today's cellular ecosystem (dominated by 4G and 5G), we next examine in depth the key factors that lead to the dilemma, regarding spectrum refarming and received signal strength.

Spectrum Refarming. As shown in Table 2, five bands are used by the four ISPs for 5G deployment in China, dubbed N1, N28, N41, N78 and N79 according to 3GPP's specifications [2]. All these bands are sub-6 GHz and three of them (N1, N28 and N41) are in fact refarmed from the three LTE bands (Band 1, Band 28 and Band 41) respectively. N78 and N79, on the other hand, are dedicated to 5G usage, among which N78 is the core band that provides most of 5G's service capacity[2] while N79 is still under test deployment. There are only three N79-related tests in our measurement, so we will exclude N79 from our analysis to avoid bias. We list the average access bandwidth of each 5G band in Figure 8, and the number of access bandwidth tests conducted on each band in Figure 9.

As shown, there exists a significant discrepancy among the average bandwidths of the three refarmed bands. Specifically, the average 5G bandwidth on N41 is 312 Mbps, which is comparable to that of 5G's core band N78 (332 Mbps). In contrast, the results on the other two refarmed bands (N1 and N28) are much lower, *i.e.,* 103 Mbps and 113 Mbps. A deeper investigation clears the mystery—a 100-MHz contiguous spectrum (2515–2615 MHz) from Band 41 has been refarmed into N41, which is quite wide to support relatively high bandwidth. In contrast, the refarmed contiguous spectrum from Band 1 and Band 28 is rather thin (*i.e.,* 60 MHz and 45 MHz), leading to undesirable bandwidth. Thus, we conclude that refarming is a major contributor to the decline of 5G's average access bandwidth.

Diurnal Pattern. We also examine the number of 5G tests and the average 5G access bandwidth at different times of the days during our measurement period (Aug. to Nov. 2021). Figure 10 demonstrates the data collected in a typical day. We observe that in most cases, the average 5G bandwidth is negatively correlated with the number of tests. This is because more bandwidth tests performed usually

[2] ISP-3 uses lower-frequency spectrum in N78, offering wider coverage while not sacrificing bandwidth, so has higher signal strength and bandwidth.

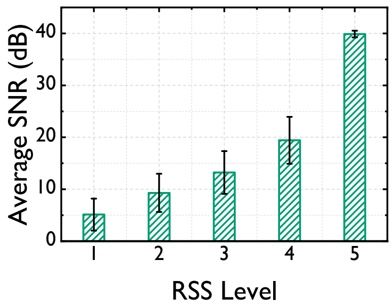

Figure 11: Correlation between 5G RSS level and the average SNR.

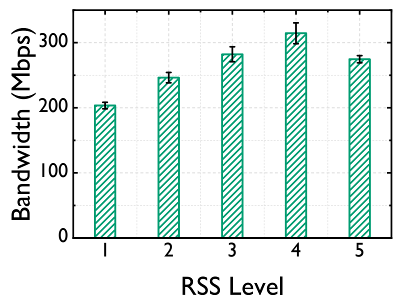

Figure 12: Correlation between 5G RSS level and average bandwidth.

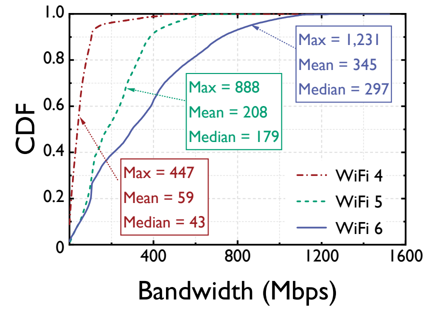

Figure 13: Bandwidth distribution for WiFi access in our measurement.

indicate that more users are sharing the access network, leading to heavier workloads and resource contention on the BSes.

Nevertheless, we find that the average bandwidth hits the bottom (276 Mbps) between 21:00 and 23:00, during which the number of tests is as small as 362 per hour. In contrast, even with 25% more tests performed per hour from 15:00 to 17:00, the average bandwidth in that time period is 10% higher (308 Mbps). Deeper investigations show that the above phenomenon stems from the sleeping strategy of 5G BSes, in which ISPs selectively turn off the active antenna processing units of 5G BSes from 21:00 to 9:00 to reduce energy consumption [16, 50, 61]. Notably, we observe that despite the sleeping strategy, the average bandwidth in fact reaches the peak (334 Mbps) between 3:00 and 5:00, since very few people are using the network during this period (46 tests per hour).

In comparison, for 4G networks, we find that the average bandwidth at different times of the days is in general positively correlated with the number of tests conducted by users. This is because an LTE BS consumes much less energy and thus does not adopt the sleeping strategy of 5G BSes.

Received Signal Strength (RSS). In common sense, an excellent RSS usually implies a higher SNR, and hence a higher access bandwidth [64]. While our data show that RSS and SNR are indeed positively correlated (Figure 11), a counter-intuitive finding is that RSS and 5G access bandwidth are not. Figure 12 clearly depicts that as the RSS rises from level-1 to level-4, the average 5G bandwidth monotonously grows from 204 Mbps to 314 Mbps. However, when the RSS becomes excellent (level 5), the average 5G bandwidth sharply drops below that with level-3 and level-4 RSS. The situation is similar when we examine the median 5G bandwidth.

To understand the above, we notice that the abovementioned excellent-RSS 5G bandwidth tests are mostly performed in crowded urban areas, where 5G BSes in close proximity tend to yield consistently low bandwidth. Heavy population in such areas often requires dense deployment of 5G BSes (termed gNodeBs) [54]. Although this can provide higher signal strength, improper gNodeB placement and antenna configurations can easily lead to cross-region coverage [3], *i.e.,* overlaps of different gNodeBs' signal coverage, which can aggravate the already complex multi-path and co-channel interference [49, 75] in urban areas with dense buildings, as well as the various load balancing issues and poor handover problems [17, 48, 65]. This may especially be the case given that current 5G technology and deployment are rather immature. In comparison, we do not observe such a phenomenon on 4G access, given its much more mature, well-provisioned infrastructure deployed for 10+ years.

3.4 WiFi Access Bandwidth

As another widely-deployed mobile access technology, WiFi mainly works in home and enterprise environments. In this part, we dig into the access bandwidth of WiFi across its 4th, 5th, and 6th generations of technical standards.

In our dataset, WiFi 4, 5 and 6 account for 57.2%, 31.3% and 11.5 % of the WiFi bandwidth tests, respectively. Figure 13 depicts their bandwidth distributions. With the evolution of WiFi technologies (4→5→6), the average bandwidth appears to substantially increase (59 Mbps→203 Mbps→345 Mbps). In more detail, given that WiFi 5 only uses the 5 GHz band, we look at the 2.4 GHz and 5 GHz WiFi bands separately (see Figures 14 and 15). We are surprised to find that the average bandwidths of WiFi 4 and WiFi 5 are in fact fairly close over the 5 GHz band—195 Mbps *vs.* 208 Mbps. This suggests that the overall bandwidth improvement from WiFi 4 to WiFi 5 is mostly because WiFi 4 users are also using the 2.4 GHz band, rather than benefiting from the technical advances introduced in WiFi 5, such as beamforming and downlink multi-user MIMO.

Delving deeper into the bandwidth distribution of WiFi 4/5/6, we notice that for each generation, WiFi bandwidths tend to cluster around certain 100× values, *e.g.,* 100 Mbps, 300 Mbps, and 500 Mbps for WiFi 5 (as shown in Figure 16). Interestingly, we find that these 100× values well match the promised bandwidths of ISPs' typical fixed broadband plans [53, 69, 70]. In a sense, they reflect the distribution of WiFi users' purchased fixed broadband plans. Based on this heuristic, the fixed broadband plans of ISPs, and other public reports on the bandwidth distribution of fixed broadband in China [7, 8], we can now roughly infer that ~64% of the WiFi users are still using ≤200-Mbps fixed "broadband" Internet access. Consequently, the technical advantages of WiFi 5 are in fact largely offset by the tardy evolution of wired Internet access.

