Through the Telco Lens: A Countrywide Empirical Study of Cellular Handovers

Michail Kalntis
Delft University of Technology
Delft, the Netherlands
m.kalntis@tudelft.nl

José Suárez-Varela
Telefónica Research
Madrid, Spain
jose.suarez-varela@telefonica.com

Jesús Omaña Iglesias
Telefónica Research
Barcelona, Spain
jesusalberto.omana@telefonica.com

Anup Kiran Bhattacharjee
Delft University of Technology
Delft, the Netherlands
a.k.bhattacharjee@tudelft.nl

George Iosifidis
Delft University of Technology
Delft, the Netherlands
g.iosifidis@tudelft.nl

Fernando A. Kuipers
Delft University of Technology
Delft, the Netherlands
f.a.kuipers@tudelft.nl

Andra Lutu
Telefónica Research
Madrid, Spain
andra.lutu@telefonica.com

Abstract

Cellular networks rely on handovers (HOs) as a fundamental element to enable seamless connectivity for mobile users. A comprehensive analysis of HOs can be achieved through data from Mobile Network Operators (MNOs); however, the vast majority of studies employ data from measurement campaigns within confined areas and with limited end-user devices, thereby providing only a partial view of HOs. This paper presents the first countrywide analysis of HO performance, from the perspective of a top-tier MNO in a European country. We collect traffic from approximately 40M users for 4 weeks and study the impact of the radio access technologies (RATs), device types, and manufacturers on HOs across the country. We characterize the geo-temporal dynamics of horizontal (intra-RAT) and vertical (inter-RATs) HOs, at the district level and at millisecond granularity, and leverage open datasets from the country's official census office to associate our findings with the population. We further delve into the frequency, duration, and causes of HO failures, and model them using statistical tools. Our study offers unique insights into mobility management, highlighting the heterogeneity of the network and devices, and their effect on HOs.

CCS Concepts

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

Keywords

Mobile Networks; Network Topology; Traffic Analysis; Measurements; Network Performance

1 Introduction

The advent of the 5G New Radio (NR) technology marked a shift in the telecommunications landscape, offering unprecedented speed and widespread connectivity to a vast array of devices [38]. With these developments, end-user expectations have surged, driven by the promise of higher bandwidth, lower latency, and importantly, ubiquitous connectivity for fast-moving User Equipment (UE) through effective handovers (HOs). However, like any emerging technology, the pace of real-world deployments does not instantly match the pace of innovation [32, 41], resulting in multiple generations of technology operated simultaneously to balance the trade-offs between OPEX/CAPEX and the stringent needs for high availability, reliability, and capacity. In this complex arena, HOs have become more intricate yet crucial for maintaining seamless connectivity.

Prior works study the intricacies behind HO management [23, 26, 52], yet the majority of them [13, 18, 23, 26, 27, 34, 36, 37, 42, 43, 50–52] analyze HOs based on limited user-side measurement campaigns and, as such, are confined to specific mobility scenarios, geographic areas, or UE manufacturers. These limitations restrict the conclusions we can draw and underscore the need for a detailed, large-scale analysis that captures the complexity and heterogeneity behind real MNO deployments.

The goal of this paper is to fill the existing gap by presenting the *first, to our knowledge, countrywide analysis of mobility management, from the perspective of a top-tier Mobile Network Operator (MNO) in Europe.*[1] We have recorded *all mobility events* over a period of

---

[1]To maintain the confidentiality of the operator, we are only able to disclose general location information for the study.

Table 1: Dataset statistics.

Feature	Value
Area covered	Country in Europe (300+ districts)
# of cell sites	24k+
# of radio sectors	350k+
# of UEs measured	≈40M
# handovers (daily)	1.7B+
Measurement duration	4 weeks (28 days)
Trace size (daily)	≈8 TB

4 weeks at millisecond granularity, across the entire country. Our datasets (see Table 1) include all HOs and HO failures (HOFs) that occurred over the observed period, enabling the analysis of network dynamics at a crucial moment when – at the time of writing – *all* digital radio access technologies (RATs) developed during the last three decades are concurrently operational within the same network. We merge this data with: (*i*) information from the MNO's deployment, to study HO performance across its topology and supported RATs, (*ii*) device-specific information, to associate HOs and HOFs with specific UE types and manufacturers, and (*iii*) the data from the country's official census office, to account for the geodemographic distribution of HOs across 300+ districts with various population density. At the time of the study, the MNO was just initiating its commercial 5G-Standalone (SA) deployment, so we measured only the 5G-Non-Standalone (NSA) deployment to avoid any early-stage issues with SA [29].

To capture the complexity and heterogeneity behind the studied network-side datasets, we define three main dimensions that significantly affect HOs and mobility management: (*i*) the heterogeneity of RATs in the MNO, namely 2G, 3G, 4G, and 5G,[2] (*ii*) the heterogeneity of UEs, and (*iii*) the geodemographic diversity. We analyze the spatio-temporal dynamics of horizontal (intra-RAT), and vertical (inter-RAT) HOs, at the district-level and with msec granularity, and characterize their pattern across the country to identify regional trends. Furthermore, we dissect the impact of UE types (smartphones, M2M/IoT[3] devices, low-tier feature phones) and manufacturers on HOs, HOFs, and mobility/performance metrics. We also analyze the causes behind HOFs, using 3GPP-based and vendor-specific failure descriptions. Lastly, we leverage statistical methods to model how the coexistence of multiple RATs affects HO performance, especially when UE connections are downgraded to older technologies (e.g., 2G, 3G). Below we present the key findings and contributions of this work.

• **Heterogeneity & Complexity of HOs (§4)**. In the MNO's deployment, 5G sectors make up 8.4% while 4G accounts for 55%, with 2G and 3G sectors covering the remaining ≈36% and handling 18% of user connectivity time. Despite this, older RATs carry only 5.23% of the uplink (UL) and 2.07% of the downlink (DL) data flowing through the network. Among all UEs, 59.1% are smartphones, primarily from Apple (54.8%) and Samsung (30.2%), from which 51.5% do not support 5G, relying instead on 4G. Additionally, over 32% of the UEs, mainly M2M/IoT devices and feature phones, support

only up to 3G. This blend of technologies highlights the challenges of phasing out older RATs, particularly in an environment where IoT manufacturers still rely on 3G/2G for devices with limited connectivity needs. Our geodemographic analysis points to a large disparity between the density of HOs in urban centers with larger population density (2.1M HOs per sq. km), and less populated rural areas (60 HOs per sq. km); in a network deployment that registers on average 13.1k HOs per sq. km.

• **HO Analysis (§5)**. Taking as a reference the HOs registered in the 4G EPC, approximately 94% of HOs are horizontal (between 4G/5G-NSA radio sectors), complete within 90 ms (median of 43ms), and correspond to smartphone activity. M2M/IoT UEs and low-tier feature phones – accounting for >40% of the device population – share the remaining 6% of HOs. By investigating the top-5 smartphone manufacturers (Apple, Samsung, Motorola, Google, Huawei), we discover similar patterns in terms of HOs (±10% of variation between them) and low HOF rates (Google exhibits -27% of HOFs w.r.t. other UEs, but with higher variability). Moreover, we find some smartphone manufacturers outside the top-5 (e.g., KVD) that exhibit higher HOF rates (up to +600% w.r.t. other UEs) and HO signaling (up to +293%).

• **HOF Analysis & Modeling (§6)**. Rural areas (with sparser deployments) experience 32.4% more HOFs during peak hours [7:00–8:00) than urban areas. Moreover, HOF rate is close to zero for the majority of the UEs; for the ones with high mobility metrics (>100 visited sectors, >100km radius of gyration), which are mostly smartphones (85%), HOF rate rises up to 0.4% (pct-75).

Furthermore, we dissect the reasons why HOs fail by using 1k+ 3GPP and vendor-specific descriptions that explain the causes. Interestingly, we find that 92% of the HOs in the entire country fail with solely 8 causes; and from the studied failures, 75% (0.03%) occur in HOs to 3G (2G), and 25% of them are due to an excessive load in the target sector (Cause #4). Moreover, we measure the duration of these 8 causes and highlight that the ones related to specific cancellations (Cause #1) and timeouts (Cause #8) require on average >2s to complete, reaching up to 10s in the latter case.

On top of the previous analysis, we aim to unveil what network-related features correlate with HOFs. Specifically, we test whether the HO type (intra/inter-RAT) is a good predictor for the HOF rate. Statistical analysis verifies that, although they infrequently occur (only 6% of all HOs are to 2G/3G), HOs to 3G (2G) increase the HOF rate by 166% (915%) compared to HOs between the newer RATs (intra 4G/5G-NSA).

Organization. The remainder of this paper is organized as follows. §2 covers relevant handover concepts, while §3 details the datasets and metrics used in our study. §4 introduces the three main axes of analysis that we use in the paper, and §5 shares insights on HOs based on these axes; §6 examines the impact and causes of HOFs. Additionally, §7 reviews related literature, providing context for our findings. Lastly, §8 and §9 discuss the study's limitations and avenues for future work, and conclude with a summary of the main findings.

---

[2]The Second, Third, Fourth, and Fifth Generation networks, and their respective RATs, are henceforth referred as 2G, 3G, 4G, and 5G, respectively.
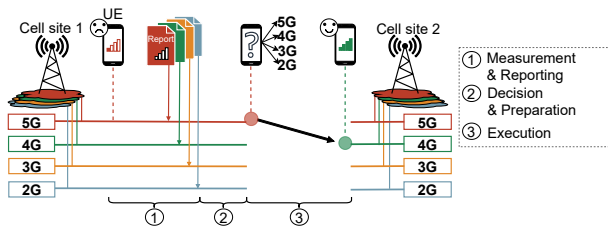[3]See Appendix C for the meaning of all abbreviations (e.g., M2M/IoT).

Figure 1: High-level description of HO procedure [6].

