How Mature is 5G Deployment? A Cross-Sectional, Year-Long Study of 5G Uplink Performance

Imran Khan[a], Moinak Ghoshal[a], Joana Angjo[b], Sigrid Dimce[b], Mushahid Hussain[c], Paniz Parastar[d], Yenchia Yu[e], Xueting Deng[a], Sumit Hawal[a], Shirui Huang[a], Ameya Rane[a], Yin Wang[a], Claudio Fiandrino[f], Charalampos Orfanidis[g], Shivang Aggarwal[h], Ana C Aguiar[c], Ozgu Alay[d], Carla Fabiana Chiasserini[e], Falko Dressler[b], Y. Charlie Hu[i], Steven Y. Ko[j], Dimitrios Koutsonikolas[a], Joerg Widmer[f]

[a]*Northeastern University, USA*
[b]*TU Berlin, Germany*
[c]*University of Porto, Portugal*
[d]*University of Oslo, Norway*
[e]*Politecnico di Torino, Italy*
[f]*IMDEA Networks Institute, Spain*
[g]*Technical University of Denmark, Denmark*
[h]*Hewlett Packard Labs, USA*
[i]*Purdue University, USA*
[j]*Simon Fraser University, Canada*

Abstract

After a rapid deployment worldwide over the past few years, 5G is expected to have reached a mature deployment stage to provide measurable improvement of network performance and user experience over its predecessors. In this study, we aim to assess 5G deployment maturity via three conditions: (1) Does 5G performance remain stable over a long time span (1 year)? (2) Does 5G provide better performance than its predecessor Long-Term Evolution (LTE)? (3) Does the technology offer similar performance across diverse geographic areas and cellular operators? We answer this important question by conducting two year-long measurement campaigns of 5G uplink performance leveraging a custom Android app: one crowd-sourced, cross-sectional campaign spanning 8 major cities in 7 countries and two different continents (Europe and North America), and one controlled campaign focusing on mmWave deployment at a fixed location in the downtown area of Boston, MA. Our datasets show that 5G deployment in major cities appears to have matured, with no major performance improvements observed over a one-year period, but 5G does not provide consistent, superior measurable

Article published in Elsevier Computer Communications

performance over LTE, especially in terms of latency, and further there exists clear uneven 5G performance across the 8 cities. Our study suggests that, while 5G deployment appears to have stagnated, it is short of delivering its promised performance and user experience gain over its predecessor.

---

1. Introduction

The most recent generation of cellular networks, 5G, promises ultra-high bandwidth and ultra-low latency, far surpassing the performance of 4G LTE, via a combination of PHY layer innovations such as higher modulation schemes, beamforming, (massive) MIMO, and wider channels. Such high data rates, combined with low latency, hold the promise to finally support *latency-critical* applications such as Augment Reality (AR), Mixed Reality (MR), Connected Autonomous Vehicles (CAVs), and the Metaverse, often dubbed as "5G killer" apps, which demand ultra-high network bandwidth and low network latency to support offloading of compute-intensive tasks to the edge cloud.

5G rollout started in 2019 and the wide-scale deployment has been rapid and aggressively marketed by all mobile network operators. As such, after a rapid deployment worldwide over the past few years, it is highly anticipated that 5G has reached a deployment stage mature enough to significantly improve the performance of mobile networks and, more importantly, the user experience, in particular, when running the class of latency-critical apps that could not be supported by LTE.

To answer this question, there have been a number of measurement studies of 5G networks in recent years [1–6], [7–15], [16–24].. However, most of these studies have focused on measuring the 5G *downlink* performance while the *uplink* performance of 5G networks remain largely unknown. Understanding the 5G uplink performance is important, since most latency-critical "5G killer" apps distinguish themselves from legacy apps for their heavy, bursty *uplink* data transfers, and 5G, similar to all its predecessors, has provisioned much higher downlink bandwidth than uplink bandwidth.

This paper, which extends our previous work [25], fills this gap by answering two questions: (1) *How mature is today's 5G deployment?* and (2)*Is today's 5G uplink performance sufficient to enable latency-critical uplink-oriented apps such as AR or CAVs?* We consider that a technology deployment is "mature" when the following three conditions are satisfied: (i) *Its performance remains stable over a long time span (1 year).* Previous works performed measurements within a short time span, ranging from a few days

Table 1: Overview of the collected data

City (Country)	**Operator**	**Tests**	**Duration**	**Cell IDs**	**Radius of Gyration (km)**
Berlin (Germany)	Telekom	341	11/22-09/23	194	6.156
Turin (Italy)	TIM, WINDTRE	90	11/22-09/23	41	6.819
Oslo (Norway)	Telenor, Telia	1429	09/22-09/23	276	2.179
Porto (Portugal)	MEO	241	01/23-08/23	57	1.191
Madrid (Spain)	Vodafone	7096	10/22-09/23	525	8.734
Vancouver (Canada)	Bell, Shaw Comm.	561	11/22-09/23	206	14.516
Boston (USA)	ATT, Verizon T-Mobile	328	07/22-04/23	93	8.71
Bay Area (USA)	T-Mobile	80	07/22-07/23	30	6.34
Total	-	10166	07/22-09/23	1422	-

up to a couple of months. However, any findings from such studies might be short-lived and lead to wrong conclusions about the potential of 5G in the long term. (ii) *The technology offers higher coverage and better performance than its predecessor*. In its mature stage, 5G should offer extended coverage replacing LTE and significantly higher throughput and lower latency than LTE, as promised. (iii) *The technology offers similar coverage and performance across diverse geographic areas and cellular operators* (in the same frequency band). Several previous works performed studies limited to one or a couple of cities or with a single operator. Such studies only provide a partial view of 5G performance, as hardware, configurations, and policies can differ not only across operators but also across cities for the same operator [10, 19]. Consequently, these two questions cannot be answered without a detailed, *longitudinal and cross-sectional* study of 5G uplink performance.

To answer these questions, in this work, we conduct a year-long cross-sectional measurement study of 5G uplink performance via two measurement campaigns. In the first measurement campaign, leveraging an Institutional Review Board (IRB) approved custom Android app, we collected a large *crowd-sourced* dataset of 5G performance (uplink throughput and latency) along with various metadata (cell IDs, handovers, GPS, coordinates, signal strength, mobility status, etc.). Our dataset, summarized in Table 1, spans 8 major cities in 7 different countries and 2 different continents – Berlin (Germany), Turin (Italy), Oslo (Norway), Porto (Portugal), and Madrid

(Spain) in Europe, Boston (USA), Bay Area (USA), and Vancouver (Canada) in North America – and 12 operators. In each of these cities, volunteers used our app to perform weekly measurements at their convenience. As such, our dataset reflects the average performance experienced by a user at home, work, or during their regular commute over a whole year. *Our dataset and scripts are publicly available.*[1]

Leveraging this unique dataset, we first look at the evolution of 5G performance in each city over the past year, in terms of uplink throughput and latency. We then look at 5G performance (throughput and latency) in each city and compare it with the corresponding LTE performance. Our main findings are as follows:

- Somewhat surprisingly, we do not observe any increasing or decreasing trend over the past one year for either metric, suggesting that condition (i) for maturity is satisfied; 5G deployment has reached a stable stage, with no major updates over the past one year.

- Surprisingly, our analysis reveals that 5G does not always yield better performance than LTE, suggesting that condition (ii) for maturity is not met. Specifically, 5G throughput is lower than LTE throughput in one city, and the 5G-LTE throughput gap across the remaining seven cities varies significantly, ranging from 2.36 Mbps to 52.23 Mbps in the median case.

- More importantly, 5G latency is lower than LTE latency in only three out of eight cities, and higher in three out of eight cities, suggesting that 5G does not consistently deliver lower latency than LTE.

- Additionally, our dataset reveals very diverse 5G performance across the eight cities, indicating that condition (iii) for maturity is not met either.

Overall, our study suggests that, while 5G deployment appears to have stabilized, it has yet to deliver the promised performance improvements over LTE. Consequently, it is not yet ready to fully support the next generation of latency-critical applications.

In the second measurement campaign, which was not included in [25], we use the same app along with XCAL Solo [26], a commercial tool that captures lower layer Key Performance Indicators (KPIs) and signaling messages, and an

[1] `https://github.com/NUWiNS/ifip2024_year_long_5G_uplink_study`

edge server, to conduct an in-depth study of 5G mmWave uplink performance via *controlled* measurements at a fixed location in downtown Boston. The 5G mmWave technology provides much higher data rates compared to low band or midband 5G [2, 3, 9], however, its limited deployment (only in the downtowns of select cities in the US and Japan) and sensitivity to blockage and mobility make it hard to assess its performance via crowdsourced measurements.

Our findings are summarized as follows:

- 5G mmWave uplink throughput exhibits no significant trends over one year, aside from seasonal fluctuations, and latency remains stable with minimal variation, suggesting again that condition (i) for maturity is satisfied.
- The mmWave throughput/latency values are significantly higher/lower than the LTE and 5G-low/mid counterparts obtained via our crowdsourced measurements in Boston over the same one-year period, suggesting that 5G mmWave and edge computing are both critical to boosting the performance of latency-critical, uplink-oriented 5G killer apps.
- Our study on the operator's resource-sharing policy among multiple users shows that it remains consistent over the one-year measurement period, allocating resources fairly to two backlogged flows.

2. Related Work

Since the initial 5G rollout in 2019, a large number of studies have measured various aspects of 5G performance [1–15, 17–24]. Table 2 summarizes these works based on their time span and geographic scope. As one can be observe from the table, most studies have a limited geographic coverage, conducting measurements in one [1, 3, 5, 7, 8, 13, 14, 18, 24] or a few cities [2, 4, 6, 17, 19, 20, 22]. While these studies provide valuable insights into 5G performance within localized areas, they may not necessarily reflect worldwide performance trends. Additionally, most of these studies conduct measurements over a limited time span, from a few days to a couple of weeks [1–4, 6, 8–13, 17, 19, 22–24] and they do not investigate the evolution of 5G performance over extended periods. Finally, most of them (with the exception of [4, 9, 22]) focus primarily on downlink performance.

A small number of studies conduct measurements over a larger span of geographic locations. The work in [9] performs a measurement study of 5G

Table 2: Classification of Studies by Duration and Geographic Scope.

Group	Duration	Geographic Scope	Studies Included
1	Short	Limited	[1-4], [6], [8], [10], [13], [17], [23], [24], [26]
2	Short	Broad	[7], [9], [11], [12], [16], [19], [20], [21], [23]
3	Long	Limited	[5], [14], 15], [22]
4	Long	Broad	[25]

and LTE performance during a cross-country drive from Los Angeles, CA to Boston, MA and the work in [11] analyzes 5G handovers via a similar cross-country drive. The work in [12] measures roaming performance in the EU while driving across four European countries. The work in [19] compares midband performance in the US and Europe via measurements in a total of 5 cities over multiple cellular operators. However, these studies limit their measurement campaigns within a short time span of at most a few weeks and do not analyze performance evolution over time. The works in [7, 16, 20, 21, 23] consider much longer time periods, from several months up to 3 years, but use datasets collected over short time intervals during those long periods. For example, the work in [16] studies the evolution of 5G performance from a mobile operator perspective over a 3-year period, but it uses a dataset spanning 3 weeks in 2020 and one week in 2022, which only offers two snapshots of the observed performance during the 3-year period.

On the other hand, the works in [5, 14, 15, 22] conduct studies over a longer time span, from several weeks up to two years. The work in [5] studies 5G performance on public transit systems over a 3-month period in Madrid and the work in [14] studies 5G NSA performance over a 7-week period in Rome. The work in [22] performs a cross-layer measurement study of commercial 5G networks under different mobility scenarios over a 10-month period. The work in [15] studies mobile access bandwidth in China by collaborating with a commercial bandwidth testing app, and, similarly to the work in [7], shows a decrease in both LTE and 5G throughput in 2023 compared to 2022. In our work, we do not observe any notable differences in performance during 2023. In contrast to our work, all these works focus on a single city or country.

In summary, *this work, to our best knowledge, is the first to perform a longitudinal and cross-sectional measurement campaign of 5G performance, spanning 8 cities in 7 countries and 2 continents.*

Table 3: List of metrics collected with our Android app

Metric	Description
GPS	User's City, Country
Network Type	5G (mmWave) /5G (sub-6 GHz)/ LTE
RSRP	Reference Signal Received Power
RSRQ	Reference Signal Received Quality
RSSI	Received Signal Strength Indicator
Cell-ID, EARFCN/ARFCN	Connected cell id and frequency
Operator	User's cellular operator

3. Methodology

Measurement servers. To enable throughput and latency measurements, we deployed three AWS (Amazon Web Services) Cloud servers, two in the US (Northern Virginia and Oregon) and one in Europe (Frankfurt, Germany). Additionally, for measurements in Boston with Verizon, we deployed an AWS Wavelength server in Boston. Wavelength servers are located *inside* Verizon's network in selected cities and specially designed for edge computing.

Measurement app. Our Android measurement app, NextG-UP [27], has two main functionalities. It measures uplink Transmission Control Protocol (TCP) throughput and Round Trip Time (RTT) while collecting various cellular network metrics. We leverage Android-provided APIs to retrieve the network metrics that require the users to grant permission to access certain data on the phone (TELEPHONY, GPS, etc.). A detailed list of the collected metrics is shown in Table 3.