For WiFi 6 that manifests the highest average bandwidth, there are fewer (~39%) users using the ≤200-Mbps fixed broadband, indicating that WiFi 6 users are more likely to live in urban areas (shown by their IP addresses) where wired broadband infrastructure evolves more quickly. In particular, we notice that for an ISP (ISP-3) that has made heavy investments in its fixed broadband infrastructure,

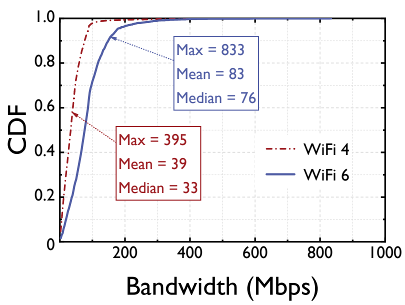

Figure 14: Bandwidth distribution for WiFi access using the 2.4 GHz band.

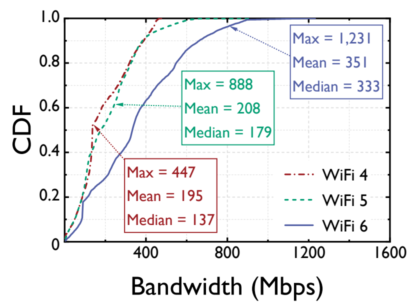

Figure 15: Bandwidth distribution for WiFi access using the 5 GHz band.

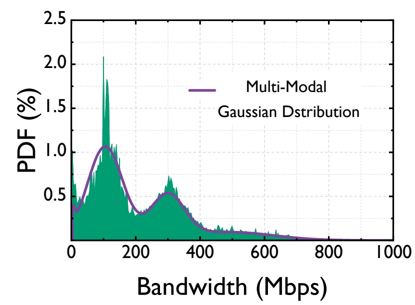

Figure 16: Probability distribution for WiFi 5 access bandwidths in our study.

the corresponding WiFi access bandwidth is also the highest among the studied four ISPs. Nevertheless, the average bandwidth of WiFi 6 is still far below its advertised capability, leading to significant under-utilization of WiFi 6's superiority.

4 IMPLICATIONS

The analysis results of 4G, 5G, and WiFi 4/5/6 in §3 depict a complete picture of today's mobile access bandwidth. On one hand, we witness the power of emerging mobile access technologies such as 5G and WiFi 6—in certain cases, they do yield very high bandwidth that is hardly achievable by their predecessors. On the other hand, we note that traditional technologies (*e.g.,* 4G LTE and WiFi 4/5) are still serving the majority of mobile users and bearing more mature deployment and stable performance. Most importantly, the developments of these technologies are closely intertwined (*e.g.,* 4G and 5G compete for limited radio resources), forming an intricate and complex ecosystem.

Our study illustrates the aggressive migration of radio resources from 4G to 5G, and quantitatively reveals its side effect on both 4G and 5G access bandwidths. Although spectrum refarming is usually considered inevitable during the evolution of cellular technologies, it should be carried out in a moderate and strategic manner so as to avoid or minimize the side effect. In particular, one should pay special attention to the fact that the current LTE spectrum resources are severely fragmented, which can be ascribed to two major reasons. First, the spectrum is often statically segmented among different ISPs and regions [57], with necessary spectrum spacing (called *guard bands* [68]) between adjacent bands to prevent signal interference [12]. Second, as different mobile telecommunication technologies (*e.g.,* CDMA, WCDMA, GSM, LTE, and NR) often work in a same band with non-overlapping spectrum [41], their heterogeneous requirements in channel bandwidth (*e.g.,* CDMA needs 1.25 MHz bandwidth while LTE needs 20 MHz bandwidth) can easily fragment the spectrum.

As a result, few of the LTE bands can provide sufficient contiguous spectrum for refarming, while 5G usually requires nearly 100 MHz contiguous spectrum to enable a high data rate and low signal interference among user devices. We therefore advocate more effective band defragmentation and refarming strategies, *e.g.,* dynamic spectrum allocation [60] and flexible band trading [43], to facilitate better utilization of spectrum resources given the fast evolution of cellular networks.

Since 4G and 5G will coexist for a very long time [19, 20], our findings also call for strengthening existing LTE infrastructure in a cost-effective manner. For instance, we suggest widening the deployment of the LTE-Advanced technology, which has yielded up to 813 Mbps bandwidth in our measurement (comparable to the typical bandwidth of commercial 5G). Moreover, LTE-Advanced's carrier aggregation feature can help combine non-contiguous channels (*i.e.,* signal carriers) among fragmented bands into a single wide channel to realize high data rate [78], leading to effective mitigation of the spectrum fragmentation and improvement to the effect of refarming.

Our findings also bring implications for other players in the mobile ecosystem. For example, in §3.4 we notice that slow broadband access can largely retard the access bandwidth of WiFi 4/5/6, so wired ISPs and content providers should take into account the emerging wireless technologies when budgeting their network infrastructure. Moreover, we encourage customers to be rational when being bombarded with ISPs' and phone vendors' 5G advertisement campaigns. They should be informed of 5G's actual performance in everyday usage, and that up-to-date mobile OS (rather than solely superior hardware) is crucial to the access bandwidth of a mobile device, as we discover in §3.1.

5 FAST AND LIGHT BANDWIDTH TESTING SERVICE AT SCALE

As a fundamental tool for understanding end users' Internet access bandwidths, bandwidth testing services (BTSes) not only facilitate large-scale characterization of network performance in the wild, but also serve as a core component of many emerging bandwidth-hungry applications (*e.g.,* UHD streaming and AR/VR). However, mainstream BTSes typically follow the "probing by flooding" approach, which brings heavy burden on end users in terms of both test time and data usage, as well as on service providers in terms of backend server cost (*cf.* §2). Worse still, these issues have aggravated in recent years given the advent of high-bandwidth 5G and WiFi 6 networks, raising considerable concerns from BTS-APP's operation team.

To tackle this, in this section we use our measurement insights to rearchitect the BTS-APP system, so that it can be much faster and lighter to accommodate the upcoming heavier workload. We first take a data-driven approach to optimizing the bandwidth probing

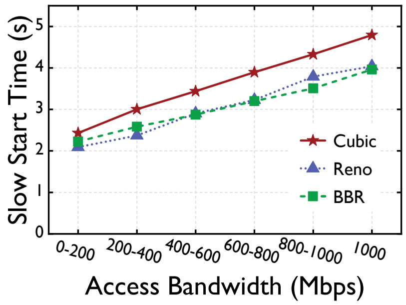

Figure 17: TCP slow start time for different congestion control algorithms.

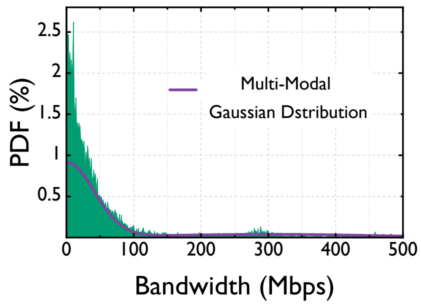

Figure 18: Probability distribution for 4G access bandwidths.

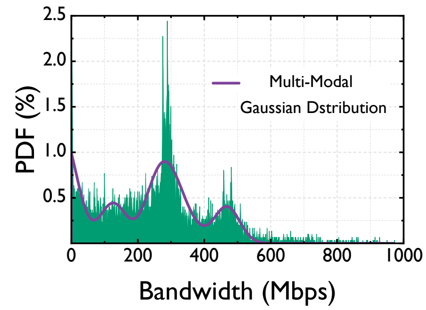

Figure 19: Probability distribution for 5G access bandwidths.

process (§5.1), and then devise a principled method to achieve cost-effective server deployment (§5.2). Afterwards, we conduct real-world implementation and evaluation for the new design (§5.3).

