Spotlight on 5G: Performance, Device Evolution and Challenges from a Mobile Operator Perspective

Paniz Parastar
University of Oslo, Norway

Andra Lutu
Telefonica Research, Spain

Özgü Alay
University of Oslo, Norway
Karlstad University, Sweden

Giuseppe Caso
Ericsson Research, Sweden

Diego Perino
Telefonica Research, Spain

***Abstract*—Fifth Generation (5G) has been acknowledged as a significant shift in cellular networks, expected to run significantly different classes of services and do so with outstanding performance in terms of low latency, high capacity, and extreme reliability. Managing the resulting complexity of mobile network architectures will depend on making efficient decisions at all network levels based on end-user requirements. However, to achieve this, it is critical to first understand the current mobile ecosystem and capture the device heterogeneity, which is one of the major challenges for ensuring the successful exploitation of 5G technologies.**

In this paper, we conduct a large-scale measurement study of a commercial mobile operator in the UK, focusing on bringing forward a real-world view on the available network resources, as well as how more than 30M end-user devices utilize the mobile network. We focus on the current status of the 5G Non-Standalone (NSA) deployment and the network-level performance and show how it caters to the prominent use cases that 5G promises to support. Finally, we demonstrate that a fine-granular set of requirements is, in fact, necessary to orchestrate the service to the diverse groups of 5G devices, some of which operate in permanent roaming.

***Index Terms*—5G, 5G measurement, Network slicing,**

I. Introduction

Since 2019, we have been witnessing rapid deployment efforts of 5G networks all around the world [1]–[3]. The research community has devoted attention to these operational networks [4]–[9], studying different user-side performance metrics, including throughput, application performance, power consumption, or the impact of extreme mobility. These studies focus on measurements from the end-user perspective, often capturing only a partial view of the network and mostly considering smartphone devices only. However, 5G mobile networks set a high bar for the entire cellular ecosystem. On the one hand, these networks aim to connect entire industries by providing smart connectivity to a dazzling number of heterogeneous devices that operate globally in different environments [10]. On the other hand, they aim to accommodate the requirements of a wide range of applications [11] and do so with stringent service-level agreements.

In this paper, we perform, to the best of our knowledge, the first country-wide thorough study of a 5G radio deployment and the end-user demand, from the perspective of a commercial Mobile Network Operator (MNO) operating in the UK. Building on substantial measurement data we collected from the MNO's infrastructure, we shed light on the current mobile ecosystem by investigating the network deployment and the network performance evolution within the landscape of multiple co-existing generations of radio technologies. We further present a detailed study of the user adoption of 5G devices, traffic patterns, and mobility patterns as part of the emerging profiles within a population of more than 30M devices connecting to this mobile network.

Within this context, our work validates that three services we investigated (smart meters, connected cars, and smartphones), which belong to the three key areas for 5G (i.e., massive Machine Type Communication (mMTC), ultra-reliable low-latency communication (URLLC) and enhanced Mobile Broadband (eMBB)), present wildly divergent profiles. However, even in the same area (e.g., connected cars), our analysis shows that there exists several different profiles. To this end, our work highlights the operators' need for fine-granular knowledge about the device population and their traffic patterns to fully exploit the innovations 5G introduces. Such fine-granular wisdom is essential to provide efficient network management and fulfill the requirements of different services. Beyond accounting for the dazzling variety of end-user devices, this is particularly challenging given that Internet of Things (IoT) verticals prefer a global deployment model for their devices [10], which the 5G architecture did not foresee. In this context, we investigate the operational reality and the actual heterogeneity of the device population connecting to the radio network of this commercial MNO.

The contribution of this paper is threefold:

(i) We provide an overview of the operator's 5G radio network deployment evolution (Sec. IV). Given a unique view on the country-wide deployment of this UK commercial operator during the last two years (2021-2022), we further compare the 5G network performance with that of the Fourth Generation (4G) network (Sec. V). We discuss the network utilization levels, the average number of active users, Downlink (DL)/Uplink (UL) traffic volumes and throughput.

(ii) We study the device population and service evolution (Sec. VI). We analyze the 5G-capable device adoption during a two-year period and highlight the shifts in the device population statistics. We then dive into device traffic analysis, where we study service connectivity, user data traffic, control traffic, inbound roamers, and user activity. We also discuss the impact of mobility patterns of different devices.

(iii) Finally, we investigate the emerging variety of network user behavior beyond the three key use case areas envisioned for 5G (namely, mMTC, URLLC, or eMBB) (Sec. VII). We show that, even within the same IoT vertical (namely, connected cars), different patterns emerge, which translate into different requirements in terms of network service for those devices. We thus demonstrate that a fine-granular set of requirements is necessary in order to optimally orchestrate the service to meet the needs of highly diverse groups of 5G devices, some of which operate in permanent roaming.

II. Background and Related Work

In this section, we provide an overview of relevant related work in terms of 5G network deployment and performance as well as device and service evolution.

Network deployment and performance: Initial 5G standardization efforts have led to two main deployment options, i.e., NSA and Standalone (SA). In NSA, the 5G New Radio (NR) cells are connected to the 4G core network for control plane communication and management; in SA, the deployment includes a dedicated 5G core network, and the 5G system is independent of its 4G counterpart. Early work in 5G focuses on performance aspects of both NSA and SA deployments from User Equipment (UE) perspective. In [12], the authors analyze throughput, latency, and application performance over millimeter wave (mmWave) deployments of three main US operators. In the same context, Narayanan et al. [13] define Lumos5G, and propose an Machine Learning (ML)-based prediction of mmWave 5G throughput based on coverage parameters (e.g., signal power). Liu et al. [14] investigate UE and Internet backbone as possible bottlenecks for the performance of 5G deployments in China. Results show that, in some scenarios, the bottlenecks of downlink performance are the UE (application) and the Internet infrastructure, rather than the radio link. Recent studies have also covered power and application aspects. A recent study of NSA deployment in a dense urban area in China [4] discussed UE power consumption, handover, throughput, latency, and other application-specific metrics. Another analysis for two US operators [5] considered performance, power, and QoE implications in NSA/SA mmWave and NSA low-band deployments. Similar investigations on throughput and latency performance were also carried out on 5G NSA deployments in Europe [8], [9].

Differently from previous works, this paper provides a unique network-side study by analyzing the country-wide deployment of a commercial cellular network operating in the UK. Our results provide key insights on the evolution of the cellular ecosystem, including but not limited to the profiling of 5G network performance and end-user demand.