The app initially collects the user's location in the background and selects the nearest server based on this information. Subsequently, the user is prompted to choose between three test types: static, walking, or driving. The app checks whether the UE's WiFi is turned off, exclusively focusing on cellular network performance. Once these checks are completed, the application measures uplink TCP throughput using nuttcp-8.1.4 over a 10-second period. Following this, the app initiates an RTT test using the ping utility, sending 11 Internet Control Message Protocol (ICMP) packets spaced 200 ms apart. The workflow of the application is presented in Fig. 1.

The app features a lightweight design, with an image size of 6.5 MB and utilizing less than 250 MB of memory while running, ensuring efficient performance and minimal resource consumption on user devices.

Crowd-sourced Measurements. We reached out to our research community to recruit volunteers to participate in the measurement study for a one-year period. We received responses and data from 16 countries. However,

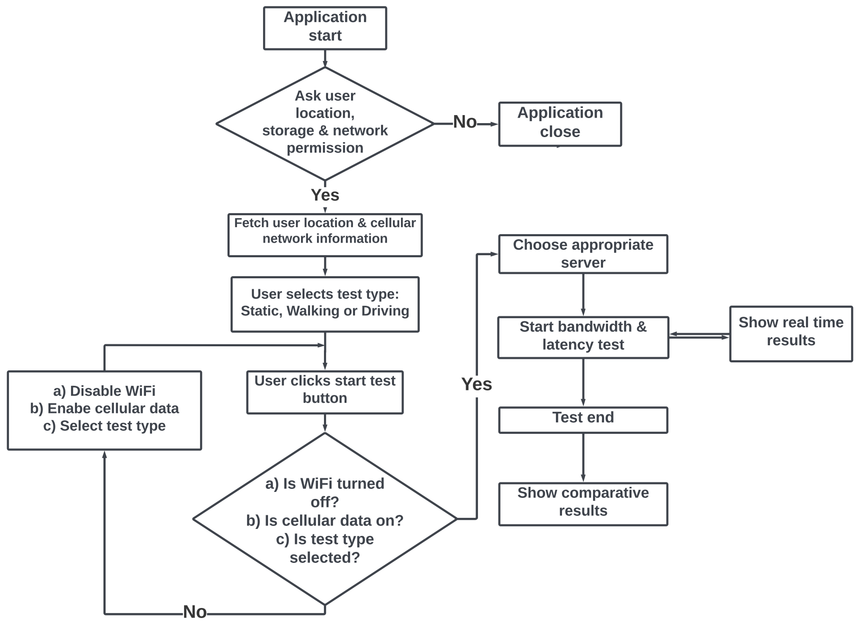
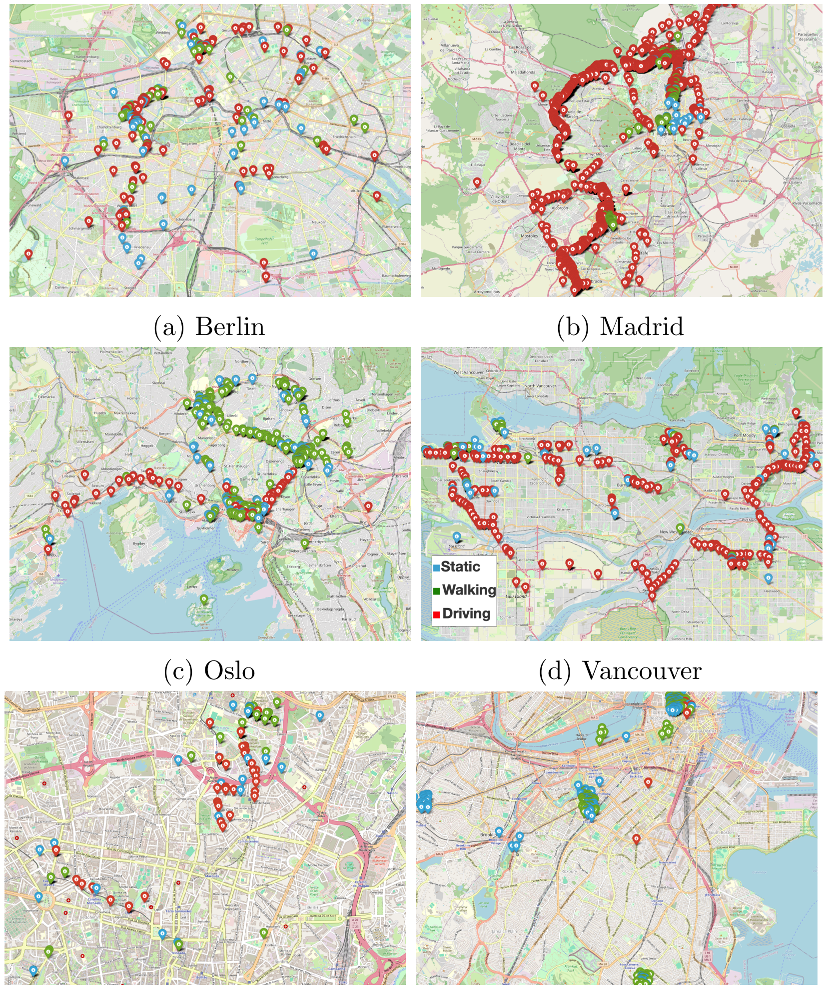


Figure 1: Workflow of the NextG-UP application.

several volunteers stopped performing tests after a couple of months. We omit such data, as they are not sufficient for a longitudinal study. Our final dataset, summarized in Table 1, consists of data from 8 cities in 7 different countries across Europe and North America, 1422 unique cell IDS, and 12 different operators.

In each city, one or two volunteers used our app to perform measurements with different mobility modes (static, walking, driving). Our dataset captures the average performance experienced by a user during their daily routine at home, office, or during their regular commute. The volunteers were asked to use all three mobility modes and perform at least a few measurements every week, however, they performed the tests at their convenience. As such, the total number of tests, their geographic spread (expressed as the radius of gyration[1] [28]), and the number of tests for each mobility mode vary significantly across cities (see Table 1). Fig. 2 shows the geographic distribution of measurement tests in 6 cities (we omit Turin and Bay Area,

[1]The radius of gyration measures the spatial spread of geographic locations around a central point, typically the mean or centroid of all locations. A larger radius indicates a wider spread of locations, while a smaller radius suggests more localized activities.



(a) Berlin (b) Madrid

(c) Oslo (d) Vancouver

(e) Porto (f) Boston

Figure 2: Geographic distribution of measurement test locations (Static: blue, Walking: green, Driving: red) in six cities based on mobility mode.

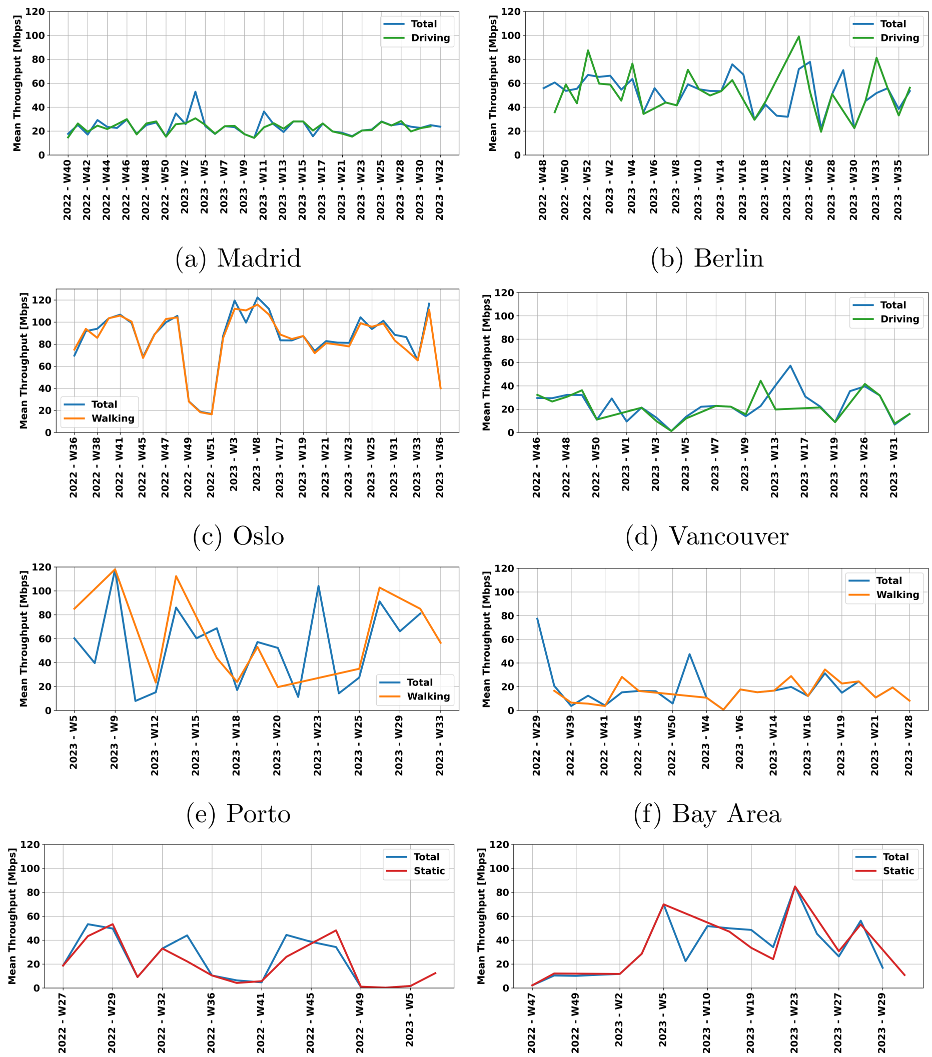
the two cities with the smallest number of tests). In some cities, e.g., Berlin (Fig. 2a), the number of tests is roughly balanced across the three mobility modes; in others, e.g., we observe a dominant mobility mode, e.g., driving in Madrid (Fig. 2b) and Vancouver (Fig. 2d) or walking in Oslo (Fig. 2c).

Controlled Measurements. In addition to the crowd-sourced measurement campaign, we conducted a year-long measurement campaign focused on 5G mmWave technology in Boston, USA. The 5G mmWave technology provides much higher data rates compared to low band or midband 5G [2, 3, 9], however, its limited deployment and sensitivity to blockage and mobility make it hard to assess its performance via crowdsourced measurements (indeed, our crowdsourced dataset includes only a small fraction of 5G mmWave data collected in 1 out of 8 cities, see Table 6). For this measurement campaign, we used Samsung S21 phones connected to XCAL Solo devices [26]. XCAL Solo taps into the Qualcomm diagnostic (Diag) interface of the smartphone and extracts lower layer KPIs and control-plane signaling messages, which are not available via the Android API, allowing for a more in-depth study. All our measurements were conducted over Verizon with a Wavelength server deployed in Boston.

We performed two types of tests, static and mobile, twice a week over the one-year period – on a weekday (Wednesday) and on a weekend day (Sunday) – at a designated spot in the downtown area of the city. All static tests were performed in front of the same 5G mmWave base station, with the phone facing towards and away from the base station. Each static test was repeated 5 times. Similarly, all the mobility tests were performed on a fixed route near the same base station; the user walked from a point A to a point B laterally to the base station and returned to point A. Performing all the tests at the same location with the same base station allows us to closely track any infrastructure changes and performance upgrades by the operator over the one-year period. We also repeated both the static and mobile tests using two phones simultaneously sending backlogged TCP traffic to track changes in the resource sharing policy used by the operator.

4. Longitudinal study

In this section, we use our crowdsourced dataset to explore the first condition for calling a technology mature, as defined in §1: does the 5G performance remain stable over a long time, without an increasing trend? To answer this question, we perform linear regression on the LTE and 5G throughput and latency values (averaged over each week) over time (the time unit is weeks) and show the results (slope and p-value) in Tables 4 and 5, respectively. Here, the slope, measured in Mbps/week (throughput) or ms/week (latency), represents the rate of change in the weekly averaged throughput/latency. A low p-value (typically <0.05) suggests strong evidence against the null hypothesis, implying a statistically significant relationship between throughput/latency and time, with the null hypothesis stating that there is no relationship between weekly averaged throughput/latency and time.



(a) Madrid (b) Berlin

(c) Oslo (d) Vancouver

(e) Porto (f) Bay Area

(g) Boston (h) Turin

Figure 3: Evolution of 5G in terms of throughput. Throughput samples collected over a week are averaged and the average throughput over each week is plotted as a timeline. Blue represents the total samples collected by static, walking, and driving measurements combined. Red represents static, orange represents walking, and green represents driving samples only.

We observe that the slopes for both technologies and both metrics are very close to 0 in all cities, indicating no increasing/decreasing trend of throughput and latency over the one-year period we consider in our study.