2 Background

MNOs install *cell sites* in various locations to ensure widespread service availability. These sites are typically equipped with multiple antennas associated with *radio sectors* (or simply, *sectors*) that support different RATs and serve the UEs residing in a bounded geographical area [46]. This design principle mandates that, as UEs move, they must transition, or *handover*, across different sectors. To offer a seamless network connectivity experience for mobile users, it is essential for HOs to execute timely and reliably.[4]

Configuring the cell sites and their sectors to optimally handle HOs is crucial: frequent HOs may lead to excessive signaling and unneeded service interruptions, whereas scarce HOs might result in poor signal reception or temporary lack of connectivity. Moreover, with the advent of 5G and its two deployment strategies [29], NSA and SA, HO management has become more intricate [37] due to the coexistence and integration of multiple generations of mobile technologies;[5] at the same time, a crucial goal of 5G-SA is to provide support for Ultra Reliable Low Latency Communications (URLLC), which mandates the provision of reliable and fast HOs.

Handover Mechanism. Every UE relies on its *primary sector* (i.e., the sector it is connected to), serving as the pivotal link for control-plane signaling and HO management. Figure 1 depicts the HO process from a source (i.e., primary) to a target radio sector [5]. When a UE attaches to a new sector, it receives a set of mobility management configurations, including parameters for the triggering of HO events (e.g., hysteresis, offsets, etc.). Based on these configurations, the UE performs signal strength/quality measurements – e.g., Reference Signal Received Quality (RSRQ) – of the source and neighboring sectors, and sends a Measurement Report (MR) to the source periodically, or if any of the mobility management criteria is met. For instance, in 4G and 5G NR, a HO triggering event typically occurs when either the serving sector's signal falls below a threshold (A2 event) or when the signal of a neighboring sector becomes offset better than the serving sector (A3 event) [4, 7]. Based on the MR, the source identifies the best target sector and initiates the HO. After the target sector accepts the request, the source transmits a HO command to the UE. For example, in 4G/5G NR, the source sends a Radio Resource Control (RRC) Connection Reconfiguration message to the UE to begin its cell synchronization with the target sector and the Random Access Channel (RACH) procedure. After

the UE reports successful access to the target sector, the source releases its resources. More details are available in [6, 33].

3 Datasets & Methodology

In this section, we present our measurement infrastructure in a large MNO in the studied country. We explain the three datasets we built for this study and introduce the official census dataset we used to complement our analysis. Finally, we detail the employed performance and mobility metrics.

3.1 Network Data Collection

Measurement Infrastructure. We collect passive measurements using commercial tools integrated into the MNO's infrastructure, see Figure 2. In a nutshell, the cellular network architecture can be divided into three primary components: (*i*) the devices accessing the network, (*ii*) the Radio Access Network (RAN), responsible for managing wireless communication, and (*iii*) the Core Network (CN), which provides the required services and functions for the network operation (e.g., user authentication and mobility management). This is consistent for all the different radio technology generations that coexist in the network. Our monitoring locations, which we depict with red pins in Figure 2, focus on key components of the core network, including the Mobile Management Entity (MME tracks and manages the mobility of devices in 4G and 5G-NSA), the Mobile Switching Center (MSC is responsible for communication switching functions), the Serving GPRS Support Node (SGSN manages data routing for 2G/3G), the Serving Gateway (SGW routes packages between RAN and the CN), and the cell sites in the RAN.

Data is collected in a private cloud environment for a given retention period, and is already anonymized before we process it. Particularly, we organize the collected data into three datasets, providing various information at the radio sector and UE-level.

Mobility Management Signaling Dataset. The captured data spans from 29-Jan-2024 to 25-Feb-2024 for the entire country (see Table 1). We analyzed the activity of users in the control plane for all RATs supported by the MNO. For each RAT, the dataset includes the (control plane) signaling messages related to events such as service requests, HOs, attach/detach, paging, and Tracking Area Update (TAU). We direct our attention to HOs, for which we capture six main variables that enable an in-depth analysis: (*i*) *timestamp*, with millisecond granularity, (*ii*) *HO result* (i.e., success/failure), (*iii*) *HO duration* (msec granularity), (*iv*) *cause codes for HO failures based on 3GPP* [3, 5], which are enriched with *sub-cause descriptions* specified by the antenna vendors, (*v*) *anonymized user ID*, based on the International Mobile Subscriber Identity (IMSI) and the International Mobile Equipment Identity (IMEI),[6] and (*vi*) source and target radio sectors along with their RATs. As mentioned before, due to the early stages of 5G-SA deployment in the studied MNO, we base our analysis on 5G-NSA, which relies on the 4G EPC for mobility management.

Radio Network Topology. We utilize this dataset to integrate in our analysis the upgrades in the MNO's network deployment footprint (e.g., newly deployed sites). We capture this dataset daily

---

[4]The HO procedure is different from *cell (re-)selection*, which happens when users do not maintain an active data connection but still need to change sectors to ensure the reception of signaling messages [48].

[5]From a mobility management perspective, 5G-NSA and 4G are identical, since the former relies on the 4G Evolved Packet Core (EPC) functions.

[6]The first 8 digits of the IMEI represent the Type Allocation Code (TAC), which we use later to classify devices.
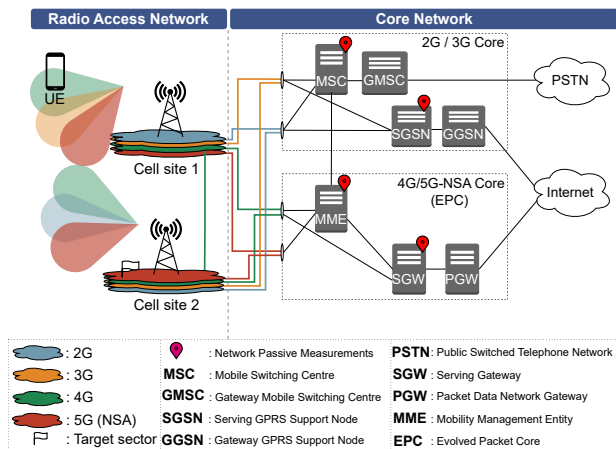
Figure 2: Network architecture & measurement points.

during the period of analysis; it contains information on each deployed radio sector, such as geographic location (longitude and latitude), the postcode of the area, and the supported technologies (i.e., 2G, 3G, 4G, 5G).

Devices Catalog. We leverage a daily commercial database, provided by the Global System for Mobile Communications (GSM) Association (GSMA) to examine correlations of device-specific characteristics with HOs. This catalog associates the TAC of each device with attributes such as the supported radio bands and RATs, the manufacturer, and the device type. We apply a heuristic to classify the devices into three types: smartphones, M2M/IoT devices, and low-tier feature phones [31]. For this, we rely on the observation that the Access Point Name (APN) configured for the UEs may contain keywords associated with IoT verticals (e.g., "m2m", "smart-meter"), and combine the information from the APN with the device characteristics of our daily commercial GSMA database.

3.2 Census Dataset

We leverage open datasets published by the official census office in the studied European country to enrich our mobility study with the geodemographic characteristics of different areas. Specifically, we take as a reference the 300+ districts defined by the census office, and collect the population density and the postcodes within each of them. Then, based on census information we classify postcode areas into two main categories (*urban* and *rural*), which correlate with population density (more than 10k and less than 10k residents) and also serve as a proxy for areas with denser and sparser RAN deployments, respectively.

3.3 Performance & Mobility Metrics

Performance Metrics. In line with prior works [9, 48], we focus on the following metrics.

• *HO count* represents the number of HOs over a time interval, which we usually set to either 30 minutes, 60 minutes, or one day. We use this metric to show how users' mobility fluctuates over time and space, and how it differs per RAT, device type, and manufacturer.

• *HO duration* represents the time interval (in msec) to complete the HO, see [18–20, 22]. Minimizing this interval is crucial for seamless connectivity and improves the users' Quality of Experience (QoE).

• *HOF rate* refers to the number of HO failures (HOFs) divided by the total number of triggered HOs.[7] HOFs dramatically affect the users' experience and typically happen due to poor signal strength, configuration and synchronization errors, or capacity issues in the network. In §6, we uncover the reasons why these failures occur and emphasize that a comprehensive understanding of HOFs can only be achieved by incorporating the perspective of MNOs.

Mobility Metrics. To showcase the mobility characteristics of users, we focus on two metrics from the MNO's perspective, as follows.

• *Number of sectors* quantifies the number of distinct radio sectors that a user successfully communicates with, per day. We highlight that this metric does not necessarily translate to the distance traveled by users in a given area, as the density of radio sectors in the area also plays a role. For instance, urban areas typically have denser deployments and, as a result, users connect to a larger number of sectors even if they travel the same distance as in less populated areas (e.g., rural) with sparser deployments.

• *Radius of gyration* complements the previous metric by capturing the traveled distance for the UEs [21]. It is defined as the root mean squared distance between each visited sector (weighted by the time spent there) and the center of mass. The radius of gyration is defined as follows:

$$g = \sqrt{\frac{1}{N} \sum_{j=1}^{N} (t_j \mathbf{l}_j - \mathbf{l}_{cm})^2},$$

where $\mathbf{l}_j$ represents the location of the jth visited cell site, t_j represents the time spent in the jth visited cell site and $\mathbf{l}_{cm}$ represents the location of the user's center of mass, calculated as $\mathbf{l}_{cm} = \frac{1}{N} \sum_{j=1}^{N} \{t_j \mathbf{l}_j\}$, where N is the total number of cell sites visited by the user. A high radius of gyration indicates that the user travels far and wide (i.e., their moves span a large geographical area). Conversely, a lower radius of gyration points to more localized movements, relatively close to a central location.

4 Exploring Data Heterogeneity

Our datasets capture the heterogeneity and complexity of HOs across the entire MNO's deployment in the studied country, which includes diverse deployment densities and RATs, as well as a broad spectrum of UEs (e.g., smartphones, M2M/IoT, etc). In this section, we explore the heterogeneity of these datasets along three particularly interesting axes, from the network's perspective: (*i*) heterogeneity of RATs, (*ii*) heterogeneity of UEs, and (*iii*) geodemographic complexity.