5.1 Data-Driven Bandwidth Probing

Recall in §2 that upon a bandwidth test request, BTS-APP probes the user's access bandwidth for 10 seconds by downloading large files from a test server (more servers will be added if necessary), and collects bandwidth samples in the meantime to generate the final result. Through the in-depth measurement, we identify important optimization opportunities for such a bandwidth probing process, and adjust the bandwidth testing logic accordingly.

Negative Effect of TCP Slow Start. When examining the recorded bandwidth samples of BTS-APP, we notice that as the user's access bandwidth increases, TCP slow start takes longer time in a bandwidth test. As TCP slow start is in fact introduced by TCP congestion control, we quantitatively investigate the effect of mainstream congestion control algorithms (*i.e.,* Cubic [30], Reno [11], and BBR [9]) on 15 test servers randomly picked from BTS-APP's server pool for two weeks after the measurement study. We manually configure the congestion control kernel modules on these servers, and monitor the duration of slow start with `tcp_probe`. For user devices, we employ 10 mobile phones with diverse hardware configurations, Android versions, and mobile access technologies (which turn out to have little impact on this experiment).

As shown in Figure 17, Cubic obviously incurs longer slow start time, while BBR behaves a little better than Reno. Even with the emerging BBR algorithm, TCP slow start takes an average of around 2 seconds and 4 seconds during the access bandwidth tests of 100-Mbps and 1-Gbps mobile networks, respectively, accounting for a significant fraction of the total test time (10 seconds). However, the bandwidth samples generated during the slow start phase are in fact "noises" that do not contribute useful information to the final result. To make matters worse, they can sometimes affect the accuracy of test results if not properly filtered, particularly when they can be triggered by spurious packet losses during transmission (common in cellular networks [72]).

Solutions of Existing BTSes. To combat network noises introduced by TCP slow start and common network fluctuations, today's BTSes employ diverse methods at different stages (including server selection, bandwidth probing, and bandwidth estimation) of a bandwidth test. First, during server selection, existing BTSes often choose test servers with the lowest latencies to the client to reduce network noises. Then, during the bandwidth probing stage, some BTSes (*e.g.,* Speedtest [27] and FAST [21]) perform large data transfers using parallel connections to saturate the client's bandwidth as quickly as possible. Finally, when performing bandwidth estimation using the bandwidth samples collected during the probing stage, existing BTSes use dedicated algorithms to rule out noise samples and generate the final result. For instance, Speedtest [27] adopts a static algorithm to filter out the top 10% and bottom 25% bandwidth samples, and then averages the remaining ones as the final result; in comparison, FastBTS [77] uses the crucial interval-based sampling algorithm to calculate the bandwidth interval with the highest *concentration*, *i.e.,* the product of sample density and quantity.

All the above efforts can mitigate the impact of network noises in practice. However, they cannot reduce the long duration caused by TCP slow start, especially under the network environments with a high bandwidth-delay product. For example, the slow start time may increase to ~6 seconds for 10 Gbps networks with 100 ms RTT.

Bandwidth Probing with Statistical Guidance. Essentially, the long duration of TCP slow start stems from its gradual, cautious probing of a user's access bandwidth. To address this, our key finding is that for a given access technology, its access bandwidth (X) in fact follows a multi-modal Gaussian distribution

$$\mathbb{P}(X) = \sum_{i=1}^{k} w_i \mathcal{N}(X|\boldsymbol{\mu_i}, \boldsymbol{\sigma_i}), \tag{1}$$

which means that the occurrence probability of a certain bandwidth value is a weighted (w_i) combination of several independent Gaussian distributions ($\mathcal{N}(X|\boldsymbol{\mu_i}, \boldsymbol{\sigma_i})$). For example, in Figure 18 and Figure 19, 4G and 5G networks' bandwidth probability distributions can both be described as multi-modal Gaussian distributions. Here one "mode" (*i.e.,* $\boldsymbol{\mu_i}$ of an individual Gaussian distribution) manifests as a peak in the bandwidth probability distribution. Similar situations are also identified for each WiFi technical standards, *e.g.,* WiFi 5 as shown in Figure 16.

It is worth noting that such statistical bandwidth distribution patterns of different mobile access technologies are not coincidences. As a matter of fact, they are produced by the joint impact of different access technologies' bandwidth limits, infrastructure status, and ISPs' data plans, *e.g.,* due to ISPs' fixed broadband plans in the case of WiFi 5 (refer to §3.4). More importantly, we observe that these factors and the resulting distributions are quite stable on a moderate time scale (*e.g.,* within a month). Therefore, by updating the statistical model periodically, we can leverage it to guide the selection of the initial data rate for bandwidth probing, thus avoiding the lengthy ramp-up in TCP slow start.

Upon a bandwidth test request with respect to a certain mobile access technology, we set the initial probing data rate as the most probable bandwidth (*i.e.,* the most significant mode) of the access technology, based on its multi-modal probability distribution of access bandwidths. Then, we select a reasonable number of nearby test servers (with the lowest PING latencies among all the test servers) as the actual test servers, whose total uplink bandwidth slightly exceeds the initial probing data rate (since the uplink bandwidth of a test server is typically an integral multiple of 100 Mbps).

During the actual probing process, we acquire a bandwidth sample every 50 milliseconds (as the original BTS-APP does), and determine whether the client's access bandwidth is fully saturated by examining whether the latest bandwidth sample falls below the current probing data rate. If not (saturated), we further tune the probing data rate to one of the larger "modal" bandwidth values (more servers will be added if necessary)—similarly, we use the most probable one among these larger "modal" bandwidth values as the next probing data rate. Otherwise, we keep the current probing data rate. Finally, if the latest ten bandwidth samples converge, we stop the test and calculate the mean of the ten samples as the final test result. Here we regard the samples as convergent if the difference ratio between their maximum and minimum values is ≤3%, following the design of FAST [21] (also a state-of-the-art BTS).

To instantiate the above design under large-scale scenarios, we alter the transmission protocol from TCP to UDP, which is practically feasible with the support of BTS-APP. Thereby, we can implement the customized bandwidth probing mechanism from scratch at the application layer without tampering the kernel network stack. Further, while our more aggressive bandwidth probing mechanism may raise the concern of network fairness, our network flows are in fact very short-lived (∼1 second as to be shown in §5.3). In addition, wireless networks have separate mechanisms for ensuring fairness at lower layers (*e.g.,* proportional-fair scheduling performed by BSes [40]). Thus, we believe that fairness should not be a concern in our context.

5.2 Cost-Effective Server Deployment

As introduced in §1, due to the rapid increase of mobile access bandwidth, the BTS infrastructure cost has become a key concern of BTS-APP's operation team. To address this, we carefully analyze the workload traces of BTS-APP's test servers. We observe that in most (98%) time, the required bandwidth (*i.e.,* the aggregated bandwidth of all the users who are running tests at a same time) for bandwidth testing does not reach even 5% of the total available bandwidth of BTS-APP's 352 test servers.

The rationale behind BTS-APP's excessive server deployment is to provide low latency for geographically distributed users. However, some state-of-the-art BTSes (*e.g.,* FastBTS) have demonstrated that low latency enabled by nearby servers is in fact not necessary for accurate bandwidth testing [77]. Recall from §5 that our proposed approach uses UDP to avoid the slow start phase, and thus is not sensitive to latency. This implies severe resource over-provisioning in BTS-APP's current Speedtest-like architecture. Therefore, we can instead strategically deploy a much smaller number of geo-distributed, budget servers to properly accommodate the expected bandwidth testing workload with margins. The workload can be practically estimated by jointly considering recent user scale and their access bandwidths reflected in our data.