Device and service evolution: A few studies have recently focused on device and service characteristics in the cellular ecosystem from the mobile operator perspective and using large-scale measurement campaigns. They investigate cellular

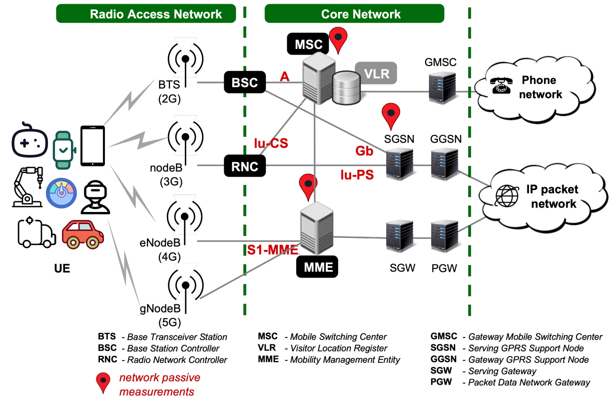


Fig. 1. High-level architecture of the measurement infrastructure integrated in the cellular network.

IoT services and use cases, such as connected cars [15], wearables [16], and industrial IoT [17]. The authors of [18] study Machine-to-Machine (M2M) devices roaming and show an increased impact on the visited network due to their rapid deployment. Geissler et al. [19] identify five clusters of IoT devices based on their signaling behavior, showing a significant heterogeneity among such devices even when a limited number of characteristics is considered. The authors of [20] classify IoT devices into five categories (i.e., monitoring, locating, portable, point of sale, and vehicle), separating them from smartphones. They rely on traffic and mobility behaviors, device properties, and the location of base stations. To apply supervised classification, they use the taxonomy provided by Global System for Mobile communications Association (GSMA) as device categories. However, this taxonomy may result incomplete (Sec. III-B) since devices are labeled as modems/modules with no information on the actual verticals (e.g., smart metering).

Our work extends the device/service evolution analysis to provide (i) a long-term analysis that shows how such devices are changing over time, (ii) a characterization of IoT devices connected to cellular networks to highlight the main differences with conventional devices (i.e., smartphones), and (iii) a better understanding of the complexity created by the heterogeneity of users and thus the need for a significant shift in network management.

III. Dataset

We collect our dataset from the network of a large MNO in the UK. In this section, we detail the measurement infrastructure, the data we collected, and the metrics we extracted. In Fig. 1, we depict a high-level schema of the MNO architecture. In a simplified form, it consists of three main domains: (i) the cellular device, (ii) the Radio Access Network (RAN), and (iii) the Core Network (CN). We monitor the Mobility Management Entity (MME), the Message Sequence Chart (MSC), the Serving GPRS Support Node (SGSN), the Serving Gateway (SGW) network elements, and the radio sectors, which we marked with red pins in Fig. 1. Using this measurement infrastructure, we collect radio sector level Key Performance

Indicators (KPIs), as well as control plane information for the devices the radio network serves.

A. Radio Network Dataset

We capture the following data feeds at the radio sector level.

Network Deployment: To capture the network deployment, we rely on a daily snapshot of the radio network topology. Cell sites represent the locations where the operator deploys their radio antennas and network access equipment. Every cell site hosts multiple radio sectors for different Radio Access Technology (RAT). For every cell site, we collect detailed information such as geolocation, available RATs, number of radio sectors, and their status (active/inactive).

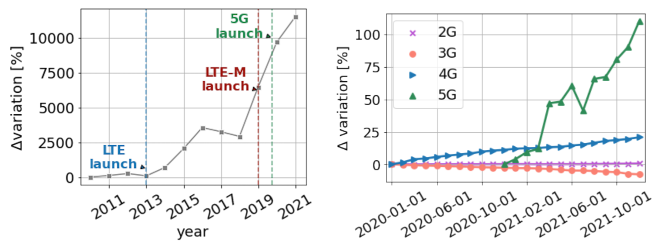
Network Performance: The MNO deploys a commercial solution to collect the radio network performance metrics at the radio sector level. This dataset includes various KPIs collected per hour per sector, including the average number of users in Radio Resource Control (RRC) CONNECTED state and the average user throughput per sector. We also monitor the DL/UL data volume, which is the sum of volume (Service Data Unit (SDU) at Packet Data Convergence Protocol (PDCP) layer) on data radio bearers that has been transferred in the DL/UL direction. DL/UL data volume itself implies total per hour, while total DL/UL data volume represents the summation over the entire day. We assess the load of the radio sectors (i.e., sectors resource utilization) according to the percentage of time at least one connected user has data to receive in the DL direction (i.e., number of Transmission Time Intervals (TTIs)) per hour.

B. Devices Catalog

In the devices catalog, we build a comprehensive daily view of the activity of each device in the entire population of devices connected to the radio network of the MNO. The devices catalog spans *from the 14th of February until the 4th of March in 2020*, and *from the 16th until 23th of February in 2022*, and contains more than 30 million devices per day. It includes the following data feeds aggregated per user per day:

Mobility management control data: We capture the activity of the users in the control plane for each RAT supported by the cellular provider. We take International Mobile Equipment Identity (IMEI) and overall statistics, such as used RAT and the number of failed/successful events. The RAT metrics are summarized into the boolean radio-flags, which show whether the device has a successful event on the data interfaces (2G/3G/4G). Note that due to the 5G NSA deployment, it is impossible to distinguish mobility management events corresponding to a 5G-capable device connecting to a 5G cell.

Further, we extract mobility metrics: radius of gyration and number of sectors. Gyration is a mobility metric that models the distance that devices travel during the day in terms of how far they are located from the center of mass [21]. It is defined as the root mean squared distance between all visited radio sector locations and their center of mass. The number of sectors represents the distinct number of radio sectors with successful events per device per day.



(a) active cell sites per year. (b) active sectors per RAT.

Fig. 2. Longitudinal analysis of the radio network deployment: we show (a) the evolution of the number of cell sites over the last decade; and (b) the Δvariation in the total number of active sectors per RAT, compared to the reference date (1 January 2020). For the latter, we show the first day of each month on the x-axis, for a period of 14 month (until 1 December 2021).

Data usage: We collect the daily data transferred per user from eXtended Detail Record (xDR), including bytes consumed, roaming status, and Access Point Name (APN) strings. APN strings enable us to uncover information about the devices' specific application (e.g., IoT verticals such as smart meters).

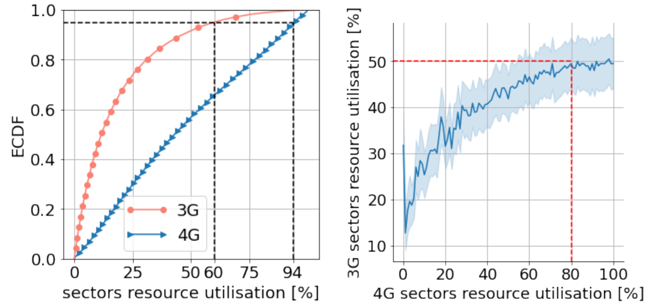
Device properties: To attain device properties, we utilize a commercial database provided by GSMA to map the Type Allocation Code (TAC) (i.e., the first 8 digits of the device IMEI) to its model specification. Some of the features obtained in this way include device manufacturer, model name, operating system, production year, and supported radio bands.