4.1 RAT Deployment & Usage

Figure 3a shows the deployment evolution in the network from 2009 to 2023. The number of sectors (solid pink line) has increased at an exponential pace in the last 15 years, with an average growth of 59% during the last 5 years (2018-2023). Throughout these 15 years, different RATs have been coexisting with a varying mix. The latest

---

[7]We primarily focus on HO failures rather than explicitly detailing HO successes; however, successes and failures are complementary to each other.
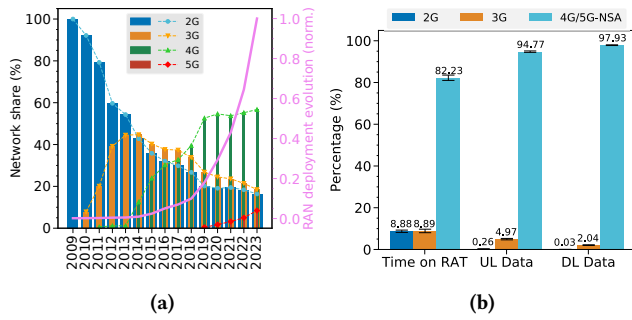
Figure 3: (a) Evolution of network deployment in a commercial MNO. The left y-axis corresponds to all bars and lines, except the pink line (right y-axis). (b) Average daily RAT use. Error bars show the min/max daily values over 4 weeks.

major network upgrade occurred in 2019 with the deployment of 5G-NR, which accounted for 8.4% of the sectors by the end of 2023. At the same time, we observe the gradual decommissioning of 2G and 3G cells (≈18% each in 2023), while 4G is still the dominant RAT (≈55%) in terms of infrastructure. This heterogeneity does not come as a surprise, since decommissioning legacy RATs is a challenging process that needs to account for various techno-economic factors, such as the turnover rate of customers or the radio coverage [14]. Nonetheless, it compounds network management and affects both the number and the success of HOs as we present in §6, and as prior studies have also identified [27, 53].

To further understand the use of the RATs, we compute the overall time that UEs spend on each of them by using the timestamps of mobility events in the dataset. With the current 5G-NSA deployment, we do not distinguish from the events captured in the core network (i.e., MME) when devices are served by a 4G or a 5G-NR sector (see §2); thus, we use the term "4G/5G-NSA". In Figure 3b, we notice that UEs rely mostly on 4G/5G-NSA (≈82% of the time on average), while 2G and 3G serve users during a non-negligible 8.9% of the time each. In terms of aggregated data volumes, the share for 4G/5G-NSA rises up to 94.77% and 97.93%, respectively, for UL and DL traffic, leaving marginal values for 2G and 3G. Yet, these legacy RATs still serve a noteworthy number of UEs that support only these technologies (see §4.2).

The heterogeneity of the network appears also in terms of the antenna vendor. Four principal vendors (V1, V2, V3, V4) employ antennas (and thus, RATs) for this network, with their deployment distributed asymmetrically across different regions. All vendors support 4G/5G-NSA and 3G RATs, and accommodate nearly the full spectrum of devices. Details are provided in Appendix B.

Key takeaways: *The cellular network we measure includes all RATs (2G-5G), where 2G and 3G radio sectors account for 36% of the total deployment. These RATs (2G & 3G) connect users on average for 18% of their up-time, while UEs generate only 5.23% (2.07%) of the UL (DL) data over them.*

4.2 User Equipment

The number of devices accessing the network over the 4-week period is ≈40M. We classify these devices into three types based on
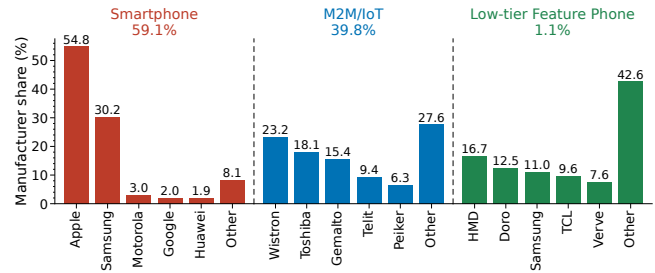
(a) Manufacturers
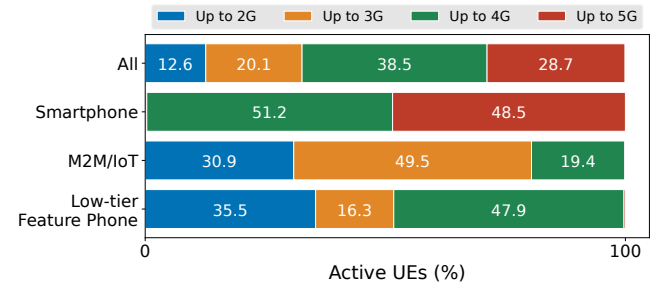
(b) RAT support

Figure 4: Percentages of (a) manufacturers and device types, and (b) the RATs they support (excluding <0.5%).

their capabilities, namely, **smartphones**, **M2M/IoT devices**, and **low-tier feature phones**, accounting for 59.1%, 39.8%, and 1.1% of UEs, respectively. Figure 4a shows the top-5 manufacturers in the three types of devices. In the larger category – smartphones – we observe that most devices are manufactured by Apple (54.8%) or Samsung (30.2%). For M2M/IoT UEs, we find a diversified set of manufacturers; namely, over 27% of these UEs are from manufacturers outside the top-5.

We infer the connectivity capabilities of mobile devices from the GSMA device catalog (see §3). We find that 12.6% of all UEs support only 2G and 20.1% up to 3G (see Figure 4b), which partially explains the slow pace of decommissioning legacy RATs. These legacy devices are mostly M2M/IoT devices or feature phones, where > 80% and > 50%, respectively, support at most 3G. The overall number of devices that support 4G or 5G adds up to 67.2%. The majority of these devices are smartphones: 51.4% of smartphones support up to 4G, and 48.5% are 5G-capable.

Key takeaways: *Over 32% of all devices support only up to 3G – predominantly M2M/IoT UEs and feature phones – and 51.5% of smartphones do not support 5G yet (the majority relies on 4G). These factors contribute to the presence of a mixture of old and new RATs in current deployments, stressing the challenges associated with decommissioning the older ones.*

4.3 Geodemographic segmentation

Population Sampling. This section demonstrates that the dataset we collect through the commercial MNO is representative of the country's overall population. Figure 5 shows the population according to census (y-axis) and the population we inferred from the MNO
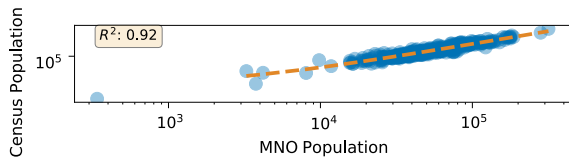
Figure 5: Comparison between the inferred population from the MNO data and the actual population from the census data (district-level).

Figure 6: Daily HOs per sq. km in the country (district level).

(x-axis), where each data point refers to the districts in the country (see §3). We derive the end-user's home location at postcode granularity from their connectivity patterns during nighttime [8]. To achieve this, we consider the main cell site the user connects to between 00:00 and 08:00 (i.e., night hours) for at least 14 days (not necessarily consecutive) during February 2024. We then aggregate their mapped home postcode at the district level. These results show a clear linear relationship ($R^2 = 0.92$) between the census data and the MNO user base, which reinforces that our dataset accurately captures the country's population distribution [40]. This renders our dataset especially interesting for analyzing mobility in the entire country, including regions with diverse population dynamics.
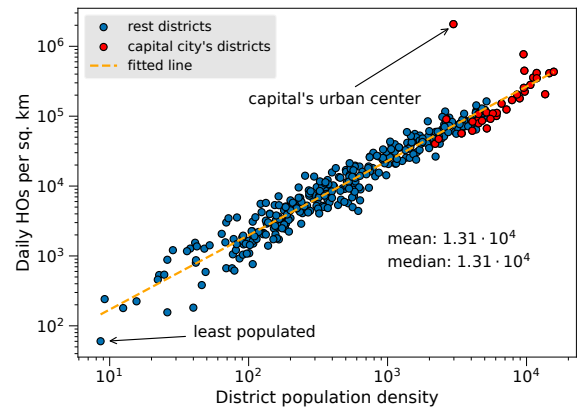
Mobility & Geodemographics. We investigate the distribution of mobility events by examining the number of HOs across districts. Figure 6 shows the number of daily HOs per sq. km in each district, together with the population density there (residents per sq. km). This analysis facilitates the characterization of mobility patterns across distinct geodemographic segments (e.g., densely populated urban areas or less populated rural areas). Overall, our findings indicate a strong positive correlation (Pearson correlation of 0.97), between the number of HOs per day and the residential population density in the corresponding district.

As anticipated, dense urban areas exhibit a high number of HOs per square km. For instance, in the district that covers the urban center of the capital, we observe approximately 2.1M HOs per square km each day. In this city the studied network's infrastructure itself comprises more than 500 radio sectors per square km. Conversely, in less populated areas the intensity of HOs is significantly lower (60 HOs per sq. km in the least densely populated district). This value is more than 200× lower compared to the district-level mean in the country (13.1k HOs per sq. km daily), reflecting the stark contrast in mobile network activity between highly urbanized and more remote areas.

Key takeaways*: We infer the home locations of approximately 40M UEs across the studied country to ensure that our data accurately represents the entire population ($R^2 = 0.92$ with census data). By analyzing HOs per square km at the district level, we observe significant disparities – from 2.1M daily HOs per sq. km in the in the center of the capital city to 60 HOs per sq. km in remote areas – highlighting the complexity of managing HOs across different regions.*

5 Characteristics of Handovers

Analyzing mobility patterns is crucial for various purposes, including urban planning, social policy design, and optimizing network infrastructure [26, 54]. In this section, we take as reference the three axes of heterogeneity from §4, and characterize geo-temporal cellular mobility patterns through HOs. We examine the horizontal and vertical HOs across UE types and districts, and investigate how mobility and UE manufacturers relate to HO performance.