Table 4: Trend Analysis of Weekly Average Throughput Over Time Using Linear Regression. The unit of slope is Mbps/week

City	**slope**		**p-value**	
	LTE	**5G**	**LTE**	**5G**
Berlin	5.3e-07	-5.6e-07	0.08	0.05
Turin	1.4e-06	1.3e-06	3e-4	4e-4
Oslo	6.6e-07	6.0e-07	0.008	0.04
Porto	-7.6e-07	9.9e-07	0.006	0.1
Madrid	1.77e-07	-3.2e-08	1.56e-13	0.44
Vancouver	-1.6e-07	6.6e-08	0.22	0.62
Boston	-6.4e-07	-1.4e-06	0.12	0.03
Bay Area	-4.1e-07	-7.7e-08	0.68	0.74

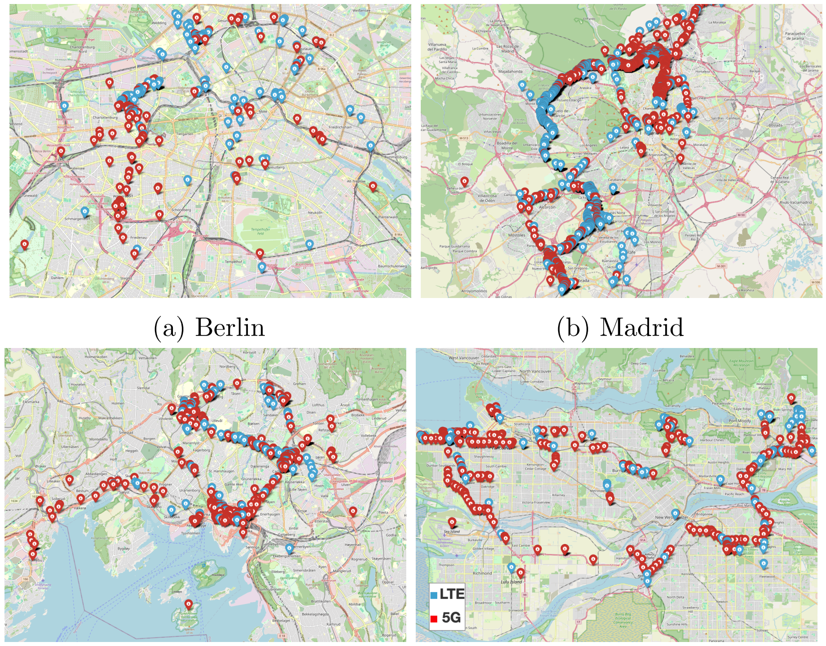
Table 5: Trend Analysis of Weekly Average Latency Over Time Using Linear Regression. The unit of slope is ms/week

City	**slope**		**p-value**	
	LTE	**5G**	**LTE**	**5G**
Berlin	-1.5e-06	-9.8e-07	0.39	0.13
Turin	-1.6e-06	4.5e-07	0.09	0.22
Oslo	-4.3e-07	2.5e-07	0.7	3.9e-19
Porto	-8.6e-07	8.6e-07	0.01	0.18
Madrid	-1.9e-06	-3.9e-06	0.29	0.01
Vancouver	1e-05	3.9e-06	0.01	0.3
Boston	6.1e-06	4.4e-06	3.1e-12	2.6e-06
Bay Area	-1.7e-07	-1.2e-07	0.8	0.8

Similarly, p values are typically (much) higher than 0.05 meaning that the throughput and latency do not show a statistically significant relationship with time. While this is expected for LTE (a mature technology), it is rather surprising for 5G four years after its initial rollout.

We further show the evolution of 5G throughput over time for each city in Fig. 3. For each city, we plot the average throughput per week over all the tests and over the dominant mobility mode – driving in Madrid, Vancouver, and Berlin, walking in Oslo, Porto, and Bay Area, static in Turin and Boston. The plots confirm our conclusions from the linear regression study. While throughput can vary significantly from one week to the next, we observe no increasing trend.

Overall, our results show that the first condition for maturity is satisfied: *5G deployment appears to have reached a mature stage in major cities in Europe and North America with no major performance improvements over the past one year.*



(a) Berlin (b) Madrid

(c) Oslo (d) Vancouver

Figure 4: Geographic distribution of measurement test locations in four cities based on cellular technology (LTE: blue, 5G: red).

5. Cross-sectional study

We now turn our attention to the remaining two conditions for maturity: does 5G offer higher coverage and better performance than LTE? Is the 5G coverage and performance similar across diverse geographic locations and operators? We study coverage in §5.1 and performance in §5.2–§5.6 using our crowd-sourced dataset.

5.1. 5G coverage

Coverage for a particular technology (5G or LTE) in terms of throughput is calculated as the fraction of throughput samples collected over that technology out of the total number of throughput samples. Similarly, RTT coverage for a particular technology is the fraction of RTT samples collected over that technology out of the total number of RTT samples. A sample equation for calculating 5G coverage for throughput is provided in Eq. 1.

$$\text{Coverage 5G}_{\text{throughput}} = \frac{\text{Number of throughput samples over 5G}}{\text{Total number of throughput samples}}. \quad (1)$$

Table 6: Technology coverage, expressed as the fraction of the number of throughput/RTT samples over a particular technology out of the total number of samples

City (Country)	Throughput LTE	Throughput 5G			RTT LTE	RTT 5G			Mobility Mode (LTE / 5G) Static	Walking	Driving
Berlin (Germany)	0.53	0.47			0.53	0.47			0.17/0.14	0.12/0.10	0.24/0.23
Turin (Italy)	0.68	0.32			0.66	0.34			0.36/0.14	0.29/0.16	0.02/0.03
Oslo (Norway)	0.36	0.64			0.31	0.69			0.1/0.16	0.24/0.45	0.02/0.03
Porto (Portugal)	0.18	0.82			0.16	0.84			0.27/0.05	0.06/0.30	0.07/0.25
Madrid (Spain)	0.53	0.47			0.55	0.45			0.01/0.01	0.03/0.12	0.51/0.32
Vancouver (Canada)	0.41	high -	mid 0.57	low 0.02	0.38	high -	mid 0.60	low 0.02	0.13/0.13	0.06/0.13	0.21/0.34
Boston (USA)	0.60	high 0.11	mid 0.26	low 0.03	0.62	high 0.24	mid 0.13	low 0.01	0.33/0.12	0.17/0.25	0.11/0.02
Bay Area (USA)	0.08	high -	mid 0.88	low 0.04	0.10	high -	mid 0.86	low 0.04	0.30/0.03	0.03/0.62	0.01/0.01
Total	0.52	0.48			0.52	0.48			0.08/0.06	0.08/0.17	0.37/0.24

Table 6 shows the results for each city as well as the overall results. We observe that the results are very similar with both metrics; hence, we focus on the throughput results in the remainder of this section.

Table 6 shows that *the overall 5G coverage is moderate*; in total, 52% of the throughput samples were collected while the UE was connected to a 5G cell. However, *coverage varies significantly across cities and operators*. The largest 5G coverage is observed in the Bay Area with T-Mobile (92%) and Porto with MEO (82%), and the lowest in Turin with TIM and WINDTRE combined (only 32%). Interestingly, the two US cities exhibit very different 5G coverage – 92% in the Bay Area with T-Mobile vs. 40% in Boston with all three major US operators combined.

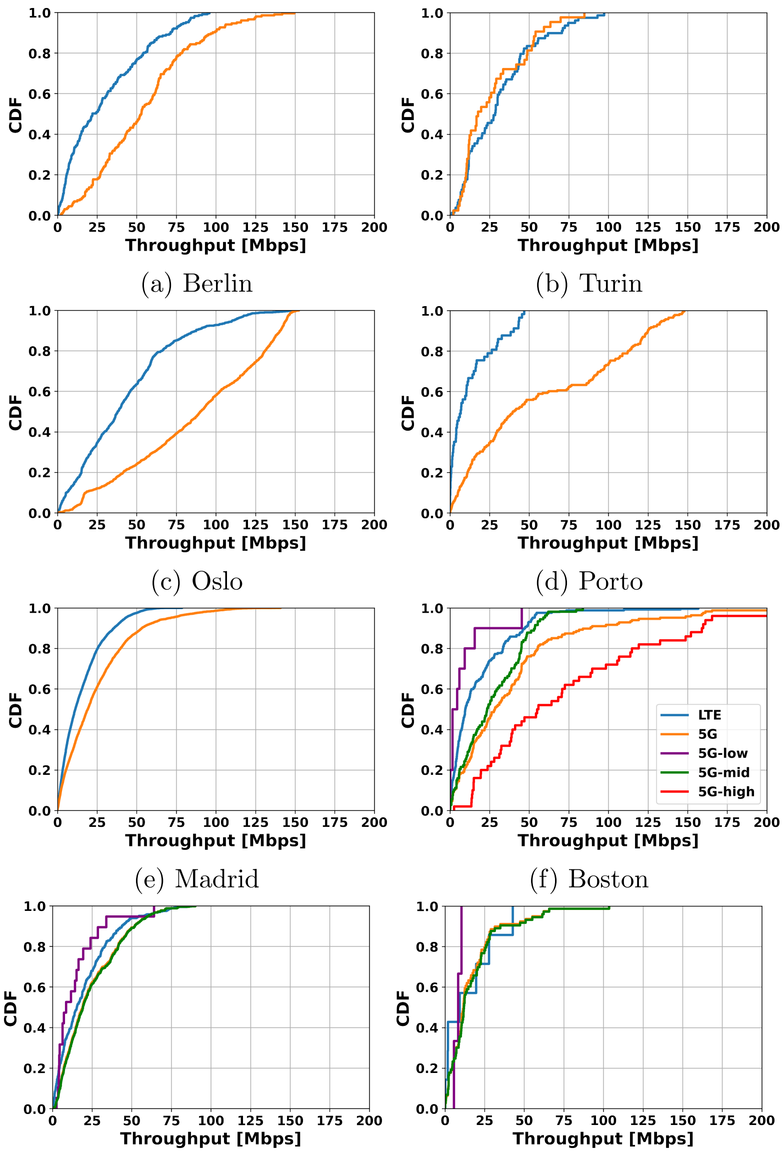
We also break down the 5G coverage based on the frequency band in 5G-low, 5G-mid, and 5G-high (mmWave) using the Absolute Radio Frequency Channel Number (ARFCN), recorded by our app. Unfortunately, the Android API that returns the ARFCN failed in all the tests conducted in Europe; hence, this information is only available for tests in North America. Nonetheless, 5G in Europe is primarily deployed in the midband (band n78) [29]. When we compare the North American locations, we observe almost exclusively 5G-midband in the Bay Area with T-Mobile and Vancouver with Bell and Shaw Comm., as these operators do not use mmWave. On the other hand, in Boston, we observe 11% of the 5G throughput samples and 24% of the RTT samples over 5G-high, mainly with Verizon and AT&T. This result is in sharp contrast with a recent study [9] that reported a significant 5G-low coverage, mainly with T-Mobile and AT&T during a cross-country drive, suggesting that 5G-low is mainly used in highways thanks to its longer coverage, while the mid and high bands are preferred in cities to provide high throughput.

We next look at the geographic coverage of the two technologies, focusing on the four cities from which we collected the largest number of measurements in Fig. 4. In Berlin, which has a balanced coverage for the two technologies (53% LTE, 47% 5G), interestingly, we observe a large aggregation of tests over 5G southwest of the city center, while most of the tests around the city center were done over LTE (Fig. 4a). In Madrid (Fig. 4b), with similar 5G coverage as Berlin, we observe two major areas of high 5G coverage and one area with mostly LTE coverage, but also areas with both technologies present. In contrast, in Oslo (Fig. 4c) and Vancouver (Fig. 4d), where 5G coverage is significantly higher compared to Berlin and Madrid (64% and 59%, respectively), we observe no area where LTE is the prevalent technology. In areas with both technologies present, we observe tests over different technologies at locations geographically very close to each other.

We also explore the relationship between 5G coverage and the geographic spread of the measurements in each city. Tables 1 and 6 show that the two cities with the shortest radius of gyration (Oslo and Porto) have the 2$^{\text{nd}}$ and 3$^{\text{rd}}$ highest 5G coverage among the 8 cities (82% and 64%, respectively). However, we also observe cities with similar radius of gyration (Berlin, Turin, Bay Area), where the 5G coverage varies significantly (from 32% to 88%). We also note that the city with the largest radius of gyration (Vancouver) has much higher 5G coverage (59%) than other cities with much smaller radius. Overall, we do not observe any clear relationship between 5G coverage and the the geographical spread of the measurements.

We finally explore the impact of the user's mobility mode on coverage. Table 6 shows that the coverage for a given mobility mode typically follows the same trend as the overall coverage. The only exception is Madrid, where 5G coverage is higher than LTE coverage during walking but lower during driving. While the same is also true for Boston, 5G coverage is also much lower than LTE coverage in Boston for static scenarios, suggesting that the user speed is not a critical factor.

In summary, our results in this section show that conditions (ii) and (iii) are not satisfied with respect to coverage across the 8 cities in our study. *Users are still connected to LTE about 50% of the time on average and coverage is very different across different locations and operators*, ranging from an impressive 92% to a disappointing 32%.



(a) Berlin (b) Turin

(c) Oslo (d) Porto

(e) Madrid (f) Boston

(g) Vancouver (h) Bay Area

Figure 5: Throughput comparison across different cities. LTE: blue, 5G (5G-low/mid/high combined): orange, 5G-low: purple, 5G-mid: green, 5G-high: red. Not all 5G bands are available in every North American city. For instance, Vancouver lacks mmWave (5G-high) base stations, and T-Mobile in the Bay Area does not offer 5G mmWave service.