To realize this goal, another challenge is the selection from heterogeneous server purchase plans, involving a variety of (egress) bandwidth configurations, numbers of available servers, and the sale prices across different server providers, particularly the virtual machine (VM) server providers such as Amazon EC2, Aliyun ECS, and OneProvider (the infrastructure provider of Speedtest). For example, on OneProvider (as of Jan. 2022) we can find 336 server configurations for purchase, with bandwidth ranging from 100 Mbps to 10 Gbps and price lying between \$10.41/month and \$2609/month; also, these configurations have different numbers of servers available.

To make a cost-effective server purchase plan, for each server configuration i ($0 \le i \le k-1$) with a_i available servers, we first need to decide its number of servers to be purchased (denoted as n_i) under the constraint that $0 \le n_i \le a_i$. Also, the total bandwidth of all the purchased servers should slightly (typically 5%–10% according to BTS-APP operation team's long-term experiences) exceed the estimated overall bandwidth for serving user requests, so that these servers are capable of handling the possible bursty workload.

Under the two (linear) constraints, we then set our goal to minimize the total purchase expense while accommodating the estimated workload. Given that all the variables to be decided (n_i) should be integers, the desired purchase plan can be formalized into an *integer linear programming problem* [6]. Since the problem is NP-hard, we choose to follow the common practice of "sweet spot" balancing, *i.e.,* to find a near-optimal solution with acceptable time complexity ($O(k^2)$) using the branch-and-bound algorithm [42].

When the server purchase plan is determined, the placement of test servers is pivotal to the BTS performance. In terms of Internet data exchange, China Mainland consists of eight domains, each containing a core IXP (Internet eXchange Point) located at Beijing, Shanghai, Guangzhou, Nanjing, Shenyang, Wuhan, Chengdu, and Xi'an, respectively. Thus, the servers should be evenly placed in these domains and as close to the core IXPs as possible.

5.3 Implementation and Evaluation

We have implemented the above-described bandwidth probing and server deployment schemes in a new system called Swiftest. We realize both the client-side and server-side components as Android/Linux user-space modules (using ∼1,200 lines of code) that can be dynamically loaded and run by the original BTS-APP system.

To understand the real-world effectiveness of Swiftest, once again we collaborate with BTS-APP's operation team to conduct back-to-back comparative experiments. Specifically, we upgrade BTS-APP to include Swiftest's client module and invite its users to participate in the new study. From Dec. 20th, 2021 to Jan. 20th, 2022 (lasting a whole month), a total of 0.2M users opted in, conducting 0.31M back-to-back bandwidth tests (*i.e.,* ∼10K tests every day).

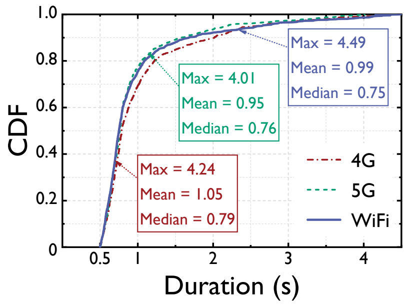

Figure 20: Test time of Swiftest for different access technologies.

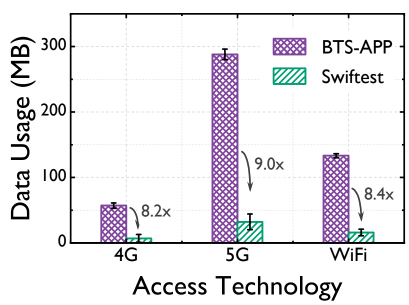

Figure 21: Average data usage per test by BTS-APP and Swiftest.

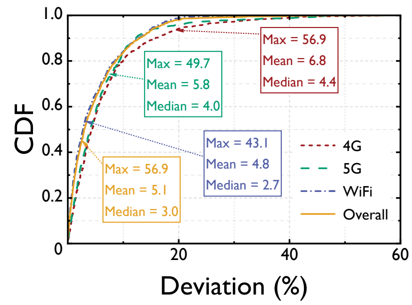

Figure 22: Test result deviation between BTS-APP and Swiftest.

At the client side, if a user chooses to opt in, she will conduct sequential (back-to-back) bandwidth tests using Swiftest and BTS-APP (referred to as a *test pair*), with one-second cooldown in between to avoid mutual interference. The execution order of Swiftest and BTS-APP is randomized in each test for a fair comparison. As compared to only using the original BTS-APP, the repeated testing process incurs little extra data usage (~30 MB on average for even a 5G test) on the opt-in users, given the fast and lightweight testing logic of Swiftest as to be demonstrated shortly.

At the server side, to support ~10K tests per day (~5% of BTS-APP's daily workload, *i.e.,* 0.2M tests), BTS-APP's operation team also allocate 5% of their overall server network capacity (measured by the total bandwidth of their server pool), involving 50 1-Gbps servers (providing a total of 50 Gbps network capacity). In practice, we find such a capacity to be mostly sufficient, only incurring a bit (~10 *ms*) increase of PING latency. As to Swiftest's infrastructure, we only purchase 20 100-Mbps servers from OneProvider (providing a total of 2 Gbps network capacity), as dictated by our ILP-based server selection strategy described in §5.2.

Test Time and Data Usage. Figures 20 and 21 depict the bandwidth test time of Swiftest and the corresponding data usage in 4G, 5G and WiFi networks. Owing to our data-driven bandwidth probing mechanism that accelerates the convergence of tests, the average (median) test time (excluding the initial PING latency) for 4G, 5G, and WiFi is 1.05 sec (0.79 sec), 0.95 sec (0.76 sec) and 0.99 sec (0.75 sec), respectively, greatly outperforming BTS-APP's 10-second fixed time and Speedtest's 15-second fixed time. In fact, even the longest test time (4.49 seconds) is well below 10 seconds.

Note that the client of Swiftest currently PINGs all the 10 test servers during the server selection phase, producing an average of additional 0.2 second test time. Even when this PING latency is taken into account, Swiftest requires merely 1.19 seconds on average to accomplish a bandwidth test, and the majority (55%) of tests are finished within one second.

Accompanying the considerable reduction of test time, the data usage is reduced substantially by 8.2×–9× for 4G, 5G, and WiFi. Most notably, even for 5G tests, the average data usage is only 32 MB, while BTS-APP consumes 289 MB.

Test Accuracy. BTS-APP adopts a Speedtest-like architecture, whose robustness and accuracy have been extensively evaluated and confirmed in the real world. Thus, we take BTS-APP's test results as the approximate ground truth, and calculate the difference between the results of Swiftest and BTS-APP, which are generated back-to-back, to estimate Swiftest's accuracy. In detail, the difference is computed as $\frac{|R_{BTS-APP}-R_{Swiftest}|}{max\{R_{BTS-APP},R_{Swiftest}\}}$, where $R_{BTS-APP}$ and $R_{Swiftest}$ denote the results generated by BTS-APP and Swiftest, respectively, in each test pair. Figure 22 shows the distribution of their results' deviations. We can see that both the average and median deviations are quite small: 5.1% and 3.0%, respectively.

On the other hand, we do notice that in a small (16%) portion of tests, the deviation can exceed 10%. Carefully examining the bandwidth samples collected during such tests, we find that the some user devices should be experiencing severe network fluctuations then, where the bandwidth samples collected by BTS-APP suddenly dropped oftentimes. Therefore, these large deviations are in fact the reflections of high network dynamics during the back-to-back tests. In a minor portion (0.7%) of tests, the deviation can exceed 30%. This is most probably due to the traffic shaping exerted by certain BSes or WiFi APs, because the network dynamics exhibit clear patterns. In fact, this issue is pervasive in all BTSes and is more noticeable for shorter test time.