5.1 Geo-temporal Analysis

First, we analyze HO patterns as a function of geodemographic factors, focusing on the difference between urban and rural areas, classified at the postcode level (§3.2).[8] This broader urban/rural classification enables us to robustly capture variations in HO dynamics across areas with different demographic characteristics.

HO Patterns. Figure 7 (top) shows the weekly temporal evolution (with *30-min granularity*) of HO counts over the 4-week period (shadows show the min/max values). To adhere to privacy and security guidelines of the MNO, we normalize the HO counts by the max value (over 30-min intervals) of the studied period. The total number of HOs in urban areas represents on average 78% of all HOs, which is consistent with findings from other studies [55]. Namely, we discover that 80% of the sectors are installed in urban areas, which cover only 49.6% of the total territory of the country.

From the daily HO patterns, we observe that weekdays (Mo-Fr) experience higher number of HOs compared to weekends (Sa-Su). Concretely, we find a 33% reduction on average in the peak of HOs during Sundays compared to Fridays. Moreover, we identify the peak HO times during weekdays at 08:00–08:30 and 15:00–15:30 for both rural and urban areas. Also, weekday HO patterns exhibit notable fluctuations, with a sharp ×3 increase in the HOs observed from 06:00 to 08:00; this is in contrast to weekends, which have a single peak of mobility between 12:00 and 13:00. During weekdays, after the second peak at 15:00–15:30, the number of HOs gradually decreases (on average 11% per 30 minutes), leading to the minimum at 02:00–03:30 (or 03:00–05:00 over the weekends).

Likewise, Figure 7 (bottom) shows the number of active sectors – handling at least one HO – over 30-min intervals. As underlined in the sequel, MNOs apply dynamic energy-saving policies to switch off sectors when they are not needed to satisfy capacity demand.

---

[8]We drop from this analysis 3.1% of the postcodes due to the lack of reliable census information in these areas.
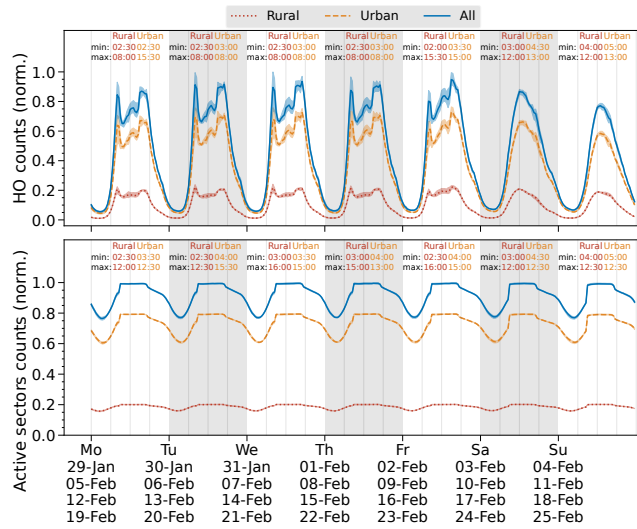
Figure 7: Temporal evolution of HOs (top) and active sectors (bottom) in urban and rural areas. Curves show the average HO volume in 30-minute intervals over the four weeks; shadows show the min/max values. All values are normalized by the max HO and sector count seen over the period.

Table 2: Statistics per handover and device type.

| | Horizontal | Vertical | | All |
	Intra 4G/ 5G-NSA (%)	4G/5G-NSA to 3G (%)	4G/5G-NSA to 2G (%)	HOs (%)
Smartphones	88.28 ± 0.77	5.84 ± 0.77	< 0.001	94.12 ± 0.77
M2M/IoT	5.73 ± 0.52	0.02 ± 0.01	< 0.001	5.75 ± 0.53
Feature phones	0.13 ± 0.05	< 0.001	< 0.001	0.13 ± 0.05
All devices	94.14 ± 1.29	5.86 ± 0.78	< 0.001	-

Figure 8: HO duration (horizontal vs vertical).

Comparing Figures 7 (top) and 7 (bottom), we see that the portion of active sectors highly correlates with the HO counts (Pearson correlation of 0.9). Weekdays and weekends present no significant differences in terms of active sectors. More precisely, after 08:00 (first peak hour) ≈99% of sectors remain active until 17:00, when a decrease of ≈1% per 30-min is observed, until midnight. As mentioned earlier, we conjecture that this decrease correlates not only with the reduced mobility of the UEs (notice the HO drop at the same hours), but also with the reduced capacity demand (i.e., less user activity) in densely deployed areas, which triggers energy-saving mechanisms to switch off sectors that act as capacity boosters [11, 47].

Key takeaways: *Handover patterns vary significantly across: (i) urban and rural areas (urban sectors account for 78% of HOs, while covering only 49.6% of the territory), and (ii) during weekdays and weekends (33% of difference during peak hours).*
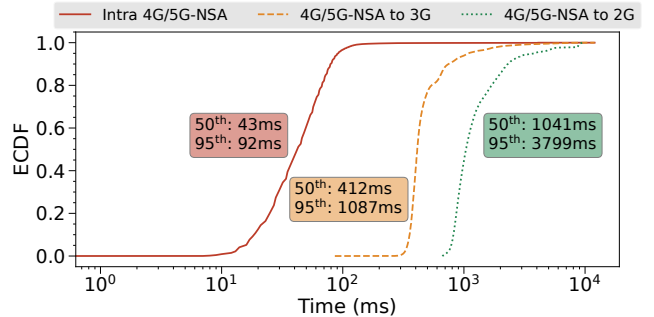
5.2 Horizontal vs Vertical HOs

To understand how devices interact with the different RATs in the network (see §4), we take as a reference the behavior of devices connected to 4G/5G-NSA (i.e., 4G and 5G-enabled devices). We differentiate three main *HO types*, namely, intra 4G/5G-NSA (horizontal), 4G/5G-NSA → 3G (vertical), and 4G/5G-NSA → 2G (vertical). Our intent is to characterize how frequently these devices still rely on older RATs, and in which circumstances.

HO Frequencies. Table 2 depicts the percentage of the different HO types we registered across UE types. The vast majority of HOs are intra 4G/5G-NSA HOs (94.14%), while vertical HOs – from 4G/5G-NSA to 3G or 2G – correspond to 5.86% and 0.001%, respectively.

Furthermore, smartphones primarily initiate intra-4G/5G-NSA HOs, contributing to 88.28% of the total, with a fallback to 3G

occurring in 5.84% of the cases. M2M/IoT devices engage mostly in intra 4G/5G-NSA HOs, with a minority transitioning to 3G, a pattern echoed by feature phones on a smaller scale (i.e., 0.13% intra 4G/5G-NSA HOs). This is particularly important given that ≈80% of M2M/IoT devices support only up to 3G (see Figure 4b). It is an artifact of the IoT vertical applications employing massive M2M deployments (e.g., smart meter applications), which often require only stationary devices with limited connectivity demands [31].

HO Duration. Figure 8 illustrates the signaling times of HOs (see definition in §3.3), revealing that 95% of intra 4G/5G-NSA HOs complete within ≈90ms (median of 43ms). These results align with previous studies [18–20]. In contrast, HOs from 4G/5G-NSA to 3G are one order of magnitude longer, with a median of 412ms and their 95th percentile exceeding 1s. The latency further increases for vertical HOs to 2G, where the median time matches the 95th percentile for HOs to 3G (≈1s), and the 95th percentile stretches beyond 3.8s. Even if these HO types rarely occur (see Table 2) their large duration reveals a clear negative impact of vertical HOs. We delve into the duration of HOFs in §6 through their causes.

HOs per District. Figure 9 provides a comprehensive view of HO dynamics across districts in the studied country. In this way, we are able to pinpoint the areas that are more dependent on newer/older RATs. Notably, densely populated urban districts – which include the districts of the capital city – exhibit a high penetration of 4G/5G-NSA (up to 99.92% of all HOs, see Figure 9a), while some less populated rural areas show more transitions to legacy RATs. For example, in the 6% least densely populated districts, HOs to 3G account for 26.5% on average of all HOs, and reach up to 58.1% for a specific remote district (Figure 9b). Likewise, the percentage of transitions to 2G remains marginal for most of the districts, with a maximum of ≈0.5% for 4 specific districts. (Figure 9c).
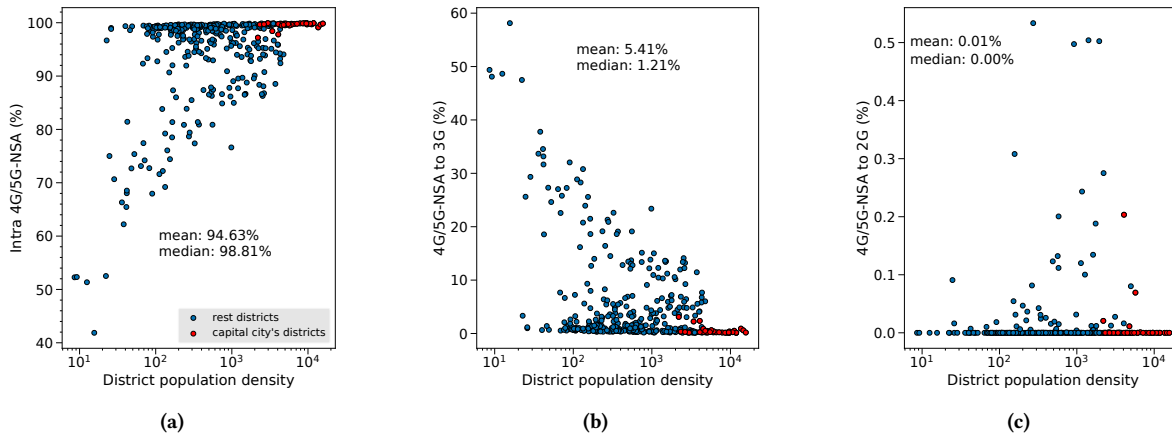
Figure 9: Distribution of (a) intra 4G/5G-NSA, (b) 4G/5G-NSA to 3G, (c) 4G/5G-NSA to 2G, HOs across districts. Y-axes have different scales.
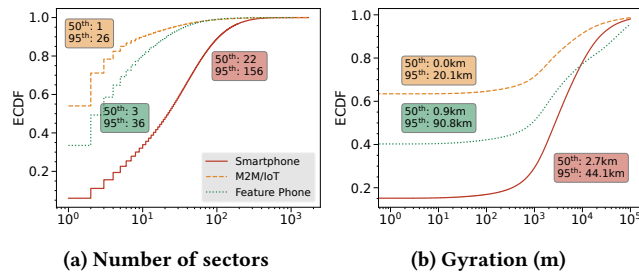
Figure 10: Mobility metrics across device types.