IoT devices: To identify IoT devices, we employ the TAC database from the GSMA. With this information, we distinguish between smartphones and IoT devices. However, GSMA tags are inadequate to identify our IoT use cases (labeled as a module, or modem, which does not imply the actual vertical, like connected cars). As a result, we further refine this classification and categorize devices based on APN strings indicating information about the IoT vertical. For example, we infer that devices using *tesla.m2m.gprs* APN are equipment installed in Tesla cars. In particular, we pursue these steps to identify connected cars (ccars) and energy smart meters to use as IoT application examples.

IV. Network Deployment

In this section, we study the evolution of the radio network deployment during the last decade. Our dataset captures the status of an operational radio network covering the entire UK with thousands of radio sectors corresponding to the different radio generations available. Using network topology information, we extract the number of cell sites (regardless of the available radio technologies) established annually over the last decade. For confidentiality reasons, we present the results in Fig. 2a in terms of the Δvariation with respect to the first day on the x-axis. Note that Δvariation represents the difference between two variables A and B, formulating as $(A/B - 1) * 100$ throughout the paper.

Overall, we observe an increasing trend in the number of new cell sites, which is an expected side-effect of the gradual commercialization of new generations of radio technologies by the operator, and the expansion of the radio coverage to



(a) ECDF. (b) Co-located 3G vs. 4G sectors.

Fig. 3. Analysis of radio sector context. We analyze radio sector resource utilization (95th percentile of TTI utilization) for 3G and 4G using data we collected over 3 days in October 2021, both in terms of (a) ECDF per technology, and (b) status comparison for co-located 3G and 4G radio sectors.

serve more of the population. Specifically, we note change points in 2013 and 2020, corresponding to when 4G and 5G became commercially available. The spike in 2019 is also a commercial response to the substantial growth in the number of new cellular-connected devices (both cellular IoT devices [22], [23] and smartphones) and year-on-year network traffic growth igniting in 2018 [24].

In every cell site location, the operator usually deploys three different radio sectors per radio technology. We further analyze the temporal evolution of the number of radio sectors per RAT, across all the active cell sites. Our goal is to capture the variation in the smallest units of geospatial radio coverage (i.e., the radio sector) per technology. In this step, we only consider radio sectors for which we recorded radio signaling activity from customers. We count the number of the active sectors on the first day of every month between *January 2020* and *December 2021*. Though the official launch date of 5G was October 2019, we consider 1st of December 2020 as a reference date for the 5G deployment analysis (when the population of 5G-capable devices was sufficiently large).

The 5G NSA setup is popularly embraced by various mobile network providers to support enhanced broadband use cases. The MNO we measure also follows the same strategy and anchors all 5G new radio sectors on 4G cell sites. In Fig. 2b, we observe that, over the last two years (2020-2021), the number of 5G active sectors has sharply increased by 110.3%. This comes with a higher energy consumption cost. The transceiver power of most of the 5G sectors is 50 W to increase throughput, while 3G/4G sectors mainly work with 43 W. Considering the significant energy consumption footprint of mobile networks [25], power consumption is becoming an urgent issue to tackle.

Even with the rapid deployment of 5G, we still observe that 4G active sectors also increase with a growth rate of 0.87 to expand access to 4G. Fig. 2b further shows that the number of active 3G sectors decreases at the rate of 0.4, which signifies that 3G is slowly being retired. At the same time, the number of 2G sectors remains unchanged (ratio = 0.01), since 2G services are still used by old IoT devices. The average lifespan of IoT devices is more than 5 years [26], and networks must support them for a long time.

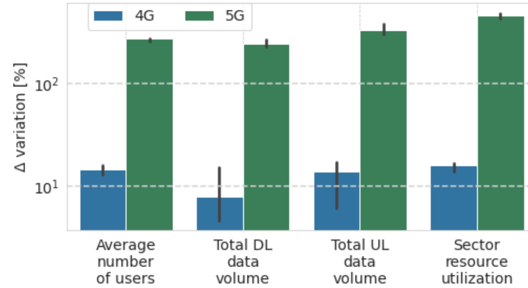


Fig. 4. Network utilization evolution - median of Δvariation of each hourly metric per radio sector for the Week 25 of 2022 (i.e., Jun 20-26) with respect to the same week in 2021 (i.e., Jun 21-27).

Fig. 3a shows the distribution of the peak radio load (i.e., 95th percentile of TTI utilization) for 3G and 4G sectors. We observe 3G sectors' relatively light occupancy, and non-outlier values are lower than 60% vs. 94% for 4G. Fig. 3b reports the median of the peak radio load of the co-located 3G and 4G sectors. The peak radio load of 4G sectors is considerably higher than 3G sectors, and 4G is preferred when 3G and 4G access are both available. Even in locations with congested 4G sectors (over 80% of TTI utilization), the median radio load of 3G sectors is still 50%. This also partially explains the commercial push towards the gradual 3G network shutdown.

Further future deployment of small-cells (e.g., femto- and pico-cells) alongside the rapidly evolving macro cells (Fig. 2b) will increase the network density even more. This is not only because of the new use cases demanding wide coverage but also because ultra-dense deployment is necessary to obtain good mmWave coverage due to its directional transmission. However, spatial network densification comes at the cost of operational challenges, such as increasing power consumption, interference, and overhead on the core network. Accordingly, smart network management is crucial for the operator to run, maintain, and optimize the network.

V. Network Performance Evolution

Given the deployment of the 5G network at the time of writing, we analyze here the commercial 5G network usage and performance as captured by the operator.

A. 5G Network Usage

To examine network usage patterns, we capture the following hourly metrics per radio sector: average number of users, sector resource utilization, and total DL/UL traffic volume. We report the median values of the metrics per 5G radio sector along with the 4G sectors usage as a baseline. For fairness purposes, wherever we compare 4G and 5G technologies, we match 4G sectors in the same locations as 5G sectors. We infer geographical location by longitude and latitude coordinates of the cell site hosting the 4G/5G radio sectors. We observe almost no 5G usage at the beginning of 2021, followed by an explosive growth during 2022. To show how the usage evolved in the last year, we compare the network status in week 25 (the last week of June) in 2022 with the same week in 2021.

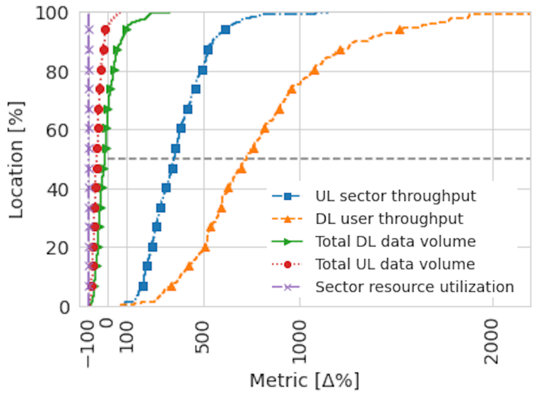


Fig. 5. Performance of co-located 5G vs. 4G sectors. Each line indicates ECDF of the Δvariation of the metric (i.e., $(5G/4G - 1) * 100$).