Comparisons with Other State-of-the-Art BTSes. Apart from extensively evaluating the performance of Swiftest in the wild, we also conduct benchmark experiments to compare its performance with that of two representative state-of-the-art BTSes: FAST and FastBTS, which have both claimed to provide fast bandwidth tests to Internet users. For the proprietary FAST BTS, we reimplement its key testing logic, which has been thoroughly reverse-engineered by prior work [77]. For FastBTS, we use its latest open-source version as of Jan. 5th, 2022. We then deploy the server side of FAST and FastBTS on the same server pool of Swiftest (detailed in §5.3) for a fair comparison. Regarding the client side, we evenly deploy ten Android 11 phones in five mega cities of China (including Beijing, Shanghai, Guangzhou, Shenzhen, and Chengdu) with the same hardware configurations (Qualcomm Snapdragon 765G CPU with 5G capability, 6 GB memory, and 128 GB storage).

We denote one *test group* as the back-to-back bandwidth tests conducted on the same smartphone using the three BTSes (Swiftest, FAST, and FastBTS) and BTS-APP in a random order, where BTS-APP's test results are used as the approximate ground truth for evaluating the test accuracy of the other three BTSes. In total, we perform 13,500 groups of tests, *i.e.,* 10 phones × 15 days (from Jan.

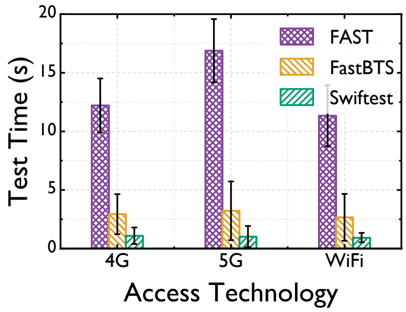

Figure 23: Average test time of the three evaluated BTSes.

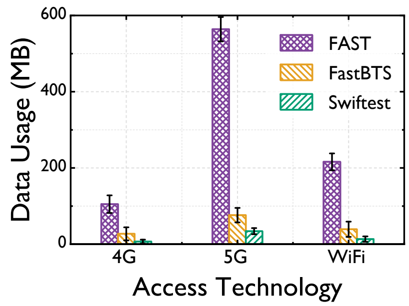

Figure 24: Average data usage per test of the three evaluated BTSes.

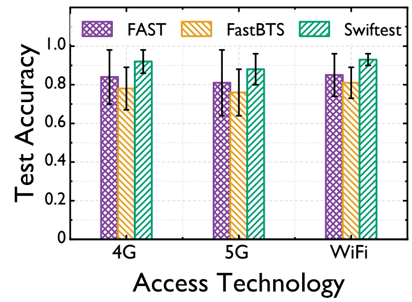

Figure 25: Average test accuracy of the three evaluated BTSes.

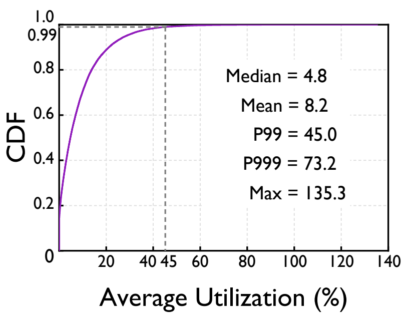

Figure 26: Average bandwidth utilization of Swiftest's servers during the one-month evaluation.

5th to Jan. 20th in 2022) × 3 different times in each day (0:00, 8:00, 16:00) × 3 different access types (4G, 5G and WiFi) × 10 repetitions.

We compare the test time, data usage, and test accuracy of the three BTSes in 4G, 5G, and WiFi networks. As illustrated in Figures 23–25, Swiftest outperforms FAST and FastBTS by yielding 2.9×–16.5× shorter test time, 3×–16.7× smaller data usage, and 8%–12% higher accuracy. In particular, we observe that although FAST adopts a bandwidth estimation algorithm similar to Swiftest (§5.1), its TCP-based bandwidth probing is easily affected by noises introduced by TCP slow start and congestion control, leading to its long test duration (13.5 seconds on average) and high data usage (295 MB on average), especially under high-speed networks. Meanwhile, FastBTS has relatively short test time and small data usage as compared to FAST, but it yields the worst accuracy (0.79 on average) in practice. This is because its crucial interval-based bandwidth estimation algorithm (refer to §5.1) tends to prematurely generate the test result before the user's bandwidth is fully saturated, thus underestimating the access bandwidth.

Infrastructure Cost. As described before, BTS-APP's operation team allocate a proportional amount of server capacity, *i.e.,* 50 1-Gbps servers, to support the evaluation workload (~10K bandwidth tests per day). In comparison, based on our cost-effective server planning, we purchase 20 100-Mbps VM servers distributed around China to support the same workload with Swiftest. It turns out that our purchased fewer budget servers can well accommodate the workload with considerable margins. In detail, Figure 26 shows that in 99% cases, the average bandwidth utilization of Swiftest's servers is ≤45%. In total, the backend infrastructure expense is reduced by 15× by Swiftest as compared to that of BTS-APP.

6 RELATED WORK

This section reviews prior measurement studies on mobile bandwidth and existing approaches to realizing bandwidth testing services. We also compare them to our study and the resulting new BTS.

Mobile Bandwidth Measurements. In the past 15 years, the research community have conducted a plethora of studies to understand realistic cellular and WiFi bandwidths through either field measurements or crowdsourcing. For example, Huang et al. perform crowdsourced measurements of 3G [37] and 4G LTE [35, 36] bandwidths in various application scenarios. Sommers et al. compare the cellular and WiFi bandwidth from different aspects in metro areas [66]. More recently, as 5G makes its debut, Narayanan et al. measure 5G bandwidth through controlled experiments and drive tests [54, 55]; a similar characterization is performed by Xu et al. [74].

Some other studies focus on mobile bandwidth in particular contexts such as multipath [13], high-speed train [44, 72], mobile virtual operators [45, 73], cellular upload [29], and crowded events [63], to name a few. Complementing academic publications, the industry have also published whitepapers and reports on mobile bandwidths [23–25]. In a broader scope, there is a number of work on estimating mobile bandwidth [36, 62, 77], incorporating bandwidth-awareness into application design [15, 46, 52, 67, 79], and saving mobile bandwidth on metered links [10, 31, 39, 47].

Compared to the above, our measurement study features a much larger scale, a special cross-technology (covering WiFi 4/5/6 and 4G/5G) perspective, and a variety of new insights.

Bandwidth Testing Approaches. Almost all commercial BTSes, such as SpeedTest [27], XFinity [28], and SpeedOf [22], take a "probing by flooding" approach to fully saturate the access bandwidth. In spite of its accuracy, this approach may consume considerable (metered) bandwidth and device energy when applied to mobile networks. In the literature, there are also much less invasive approaches, such as IGI [34], TOPP [51], and pathChirp [59], that use strategically crafted packets to probe the bandwidth. However, they are known to suffer from high measurement errors in particular over high-speed wireless links [33, 77].

Recently, some BTSes such as FAST [21] and FastBTS [77] take a more balanced strategy that reduces the probing traffic while maintaining high accuracy. Despite these efforts, we find that when serving real-world cellular and WiFi customers, they still incur various limitations such as premature convergence during slow start and bandwidth re-probing for progressively added servers. Our proposed Swiftest BTS addresses these limitations by leveraging the statistical wisdom from big data, and demonstrates its effectiveness by real-world deployment serving a great number of users.

7 DISCUSSION

Global Applicability of Measurement Results. Among our analysis results of sub-6GHz 5G access, the received signal strength (RSS)-related ones are also applicable to mmWave 5G. Specifically, mmWave 5G requires dense deployment of base stations (BSes), which could also easily cause improper BS placement, antenna configurations, and cross-region coverage problems. Consequently, similar to sub-6GHz 5G, mmWave 5G may also yield undesirable access bandwidth even with excellent RSS. In contrast, our analysis results regarding frequency bands are not applicable to mmWave 5G, since it works on higher frequency bands with contiguous spectrum, rather than fragmented mid-bands.