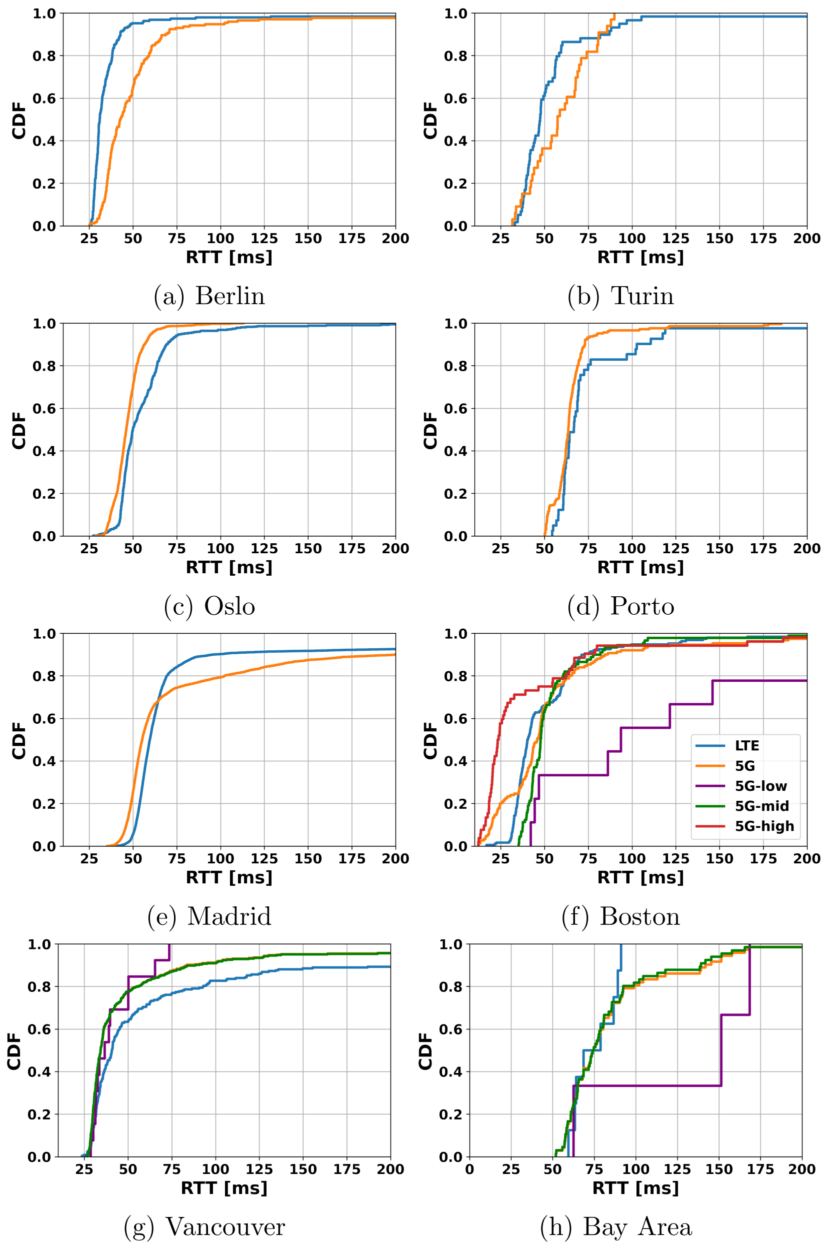
5.2. Throughput

Fig. 5 plots the CDFs of uplink 5G and LTE throughput in each of the 8 cities. For the 3 cities in North America, we further break down the 5G throughput into 5G-low, 5G-mid, and 5G-high. We observe that *5G offers higher throughput than LTE in 7/8 cities*. However, *the median gain varies significantly across cities*, from 2.36 Mbps in the Bay Area to 52.23 Mbps in Oslo, showing that *four years after its initial rollout, 5G does not always deliver the high throughput gains it promised*. Interestingly, in these two cities, the maximum 5G throughput is similar to the LTE throughput. In all the other cities (with the exception of Turin), the maximum 5G throughput is higher than the maximum LTE throughput, typically by several tens of Mbps up to 100 Mbps.

Two exceptions are worth noting – Bay Area and Turin.[2] In Bay Area, the location with the highest 5G coverage (92%), 5G throughput is largely similar to LTE throughput, although it exhibits a much longer tail, indicating that better coverage does not necessarily translate to better user experience. Even more surprisingly, in Turin, 5G offers lower throughput than LTE. After contacting the volunteer in Turin, we found out that initially they used WindTre with a 5G subscription of a maximum rate of 10 Mbps throughput, and later they switched to using TIM as an operator, with an unlimited subscription. While the rate limiting imposed by WindTre explains the lower 30% of the samples in Fig. 5b, the remaining samples also exhibit very low throughput values of at most 85 Mbps.

Among the three different 5G bands in North America, 5G mmWave offers the highest throughput, followed by 5G-mid and then by 5G-low, as expected. Interestingly, our small number of 5G-low samples exhibit lower median and maximum throughput than LTE in all three cities. Further, even though 5G midband is viewed as the band that offers the best tradeoff between range and performance, our results show that the gains over LTE in the uplink direction are quite low – 2.69 Mbps in the Bay Area, 3.91 Mbps in Vancouver, and 13.65 Mbps in Boston in the median case. Interestingly, in Boston, we observed a maximum 5G-mid throughput of 80 Mbps while the maximum LTE throughput exceeded 150 Mbps.

[2]Note that these are the two locations with the smallest number of runs, and hence, the results may not be fully representative.



(a) Berlin (b) Turin (c) Oslo (d) Porto (e) Madrid (f) Boston (g) Vancouver (h) Bay Area

Figure 6: Latency comparison across different cities. LTE: blue, 5G (5G-low/mid/high combined): orange, 5G-low: purple, 5G-mid: green, 5G-high: red. Not all 5G bands are available in every North American city. For instance, Vancouver lacks mmWave base stations, and T-Mobile in the Bay Area does not offer 5G mmWave service.

5.3. Latency

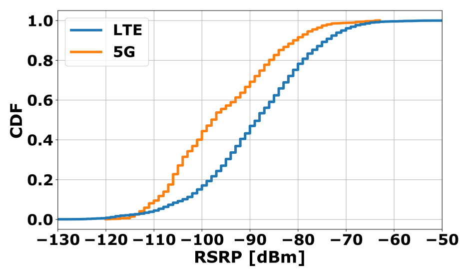
Fig. 6 plots the CDFs of uplink 5G and LTE latency in each of the 8 cities. For the 3 cities in North America, we further break down the 5G throughput into 5G-low, 5G-mid, and 5G-high. *Although 5G promises a significantly lower latency than LTE, our results in Fig. 6 surprisingly show that this is typically not the case.* 5G offers lower latency than LTE only in 3 out of 8 cities and the improvements are marginal. The median values for 5G vs. LTE latency in these three cities are – 46 ms vs. 50 ms in Oslo, 64 ms vs. 67 ms in Porto, and 34 ms vs. 41 ms in Vancouver. In the remaining 5 cities, the 5G latency is similar to or higher than the LTE latency. In Boston, latency is similar for the two technologies, although 5G offers lower best-case latency (25 ms vs. 34 ms at the 20-th percentile). In Madrid, 5G offers lower latency than LTE in the median case (55 ms vs. 60 ms) but significantly higher at the 80-th percentile (102 ms vs. 69 ms). In the Bay Area, latency is similar for the two technologies, but 5G has a much higher worst-case latency (e.g., 142 ms vs. 90 ms at the 90-th percentile). Finally, in Berlin and Turin, 5G latency is higher than LTE latency – 43 ms vs. 31 ms and 57 ms vs. 47 ms in the median case, respectively. In fact, in Berlin, the upper quartile of the LTE latency is equal to lower quartile of the 5G latency.

We ran a few traceroute tests to the AWS Frankfurt server from Berlin (the city with the largest gap between 5G and LTE latency) over 5G and LTE and found that the path is the same over both technologies. This suggests that the root cause for the higher 5G latency lies in the RAN. We plan to further investigate this as part of our future work.

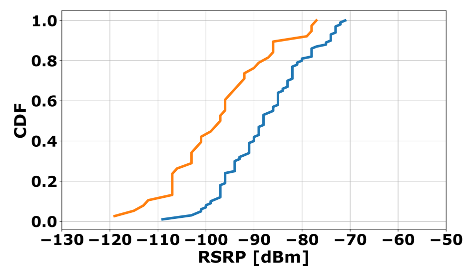
When we compare the three different bands in North America, we observe that 5G-high in Boston over Verizon, combined with an edge AWS Wavelength server, offers significantly lower latency than all the other technologies and is responsible for the lowest 10-th percentile of the overall 5G latency in Boston in Fig. 6f. On the other hand, the 5G-low and 5G-mid latency is higher than the LTE latency in the two US locations but lower in Vancouver. In particular, the 5G-low latency is very high in Boston and Bay Area, but given the very small number of samples, it does not contribute significantly to the overall latency, which is mainly affected by the 5G-mid samples.

5.4. Impact of signal strength

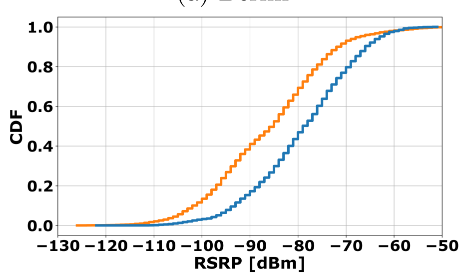
In this section, we compare the signal strength of the two technologies and their correlation with performance. Fig. 7 plots the CDFs of the Reference Signal Received Power (RSRP) for 5G and LTE in each of the 8 cities. We



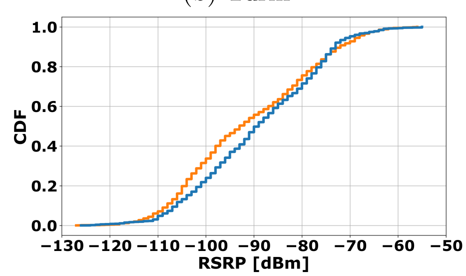
(a) Berlin



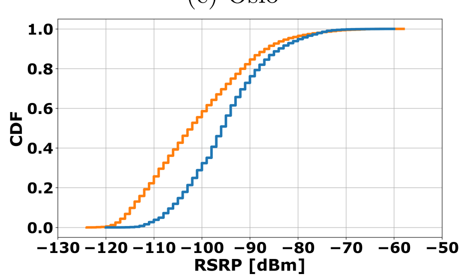
(b) Turin



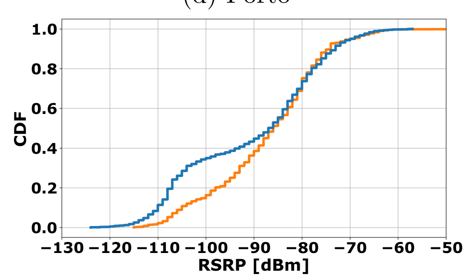
(c) Oslo



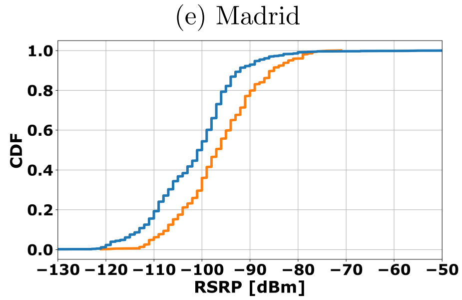
(d) Porto



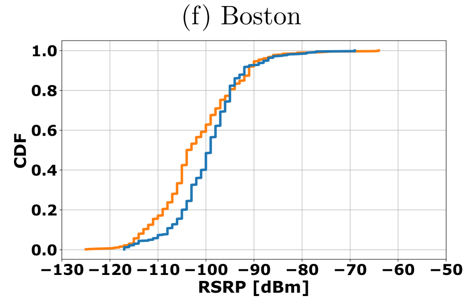
(e) Madrid



(f) Boston



(g) Vancouver



(h) Bay Area

Figure 7: RSRP comparison across different cities (LTE: blue, 5G: orange).

observe that the RSRP is lower over 5G than over LTE in 7/8 cities; the gap varies from -4 dB (Porto) to -10 dB (Turin) in the median case. Boston and

Vancouver are the only two exceptions. However, the impact of RSRP on throughput and latency is different across different cities.

Throughput. In Boston and Vancouver, the two cities where RSRP is higher over 5G than over LTE, the 5G throughput is also higher than the LTE throughput (Figs. 5f, 5g). Among the remaining 6 cities (Figs. 5a, 5b, 5c, 5d, 5e, 5h), 5G yields higher throughput than LTE in four of them (Berlin, Oslo, Porto, Madrid) but lower or similar in the other two (Turin, Bay Area). The availability of wider channels in 5G NR than in LTE is the main reason for the overall higher throughput observed with 5G than with LTE in spite of the lower signal strength. 5G NR channel bandwidths of the operators under analysis are at least four times larger than the maximum LTE channel bandwidth (i.e., 20 MHz). For example, previous measurement studies in Spain, France, Germany, and Italy show channel bandwidths in the range 80-100 MHz [19]. As operators try to allocate the maximum number of frequency resources per user with bulk transfers like our throughput experiments [19], the use of robust modulation schemes is sufficient to explain the reason behind the reported higher throughput with 5G despite a lower signal strength.