Key takeaways: *(i) 94% of HOs are intra 4G/5G-NSA, and are triggered by smartphones. (ii) HOs to 3G/2G take up to 3.8 seconds (pct-95) to execute and still represent 6% of all HOs. (iii) The most densely populated urban areas rely almost exclusively on 4G/5G-NSA for HOs (>99%); less densely populated rural areas still use older RATs (HOs to 3G are up to 58.1% in a remote area and on average 26.5% in the least densely populated districts). This analysis helps the MNO to identify areas where a great volume of 4G and 5G-capable devices are frequently using legacy RATs, thus building a realistic strategy towards their decommissioning.*

5.3 Mobility across Device Types

This section examines the relationship between UEs' mobility and their HO performance. We first characterize mobility metrics across different device types. Then, we analyze the relation between the mobility metrics of the UEs and the HOF rate that they experience, serving as an indicator of how these UEs suffer from service disruptions.

We take as a reference the two mobility metrics described in §3.3: radius of gyration and number of sectors. Figure 10 shows the empirical cumulative distribution function (ECDF) of both mobility metrics across device types. Overall, we observe that smartphones are considerably more mobile than the two other types, exhibiting a median of 22 distinct visited sectors per day, and a median radius

of gyration of 2.7km. Conversely, the majority of M2M/IoT devices and low-tier feature phones are more static, with median values of 1 and 3 visited sectors per day, respectively, and a median gyration of 0.0km and 0.9km. This reflects that these UEs are mostly static, and the few HOs that these devices experience are typically between sectors in the same sites.

Given the heterogeneity of M2M/IoT vertical applications, there are devices in the 95th percentile that show high mobility, with gyrations of 20.1km for M2M/IoT devices (see Figure 10b). These UEs mainly correspond to modems and routers that are deployed in fast-moving vehicles (e.g., trains), integrated into modern cars, embedded in industrial equipment, or wearable IoT devices carried by users who typically travel long distances. While feature phones (green line) surpass smartphones (red line) at around the 80th percentile, the former comprise only about 1% of the total UEs, while the latter makes up approximately 60% of UEs (see Figure 4a).

Manufacturer Impact. We assess whether higher HO counts and HOF rate correlate with specific UE manufacturers (e.g., due to a suboptimal mobility management implementation). We observe that the distribution of UEs is remarkably unbalanced across the studied country, e.g., Samsung and Apple smartphones are considerably more common in densely populated areas. To make a fair comparison and account for potential deviations due to the area itself (e.g., population, deployment density) – see Figure 6 – we create a metric that makes a unified comparison of UE manufacturers at the district level. That is, in each district we get the average HOs per UE for a specific manufacturer and divide it by the average HOs per UE including all manufacturers within that district (i.e., *normalized district-level HO*);[9] and similarly for the HOF rate (i.e., *Normalized district-level HOF rate*). A value greater than 1 indicates that UEs of a specific manufacturer generate more HOs (or HOF rate) on average than the total population of UEs in the same district.

Figure 11 shows the results for the top-5 manufacturers in the studied country (see Figure 4a), as well as the 5 manufacturers

---

[9]Some manufacturers have few devices in specific districts. We exclude district-manufacturer pairs that account for <1k devices.
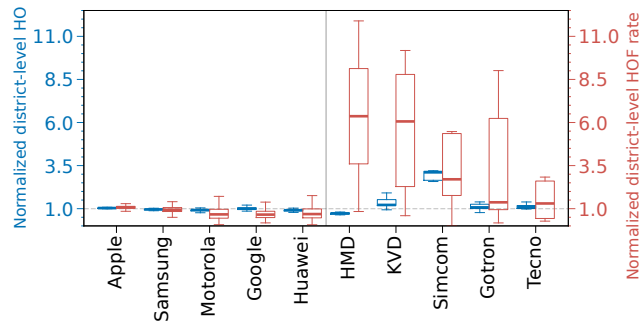
Figure 11: Normalized district-level HOs (left) and HOFs (right) per UE manufacturer. Boxplots include top-5 UE manufacturers and the top-5 UE manufacturers with greater median HOF values.

Figure 12: HOF counts per hour in urban and rural areas, normalized separately with the number of active sectors in each class (i.e., urban/rural).

exhibiting the highest *Normalized district-level HOF rate*, based on the median behavior across all districts (see boxplots). For the top-5 manufacturers ratios are close to 1, which means that devices behave similarly to their peers in the same district, both in terms of HOs and HOF rate. Specifically, we observe that Apple smartphones, the most popular ones (≈32% of all UEs), generate slightly more HOs per UE and HOF rates than other devices (respectively +4% HOs and +8% HOF w.r.t. their peers). Likewise, Google smartphones are the ones that experience the smallest HOF rates (-27% w.r.t. their peers). Moreover, we find that some manufacturers show high HOF rates, such as KVD smartphones or HMD feature phones (+600% HOF rate), as well as others that generate higher HOs per UE, such as Simcom M2M/IoT (+293% HOs per UE).
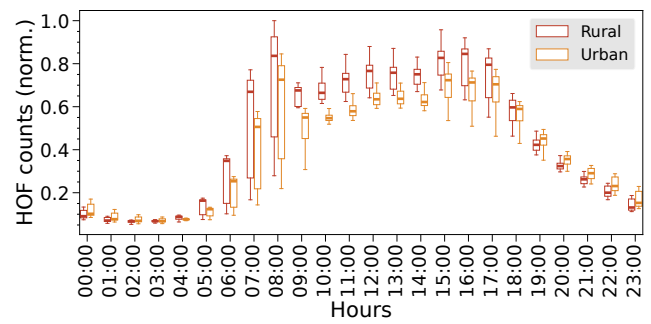
Key takeaways: *(i) Different UE types exhibit different mobility patterns; smartphones are, on median, the ones connecting to more distinct sectors (22 sectors per day), with a daily median radius of gyration of 2.7km. (ii) The most popular device manufacturers exhibit a consistent behavior in terms of HOs (±10% of variation between them). While HOF rates are considerably small, some manufacturers (e.g., Google) exhibit lower HOF rates (-27%) than other manufacturers. For some niche manufacturers, we find high HOF rates (up to +600%) and HO counts (up to +293%). Based on these results, we conjecture that manufacturer-specific mobility management implementations and application-specific usage correlate with HO performance.*

6 Handover Failure Analysis

This section provides an in-depth analysis of HOFs. Initially, we examine the daily patterns of HOFs and their correlation with key mobility metrics. Next, we explore the causes of HOFs from the network's perspective and present modeling techniques that assess how network features at the radio sector level influence the HOF rate. Our analysis puts the spotlight on the need to reduce the network's complexity by decommissioning legacy RATs.

6.1 HOF Patterns & Impact

HOF Patterns. We analyze the temporal evolution of HOF rate (see §3.3) along the day, aggregating data over the 4-week period. Figure 12 presents the hourly evolution of HOFs, where boxplots aggregate data from all active radio sectors at a specific hour. To comply with the privacy policies of the MNO and account for the different distribution of sectors in rural and urban areas, we have separately normalized the hourly HOFs for rural (urban) areas with the number of active sectors observed in rural (urban) settings (see Figure 7, bottom). Overall, we observe that HOFs reach a local peak during the morning commuting time [7:00−9:00], and a lower local peak can be observed during the afternoon commuting time [15:00−18:00]. Moreover, urban areas experience fewer HOFs compared to rural ones, especially during peak hours; e.g., the median HOF count is 32.4% higher in rural areas than in urban ones during [7:00−8:00]. We conjecture that this pattern is likely due to the more limited 4G/5G coverage in these areas, which makes 4G and 5G-capable devices fall back more frequently on older RATs (i.e., 2G, 3G) to keep connectivity. We further delve into this aspect in §6.3 by modeling the negative impact of vertical HOs on HOFs and inspecting the causes of such failures.

HOFs & Mobility. We explore the association of radius of gyration and number of sectors with the HOF rate. In Figure 13, the left y-axis shows the daily average HOF rate for the UEs according to the number of sectors (Figure 13a), or radius of gyration (Figure 13b). Meanwhile, the right y-axis displays the ECDF for the number of UEs along the bins in the x-axis (in log scale).

Concretely, Figure 13a shows that the HOF rate is close to zero for 87% of the UEs, which connect to 100 or less sectors per day. For the remaining 13% of the UEs (traveling >100 sectors), the HOF rate slightly increases (up to 0.4% for pct-75), but the median is still close to zero; except for <0.0001% of the UEs that connect to >1k sectors and have a median HOF rate of 0.1%. Similarly, from Figure 13b, HOFs mainly occur in devices that move within a radius higher than 100km (which is the case for 0.007% of the devices, see the right y-axis), with the HOF rate reaching up to 0.4% (pct-75). Yet, the median HOF rates remain close to zero for all bins. We observe that the devices with increased mobility (>100 visited sectors, >100km radius of gyration) are mostly smart/feature-phones (90%) and M2M devices (10%) – such as modems, routers and IoT wearables – attached or carried in fast-moving vehicles, like trains. It is interesting to note that UEs with <10km radius of gyration and <50 visited sectors, which show almost zero HOF rate, include a very similar share of UE types (85% smart/feature-phones and 15% M2M); which confirms that the increase in HOFs in UEs with higher

(a) Number of sectors.