With respect to spectrum refarming from 4G to 5G, we note that different ISPs across the globe adopt diverse refarming methods. For example, Chinese ISPs mainly adopt static methods that divide the existing spectrum into several pieces for different services [14]. In comparison, ISPs in the US usually adopt dynamic methods to enable different services within the same frequency band [5, 71]. In practice, both methods could incur bandwidth degradation in 4G and 5G networks [4], thus calling for more effective band defragmentation and utilization strategies.

Design Choices of Swiftest. We devise Swiftest as a UDP-based BTS to eliminate the impact of TCP slow start. However, we should clarify that the adoption of UDP is just one of the feasible design choices, and similar benefits can also be achieved by not giving up TCP. For example, we can customize the TCP congestion control algorithm to realize in part the data-driven bandwidth probing mechanism (§5.1), while retaining TCP's fairness properties. However, this approach involves heavy modifications to the congestion control of TCP, and requires many efforts in adapting to the other mechanisms of TCP (*e.g.,* TCP retransmission).

8 CONCLUSION

As 5G and WiFi 6 flourish over the past two years, this paper presents a timely study on the status quo, evolution, and optimization opportunities of mobile access bandwidth. Our study is featured by its cross-layer and cross-technology measurement at scale in the wild, which is enabled by our collaboration with a major mobile bandwidth testing app that serves around 0.2M user requests per day. Based on the fine-grained data we collected from 23.6M bandwidth tests, we discover critical performance gaps between the advertised mobile access bandwidth and what is actually delivered in the wild. For the first time, we reveal the root causes of these gaps by jointly considering the impact of user devices, ISP infrastructure investment, radio resource allocation and migration, and recent advances in cellular technology, with potential solutions to filling these gaps. Furthermore, our study provides insights on building efficient bandwidth testing services; real-world deployment demonstrates the remarkable improvement in test time and infrastructure cost.

ACKNOWLEDGEMENTS

We appreciate Li Ma and Guangbiao Liu for their help in data collection and system deployment. We sincerely thank our shepherd Jia Wang and the anonymous reviewers for their valuable feedbacks and constructive suggestions. We also appreciate Jiaxing Qiu, Wei Liu, Yang Li, and Jianwei Zheng for their help in improving the presentation of this paper. This work is supported in part by the National Key R&D Program of China under grant 2021YFB2900100, the National Natural Science Foundation of China (NSFC) under grant 61902211, and Microsoft Research Asia.

9 ARTIFACT APPENDIX

Abstract

Swiftest's artifacts are publicly available at GitHub. To facilitate a better understanding of Swiftest, we provide detailed instructions on how to build, deploy, install and use Swiftest, as well as how to compare Swiftest with other state-of-the-art bandwidth testing services. Please refer to our README file at https://github.com/mobilebandwidth/Artifacts for details.

Scope

The artifacts can be used to reproduce all the major results of Swiftest.

Contents

The artifacts include the client-side and server-side source code of Swiftest, binaries of both Swiftest and BTS-APP, and our evaluation data/figures.

Hosting

Both the code/binary and data are hosted in the `main` branch of the `Artifacts` repository.

- **Swiftest Code and Binary.**
  https://github.com/mobilebandwidth/Artifacts/tree/main/Swiftest.
- **BTS-APP Binary.**
  https://github.com/mobilebandwidth/Artifacts/tree/main/BTS-APP.
- **Evaluation Data and Figures.**
  https://github.com/mobilebandwidth/Artifacts/tree/main/plots.
- **DOI for the Artifacts.**
  10.5281/zenodo.6782121.

REFERENCES

[1] 3GPP. 2008. Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) Radio Transmission and Reception (Release 8). *The 3rd Generation Partnership Project* (2008).
[2] 3GPP. 2020. 5G; NR; User Equipment (UE) Radio Transmission and Reception; Part 1: Range 1 Standalone (Release 16). *The 3rd Generation Partnership Project* (2020).

[3] Md Maruf Ahamed and Saleh Faruque. 2021. 5G Network Coverage Planning and Analysis of the Deployment Challenges. *Sensors* 21, 19 (2021), 6608.

[4] Amr Ashraf. 2021. Dynamic Spectrum Sharing for 4G LTE and 5G NR Deployments. https://digis2.com/dss/. (2021). (Accessed on May 25, 2022).

[5] AT&T. 2020. Technology Report on Dynamic Spectrum Sharing: From Labs to Live. https://about.att.com/innovationblog/2020/05/dynamic_spectrum_sharing.html. (2020). (Accessed on May 25, 2022).

[6] Stephen Boyd, Stephen P Boyd, and Lieven Vandenberghe. 2004. *Convex Optimization*. Cambridge University Press.

[7] CAITC. 2020. White Paper on Broadband Development in China (in Chinese). https://pdf.dfcfw.com/pdf/H3_AP202010161421652745_1.pdf?160284701000. (2020). (Accessed on May 27, 2022).

[8] CAITC. 2021. White Paper on Broadband Development in China (in Chinese). https://pdf.dfcfw.com/pdf/H3_AP202109291519270217_1.pdf?1632921062000.pdf. (2021). (Accessed on May 27, 2022).

[9] Neal Cardwell, Yuchung Cheng, C Stephen Gunn, Soheil Hassas Yeganeh, and Van Jacobson. 2017. BBR: Congestion-Based Congestion Control. *Commun. ACM* 60, 2 (2017), 58–66.

[10] Li Chen, Justinas Lingys, Kai Chen, and Feng Liu. 2018. Auto: Scaling Deep Reinforcement Learning for Datacenter-Scale Automatic Traffic Optimization. In *Proc. of ACM SIGCOMM*. 191–205.

[11] Ben Cox, Jan G Laufer, Simon R Arridge, Paul C Beard, A Jan G Laufer, and A Simon R Arridge. 1984. Long Range Dependence: A Review. In *Iowa State University*.

[12] Erik Dahlman, Stefan Parkvall, and Johan Skold. 2016. *4G, LTE-advanced Pro and the Road to 5G*. Academic Press.

[13] Shuo Deng, Ravi Netravali, Anirudh Sivaraman, and Hari Balakrishnan. 2014. WiFi, LTE, or Both? Measuring Multi-Homed Wireless Internet Performance. In *Proc. of ACM IMC*. 181–194.

[14] DVBCN. 2021. China Spectrum Division Policies (in Chinese). https://www.yunxihuixiang.com/2021/0910/90028.html. (2021). (Accessed on May 25, 2022).

[15] Jeongyoon Eo, Zhixiong Niu, Wenxue Cheng, Francis Y Yan, Rui Gao, Jorina Kardhashi, Scott Inglis, Michael Revow, Byung-Gon Chun, Peng Cheng, and Yongqiang Xiong. 2022. OpenNetLab: Open Platform for RL-based Congestion Control for Real-Time Communications. In *Proc. of APNet*.

[16] Ericsson. 2021. 5G Energy Consumption: What's the Impact of 5G NR in Real Networks? https://www.ericsson.com/en/blog/2021/10/5g-energy-consumption-impact-5g-nr. (2021). (Accessed on May 15, 2022).

[17] Amina Gharsallah, Faouzi Zarai, and Mahmoud Neji. 2018. SDN/NFV-Based Handover Management Approach for Ultradense 5G Mobile Networks. *International Journal of Communication Systems* 32, 17 (2018), e3831.

[18] Amitava Ghosh, Rapeepat Ratasuk, Bishwarup Mondal, Nitin Mangalvedhe, and Tim Thomas. 2010. LTE-Advanced: Next-Generation Wireless Broadband Technology. *IEEE Wireless Communications* 17, 3 (2010), 10–22.

[19] CNMO Group. 2022. 5G And 4G Will Coexist for A Long Time (in Chinese). https://www.sohu.com/a/353383047_115831. (2022). (Accessed on Jan. 20, 2022).

[20] Media Tek Group. 2022. 5G and 4G Are Complementary Technologies in The Context of Latin America. https://impactotic.co/en/latam-4g-5g-complementary-chipset-in-latam-mmlc21/. (2022). (Accessed on Jan. 20, 2022).