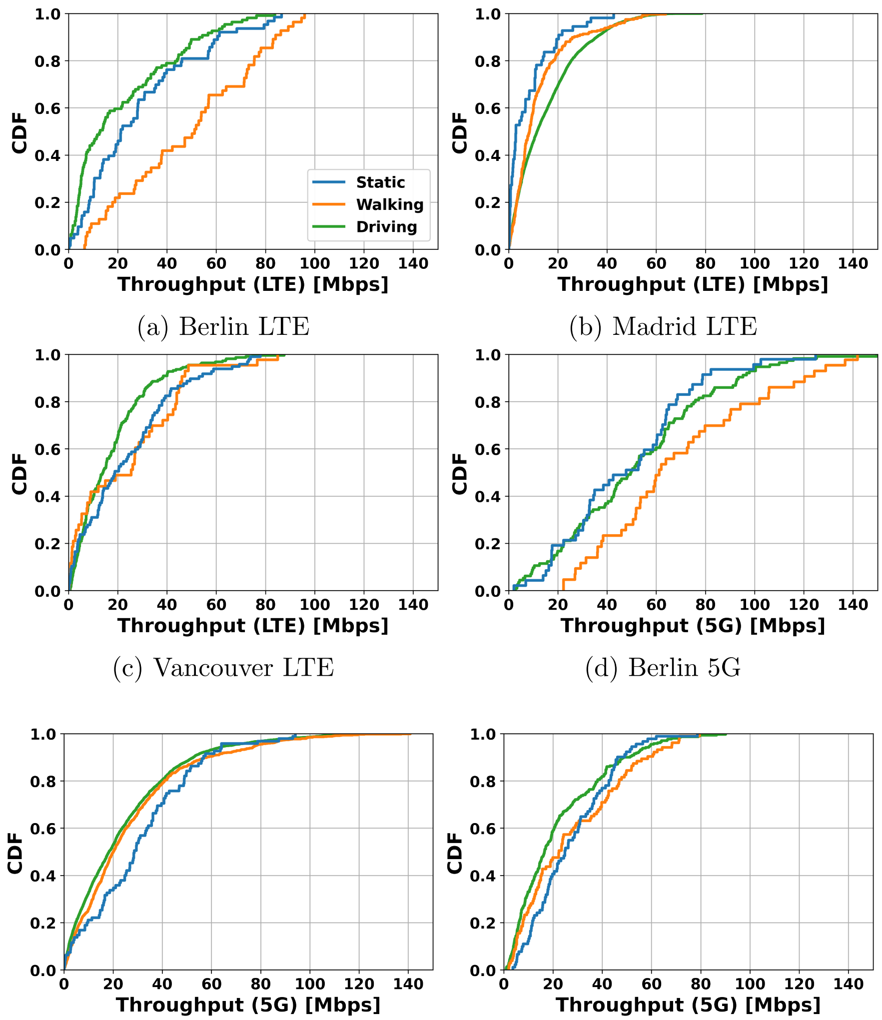
Latency. The higher 5G RSRP results in lower 5G latency in Vancouver (Fig. 6g), but only improves the worst-case 5G latency compared to the LTE latency in Boston (Fig. 6f). Note that in Boston (Fig. 7f) the 5G RSRP is higher than the LTE RSRP only for the lower half of the CDFs. Among the remaining 6 cities (Figs. 6a, 6b, 6c, 6d, 6e, 6h), the latency is lower over 5G than over LTE in two of them (Oslo, Porto), but similar or worse in the remaining four (Berlin, Turin, Madrid, Bay Area).

Overall, we observe that *RSRP has a weak correlation with performance but it appears to affect the latency more than the throughput.*

5.5. In-depth analysis of select cities

In this section, we analyze in depth the performance in three cities and explore the impact of mobility mode. We select Madrid and Vancouver, the two cities in Europe and North America, respectively, with the largest number of measurement tests, and Berlin as an example of a city with a good balance of tests with each mobility mode. Figs. 8 & 10 plot the technology-wise CDFs of throughput, latency, and RSRP, respectively, for each mobility mode in these three cities.

Berlin. Figs. 8a, 8d and 9a, 9d show that in Berlin both LTE and 5G exhibit the best performance (highest throughput and lowest latency) under walking. In contrast, the performance under static conditions is poor with both technologies and similar to that under driving, especially over 5G. Although in the previous section we concluded that RSRP alone cannot



(a) Berlin LTE (b) Madrid LTE

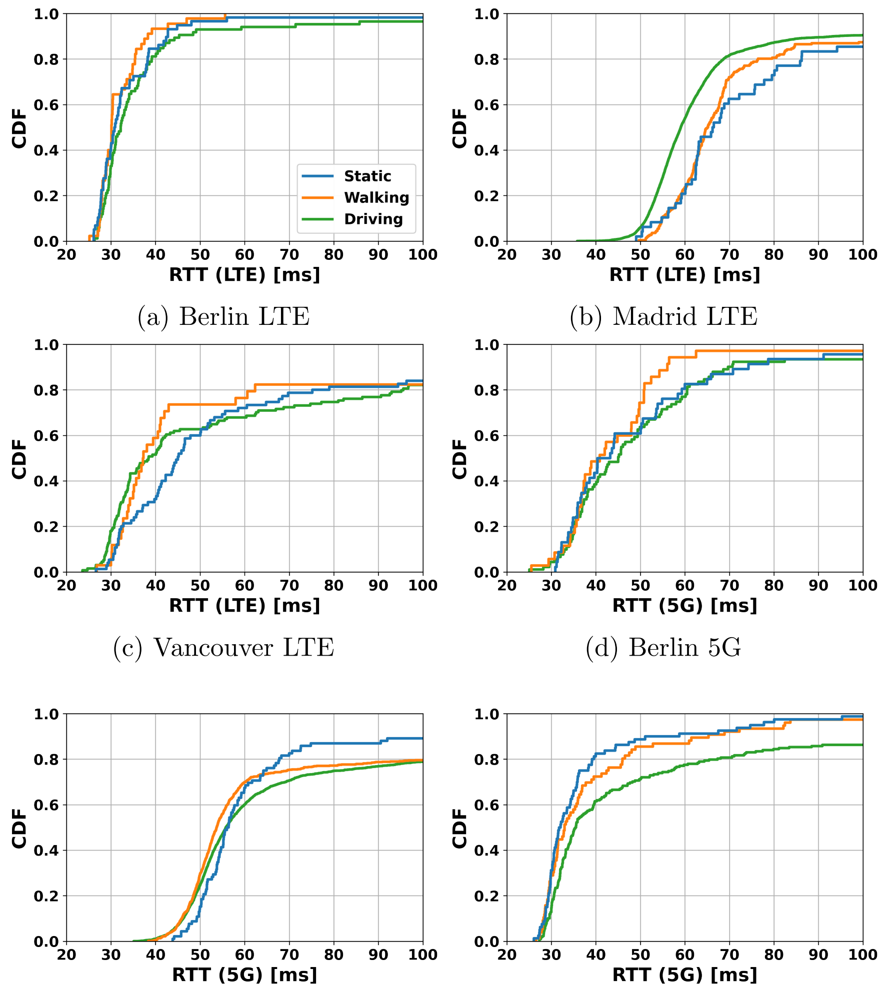
(c) Vancouver LTE (d) Berlin 5G

(e) Madrid 5G (f) Vancouver 5G

Figure 8: Throughput Comparison across different mobility modes (Static: blue, Walking: orange, Driving: green).

explain the performance difference between the two technologies, Figs. 10a, 10d show that RSRP can explain the performance for a given technology. These figures show that in Berlin, RSRP was high during walking tests and low during static and driving tests. Our volunteers in Berlin did the majority of the static tests indoors, which explains the low RSRP values and the low performance in static conditions.

Madrid. Figs. 8b, 8e and 9b, 9e show that in Madrid, static tests exhibit the



(a) Berlin LTE

(b) Madrid LTE

(c) Vancouver LTE

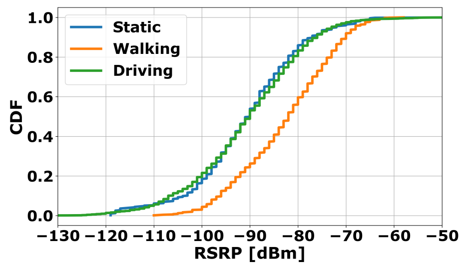
(d) Berlin 5G

(e) Madrid 5G

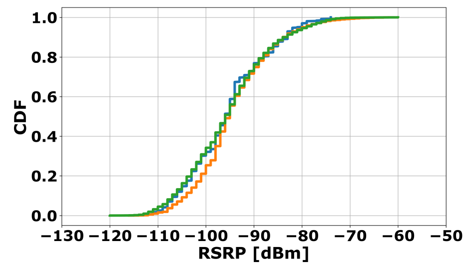
(f) Vancouver 5G

Figure 9: RTT Comparison across different mobility modes (Static: blue, Walking: orange, Driving: green).

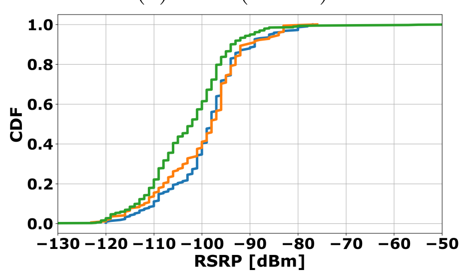
worst performance over LTE but the best performance over 5G. Interestingly, driving exhibits the best performance over LTE but the worst over 5G. Walking also exhibits poor performance – worse than driving over LTE and similar to driving over 5G. However, the RSRP in Madrid is similar across all three mobility modes for each technology (Figs. 10b, 10e), and hence, it cannot explain the performance, unlike in Berlin. Several 5G walking tests were run outdoors around the volunteer's apartment building where there is



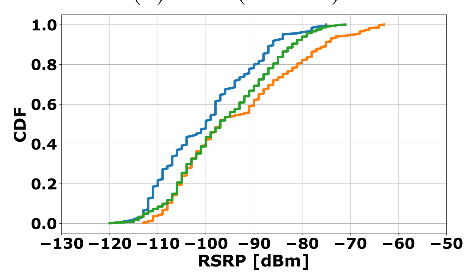
(a) LTE (Berlin)



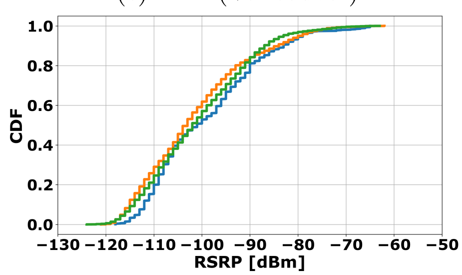
(b) LTE (Madrid)



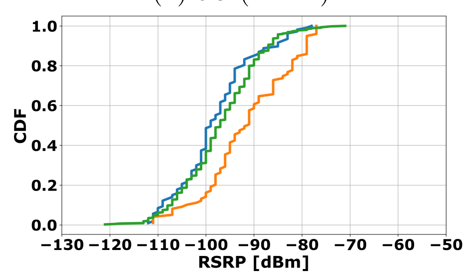
(c) LTE (Vancouver)



(d) 5G (Berlin)



(e) 5G (Madrid)



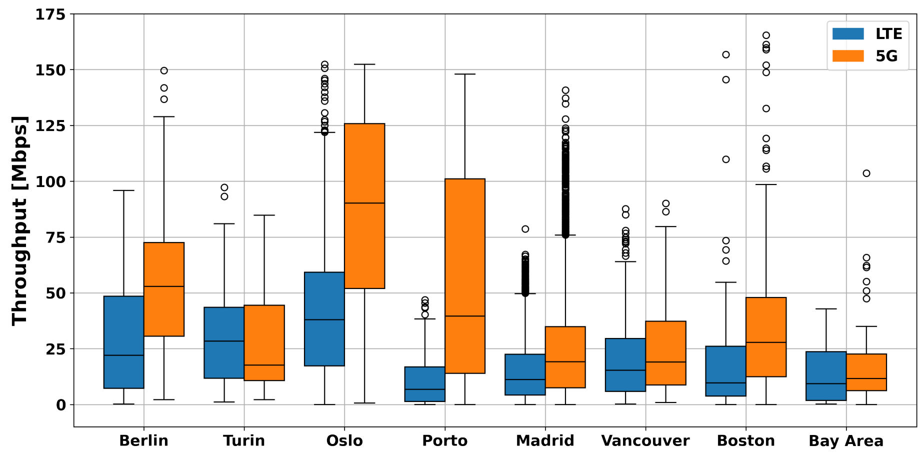
(f) 5G (Vancouver)

Figure 10: RSRP comparison across different mobility modes (Static: blue, Walking: orange, Driving: green).

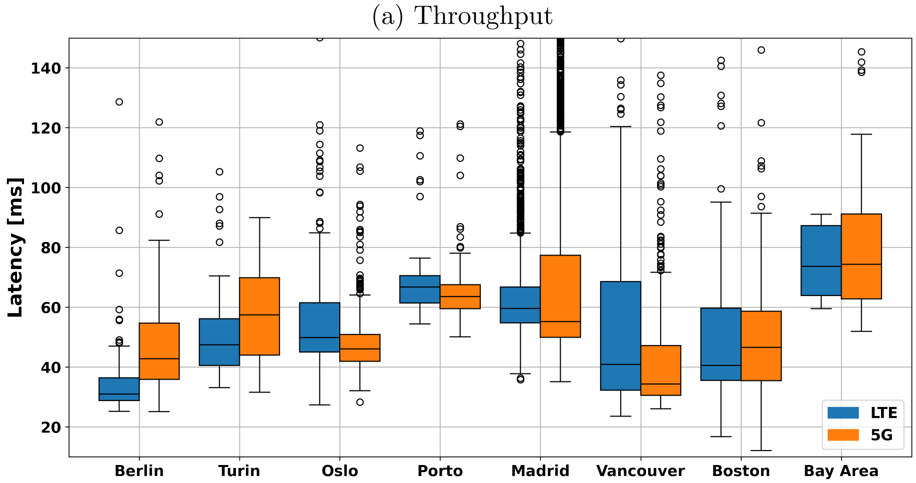
a 5G a tower installation from a different operator (Orange) than the one used for the measurements (Vodafone). Since the two operators have a RAN sharing agreement, we conjecture that interference from the other operator is responsible for the low 5G performance in that area.