In Fig. 4, we evaluate the delta variation of the four metrics over the days of the week we consider (we show median values). Overall, we observe a stable growing trend in the 4G usage while the 5G usage has almost tripled. The average number of users increases for both technologies by a factor of 14% for 4G and 272% for 5G. The total data volume of 4G sectors increases by 14% in UL and 8% in DL. Meanwhile, the total data volume of 5G sectors rises significantly at the rate of 332% in UL and 242% in DL. We conjecture this to the uptake of 5G-capable devices by end-users (which we further discuss in Sec. VI-A). We also note that UL traffic has increased more than DL traffic during the last year for both technologies. Finally, the sector resource utilization also exhibits a similar ascending trend with a rate of 463% for 5G and 16% for 4G.

B. 5G Network Performance

While 5G NSA only partially addresses URLLC and mMTC use cases, it already targets to deliver higher capacity to users towards eMBB services compared to 4G. This is achieved by exploiting the existing infrastructure (i.e., 4G core) and 4G-5G inter-operability mechanisms. In this section, we aim to measure to which extent this approach actually supports eMBB use cases. To this end, we focus on six different metrics that capture the performance of 5G sectors over May-June 2022, which we compare to 4G sectors. By taking median values over two months, our goal is to minimize the impact of traffic anomalies, special events, or the differences between days of the week. We take the same approach as in Sec. V-A and calculate median of hourly metrics over radio sectors per day and compare for co-located 4G and 5G sectors. We compute the Δvariation of the 5G to 4G median values and show the ECDF in Fig. 5. Each line corresponds to a metric, and note that we slightly clip the tail of the plot on the right side to make the jammed lines on the left legible. In terms of average DL user throughput, we observe that in all locations with 5G deployments, 5G increases the median throughput by at least 94%. 5G UL sector throughput follows the same trend with values no less than 91% higher than 4G while total UL data volume is still quite small for 5G. In 95% of locations, UL data transmitting on 5G is at least 92% lower.

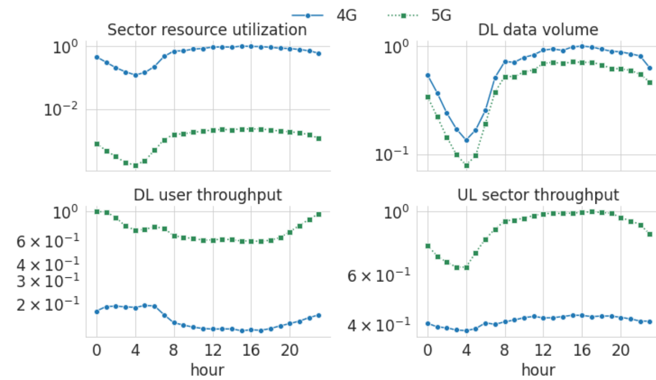


Fig. 6. Diurnal network performance nationwide. Each plot shows normalised median of the metric shown in the title of the plot over two months over all 5G sectors' geolocations.

Conversely, we observe total DL data volume is already higher for 5G than that of 4G in 37% of locations. The sector resource utilization also shows limited usage on 5G sectors compared to 4G, which is at least 98% lower in all locations. Overall, our results show that 5G outperforms 4G for eMBB use cases while the majority of the traffic is still carried out on 4G.

C. Temporal pattern

We next study the daily patterns of the network performance metrics. For all metrics, we calculate the median value per hour for each sector over the same period of two months. Note that, in this section, DL/UL data volume is median value of the metric per hour over all the days in our sample. Fig. 6 depicts the daily pattern of the values normalized by maximum for the country-wide co-located 4G and 5G sectors. High load on sectors in terms of both sector radio resource utilization and data volume between 8 AM to 6 PM identifies working hours patterns. We measure the lowest throughput values in the afternoon when the network carries a high traffic load (e.g., at 5 PM). Fig. 6 also presents a clear match between the daily utilization pattern of 4G and 5G in terms of DL data volume and sector resource utilization. We discard other metrics, such as UL data volume since they are almost identical to the reported ones. Although, on average, radio resource utilization per 4G sector is still one order of magnitude higher than 5G, DL data volume is only 38% larger for 4G. Their difference is even lower during peak hours (i.e., 8 AM to 8 PM), with only 32% larger DL data volume on 4G sectors. This indicates that 5G users generate more DL data volume even though 5G deployment is in its early stages. This might also be explained by the profile of 5G users, which includes mostly top tech smartphones, against diverse users in 4G who do not necessarily generate significant DL traffic (see Sec.VI-B for further analysis).

When we compare our analysis with previous studies [13], we find that our results are consistent in terms of observing very high values and variation at the same time for the 5G user DL throughput; the Coefficient of Variation (CV) of 5G DL user throughput is 86% during the busy time in the evening (i.e., 3-8 PM) whereas 58% for 4G. This suggests 5G throughput prediction is more complex than 4G and affected by a broad range of factors. Additionally, a high average against a low median indicates right-skewed data, with major outliers in the high-end with superior throughput values.

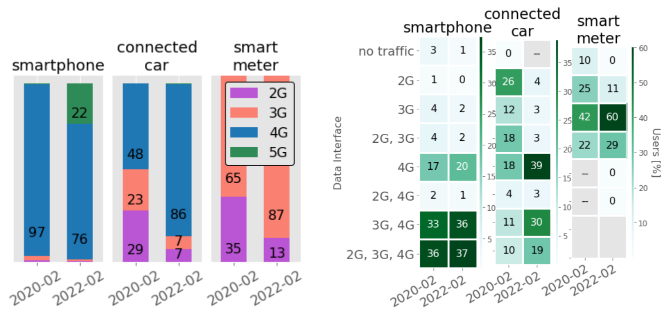
VI. Device Population

As operators prepare to deploy 5G SA networks, one of the main challenges they face is managing the heterogeneity of devices and their service demand. This is particularly dire when considering that the key promise of 5G is to enable use cases for URLLC and mMTC areas did not anticipate the success of the global operational model that IoT verticals chose (i.e., the ability to deploy their devices in different economies across the world with minimal management overhead by overloading the international mobility function of the cellular ecosystem). In this section, we investigate the diversity within the device population that connects to the network of the MNO. First, we explore the adoption of 5G-capable smartphones in the UK. Then, we analyze three representative groups of devices: connected cars as a use case of URLLC, smart meters as a representative of mMTC, and smartphones for eMBB. We analyze their spatial and temporal behaviors by computing daily metrics and also exploring their evolution over the last two years (2022 vs. 2020).