[21] Netflix Group. 2022. FAST Internet Speed Test. https://fast.com/. (2022). (Accessed on Jan. 20, 2022).

[22] SpeedOf.Me Group. 2019. SpeedOf.Me: Internet Speed Test for All Your Devices. https://www.speedof.me/. (2019). (Accessed on Dec. 15, 2021).

[23] SpeedTest Group. 2022. 5G in Europe: EU Targets Require a Rethink. https://www.speedtest.net/insights/blog/eu-5g-targets-require-rethink/. (2022). (Accessed on Jan. 20, 2022).

[24] SpeedTest Group. 2022. Growing and Slowing: The State of 5G Worldwide in 2021. https://www.speedtest.net/insights/blog/state-of-worldwide-5g-2021/. (2022). (Accessed on Jan. 20, 2022).

[25] SpeedTest Group. 2022. Speedtest Global Index Market Analyses Are Now Available for 41 Countries. https://www.speedtest.net/insights/blog/global-index-market-analyses-q4-2021/. (2022). (Accessed on Jan. 20, 2022).

[26] SpeedTest Group. 2022. SpeedTest Insights. (2022).

[27] SpeedTest Group. 2022. The Global Broadband Speed Test by Ookla. http://www.speedtest.net/. (2022). (Accessed on Jan. 20, 2022).

[28] XFinity Group. 2021. Xfinity Speed Test. http://speedtest.xfinity.com/. (2021). (Accessed on Oct. 27, 2021).

[29] Yihua Guo, Feng Qian, Qi Alfred Chen, Zhuoqing Morley Mao, and Subhabrata Sen. 2016. Understanding On-device Bufferbloat for Cellular Upload. In *Proc. of ACM IMC*. 303–317.

[30] Sangtae Ha, Injong Rhee, and Lisong Xu. 2008. CUBIC: A New TCP-Friendly High-Speed TCP Variant. *ACM SIGOPS Operating Systems Review* 42, 5 (2008), 64–74.

[31] Bo Han, Feng Qian, Lusheng Ji, and Vijay Gopalakrishnan. 2016. MP-DASH: Adaptive Video Streaming over Preference-Aware Multipath. In *Proc. of ACM CoNEXT*. 129–143.

[32] IT Home. 2018. All You Need to Know About 4G (in Chinese). https://tech.sina.cn/mobile/pc/2018-11-23/detail-ihmutuec2853641.d.html. (2018). (Accessed on May. 20, 2022).

[33] Ningning Hu, Li Li, Zhuoqing Morley Mao, Peter Steenkiste, and Jia Wang. 2004. Locating Internet Bottlenecks: Algorithms, Measurements, and Implications. *ACM SIGCOMM Computer Communication Review* 34, 4 (2004), 41–54.

[34] Ningning Hu and Peter Steenkiste. 2003. Evaluation and Characterization of Available Bandwidth Probing Techniques. *IEEE Journal on Selected Areas in Communications* 21, 6 (2003), 879–894.

[35] Junxian Huang, Feng Qian, Alexandre Gerber, Z Morley Mao, Subhabrata Sen, and Oliver Spatscheck. 2012. A Close Examination of Performance and Power Characteristics of 4G LTE Networks. In *Proc. of ACM MobiSys*. 225–238.

[36] Junxian Huang, Feng Qian, Yihua Guo, Yuanyuan Zhou, Qiang Xu, Z Morley Mao, Subhabrata Sen, and Oliver Spatscheck. 2013. An In-Depth Study of LTE: Effect of Network Protocol and Application Behavior on Performance. *ACM SIGCOMM Computer Communication Review* 43, 4 (2013), 363–374.

[37] Junxian Huang, Qiang Xu, Birjodh Tiwana, Z Morley Mao, Ming Zhang, and Paramvir Bahl. 2010. Anatomizing Application Performance Differences on Smartphones. In *Proc. of ACM MobiSys*. 165–178.

[38] UUSense Inc. 2021. UUSpeed Bandwidth Testing Services. http://uuspeed.uutest.cn/. (2021). (Accessed on May 25, 2022).

[39] Jaehong Kim, Youngmok Jung, Hyunho Yeo, Juncheol Ye, and Dongsu Han. 2020. Neural-Enhanced Live Streaming: Improving Live Video Ingest via Online Learning. In *Proc. of ACM SIGCOMM*. 107–125.

[40] Harold J Kushner and Philip A Whiting. 2004. Convergence of Proportional-Fair Sharing Algorithms under General Conditions. *IEEE Wireless Communications* 3, 4 (2004), 1250–1259.

[41] Youngsun Kwon, Duk Kyu Park, and Hongjai Rhee. 2017. Spectrum Fragmentation: Causes, Measures and Applications. *Telecommunications Policy* 41, 5-6 (2017), 447–459.

[42] Eugene L Lawler and David E Wood. 1966. Branch-and-Bound Methods: A Survey. *Operations Research* 14, 4 (1966), 699–719.

[43] Feng Li, Kwok-Yan Lam, Nan Zhao, Xin Liu, Kanglian Zhao, and Li Wang. 2018. Spectrum Trading for Satellite Communication Systems with Dynamic Bargaining. *IEEE Transactions on Communications* 66, 10 (2018), 4680–4693.

[44] Li Li, Ke Xu, Tong Li, Kai Zheng, Chunyi Peng, Dan Wang, Xiangxiang Wang, Meng Shen, and Rashid Mijumbi. 2018. A Measurement Study on Multi-Path TCP with Multiple Cellular Carriers on High Speed Rails. In *Proc. of ACM SIGCOMM*. 161–175.

[45] Yang Li, Jianwei Zheng, Zhenhua Li, Yunhao Liu, Feng Qian, Sen Bai, Yao Liu, and Xianlong Xin. 2020. Understanding the Ecosystem and Addressing the Fundamental Concerns of Commercial MVNO. *IEEE/ACM Transactions on Networking* 28, 3 (2020), 1364–1377.

[46] Zhenhua Li, Yafei Dai, Guihai Chen, and Yunhao Liu. 2016. *Content Distribution for Mobile Internet: A Cloud-based Approach*. Springer.

[47] Zhenhua Li, Weiwei Wang, Tianyin Xu, Xin Zhong, Xiang-Yang Li, Yunhao Liu, Christo Wilson, and Ben Y Zhao. 2016. Exploring Cross-Application Cellular Traffic Optimization with Baidu TrafficGuard. In *Proc. of USENIX NSDI*. 61–76.

[48] Jiajia Liu, Yuichi Kawamoto, Hiroki Nishiyama, Nei Kato, and Naoto Kadowaki. 2014. Device-to-Device Communications Achieve Efficient Load Balancing in LTE-Advanced Networks. *IEEE Wireless Communications* 21, 2 (2014), 57–65.

[49] Wanzhi Ma, Hongzhi Zhao, Ying Liu, Shihai Shao, and Wensheng Pan. 2018. A Co-Channel Interference Rejection Method for 5G Ultra Dense Heterogeneous Networks. In *Proc. of IEEE ICC Workshops*. 1–5.

[50] Sue Marek. 2022. Marek's Take: The greening of the Gs. https://www.fiercewireless.com/5g/mareks-take-greening-gs. (2022). (Accessed on May 16, 2022).

[51] Bob Melander, Mats Bjorkman, and Per Gunningberg. 2000. A New End-to-End Probing and Analysis Method for Estimating Bandwidth Bottlenecks. In *Proc. of IEEE GlobeCom*. 415–420.

[52] Zhao Ming, Xiuhua Li, Chuan Sun, Qilin Fan, Xiaofei Wang, and Victor CM Leung. 2022. Sleeping Cell Detection for Resiliency Enhancements in 5G/B5G Mobile Edge-Cloud Computing Networks. *ACM Transactions on Sensor Networks (TOSN)* 18, 3 (2022), 1–30.