Vancouver. Figs. 8c, 8f and 9c, 9f show that in Vancouver, driving exhibits the worst throughput over both LTE and 5G and the worst latency over 5G, but surprisingly not over LTE. Static and driving tests, conducted outdoors in Vancouver, exhibit similar throughput, better than driving tests over both technologies. However, that latency is the best over 5G but the worst over LTE (in the median case). Fig. 10c shows that the LTE RSRP is similar for static and walking tests and much higher than for driving tests, which explains the throughput results but not the latency results. Fig. 10f shows that 5G RSRP was the lowest under static conditions (much lower than under walking), yet static tests exhibit the best latency and similar throughput to walking tests over 5G.

Overall, we observe that *cellular performance is the result of the complex interplay among a large number of factors and cannot be explained by looking individually at a single factor*. Previous works also arrived at similar conclusions, showing a poor correlation of cellular throughput with RSRP [9, 18] and UE speed [9].



(a) Throughput



(b) Latency

Figure 11: Technology-wise comparison across different cities (LTE: blue, 5G: orange).

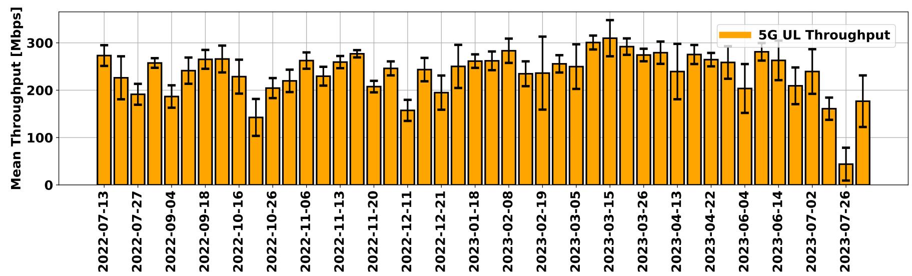
5.6. Overall performance across all cities

In the previous section, we focused on the comparison between 5G and LTE performance and showed that the second condition for maturity is not satisfied. In this section, we turn our attention to the third condition and compare the performance of a given technology across cities in Fig. 11.

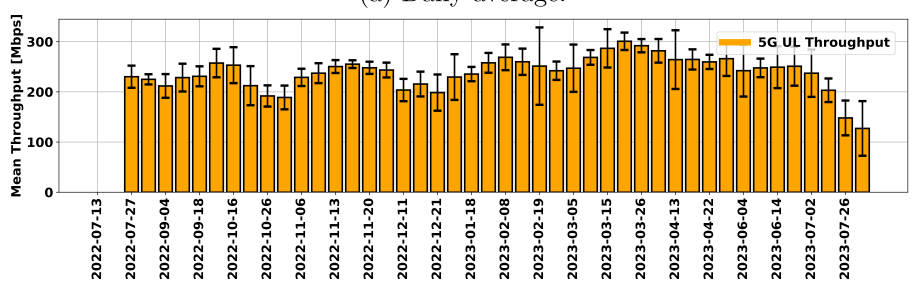
Throughput. Fig. 11a shows that Oslo has the highest overall 5G throughput across the 8 cities, with a median/75-th percentile of 88/125 Mbps. Berlin comes second in terms of median throughput (52 Mbps vs. Porto's 38 Mbps), but Porto has a much higher 75-th percentile (101 Mbps vs. 74 Mbps). On the other hand, Bay Area has the lowest 5G throughput among the 8 cities, with a median/75-th percentile of 12/23 Mbps. Note that Oslo's lower quartile of 5G throughput is higher than the upper quartile of all cities except Berlin and Porto. Fig. 11a also shows that Oslo exhibits the highest LTE throughput with a median/75-th percentile of 40/59 Mbps, followed by Berlin and Turin. Interestingly, the median LTE throughput in Oslo matches the median 5G throughput in Porto and is higher than the 75-th percentile of the 5G throughput in Turin, Madrid, Vancouver, Boston, and Bay Area.

Overall, *we observe a large disparity among the 5G throughput values across the 8 cities, suggesting that the third condition for maturity is not satisfied*. We also observe a much larger spread of throughput values for 5G compared to LTE. Oslo and Porto, the two cities with the highest 75-th percentiles also exhibit the largest IQR (74 Mbps and 87 Mbps, respectively). Note that these two cities have the lowest geographic sample spread, indicating that 5G throughput exhibits strong variations even in limited geographic areas, and further reinforcing our conclusion that the third condition for maturity is not met yet.

Latency. A direct latency comparison among different cities is challenging, as the server location has a much higher impact on RTT than on throughput. For example, it is not surprising that Berlin exhibits the lowest median and lower quartile values for both 5G and LTE latencies among the 5 European cities in Fig. 11b, given that its distance to the Frankfurt AWS server we used for the measurements in Europe is the shortest. Yet, a few interesting observations are worth noting. First, Berlin's upper quartile for the 5G latency is higher than Oslo's, even though Oslo's distance from the Frankfurt AWS server is much longer. Note that Berlin is the city with the largest disparity between 5G and LTE latency. Second, in Boston 24% of the 5G RTT measurements were done over Verizon to an AWS Wavelength server located in the same city resulting in very low latency (notice the low whisker of the 5G boxplot in Boston in Fig. 11b), but for the remaining tests to an AWS server in North Virginia, the 5G latency is higher than in Vancouver, where



(a) Daily average.



(b) Rolling (3-day) average.

Figure 12: Single UE static throughput facing the base station.

the measurements were performed to a server located in Oregon. Third, we observe again a large disparity in the IQRs among different cities. Oslo and Porto, two cities with a large distance to the Frankfurt AWS server exhibit low IQRs for both 5G and LTE, suggesting the the latency is dominated by the wired network. On the other hand, for Madrid, which is also located far from the Frankfurt server, we observe a low IQR for LTE but not for 5G.

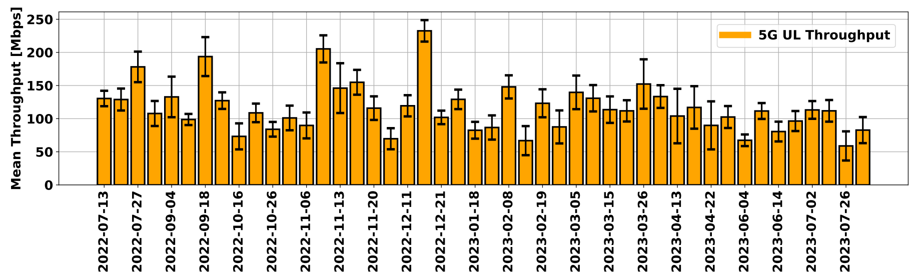
6. Case study: mmWave evolution in a single city

In this section, we use the dataset obtained from our controlled measurement campaign in Boston and analyze the evolution of 5G mmWave performance over a one-year period at a fixed location, while maintaining connectivity to the same base station(s). This analysis aims to uncover the variations in performance throughout the year, driven by potential infrastructure or policy changes by the operator at a fix location.

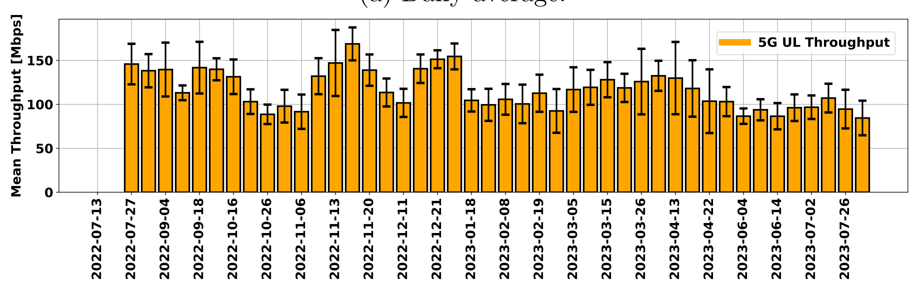
6.1. Single UE, static tests

6.1.1. Throughput

Figs. 12a, 13a show the average daily uplink throughput measured by a single phone facing towards/away from the mmWave base station, re-



(a) Daily average.



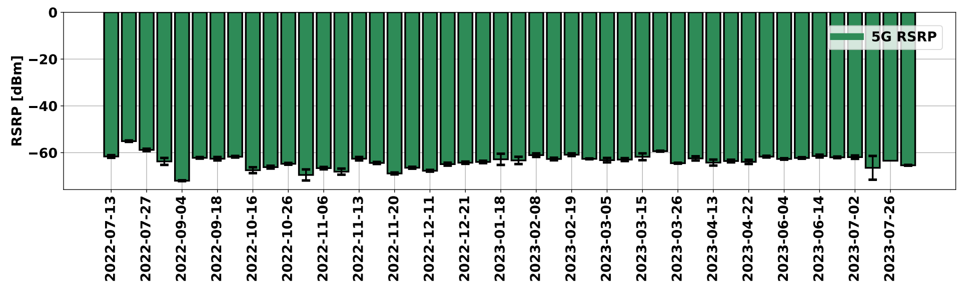
(b) Rolling (3-day) average.

Figure 13: Single UE static throughput facing away from the base station.

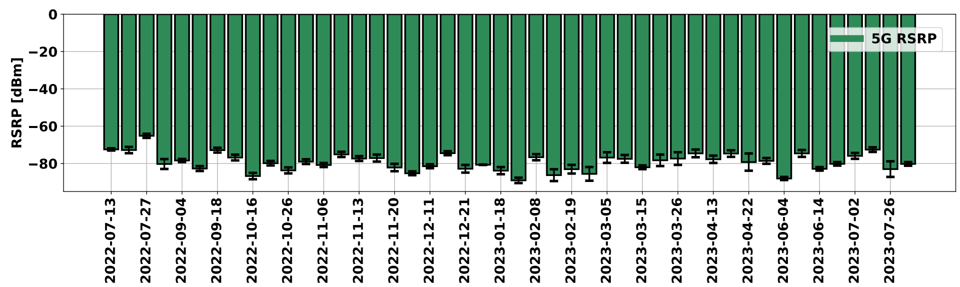
spectively, over the one-year period. We make two observations from these figures.

First, we observe that the throughput is in general higher than the LTE throughput obtained from our crowdsourced measurements in the city of Boston (50 Mbps at the 90th percentile, Fig. 5f). In particular, when the UE faces the base station, the average daily throughput is consistently higher than 100 Mbps (with the exception of one day – 07/26/2023) and can reach up to 300 Mbps. When the UE faces away from the base station, the average throughput values are lower due to blockage, but still above 50 Mbps.

Second, *we observe no clear increasing or decreasing trend during the one-year period*, which agrees with our crowdsourced results in §4. To reduce short-term fluctuations, Figs. 12b, 13b plot the rolling average throughput, calculated over a 3-day window of measurement data. Although we still observe no clear trend, we do notice seasonal dips and peaks, e.g., in October 2022 and December 2022. Fig. 12b shows that, compared to July 2022 (the beginning of our measurement campaign), the average throughput increases in March 2023. Following this increase, there is a steep decline in throughput after that from April 2023 to July 2023. A similar pattern (an increase in March 2023 followed by a decline) is also observed in Fig. 13b, although the throughput never rises above the initial value (in July 2022) and the decline



(a) UE facing the base station.



(b) UE facing away from the base station.

Figure 14: Single UE daily average RSRP under static conditions.

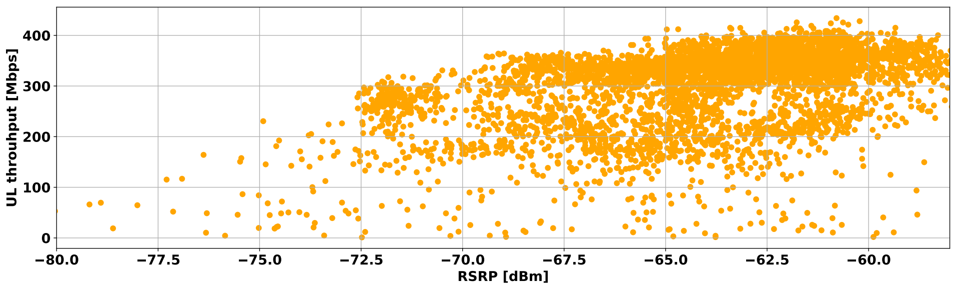
in July 2023 is not as steep as in Fig. 12b.

To understand the underlying cause of the varying uplink throughput across different days, in the following two sections, we dig deeper by analyzing lower layer metrics like RSRP and uplink Carrier Aggregation (CA).