(b) Gyration
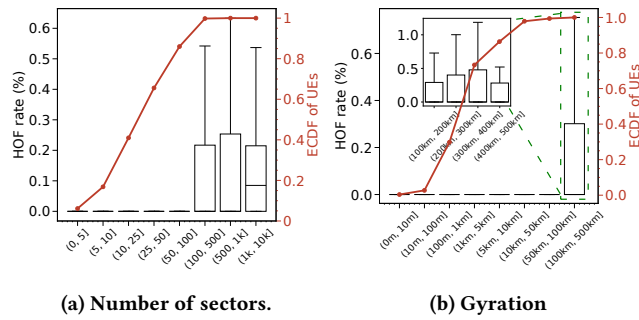
Figure 13: HOF rate (left y-axis) and ECDF for the number of UEs (right y-axis) w.r.t. binned device-level mobility metrics (log scale).

(a)

(b)

Figure 14: (a) Percentage of HOF causes w.r.t. the total HOFs; (b) Distribution of HO signaling time per cause.
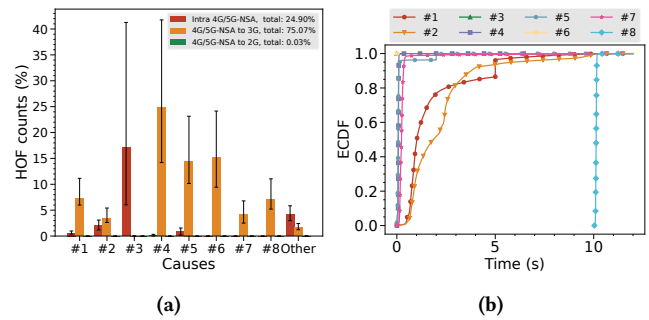
mobility metrics cannot be explained by an unequal distribution of UE types in this group.

Key takeaways: *(i) Rural areas (with sparser deployments) suffer from 32.4% more HOFs during peak hours than urban. (ii) A small number of UEs with high mobility metrics daily (> 100 visited sectors, > 100km radius of gyration) experience a non-negligible HOF rate (0.4% for pct-75); the number of visited sectors and the radius of gyration are good predictors to flag UEs that can potentially experience high HOF rates.*

6.2 Causes of HOFs

We study the HO failures using cause codes delineated by the 3GPP standards [3, 5] and the antenna vendors. In total, we collect 1k+ different causes for the failures. Our causes analysis complements prior studies that had exclusively focused on the user side, being mostly coarser, and solely for specific devices and failure types [27, 50]. In Figure 14a, we present the HOF counts in percentage, by calculating the HOF for each cause and dividing it by the total HOFs per day. We also plot alongside the minimum and maximum values observed in this period (i.e., 4 weeks). Our analysis reveals that (*i*) 92% of all HOFs occur because of 8 causes from the 1k+ that exist, and that (*ii*) 75% of all HOFs occur in transitions from 4G/5G-NSA to 3G, with the remainder (i.e., ≈25%) associated with intra 4G/5G-NSA HOs. HOFs attributable to transitions to 2G represent 0.03% of all. This distribution highlights the real-world implications of managing a layered cellular deployment, where < 6% of handovers are vertical handovers to 3G, and the remaining 95% are intra 4G/5G-NSA handovers. We present next the 8 most common handover failure causes. Additional insights for the reasons for HOFs in rural/urban areas, different smartphone manufacturers, and UE types can be seen in Figure 15.

• **Cause #1**: "The source sector canceled the HO" relates to the cancellation of an ongoing or prepared handover. HO Cancellation procedure [5] can occur for several reasons, such as timeouts on the MSC, cell site, or issues with the size of the Forward Relocation Request [3]. This cause is predominantly observed in HOs to 3G, affecting 7.3% to 11.2% of cases daily, which is significantly higher compared to intra 4G/5G-NSA and 4G/5G-NSA to 2G HOs (< 1% per day). We observe that this failure cause affects evenly all UE

types, but is 50% more prevalent in rural than in urban areas (see Figure 15).

• **Cause #2**: "The signaling procedure was aborted due to interfering S1AP Initial UE Message [5]". This error involves the interruption of the signaling process by an initial message to the MME, which includes critical user information and service requests. This issue affects 2% of intra 4G/5G-NSA HOs and 3.4% of HOs to 3G, but not HOs to 2G.

• **Cause #3**: "Signaling procedure was rejected due to invalid target sector ID" occurs when the target sector ID is not recognized or if there are configuration issues with the MME pool area (i.e., a collection of MMEs configured to serve any common part of a radio network). This is the main reason for failure in intra 4G/5G-NSA HOs, accounting for an average of 17.2% of the failures, and reaching up to 41.3%. From this cause, 59% of M2M/IoT devices fail (see Figure 15).
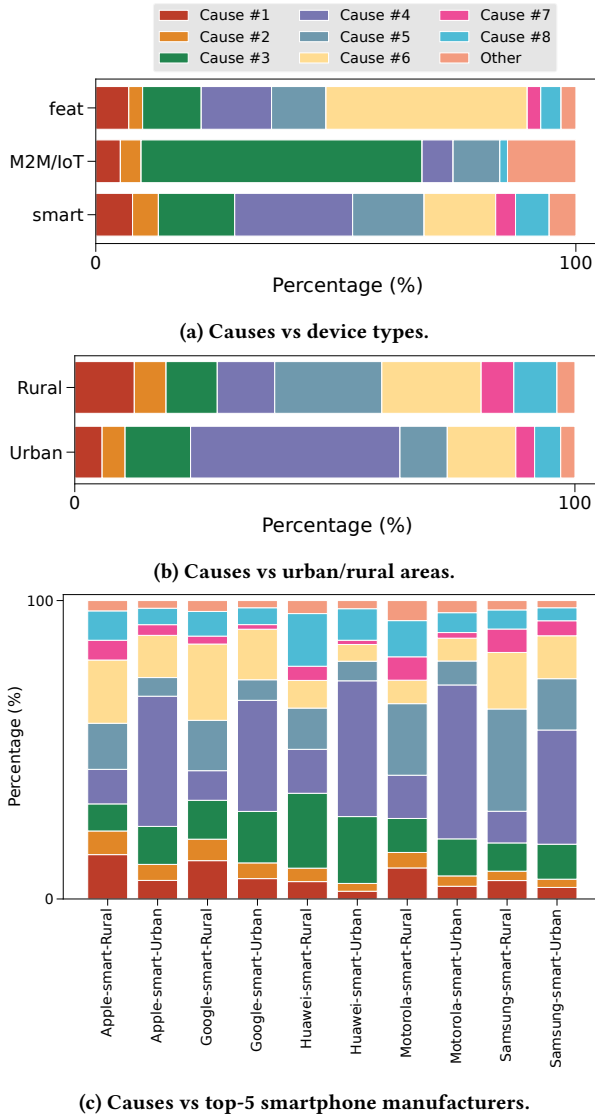
• **Cause #4**: "Load on target sector is too high" indicates that the target sector cannot accommodate the HO due to resource constraints. It is the most common reason for failure in HOs to 3G (up to 42.3% of all HOFs), affecting 25% of the failures per day, on average. It happens mainly during peak hours in dense urban areas (see Figure 7), causing 42% of the total HOFs there (see Figure 15).

• **Cause #5**: "MME detects a HO-related failure in the target MME, SGW, PGW, cell, or system"; these types of infrastructure-related outages occur for 14–23% of HOs to 3G, and for 0.8–1.6% of intra 4G/5G-NSA HOs. This cause does not pinpoint precisely the reason that the HOF occurred; however, it is important to note that this is the extent of information that is available to the MNO.

Causes #6, #7 #8 are specific to HOs from 4G/5G-NSA to 3G. We provide more information in the sequel.

• **Cause #6**: "The Single Radio Voice Call Continuity (SRVCC) service is not subscribed by the UE" affects 15.2% of HOs to 3G on average, peaking at 24.1%. SRVCC is a scheme used with VoLTE (Voice over LTE) and ensures seamless handovers of voice calls from packet-switched (PS), like 4G, to circuit-switched (CS) networks, like 2G and 3G [1, 2]. We note that this failure occurs primarily in rural areas and in feature phones, where the MNO still relies mostly on 3G to ensure the support of voice services (see Figure 15).

• **Cause #7**: Like Cause #6, Cause #7 is associated with SRVCC HOs, and it occurs when "the MSC responds with PS to CS Response with cause indicating failure" during SRVCC HO preparation; it

(a) Causes vs device types.

(b) Causes vs urban/rural areas.

(c) Causes vs top-5 smartphone manufacturers.

Figure 15: Stacked bars showing the percentages of HOF causes (each bar adds to 100%).

affects about 4.2% of all HOs [2]. We note that it affects almost no M2M/IoT device and occurs twice as often in rural than in urban areas (see Figure 15).

• **Cause #8**: "No Forward Relocation Complete or Notification was received before the max time for waiting for the relocation completion expires", affecting 7.1% of HOs. Forward Relocation Complete message is sent to the source MME/SGSN to indicate the HO has been successful. We observe this cause ×3 in M2M/IoT devices w.r.t. smartphones and feature phones (see Figure 15).

HOF Duration per Cause. Figure 14b complements the HOF analysis by delineating the HO duration associated with the 8 causes of failure. Causes #3 and #6 result in failures that prevent the initiation of the HO (i.e., signaling time equals 0ms), with the former attributed to an invalid target sector ID and the latter to SRVCC service

Table 3: Regression covariates.

Feature	Values
Number of HOs per day	≥ 0
RATs	4G/5G-NSA, 3G, 2G
District population	≥ 0
Sector Region	West, South, North, Capital area
Area Type	Rural / Urban
Antenna Vendor	4 vendors (V1, V2, V3, V4)

not being permitted for the UE. Cause #4, linked to insufficient resources in the target sector, exhibits a median duration of 81ms and a 95th percentile of 97ms. More prolonged delays in HO signaling are caused by Causes #1 and #2, where the HO is halted due to cancellation by the source or interference, respectively, leading to medians of 1–2s and 95th percentiles of 5–6s. Notably, Cause #8, associated with timeout failures, demands the most extended signaling, with a median >10s and 95% of cases occurring in < 10.2s. Our study aligns with existing works, such as [39, 50], in demonstrating the increased duration involved in HOFs.