[53] China Mobile. 2021. China Mobile Data Plan for Fixed Broadband (in Chinese). http://www.10086.cn/kdzq/bj/index_100_100.html. (2021). (Accessed on May 21, 2022).

[54] Arvind Narayanan, Eman Ramadan, Jason Carpenter, Qingxu Liu, Yu Liu, Feng Qian, and Zhi-Li Zhang. 2020. A First Look at Commercial 5G Performance on Smartphones. In *Proc. of ACM WWW*. 894–905.

[55] Arvind Narayanan, Xumiao Zhang, Ruiyang Zhu, Ahmad Hassan, Shuowei Jin, Xiao Zhu, Xiaoxuan Zhang, Denis Rybkin, Zhengxuan Yang, Zhuoqing Morley Mao, et al. 2021. A Variegated Look at 5G in the Wild: Performance, Power, and QoE Implications. In *Proc. of ACM SIGCOMM*. 610–625.

[56] China Ministry of Industry and Information Technology. 2022. Three Major ISPs in China Are Authorized to Re-Farm Mid-Band and Low-Band for 5G Use (in Chinese). https://m.mp.oeeee.com/a/BAAFRD000020201223400806.html. (2022). (Accessed on Jan. 10, 2022).

[57] Constantinos B Papadias, Tharmalingam Ratnarajah, and Dirk TM Slock. 2020. *Spectrum Sharing: The Next Frontier in Wireless Networks*. John Wiley & Sons.

[58] Faizan Qamar, MHD Nour Hindia, Kaharudin Dimyati, Kamarul Ariffin Noordin, and Iraj Sadegh Amiri. 2019. Interference Management Issues for the Future 5G Network: A Review. *Telecommunication Systems* 71, 4 (2019), 627–643.

[59] Vinay Joseph Ribeiro, Rudolf H Riedi, Richard G Baraniuk, Jiri Navratil, and Les Cottrell. 2003. Pathchirp: Efficient Available Bandwidth Estimation for Network Paths. In *In Proc. of PAM Workshop*.

[60] Rakibul Islam Rony, Elena Lopez-Aguilera, and Eduard Garcia-Villegas. 2021. Dynamic Spectrum Allocation Following Machine Learning-Based Traffic Predictions in 5G. *IEEE Access* 9 (2021), 143458–143472.

[61] Fatima Salahdine, Johnson Opadere, Qiang Liu, Tao Han, Ning Zhang, and Shaohua Wu. 2021. A Survey on Sleep Mode Techniques for Ultra-Dense Networks in 5G and Beyond. *Computer Networks* 201 (2021), 108567.

[62] Natsuhiko Sato, Takashi Oshiba, Kousuke Nogami, Anan Sawabe, and Kozo Satoda. 2017. Experimental Comparison of Machine Learning-Based Available Bandwidth Estimation Methods over Operational LTE Networks. In *Proc. of IEEE ISCC*. 339–346.

[63] Muhammad Zubair Shafiq, Lusheng Ji, Alex X Liu, Jeffrey Pang, Shobha Venkataraman, and Jia Wang. 2013. A First Look at Cellular Network Performance during Crowded Events. In *Proc. of ACM SIGMETRICS*. 17–28.

[64] Claude Elwood Shannon. 1948. A Mathematical Theory of Communication. *The Bell System Technical Journal* 27, 3 (1948), 379–423.

[65] Min Sheng, Chungang Yang, Yan Zhang, and Jiandong Li. 2013. Zone-Based Load Balancing in LTE Self-Optimizing Networks: A Game-Theoretic Approach. *IEEE Transactions on Vehicular Technology* 63, 6 (2013), 2916–2925.

[66] Joel Sommers and Paul Barford. 2012. Cell vs. WiFi: On the Performance of Metro Area Mobile Connections. In *Proc. of ACM IMC*. 301–314.

[67] Tiao Tan, Ming Zhao, and Zhiwen Zeng. 2022. Joint Offloading and Resource Allocation Based on UAV-Assisted Mobile Edge Computing. *ACM Transactions on Sensor Networks (TOSN)* 18, 3 (2022), 1–21.

[68] Techopedia. 2020. Definition of Guard Band. https://www.techopedia.com/definition/7494/guard-band-telecommunications. (2020). (Accessed on May. 23, 2022).

[69] China Telecom. 2021. China Telecom Data Plan for Fixed Broadband (in Chinese). https://www.189.cn/hd/ywdj/?intaid=fj-sy-daohang-04-ywdj. (2021). (Accessed on May 21, 2022).

[70] China Unicom. 2021. China Unicom Data Plan for Fixed Broadband (in Chinese). https://m.10010.com/queen/new-broadband-web/new-broadband-web.html?activeId=8818100954349604. (2021). (Accessed on May 21, 2022).

[71] Verizon. 2020. Verizon Completes Successful Dynamic Spectrum Sharing Technology Trials in Advance of 5G Nationwide. https://www.verizon.com/about/news/verizon-dynamic-spectrum-sharing-technology. (2020). (Accessed on May 25, 2022).

[72] Jing Wang, Yufan Zheng, Yunzhe Ni, Chenren Xu, Feng Qian, Wangyang Li, Wantong Jiang, Yihua Cheng, Zhuo Cheng, Yuanjie Li, Xiufeng Xie, Yi Sun, and Zhongfeng Wang. 2019. An Active-Passive Measurement Study of TCP Performance over LTE on High-Speed Rails. In *Proc. of ACM MobiCom*. 1–16.

[73] Ao Xiao, Yunhao Liu, Yang Li, Feng Qian, Zhenhua Li, Sen Bai, Yao Liu, Tianyin Xu, and Xianlong Xin. 2019. An In-Depth Study of Commercial MVNO: Measurement and Optimization. In *Proc. of ACM MobiSys*. 457–468.

[74] Dongzhu Xu, Anfu Zhou, Xinyu Zhang, Guixian Wang, Xi Liu, Congkai An, Yiming Shi, Liang Liu, and Huadong Ma. 2020. Understanding Operational 5G: A First Measurement Study on Its Coverage, Performance and Energy Consumption. In *Proc. of ACM SIGCOMM*. 479–494.

[75] Sheng Xu. 2018. Co-Channel Interference Cancellation for 5G Cellular Networks Deploying Radio-over-Fiber and Massive MIMO Beamforming. In *Broadband Communications Networks-Recent Advances and Lessons from Practice*. IntechOpen.

[76] Xiaolong Xu, Xihua Liu, Zhanyang Xu, Chuanjian Wang, Shaohua Wan, and Xiaoxian Yang. 2020. Joint Optimization of Resource Utilization and Load Balance with Privacy Preservation for Edge Services in 5G Networks. *Mobile Networks and Applications* 25, 2 (2020), 713–724.

[77] Xinlei Yang, Xianlong Wang, Zhenhua Li, Yunhao Liu, Feng Qian, Liangyi Gong, Rui Miao, and Tianyin Xu. 2021. Fast and Light Bandwidth Testing for Internet Users. In *Proc. of USENIX NSDI*. 1011–1026.

[78] Guangxiang Yuan, Xiang Zhang, Wenbo Wang, and Yang Yang. 2010. Carrier Aggregation for LTE-Advanced Mobile Communication Systems. *IEEE Communications Magazine* 48, 2 (2010), 88–93.

[79] Zhilong Zheng, Yunfei Ma, Yanmei Liu, Furong Yang, Zhenyu Li, Yuanbo Zhang, Jiuhai Zhang, Wei Shi, Wentao Chen, Ding Li, Qing An, Hai Hong, Hongqiang Harry Liu, and Ming Zhang. 2021. XLINK: QoE-Driven Multi-Path QUIC Transport in Large-Scale Video Services. In *Proc. of ACM SIGCOMM*. 418–432.