A. User adoption

5G. We first explore 5G-capable device adoption over the last two years (2020 - 2022). Note that we only report statistics of 5G-capable devices, but this does not necessarily imply network activity on 5G interface (Sec. III-B). Our dataset shows that the population of 5G users has increased 2.6 times since the early months of 2020. On average, the number of devices is daily increasing at the ratio of 0.29% in 2022, which we attribute to the maturity of the 5G deployment and the availability of 5G-capable smartphones. Further, we observe that DL (UL) traffic generated by 5G-capable smartphones is 4.46 (3.5) higher than that of 4G (comparing their medians). One might argue that these early adopters are technology-oriented people buying 5G smartphones. To investigate it, we compare the DL traffic of the upper 25 percentile of 4G smartphones in 2020 (as traffic-intensive candidates may have migrated to 5G) with the median of 5G-capable smartphones in 2022. We observe that the former is 49% higher than the latter. Thus, 5G smartphones are not only limited to traffic-intensive users anymore, and the user profile is not the primary reason for the enormous traffic generated by 5G smartphones. This suggests traffic growth by 5G market acquisition, which was correctly anticipated with the 5G vision [27].

Device type. We next explore the device population breakdown over the three major categories that represent the three 5G pillars. Smartphones are still the primary users of the cellular network, with a share of 63% in 2022. Still, they lost 2% compared to 2020 when they had 65% of daily users within our sample. In the last two years, while the number of smartphones has been increasing by about 12%, the number of connected cars and smart meters has been rapidly rising at an average rate of 635% and 111%, respectively. The fact that the production year of 13% (20%) of connected car (smart meters) equipment is either 2020 or 2021 is also a clear validation of their rapid adoption.



(a) Device capability (b) Data interface connectivity

Fig. 7. Device share with respect to RAT

B. Traffic Analysis

We further explore a validated sample of 14 million smartphones, 5.5 million smart meters, and 1 million connected cars for this major operator. We consider daily metrics, and compute the mean per user over all days of analysis.

Service connectivity. Figure 7a reports the radio technologies each device in our sample supports. This is based on the frequency bands reported in the GSMA TAC database for each device type. From 2020 to 2022, we observe a clear trend of migrating to newer radio technologies. In our sample for 2022, we observe a ratio of 22.4% of 5G-capable smartphones and 0.06% of 5G-capable connected cars, validating that 5G penetration is growing. Additionally, 85.6% of cars are 4G-capable in 2022. However, for smart meters in 2022, we find that 87.5% of them are still 3G-capable only, and 12.5% are only 2G-capable. This, jointly with the long life of smart meters (i.e., 4.7 years on average), suggests the operators must still consider support for legacy networks.

Figure 7b shows heatmaps of device activity on the data interfaces per radio technology. Note that, from the point of view of the core, it is impossible to detect 5G connectivity due to the current NSA deployment. We observe the evolution of service connectivity over two years. In particular, in 2022, 90% of cars are active at least once on the 4G interface per day. Also, unlike in 2020 – when 10.3% of smart meters were not active on any data interfaces (i.e., they relied only on SMS/voice communications) – all of them are active on data interfaces in 2022.

User traffic. Figure 8 shows the ECDF of UL and DL traffic, UL/DL ratio (i.e., the amount of UL traffic divided by DL traffic per user) in 2022. The distributions of both DL and UL traffic are highly skewed: even though 75 percent of users generate relatively low traffic (i.e., in order of 10-100 MB), there are a few users with huge traffic volumes (a few GB) in DL/UL direction. We notice a similar trend for connected cars, with a median of 99.5% (98%) lower than the median DL (UL) traffic of smartphones. Contrariwise, we note very small volumes of traffic for the majority of smart meters. This

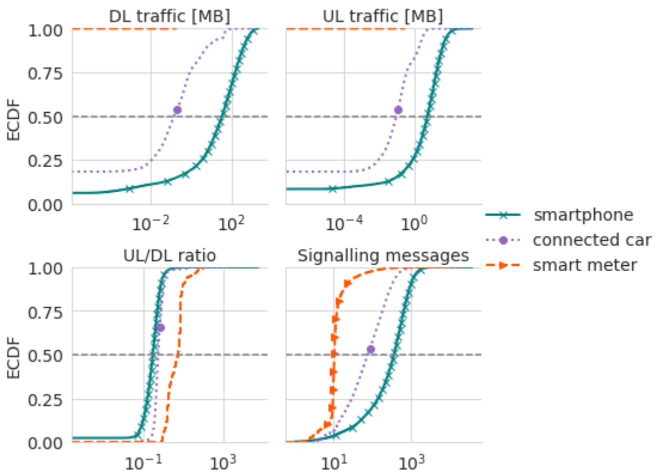


Fig. 8. Daily traffic analysis [log scale].

implies it is hard to separate their little traffic from downlink-intensive applications.

We observe that 88% of the smartphones in our sample have an UL/DL ratio lower than 1, with a share of 11% generating DL traffic 10 times higher than UL traffic. Although connected cars resemble smartphones, the median UL/DL ratio exceeds 0.58 (vs. 0.31 for smartphones). This implies a relatively larger UL than DL for connected cars compared to smartphones. However, the trend is decreasing slowly, and we note a 21% lower UL/DL ratio in 2022 compared to 2020. This suggests connected cars begin to generate more significant DL traffic than UL. On the contrary, the median UL/DL ratio for smartphones is 8% higher in 2022. Our dataset also reveals an upward trend in user traffic in both directions over the last two years, only a further rise for UL: UL traffic has grown 65%, while this has been so 50% for DL. This is consistent with the results in Sec. V-A, and suggests UL communication is becoming more significant. In a nutshell, smartphones and connected cars are DL-intensive applications, whereas those few smart meters generating traffic focus mostly on the uplink.

Control traffic. We next analyze the amount of radio signaling events from devices. We capture signaling activity from the passive monitoring of MSC and MME elements (see Fig. 1), and show ECDFs of the overall number of messages per day per device category in Fig. 8. Smartphones generate the most signaling traffic, which is 35.2 times higher than that of smart meters and 4.69 times higher than connected cars (comparing their median values). The slope of signaling ECDF for smart meters is steep around the middle, which suggests a tight range of values. Since they connect to a few sectors (with 90% of the time on the same sector), they may congest the radio network in the case of dense deployment of synchronized smart meters, which may trigger signaling storms. In this case, the MNO needs to do the whole signaling procedure to send only few bytes, and the signaling load may exceed the benefits from carrying the actual traffic. This issue appertains to the all mMTC use cases such as smart meters.