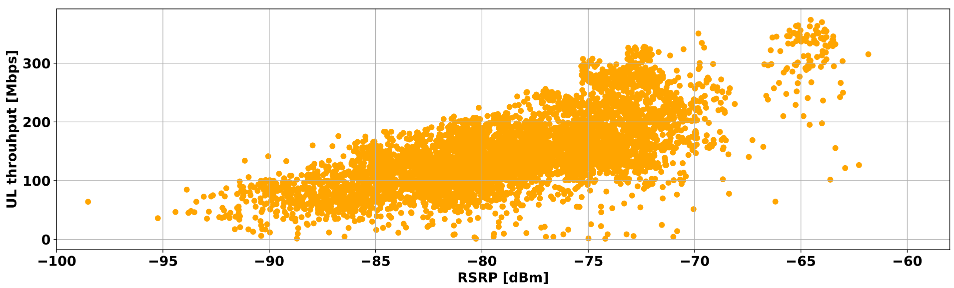
6.1.2. Impact of signal strength

Figs. 14, 14b plot the daily average RSRP, when the UE faces towards/away from the base station. We observe that the RSRP remains similar across different days, particularly when the UE faces the base station, and *exhibits a low correlation with throughput*. For example, the lowest average throughput in Fig. 12a is observed on 07/26/2023, but the RSRP is very high on that day in Fig. 14a. Similarly, the lowest RSRP value in Fig. 14b is observed on 07/27/2022, but the throughput is very high on that day in Fig. 13a.

To further analyze the RSRP-throughput correlation, Figs. 15a & Fig. 15b plot scatterplots of each 100 ms uplink throughput sample vs. the corresponding RSRP sample. The general trend shows that uplink throughput increases with RSRP; however, significant variability exists even at higher



(a) UE facing the base station.



(b) UE facing away from the base station.

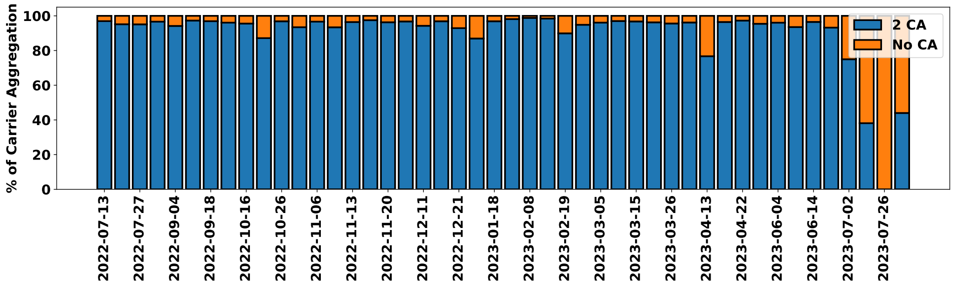
Figure 15: Single UE throughput vs. RSRP.

RSRP values. For example, in Fig. 15a we observe some high throughput samples ($\geq$ 300 Mbps) are achieved at moderate RSRP levels of -70 dBm, which is at least 10 dBm lower than the best RSRP values recorded. Conversely, low throughput values (0–100 Mbps) are also observed at strong RSRP levels ($\geq$ -65 dBm). The variance is smaller for measurements taken with the phone facing away from the base station, as shown in Fig. 15b, but it still exists.

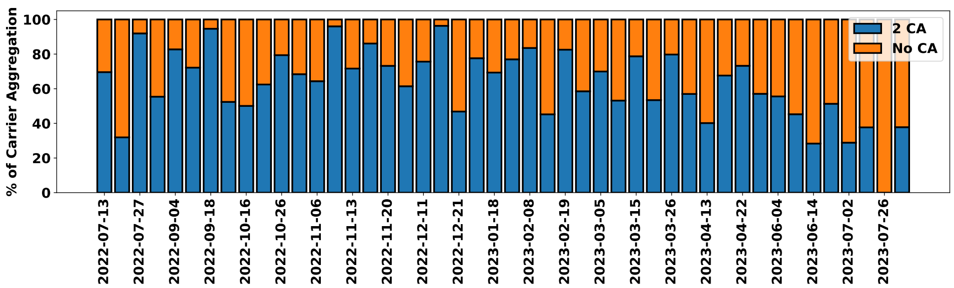
Our observations in this section align with the findings from our previous work [9], which showed that throughput in cellular networks is affected by multiple factors beyond signal strength.

6.1.3. Impact of carrier aggregation

In this section, we take a look at the uplink carrier aggregation (CA) distribution in Fig. 16. Carrier Aggregation (CA) is a technology that combines multiple channels (referred to as component carriers or CCs) within the same frequency band or across different bands to increase bandwidth and improve data speed. Its worth noting that the UE supports a maximum of 2 mmWave carriers, each with a bandwidth of 100 MHz, for uplink



(a) UE facing the base station.



(b) UE facing away from the base station.

Figure 16: Single UE CA distribution under static conditions.

communication. Similar to RSRP, Fig. 16a shows very low variability in carrier aggregation when the UE faces the base station. Two carriers are used most of the time except for the last 3 days of our measurement campaign. Out of these 3 days, the UE used 2 carriers 40-45% of the times on two days, while on 07/26/2023, the UE exclusively used 1 carrier. As a result, in Fig. 12a, we clearly observe that the average throughput is lower on these three days compared to most other days of the year. In particular, on 07/26/2023, we observe the lowest throughput of our measurement campaign in Fig. 12a. The amount of carrier aggregation drops significantly when the UE faces away from the base station, as shown in Fig. 16b, suggesting a correlation between carrier aggregation and throughput. Note again the absence of carrier aggregation on 07/26/2023 in Fig. 16b, which also corresponds to the lowest throughput observed when the UE faces away from the base station in Fig. 13a.

Overall, the results in this section show a *stronger correlation between uplink throughput and carrier aggregation compared to RSRP*. Nonetheless, we also observe some exceptions – high carrier aggregation and low throughput (e.g., on 12/11/2022 in Fig. 12a) or very different throughput values for the

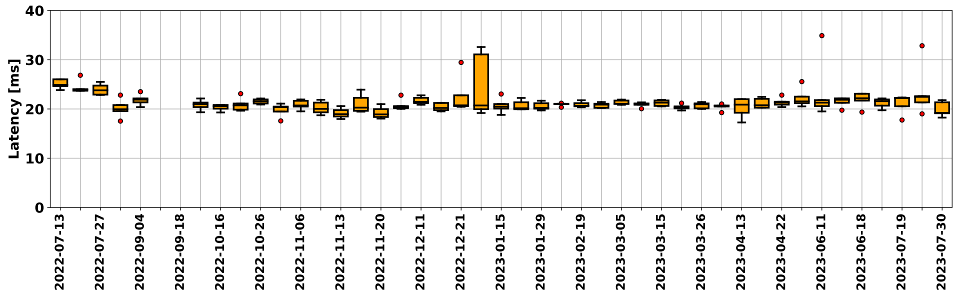


Figure 17: Single UE latency under static conditions

same level of carrier aggregation (e.g., the fraction of time when two carriers are used is about 50% on 12/21/2022 and on the last day of the measurement campaign in Fig. 16b, but the throughput is very different on those two days in Fig. 13a).

6.1.4. Latency

Fig. 17 presents a boxplot of the latency values measured with the UE facing the base station. The median latency consistently ranges between 18–21 ms, except for the first three measurement days. Notably, the measurements conducted on Jan 5, 2023, exhibit a large interquartile range. Further analysis revealed that on this day, the UE was connected to LTE for 40% of the latency tests. Overall, we conclude that 5G latency remains consistent throughout the year-long measurement period, with minimal variability.

Notice that the CDF of the 5G mmWave latency obtained from our crowdsourced measurements in Fig. 6f includes much higher values than the values in Fig. 17. Recall that the latency values in Fig. 17 were all obtained with measurements over Verizon to a Wavelength edge server located in Boston and attached to the Verizon core network, while the measurements in Fig. 6f are obtained from a mix of tests with an edge server (over Verizon) and a cloud server located in Virginia (over AT&T).

Overall, the results in §6.1.1 (Figs. 12a, 13a) and §6.1.4 (Fig. 17) show that *the 5G mmWave throughput and latency obtained from our controlled measurements with an edge server are significantly better than the LTE and 5G-low/mid values and the 5G mmWave values with a cloud server obtained via our crowdsourced measurements in Boston over the same one-year period, suggesting that 5G mmWave and edge computing are both critical to boosting the performance of latency-critical, uplink-oriented 5G killer apps.*

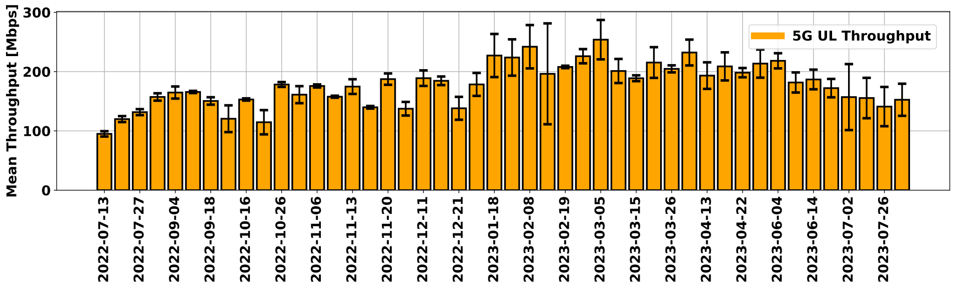


Figure 18: Single UE throughput during mobility.

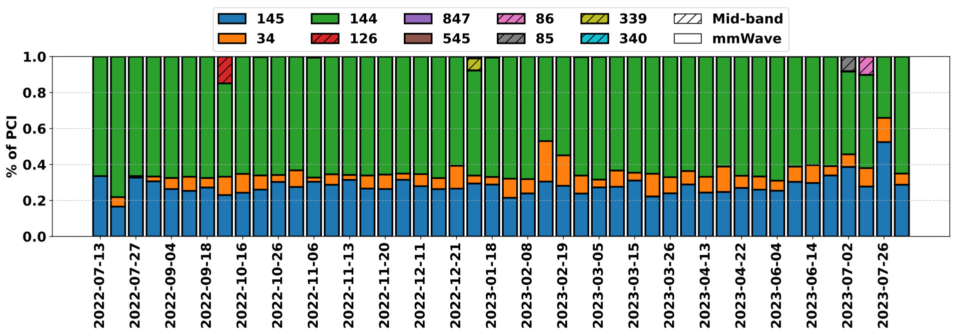


Figure 19: 5G PCI distribution during mobility. Hatched '///' bars represent 5G-midband PCIs. Clear bars represent 5G-mmWave PCIs.

6.2. Single UE, mobile tests

Fig. 18 shows the daily average uplink throughput across the one-year period under mobility, when the user walked laterally to the base station. Contrary to the static scenario, here we observe an increasing trend from the beginning of the measurement campaign till January 2023, followed by a decreasing trend from April 2023 to July 2023. The average throughput was around 100-120 Mbps in July 2022, increased to 250 Mbps at the beginning of 2023, and then dropped down to 150 Mbps in July 2023. Interestingly, we notice larger standard deviations as the daily average throughput starts increasing in 2023.

Unlike static tests, where the UE was always connected to the same base station, in mobility tests the UE experiences handovers to different base stations. Hence, we take a closer look at handovers to explain the throughput variations in Fig. 18. In our dataset, we found a negligible number of 5G to LTE handovers over the one-year period, however, we did notice a few handovers from 5G-high to 5G-low or 5G-mid, as well as handovers between

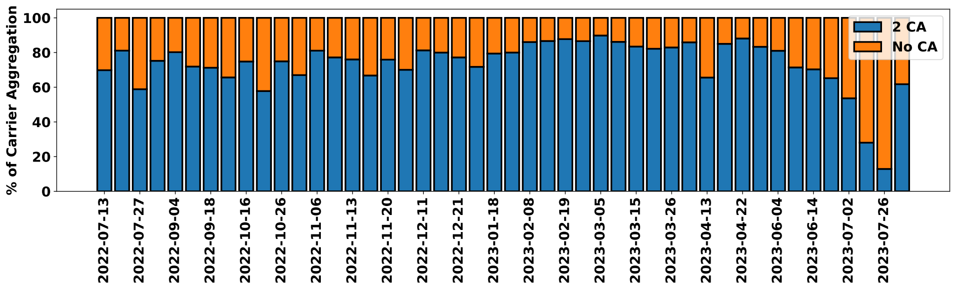


Figure 20: CA distribution during mobility.

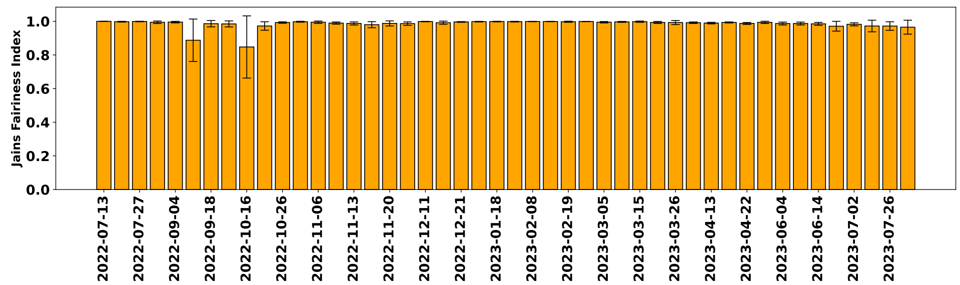
5G mmWave base stations. Fig. 19 shows the distribution of physical cell IDs (PCIs) the UE was connected to during the mobility tests. PCI is a unique identifier assigned to each cell within the network to differentiate and distinguish between neighboring cells. We observe two dominant PCIs, 144 and 145, that appear in every test. A third PCI (34) also appears in every test, although much less often. All these three PCIs belong to 5G mmWave cells; in other words, the UE was connected to a 5G mmWave cell 85-100% of the time during each test and only occasionally switched to 5G-mid cells (e.g., PCIs 126, 339, 85, 86). We also observe that the fraction of time during which the UE was connected to each of the three mmWave cells (34, 144, 145) is roughly constant every day (with the exception of 07/26/2023). Overall, our results suggest that the operator did not make any changes to the 5G mmWave infrastructure at this location over the one-year period; it did not deploy any new cells and it did not change the transmission power of the three existing cells.