Key takeaways: *(i) Despite the prevalence of 4G and 5G, 75% of all failures concern HOs from 4G/5G-NSA to 3G, and 25% of them happen due to high load in the target sector. (ii) 59% of M2M/IoT UEs and 42% of feature phones fail due to Cause #3 and Cause #6, respectively. (iii) 42% of HOFs in urban areas occur due to Cause #4, while Causes #5 and #6 account for 20% each, in rural areas. (iv) The duration of HOs that fail due to timeouts (Cause #8) or cancellations (Cause #1) exceeds on average 2s (for the former case it reaches 10s). These numbers highlight the noticeable outage duration caused by HOFs in the network.*

6.3 Modeling HOFs

We aim to understand which sector-level features contribute to HOFs, by isolating and combining the effect of various parameters and ensuring our findings in the previous sections are robust against potential biases or unaccounted variables. We reorganize the dataset using as dependent variable the *daily* HOF rate of each source sector, and use as covariates the sector-level features in Table 3. This creates 6.7M observations. Our hypothesis is that the HO type is the primary factor influencing HOF rates. Note that while HOs to 3G amount to only 5.86% of the total HOs (§5.2), they are responsible for 75% of all HOFs (§6.2).

A first look into the data demonstrates that HOs to 2G and 3G are associated with substantially higher failure rates with medians of 21.42% and 5.85% respectively, compared to 0.04% for HOs to 4G/5G-NSA; and this disparity persists even when we focus on the failed HOs and filter outliers, see Appendix B. We further perform an analysis of variance (ANOVA) test [16] (log-transforming the HOF rates) which verifies the significance of this effect ($p < 0.001$); and the same conclusion is reached using the Kruskal-Wallis test [25]. We repeat these tests, with the same findings, even when controlling for variations in the area and antenna vendor.

Accordingly, we use a generalized linear regression model (with log transformation) to quantify the effect of RAT on HOFs. We first run a univariate model to facilitate interpretation. We find that HOs to 3G (2G) increase the HOF rate by 166% (915%, respectively) compared to HOs to 4G/5G-NSA, see Table 4. We repeat this analysis

Table 4: Linear model coefficients for $\log(HOF\ rate)$.

Feature	Coef.	95% CI	P-value
Intra 4G/5G-NSA	-2.77	-2.77, -2.76	0
4G/5G-NSA→3G	5.12	5.12, 5.13	0
4G/5G-NSA→2G	6.82	6.76, 6.88	0

Table 5: Regression Summary: Linear Model, All Covariates.

Feature	Coeff.	Std Err	t value	Pr(> $|t|$)
(Intercept)	-3.10	0.0217	-143	0
HO type: 4G/5G-NSA→2G	5.48	0.118	46.4	0
HO type: 4G/5G-NSA→3G	4.77	0.00150	3169	0
Number of daily HOs	$-2.84 \cdot 10^{-5}$	0	-331	0
Area Type: Rural	0.260	0.00272	95.5	0
Area Type: Urban	0.190	0.00258	73.4	0
Antenna Vendor: V2	0.115	0.00173	66.7	0
Antenna Vendor: V3	0.719	0.0203	35.3	0
Antenna Vendor: V4	0.0629	0.0222	2.84	0.49
Sector Region: North	-0.0728	0.0216	-3.57	$4.05 \cdot 10^{-6}$
Sector Region: South	-0.0168	0.00166	-10.1	$2.28 \cdot 10^{-6}$
Sector Region: West	0.398	0.0204	19.5	$3.89 \cdot 10^{-66}$
District population	$-1.75 \cdot 10^{-7}$	0	-61.6	0

$N = 3857074$,   RMSE=1.023, $R^2 = 0.8265$,   AIC=11121590

while controlling for the other covariates and filtering the outliers (HOF rate<50%, number of HOs per day in $[50, 30k]$), finding the same result with slightly smaller intensity (coefficients of 5.48 and 4.77 instead of 6.82 and 5.12), as can be seen in Table 5. From the remaining covariates, the antenna vendor has a significant but smaller effect, which we also verify with an ANOVA test. These findings are also robust to alternative models (step-wise covariate selection and removing HOs to 2G), including also a quantile linear regression model. The details of these additional tests are deferred to Appendix B.

Key takeaways: *By modeling HOFs and investigating different covariates (see Table 3), we verify our hypothesis that the HO type is the main factor shaping the observed HOF rates: HOs to 3G (2G) increase the HOF rate by 166% (915%, respectively) compared to HOs to 4G/5G-NSA.*

7 Related Work

In terms of measurement approaches, the vast majority of studies rely on (mainly rooted) UEs and collect traces from their cellular modems [13, 18, 23, 26, 27, 34, 36, 37, 42, 43, 50, 51]. For instance, [13] and [26] build their mobility analysis upon Mobile Insight [28] with rooted phones, while [42, 43] use the G-NetTrack Pro monitoring tool [45]. These solutions are confined to certain chipset manufacturers and have limited data collection granularity (orders of seconds, instead of msec as in the current study). Other works study mobility patterns in one [52] or a few cities [37], such as Minneapolis [34, 35], Chicago [34, 36], Atlanta [34], and Rome [24]. These studies provide valuable information, yet their spatial focus does not facilitate insights across larger scales (e.g., countrywide) and in varied settings (e.g., urban/rural areas). The works of [39, 50, 51] conduct extensive 4G performance measurements on high-speed rails in China, and [18, 23] study mobility management policies in 4G/5G. The collected data in these relevant works are related to certain mobility patterns, and a subset of users, and do not contain network-side data.

Our study, on the other hand, records *all* mobility events from a commercial MNO network with ≈40M UEs connected, with msec granularity, during 4 weeks, and for the entire territory of a European country; it is not limited to specific routes, cities, mobility modes, or user types. To date, only a few works study HOs from the operator's perspective, as this involves technical challenges [13] and requires in-network measurements (see Figure 2). Namely, [15] suggests an approach to categorize and minimize undesired Ping-Pong (PP)[10] HOs based on a restricted dataset with 1.7k UEs; and [55] investigates PP HOs using 13 days of data from a network operator in a Mediterranean area. Our study differs from these works due to the scale, coverage, and granularity of measurements

(all active connections of a top-tier MNO at the country level; see Table 1), and due to the fusion of different datasets (about UEs and population) that allows drawing fresh insights, e.g., about the impact of HOs and HOFs on different RATs, device types, and areas (rural/urban, and district level).

Specifically, in terms of measurement results, our findings about the HO duration are on par with previous studies, e.g. [18, 23, 50], and provide additional insights, e.g., about the effect of RAT, finding that inter-RAT HOs are the most impactful. Several studies measured the volume of HOs [34, 37, 53], finding, e.g., horizontal HOs to be more frequent in 5G-SA and 4G and vertical HOs in 5G-NSA [53]. Here, we enrich these results by dissecting the HOs per RAT and UE manufacturer/type, analyzing their temporal pattern over 4 weeks, and their relation to the demographic distribution over the studied country, with district granularity (300+ districts), thus refining the typical urban/rural categorization of prior studies [55].

Furthermore, leveraging our unique network-side dataset, we characterize HOFs (cause *and* duration) using detailed antenna vendor-specific information. Prior studies, inhibited by their UE-side data, have mainly studied the effect of user speed on HOFs [26] or used coarser categorization, e.g., 2 possible causes [50], or analyzed general connectivity failures for specific devices [27]. Given that HOs were found to affect significantly the user-perceived network performance, our work can inform the design of new HO policies, such as [12, 30, 53], and guide the optimization of network deployment and RAT upgrades.

8 Discussion

Limitations. Datasets and actual, unnormalized, numbers in the figures cannot be published openly due to privacy guidelines of the MNO (see Appendix A). Moreover, at the time we conducted our study, the 5G-SA deployment of the MNO was still in its early stages, with a limited range of (mostly test) UEs actively using it. Thus, we focus on 5G-NSA, which relies on the 4G EPC for mobility management. In other words, we cannot explicitly capture the HOs to/from 5G radio sectors, since the EPC only sees their corresponding 4G radio sector anchor. In addition, the studied HOs have 4G/5G-NSA as the source RAT, and 4G/5G-NSA, 3G, or 2G as the target. In other words, apart from the horizontal HOs in

---

[10]PP HO occurs when a UE is handovered from a source to a target sector, and then back to the source, under a short, predefined time.

4G/5G-NSA, we focus on the specifics of how/when/why users downgrade to older RATs, and not the other way around, given that users spend more than 82% of their time and 94.5% of their traffic in 4G/5G (see §4).

Lastly, (*i*) for this study we did not have access to HO configuration parameters and policies, which are dynamically configured by proprietary solutions from equipment vendors (see §2), and (*ii*) our analysis on *HOFs & Mobility* is limited to the use of mobility metrics (number of sectors and radius of gyrations) at daily intervals, which may hide correlations that occur at finer time scales. While this paper represents a first attempt to provide an overview on a countrywide scale, we stress the importance for the community to conduct further studies on the previous aspects in order to contrast the coarse-grained correlations found in this study with specific analyses focused on establishing causal relations between various metrics, e.g., by analyzing performance degradation for specific users at the session-level just before and after a HOF occurs.

Guidelines, Implications & Future Works. The progression towards RATs beyond 4G is pivotal for realizing greater network speeds, minimizing latency, and improving performance. However, the findings of this study underline the significant challenges that older RATs present in HO performance, which directly impacts the users' Quality of Experience. Although some of these technologies are close to their sunset, they continue to play a crucial role in many operational networks (as the studied one) and hence, require attention to maintain users' satisfaction. It is crucial for network operators to monitor and report activity in the legacy RATs, so as to design realistic strategies towards fully decommissioning them. The gradual phasing out of older RATs should be carefully managed to avoid any unintended negative consequences on network reliability and user experience. This may involve transitions of certain regions or user segments to newer RATs before others, based on usage patterns and network demands.