Inbound roamers. Inbound roamers are devices operating

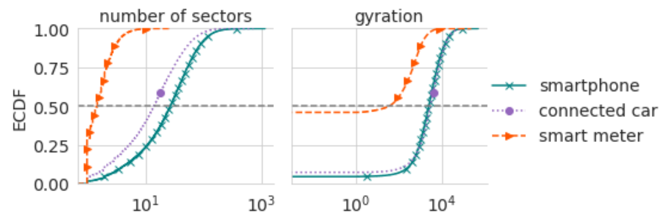


Fig. 9. Mobility of devices.

with SIM cards of a different operator rather than the MNO whose radio network is used. 12.93% of smart meters we observe in our sample dataset are inbound roamers, and only 1.82% of cars are native devices (i.e., operating with a native SIM card from the operator we analyze). This global approach for deploying different verticals brings additional complexity and challenges to the visited network, especially when considering the signaling exchange required with their home network [18], [19]. Plus, we note that the manufacturer of inbound roamer and native smart meters are divergent and might differ in configuration and agreement policies. We further observe car brands follow a centralized strategy regarding SIM card ownership. In this respect, we observe a clear winner per brand for the home country of the SIM card. For example, 98% of cars from brand 1 use SIM cards from an Austrian operator.

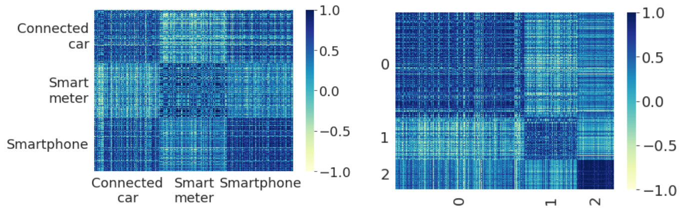
User activity. In the next step, we calculate the number of days each inbound roaming device is active on the network to analyze its roaming persistence. Overall, we find that 33% of inbound roamer smartphones are only active for one day. On the other hand, we find that 80% (18%) of roaming cars (smart meters) are active on the network for all days we capture in our sample. Moreover, 52% of roaming smart meters are connected at least half of the days. This shows that inbound roamer IoT devices are not only more in terms of numbers but also stay on the network permanently, unlike roamer smartphones. For native devices, we find that overall, 79% of smartphones, 80% of connected cars, and 87% of smart meters use the network during the whole period of analysis. This indicates smart meters are permanent on the network.

C. Mobility patterns

Use cases such as connected cars are especially challenging because of their high mobility, which might impact their experienced latency or throughput. In this section, we report on the ECDF of per device mobility metrics (gyration and number of sectors) in Fig. 9. Although connected cars follow the same mobility pattern as smartphones in terms of gyration, they connect to a lower number of sectors, which is in line with being less active on the network. Connected cars have gyration 13% higher than smartphones in non-outlier ranges (i.e., comparing their 75th percentile). Plus, connected cars travel more in 2022, with a median of gyration 1.59x higher than in 2020. Smart meters are, on the contrary, stationary devices with the low number of sectors and gyration. In fact, 75% of these devices present a gyration lower than 441m and connect to at most 2 unique sectors.

TABLE I
CLUSTERING ALGORITHM RESULTS ON A SMALL SAMPLE.

	Silhouette	Davies-Bouldin	Size(%)
K-means [28]	0.54	0.6	38, 33, 29
Gaussian Mixture	0.53	0.59	36, 32, 30
Birch [29]	**0.57**	**0.45**	62, 34, 2
Spectral [30]	0.28	0.85	42, 36, 22
Agglomerative Clustering [31]	0.53	0.58	45, 40, 15



(a) Device Category (b) Cluster

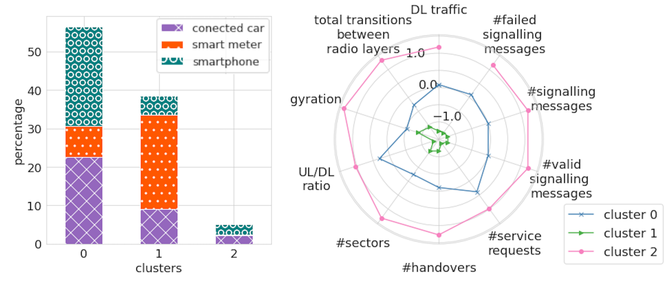
Fig. 10. RSMs between pair of users on a random sample of 1000 users

VII. DEVICE HETEROGENEITY

The empirical analysis in Sec.VI shows that some basic metrics captured the fundamentally different behavior of the three groups of devices we analyzed. This validates that 5G design correctly predicted the importance of their corresponding use case areas. However, when building the devices samples, we found that the global deployment model IoT verticals prefer is quickly becoming the norm (i.e., less than 2% of connected cars in our sample are native to the operator we analyze), reducing the visibility for the radio network operator into end-user requirements. Despite this, we observe similarities in the network behavior of devices in different categories, mainly between smartphones and connected cars. Furthermore, within each group of devices, different traffic patterns emerge, which point to the need for the network operators to apply tailored approaches in order to cater to the requirements of those devices. This requires a fine-granular understanding of the devices demand, and intelligent analytics solutions within the operator environment to match devices to their predicted requirements profiles. In this section, we take the first step in this direction and investigate the emerging patterns for these three major 5G device categories, both in terms of similarity between device categories and also the heterogeneity within a standalone use case belonging to one category.

A. Inter-category similarities

We take a random sample of 1 million devices (about 330 thousand per device category) during a typical working week and apply an explanatory clustering method to study the similarities between device categories. For this, we dissect the behavior of devices, ignoring their device category (i.e., smartphone, connected car, smart meter). To do so, we rely on 456 average weekly metrics to capture traffic, mobility, signaling, handovers, and time spent on the network as input features. We discard RAT-based metrics (e.g., the number of Attach events in 3G) since if all the devices are 5G-capable in the future, these sliced metrics will be worthless. Then, we leverage a simple feed-forward neural network autoencoder to reduce the dimensions of the feature set to 8. We use a three-layer Encoder-Decoder architecture with 128, 64, 32 hidden units and consider mean-squared-error (MSE) as the loss function. We run experiments for 50 epochs, where the batch size and learning rate are set to 128 and 0.001, respectively.



(a) (b)

Fig. 11. The results of clustering users (a) Devices breakdown with respect to the clusters. (b) The relative difference of top 10 features per cluster to overall median per feature; we normalize it by standard scaler to fit all features between [-1,1].

Then, we serve the autoencoder output as features for clustering algorithms. To test the optimal number of clusters and the clustering approach, we first take a smaller random sample of 60,000 users. We choose a number of clusters as 3 according to the elbow-method on the within-cluster sum of squared distances to the center. We compare the result of different algorithms using the Python Scikit-learn library [32] in Table I considering two scores. The model with the highest Silhouette score [33] or the lowest Davies-Bouldin score [34] represents the best approach. Although the clustering algorithms result in a different group of users, as shown in the last column of Table I, we observe they entirely fail to separate the three groups of devices; this evidences similarity among device categories that gathers them into the same cluster. Notice we exclude the results of the share of device categories in clusters due to insufficient space. Hereafter, we use Birch clustering, which yields the best score.