Finally, Fig. 20 shows the CA distribution for mobility tests. We observe a noticeable gradual increase in the usage of 2 carriers until March 2023 but a gradual decrease from April 2023 onward, which aligns with the observed changes in throughput in Fig. 18, further confirming a strong correlation between throughput and CA.

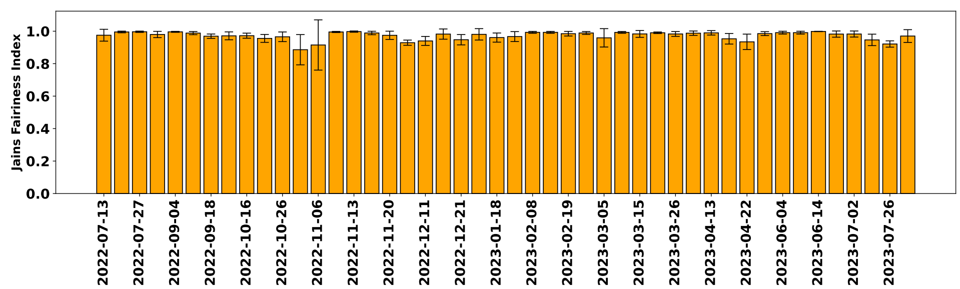
Overall, the observed drop of both throughput and carrier aggregation level at the end of our measurement campaign (during July 2023 for static tests and from April 2023 to July 2023 for mobile tests) suggest an increase in the network load during that period and the need for an upgrade in the cellular infrastructure.

6.3. Parallel tests with two UEs

We also conducted tests with two static UEs sending backlogged uplink traffic simultaneously, while both face towards or away from the base station.



(a) Both UEs face the base station.



(b) Both UEs face away from the base station

Figure 21: Fairness index for parallel tests with 2 UEs.

In this case, we are interested in the way the operator allocates resources to the two flows. Fig. 21 plots the Jain's Fairness Index for the two flows over the one-year period. We observe that the fairness index remains close to 1, indicating that the two devices generally share network resources equally, with two exceptions in Fig. 21a and two more exceptions in Fig. 21b, where the fairness index drops to 0.8-0.9 with large standard deviations. Overall, we conclude that *the operator's resource sharing policy among users remains consistent and fair throughout the one-year measurement period*.

7. Conclusion

In this paper, we conducted a cross-sectional, year-long measurement study of 5G aiming to assess its deployment maturity via three metrics: stability of its performance over a long time span, performance comparison with its predecessor LTE, and performance diversity in geographic locations and operators. Our crowdsourced measurements show that 5G deployment in major cities appears matured, with no major performance improvements observed over a one-year period, however, 5G uplink throughput often exhibits erratic and suboptimal behavior, and in some cases, is inferior to LTE. Further, 5G has not demonstrated significant improvements over LTE in terms of latency. Surprisingly, in certain cities worldwide, latency over LTE networks is comparable to or even lower than over 5G networks. Additionally, our controlled measurements over 5G mmWave show that uplink throughput exhibits no significant trends, aside from seasonal fluctuations, whereas latency remains stable with minimal variations. However, throughput and latency over 5G mmWave with an edge server are significantly better than the LTE and 5G-low/mid values and the 5G mmWave values with a cloud server obtained via our crowdsourced measurements in the same city, suggesting that 5G mmWave and edge computing are both critical to boosting the performance of latency-critical, uplink-oriented 5G killer apps. Overall, our findings suggest that, while 5G holds promise for transformative enhancements in mobile networks, its full potential has yet to be realized.

Acknowledgments

This work was supported in part by NSF grants 2112778 and 2312834 as well as by the Federal Ministry of Education and Research (BMBF, Germany) within the 6G Research and Innovation Cluster 6G-RIC under Grant 16KISK020K, and by project 6G-ELSA (PID2022-136769NB-I00) funded by MCIN/AEI/10.13039/501100011033/ and the European Union ERDF. C. Fiandrino is a Ramón y Cajal awardee (RYC2022-036375-I), funded by MCIU/AEI/10.13039/501100011033 and the ESF+. Y. Yu is supported by Leonardo SpA. C. F. Chiasserini is supported by the EU under the Italian NRRP of NextGenerationEU program "RESTART" (PE00000001). M. Hussain and A Aguiar are supported by FCT/MCTES through Portuguese funds and co-funded EU funds under the project UIDB/50008/2020, and by Next Generation EU through PRR Project Route 25 (C645463824-00000063).

References

[1] A. Narayanan, E. Ramadan, R. Mehta, X. Hu, Q. Liu, R. A. Fezeu, U. K. Dayalan, S. Verma, P. Ji, T. Li, et al., Lumos5G: Mapping and predicting commercial mmWave 5G throughput, in: Proc. of ACM IMC, 2020.

[2] A. Narayanan, E. Ramadan, J. Carpenter, Q. Liu, Y. Liu, F. Qian, Z.-L. Zhang, A first look at commercial 5G performance on smartphones, in: Proc. of ACM WWW, 2020.

[3] A. Narayanan, X. Zhang, R. Zhu, A. Hassan, S. Jin, X. Zhu, X. Zhang, D. Rybkin, Z. Yang, Z. M. Mao, F. Qian, Z.-L. Zhang, A Variegated Look at 5G in the Wild: Performance, Power, and QoE Implications, in: Proc. of ACM SIGCOMM, 2021.

[4] M. Ghoshal, Z. J. Kong, Q. Xu, Z. Lu, S. Aggarwal, I. Khan, Y. Li, Y. C. Hu, D. Koutsonikolas, An In-Depth Study of Uplink Performance of 5G MmWave Networks, in: Proc. of the ACM SIGCOMM 5G-MeMU Workshop, 2022.

[5] C. Fiandrino, D. Juárez Martínez-Villanueva, J. Widmer, Uncovering 5G Performance on Public Transit Systems with an App-based Measurement Study, in: Proc. of ACM MSWiM, 2022.

[6] M. I. Rochman, V. Sathya, D. Fernandez, N. Nunez, A. S. Ibrahim, W. Payne, M. Ghosh, A comprehensive analysis of the coverage and performance of 4G and 5G deployments, Computer Networks 237 (2023).

[7] G. Caso, M. Rajiullah, K. Kousias, U. Ali, L. De Nardis, A. Brunstrom, O. Alay, M. Neri, M.-G. Di Benedetto, An Initial Look into the Performance Evolution of 5G Non-Standalone Networks, in: Proc. of IFIP/IEEE TMA, 2023.

[8] A. Narayanan, M. I. Rochman, A. Hassan, B. S. Firmansyah, V. Sathya, M. Ghosh, F. Qian, Z.-L. Zhang, A Comparative Measurement Study of Commercial 5G mmWave Deployments, in: In Proc. of IEEE INFOCOM, 2022.

[9] M. Ghoshal, I. Khan, Z. J. Kong, P. Dinh, J. Meng, Y. C. Hu, D. Koutsonikolas, Performance of Cellular Networks on the Wheels, in: Proc. of ACM IMC, 2023.

[10] P. Dinh, M. Ghoshal, D. Koutsonikolas, J. Widmer, Demystifying Resource Allocation Policies in Operational 5G mmWave Networks, in: Proc. of IEEE WoWMoM, 2022.

[11] A. Hassan, A. Narayanan, A. Zhang, W. Ye, R. Zhu, S. Jin, J. Carpenter, Z. M. Mao, F. Qian, Z.-L. Zhang, Vivisecting mobility management in 5G cellular networks, in: Proc. of ACM SIGCOMM, 2022.

[12] R. A. K. Fezeu, C. Fiandrino, E. Ramadan, J. Carpenter, D. Chen, Y. Tan, F. Qian, J. Widmer, Z.-L. Zhang, Roaming across the European Union in the 5G era: Performance, challenges, and opportunities, in: In Proc. of IEEE INFOCOM, 2023.

[13] D. Xu, A. Zhou, X. Zhang, G. Wang, X. Liu, C. An, Y. Shi, L. Liu, H. Ma, Understanding operational 5G: A first measurement study on its coverage, performance and energy consumption, in: Proc. of ACM SIGCOMM, 2020.

[14] K. Kousias, M. Rajiullah, G. Caso, O. Alay, A. Brunstorm, L. D. Nardis, M. Neri, U. Ali, M.-G. D. Benedetto, Coverage and Performance Analysis of 5G Non-Standalone Deployments, in: Proc. of ACM WiNTECH, 2022.

[15] X. Yang, H. Lin, Z. Li, F. Qian, X. Li, Z. He, X. Wu, X. Wang, Y. Liu, Z. Liao, et al., Mobile access bandwidth in practice: Measurement, analysis, and implications, in: Proc. of ACM SIGCOMM, 2022.

[16] P. Parastar, A. L. O. O. Alay, G. Caso, D. Perino, Spotlight on 5G: Performance, Device Evolution and Challenges from a Mobile Operator Perspective, in: Proc. of IEEE INFOCOM, 2023.

[17] M. I. Rochman, W. Ye, Z.-L. Zhang, M. Ghosh, A Comprehensive Real-World Evaluation of 5G Improvements over 4G in Low- and Mid-Bands, Tech. Rep. 2312.00957[cs.NI] (2023).

[18] J. Moreno, M. Contini, A. Aguiar, 5G NSA Performance: A Measurement Study, in: Proc. of IFIP WONS, 2024.

[19] R. A. K. Fezeu, C. Fiandrino, E. Ramadan, J. Carpenter, L. C. de Freitas, F. Bilal, W. Ye, J. Widmer, F. Qian, Z.-L. Zhang, Unveiling the 5G mid-band landscape: From network deployment to performance and application QoE, in: Proc. of ACM SIGCOMM, 2024.

[20] W. Ye, X. Hu, S. Sleder, A. Zhang, U. K. Dayalan, A. Hassan, R. A. K. Fezeu, A. Jajoo, M. Lee, E. Ramadan, F. Qian, Z.-L. Zhang, Dissecting Carrier Aggregation in 5G Networks: Measurement, QoE Implications and Prediction, in: Proc. ACM SIGCOMM, 2024.

[21] F. Chen, M. Ghoshal, E. Nan, P. Dinh, I. Khan, Z. J. Kong, Y. C. Hu, D. Koutsonikolas, A Large-Scale Study of the Potential of Multi-Carrier Access in the 5G Era, in: Proc. PAM, 2024.

[22] H. Lim, J. Lee, J. Lee, S. D. Sathyanarayana, J. Kim, A. Nguyen, K. T. Kim, Y. Im, M. Chiang, D. Grunwald, K. Lee, S. Ha, An Empirical Study of 5G: Effect of Edge on Transport Protocol and Application Performance, IEEE Transactions on Mobile Computing 23 (4) (2024) 3172–3186.

[23] M. Kalntis, J. Suárez-Varela, J. O. Iglesias, A. K. Bhattacharjee, G. Iosifidis, F. A. Kuipers, A. Lutu, Through the Telco Lens: A Countrywide Empirical Study of Cellular Handovers, in: Proc. of ACM IMC, 2024.

[24] J. Hu, L. Wang, J. Wu, Q. Pei, F. Liu, B. Li, A comparative measurement study of cross-layer 5G performance under different mobility scenarios, Elsevier Computer Networks 257 (2025).

[25] I. Khan, M. Ghoshal, J. Angjo, S. Dimce, M. Hussain, P. Parastar, Y. Yu, C. Fiandrino, C. Orfanidis, S. Aggarwal, A. Aguiar, O. Alay, C. F. Chiasserini, F. Dressler, Y. C. Hu, S. Ko, D. Koutsonikolas, J. Widmer, How Mature is 5G Deployment? A Cross-Sectional, Year-Long Study of 5G Uplink Performance, in: Proc. of IFIP Networking, 2024.

[26] XCAL Solo, https://accuver.com/sub/products/view.php?idx=11;.

[27] M. Ghoshal, I. Khan, Q. Xu, Z. J. Kong, Y. C. Hu, D. Koutsonikolas, NextG-Up: A Tool for Measuring Uplink Performance of 5G Networks, in: Proc. of ACM MobiSys, 2022.

[28] M. C. González, C. A. Hidalgo, A.-L. Barabási, Understanding individual human mobility patterns, Nature 453 (7196) (2008).

[29] RADIO SPECTRUM POLICY GROUP, 5G developments and possible implications for 6G spectrum needs and guidance on the rollout of future wireless broadband networks, https://radio-spectrum-policy-group.ec.europa.eu/system/files/2023-10/RSPG23-040final-RSPG_Opinion_on_5G_developments_and_6G_spectrum_needs.pdf.