To justify the clustering approach, we use Representational Similarity Matrices (RSMs) and examine the cosine similarity between the representations of each pair of users in the same category (Fig. 10a) and in the same cluster (Fig. 10b) for a random sample of 1000 users. Square sections in Fig. 10a indicate the similarity between connected cars and smartphones. The square sections of the RSM lying along the blue diagonal in Fig. 10b correspond to the similarities between users in the same cluster, implying that the clustering approach can capture the similarities among users with the same behavior.

Ultimately, we apply clustering on the whole 1 million users and show the results in Fig. 11. As it appears in Fig. 11a, clusters contain 57%, 38%, and 5% of users. Then, we focus on finding the essential features enabling us to differentiate clusters. Fig. 11b shows the relative difference between the median of clusters and the overall median for the top ten features. The clusters and their relative deviations from the median value are depicted as lines in this radar plot. Overall, we catch three kinds of behavior: most users are in cluster 0 in

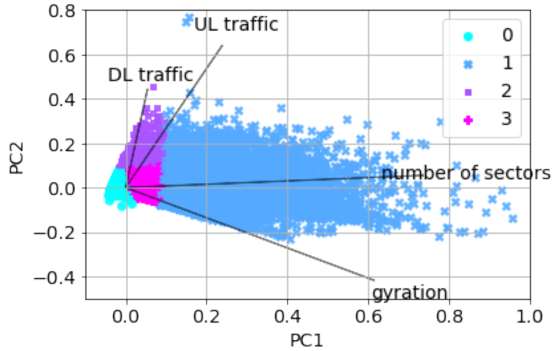


Fig. 12. The results of clustering cars. The biplot of the two main PCs along the original features.

TABLE II
CARS WITH THE TOP APN CHARACTERISTICS: THE PERCENTAGE OF DEVICES AND ΔVARIATION PERCENTAGE OF METRICS FROM THE MEDIAN OF SMARTPHONES.

	%	**DL**	**UL/DL ratio**	**gyration**	***cluster***
0	75.2	-99.48	91.98	-29.88	*low mobility low traffic*
1	5.1	-92.70	56.27	1703.14	*high mobility significant traffic*
2	1.2	853.43	-93.48	1.47	*low mobility high traffic*
3	18.4	-96.51	55.02	235.82	*high mobility low traffic*

the middle with a median of 0.5MB DL traffic, connecting to 14 sectors per day (shown as a blue line in Fig. 11b). Whereas a limited number of users in cluster 2 are highly mobile and generate 1.8x higher DL traffic and transmit 2.2x more between radio layers on the network than cluster 0 comparing their medians. The last group of users in cluster 1 is stationary devices, mainly smart meters, that do not generate much DL traffic and remain in the same place. Overall, the results not only confirm the similarity between device categories (Fig. 10a) but also contribute to the smart user classification.

B. Intra-category variation

We separate connected cars sending the top-ranked APN (i.e., covering the most significant share of 44% among cars) on a typical working day. Then, we input them into the k-means clustering by considering only traffic behavior (i.e., DL/UL traffic and signaling messages) and mobility behavior (i.e., number of sectors, gyration) as input features. We choose 4 clusters based on the elbow-method evaluating the within-cluster sum of squares. Fig. 12 shows the biplot of the two most significant PCs resulting from applying Primary Component Analysis (PCA). Each point represents one user, and the colors denote the appointed cluster. The arrows and labels represent the variance of the original features on the shown primary components. This allows for identifying the relationship between clusters and features. We find four groups of high/low mobile cars generating low or high traffic. The biplot also shows mobility and traffic are not correlated since the angle between their vectors is near 90°. On the other hand, UL and DL are likely correlated.

Table II shows the characteristic of each cluster based on the delta variation percentage of each metric from the corresponding median of smartphones. 75% of these cars are in cluster 0 with limited mobility and low traffic, whereas traffic of 1.2% in cluster 2 is higher than that median of smartphones. (i.e., 1st vs. 3rd row in Table II). The median gyration of cars in cluster 1 is 18 times higher than that of smartphone, while it is only 3.3x higher for cluster 3. This confirms that neither the device category nor the APN is a reliable approach to group devices with consistent behavior and similar traffic patterns.

VIII. CONCLUSION

In this paper, we studied a 5G NSA network deployment, its performance, and the heterogeneity of the operator's device population. Our results confirmed the growing complexity of the cellular ecosystem. We validated that the current 5G NSA deployment supports eMBB use cases. Meanwhile, we found that the fastest growing mobile device category today is IoT devices (e.g., connected cars, smart meters). With this in mind, we studied the device population, highlighting its diversity in terms of traffic patterns. Currently, transparency into device requirements is provided by the APN, IMSI ranges (full or partial), and, in some cases, the specific RAT. While it is critical to identify different services/devices for efficient 5G network slicing, we conclude that the APN approach is sub-optimal, as devices using the same APN behaved in different ways and generated divergent traffic patterns. This highlights the importance of smart user classification in the context of IoT market growth to be able to benefit from and and exploit better 5G innovations.

ETHICAL CONSIDERATIONS

The data collection and retention at network middle-boxes and elements are in accordance with the terms and conditions of the MNO and the local regulations. All datasets used in this work are covered by Non-Disclosure Agreements (NDAs), prohibiting any re-sharing with 3rd parties even for research purposes. The operator has reviewed and validated raw data with respect to General Data Protection Regulation (GDPR) compliance (e.g., no identifier can be associated with the person). Besides, even if we compute the metrics per user, data processing only extracts aggregated user information. Ultimately, neither personal nor contact information was available for this study, and none of the authors participated in the extraction or encryption of the raw data.

ACKNOWLEDGEMENTS

The work of Andra Lutu was supported by the EC H2020 Marie Curie Individual Fellowship 841315 (DICE). This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement no.101017109 "DAEMON".

References

[1] Ericsson, "Publicly announced 5g contracts 5g map," https://www.ericsson.com/en/5g/contracts.

[2] "China reaches over 1.3 million 5G base stations nationwide: Report," https://www.rcrwireless.com/20211227/5g/china-reaches-over-million-5g-base-stations-nationwide-report.

[3] Verizon, "Verizon 5G Ultra Wideband expected to cover 175 million people by the end of 2022 — About Verizon," https://www.verizon.com/about/news/verizon-5g-ultra-wideband-175-million-people-2022.

[4] D. Xu, A. Zhou, X. Zhang, G. Wang, X. Liu, C. An, Y. Shi, L. Liu, and H. Ma, "Understanding operational 5g: A first measurement study on its coverage, performance and energy consumption," in *Proceedings of the Annual Conference of the ACM Special Interest Group on Data Communication on the Applications, Technologies, Architectures, and Protocols for Computer Communication*, ser. SIGCOMM '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 479–494.

[5] A. Narayanan, X. Zhang, R. Zhu, A. Hassan, S. Jin, X. Zhu, X. Zhang, D. Rybkin, Z. Yang, Z. M. Mao, F. Qian, and Z.-L. Zhang, "A Variegated Look at 5G in the Wild: Performance, Power, and QoE Implications," ser. SIGCOMM '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 610–625.

[6] Y. Pan, R. Li, and C. Xu, "The first 5g-lte comparative study in extreme mobility," *SIGMETRICS Perform. Eval. Rev.*, vol. 50, no. 1, p. 31–32, jun 2022.

[7] A. Narayanan, M. I. Rochman, A. Hassan, B. S. Firmansyah, V. Sathya, M. Ghosh, F. Qian, and Z.-L. Zhang, "A comparative measurement study of commercial 5g mmwave deployments," in *IEEE INFOCOM 2022 - IEEE Conference on Computer Communications*. IEEE Press, 2022, p. 800–809.

[8] K. Kousias, M. Rajiullah, G. Caso, O. Alay, A. Brunstorm, L. De Nardis, M. Neri, U. Ali, and M.-G. Di Benedetto, "Coverage and Performance Analysis of 5G Non-Standalone Deployments," in *Proceedings of the 16th ACM Workshop on Wireless Network Testbeds, Experimental evaluation & CHaracterization*, 2022, pp. 61–68.

[9] K. Kousias, M. Rajiullah, G. Caso, O. Alay, A. Brunstrom, L. De Nardis, M. Neri, U. Ali, and M.-G. Di Benedetto, "Implications of Handover Events in Commercial 5G Non-Standalone Deployments in Rome," in *Proceedings of the ACM SIGCOMM Workshop on 5G and Beyond Network Measurements, Modeling, and Use Cases*, 2022, pp. 22–27.

[10] S. Alcalá-Marín, A. Raman, W. Wu, A. Lutu, M. Bagnulo, O. Alay, and F. Bustamante, "Global mobile network aggregators: Taxonomy, roaming performance and optimization," in *Proceedings of the 20th Annual International Conference on Mobile Systems, Applications and Services*, ser. MobiSys '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 183–195.

[11] C. Marquez, M. Gramaglia, M. Fiore, A. Banchs, and X. Costa-Perez, "How should i slice my network? a multi-service empirical evaluation of resource sharing efficiency," in *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 191–206.

[12] A. Narayanan, E. Ramadan, J. Carpenter, Q. Liu, Y. Liu, F. Qian, and Z.-L. Zhang, "A first look at commercial 5g performance on smartphones," in *Proceedings of The Web Conference 2020*, ser. WWW '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 894–905.

[13] 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 *Proceedings of the ACM Internet Measurement Conference*, 2020, pp. 176–193.

[14] T. Liu, J. Pan, and Y. Tian, "Detect the bottleneck of commercial 5g in china," in *2020 IEEE 6th International Conference on Computer and Communications (ICCC)*, 2020, pp. 941–945.

[15] C. E. Andrade, S. D. Byers, V. Gopalakrishnan, E. Halepovic, D. J. Poole, L. K. Tran, and C. T. Volinsky, "Connected Cars in Cellular Network: A Measurement Study," in *Proceedings of the 2017 Internet Measurement Conference*, ser. IMC '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 235–241.

[16] H. Kolamunna, I. Leontiadis, D. Perino, S. Seneviratne, K. Thilakarathna, and A. Seneviratne, "A First Look at SIM-Enabled Wearables in the Wild," in *Proceedings of the Internet Measurement Conference 2018*, ser. IMC '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 77–83.

[17] B. Finley, J. Benseny, A. Vesselkov, and J. Walia, "How does enterprise IoT traffic evolve? Real-world evidence from a Finnish operator," *Internet of Things*, vol. 12, p. 100294, 2020.

[18] A. Lutu, B. Jun, A. Finamore, F. E. Bustamante, and D. Perino, "Where things roam: Uncovering cellular iot/m2m connectivity," in *Proceedings of the ACM Internet Measurement Conference*, 2020, pp. 147–161.

[19] S. Geissler, F. Wamser, W. Bauer, M. Krolikowski, S. Gebert, and T. Hoßfeld, "Signaling traffic in internet-of-things mobile networks," in *2021 IFIP/IEEE International Symposium on Integrated Network Management (IM)*, 2021, pp. 452–458.

[20] S. Hui, H. Wang, D. Xu, J. Wu, Y. Li, and D. Jin, "Distinguishing between smartphones and iot devices via network traffic," *IEEE Internet of Things Journal*, pp. 1–1, 2021.

[21] M. C. González, C. A. Hidalgo, and A.-L. Barabási, "Understanding individual human mobility patterns," *Nature*, vol. 453, no. 7196, pp. 779–782, 2008.

[22] L. S. Vailshery, "Global IoT and non-IoT connections 2010-2025 — Statista," 3 2021, https://www.statista.com/statistics/1101442/iot-number-of-connected-devices-worldwide/.

[23] Ericsson, "IoT connections forecast – Mobility Report," https://www.ericsson.com/en/reports-and-papers/mobility-report/dataforecasts/iot-connections-outlook.

[24] ——, "Ericsson Mobility Report Q4 Update February 2019," www.ericsson.com/mobility-report.

[25] M. Masoudi, M. G. Khafagy, A. Conte, A. El-Amine, B. Françoise, C. Nadjahi, F. E. Salem, W. Labidi, A. Süral, A. Gati, D. Bodéré, E. Arikan, F. Aklamanu, H. Louahlia-Gualous, J. Lallet, K. Pareek, L. Nuaymi, L. Meunier, P. Silva, N. T. Almeida, T. Chahed, T. Sjölund, and C. Cavdar, "Green Mobile Networks for 5G and Beyond," *IEEE Access*, vol. 7, pp. 107 270–107 299, 2019.

[26] P. Parastar, "Rising cellular multimedia iot: The call for an application-aware resource management," in *Proceedings of the 12th ACM Multimedia Systems Conference*, ser. MMSys '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 388–392.

[27] "5G-PPP: Key Performance Indicators." [Online]. Available: https://5g-ppp.eu/kpis/

[28] D. Arthur and S. Vassilvitskii, "K-means++: The advantages of careful seeding," in *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, ser. SODA '07. USA: Society for Industrial and Applied Mathematics, 2007, p. 1027–1035.

[29] T. Zhang, R. Ramakrishnan, and M. Livny, "Birch: an efficient data clustering method for very large databases," in *In Proc. of the ACM SIGMOD Intl. Conference on Management of Data (SIGMOD*, 1996, pp. 103–114.

[30] Yu and Shi, "Multiclass spectral clustering," in *Proceedings Ninth IEEE International Conference on Computer Vision*, 2003, pp. 313–319 vol.1.

[31] J. H. Ward, "Hierarchical grouping to optimize an objective function," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 236–244, 1963. [Online]. Available: http://www.jstor.org/stable/2282967

[32] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[33] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *Journal of Computational and Applied Mathematics*, vol. 20, pp. 53–65, 1987.

[34] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-1, no. 2, pp. 224–227, 1979.