The implications of these findings extend to HO policies, which need to be revisited in light of the persistent issues associated with older RATs (see §6). Network operators should consider adopting more dynamic and adaptive HO algorithms (e.g., in response to failures and mobility patterns) that can handle the specific challenges posed by these legacy technologies. These solutions should be tailored to the various causes of HOFs associated with HOs to 3G/2G and intra 4G/5G-NSA HOs (see §6.2), and therefore, to the different time granularity in which these HOs occur (i.e., hundreds/thousands of msecs and tens of msecs, respectively, see §5.2).

Further investigation into the role of device manufacturers and operating systems is also essential. Optimizing network performance is not solely the responsibility of network operators; additional *coordinated* studies that examine the internal mechanisms of devices and their influence on HOs are required. This entails a deeper understanding of how manufacturers and operating systems interact with network procedures and the identification of HO improvement opportunities (e.g., where HO durations are prolonged).

Future work could also explore the impact of HOFs on performance metrics, such as throughput, voice/data accessibility, and success rates [17], from the operator's perspective. This would enable us to better understand the relation between HO performance and users' Quality of Service (QoS).

Additionally, large-scale analyses like the one presented here often face challenges in handling, storing, and processing vast datasets, underscoring the need for further research into efficient data sampling techniques.

Handover Challenges in 5G and beyond. While 5G capabilities are expansive, the co-existence of multiple RATs presents a significant challenge, particularly in the context of HO management [23, 48]. The integration of 5G with legacy systems like 4G, 3G, or 2G, introduces significant complexities in the HO process, due to the amplified differences in terms of latency, bandwidth, and signaling requirements in 5G, the new mobility features introduced (e.g., dual connectivity [49]), and the wide range of device types and new services to be supported (e.g., IoT verticals, time-critical communications). For instance, in EN-DC (EUTRA-NR Dual Connectivity) used in 5G-NSA [44], two simultaneous connections are established (a 4G master node, and a 5G secondary node) for data plane messages, but only 4G is used in the control plane. This mechanism makes remarkably more complex the HO procedure, as additional signaling messages need to be exchanged due to the presence of the secondary node, resulting in increased time complexity that could be amplified, e.g., in case of PP HOs. To mitigate these challenges, it is imperative to implement differentiated HO policies tailored to the wide range of device types and services supported by 5G; e.g., IoT and time-critical communications require distinct HO strategies to meet specific service level agreements.

9 Conclusion

This work provides the first comprehensive, countrywide, analysis of HOs, leveraging data from a leading MNO in a European country, by studying ≈40M users over four weeks. Our findings highlight the critical impact of spatio-temporal factors, RATs, device types, and manufacturers on horizontal and vertical HOs and HOFs, specifying the reasons for the latter, and modeling them with statistical methods. These findings are crucial for understanding and developing new HO mechanisms and policies, and identifying groups of UEs and areas that require enhanced support. In this way, our analysis lays the groundwork for future improvements in network performance, ensuring that the promise of 5G and subsequent generations of cellular technologies can be fully realized.

Acknowledgments

References

[1] 3GPP. 2022. *Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; 5G; Organization of subscriber data*. Technical Specification (TS) 23.008. 3rd Generation Partnership Project (3GPP).

[2] 3GPP. 2022. *Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Single Radio Voice Call Continuity (SRVCC); Stage 2*. Technical Specification (TS) 23.216. 3rd Generation Partnership Project (3GPP).

[3] 3GPP. 2023. *3GPP Evolved Packet System (EPS); Evolved General Packet Radio Service (GPRS) Tunnelling Protocol for Control plane (GTPv2-C); Stage 3*. Technical Specification (TS) 29.274. 3rd Generation Partnership Project (3GPP).

[4] 3GPP. 2023. *Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification*. Technical Specification (TS) 36.331. 3rd Generation Partnership Project (3GPP).

[5] 3GPP. 2023. *Evolved Universal Terrestrial Radio Access Network (E-UTRAN); S1 Application Protocol (S1AP)*. Technical Specification (TS) 36.413. 3rd Generation Partnership Project (3GPP).

[6] 3GPP. 2023. *General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access*. Technical Specification (TS) 23.401. 3rd Generation Partnership Project (3GPP).

[7] 3GPP. 2023. *NR; Radio Resource Control (RRC); Protocol specification*. Technical Specification (TS) 38.331. 3rd Generation Partnership Project (3GPP).

[8] Rein Ahas, Siiri Silm, Olle Järv, Erki Saluveer, and Margus Tiru. 2010. Using Mobile Positioning Data to Model Locations Meaningful to Users of Mobile Phones. *Journal of Urban Technology* 17, 1 (2010), 3–27. https://doi.org/10.1080/10630731003597306

[9] Tugce Bilen, Berk Canberk, and Kaushik R. Chowdhury. 2017. Handover Management in Software-Defined Ultra-Dense 5G Networks. *IEEE Network* 31, 4 (2017), 49–55. https://doi.org/10.1109/MNET.2017.1600301

[10] Leo Breiman. 2001. Random Forests. *Machine Learning* 45, 1 (01 Oct 2001), 5–32. https://doi.org/10.1023/A:1010933404324

[11] Antonio De Domenico, David López-Pérez, Wenjie Li, Nicola Piovesan, Harvey Bao, and Xinli Geng. 2023. Modeling User Transfer During Dynamic Carrier Shutdown in Green 5G Networks. *IEEE Transactions on Wireless Communications* 22, 8 (2023), 5536–5549.

[12] Haotian Deng, Qianru Li, Jingqi Huang, and Chunyi Peng. 2020. ICellSpeed: Increasing Cellular Data Speed with Device-Assisted Cell Selection. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking* (London, United Kingdom) *(MobiCom '20)*. Association for Computing Machinery, New York, NY, USA, Article 44, 13 pages. https://doi.org/10.1145/3372224.3419201

[13] Haotian Deng, Chunyi Peng, Ans Fida, Jiayi Meng, and Y. Charlie Hu. 2018. Mobility Support in Cellular Networks: A Measurement Study on Its Configurations and Implications. In *Proceedings of the Internet Measurement Conference 2018* (Boston, MA, USA) *(IMC '18)*. Association for Computing Machinery, New York, NY, USA, 147–160. https://doi.org/10.1145/3278532.3278546

[14] Ericsson. 2023. *Voice and SMS transformation following 2G/3G Sunset*. White Paper. Ericsson.

[15] Zoltan Feher, Andras Veres, and Zalan Heszberger. 2012. Ping-Pong Reduction Using Sub Cell Movement Detection. In *2012 IEEE 75th Vehicular Technology Conference (VTC Spring)*. IEEE, Yokohama, Japan, 1–5. https://doi.org/10.1109/VETECS.2012.6239926

[16] Ronald A. Fisher. 1970. *Statistical Methods for Research Workers* (14th ed.). Oliver and Boyd, Edinburgh.

[17] Changhan Ge, Zihui Ge, Xuan Liu, Ajay Mahimkar, Yusef Shaqalle, Yu Xiang, and Shomik Pathak. 2023. Chroma: Learning and Using Network Contexts to Reinforce Performance Improving Configurations. In *Proceedings of the 29th Annual International Conference on Mobile Computing and Networking (MobiCom '23)*. Association for Computing Machinery, New York, NY, USA, Article 42, 16 pages. https://doi.org/10.1145/3570361.36132

[18] Moinak Ghoshal, Imran Khan, Z. Jonny Kong, Phuc Dinh, Jiayi Meng, Y. Charlie Hu, and Dimitrios Koutsonikolas. 2023. Performance of Cellular Networks on the Wheels. In *Proceedings of the 2023 ACM on Internet Measurement Conference* (Montreal, Canada) *(IMC '23)*. Association for Computing Machinery, New York, NY, USA, 678–695. https://doi.org/10.1145/3618257.3624814

[19] Lucas Chavarria Gimenez, Maria Carmela Cascino, Maria Stefan, Klaus I. Pedersen, and Andrea F. Cattoni. 2016. Mobility Performance in Slow- and High-Speed LTE Real Scenarios. In *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*. IEEE, Nanjing, China, 1–5. https://doi.org/10.1109/VTCSpring.2016.7504347

[20] Lucas Chavarria Gimenez, Per Henrik Michaelsen, Klaus I. Pedersen, Troels E. Kolding, and Huan Cong Nguyen. 2017. Towards Zero Data Interruption Time with Enhanced Synchronous Handover. In *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*. IEEE, Sydney, NSW, Australia, 1–6. https://doi.org/10.1109/VTCSpring.2017.8108504

[21] Marta C. González, César A. Hidalgo, and Albert-László Barabási. 2008. Understanding individual human mobility patterns. *Nature* 453, 7196 (01 Jun 2008), 779–782. https://doi.org/10.1038/nature06958

[22] Donghyuk Han, Sungjin Shin, Hyoungjun Cho, Jong-moon Chung, Dongseok Ok, and Iksoon Hwang. 2015. Measurement and stochastic modeling of handover delay and interruption time of smartphone real-time applications on LTE networks. *IEEE Communications Magazine* 53, 3 (2015), 173–181. https://doi.org/10.1109/MCOM.2015.7060501

[23] Ahmad Hassan, Arvind Narayanan, Anlan Zhang, Wei Ye, Ruiyang Zhu, Shuowei Jin, Jason Carpenter, Z. Morley Mao, Feng Qian, and Zhi-Li Zhang. 2022. Vivisecting Mobility Management in 5G Cellular Networks. In *Proceedings of the ACM SIGCOMM 2022 Conference* (Amsterdam, Netherlands) *(SIGCOMM '22)*. Association for Computing Machinery, New York, NY, USA, 86–100. https://doi.org/10.1145/3544216.3544217

[24] Konstantinos Kousias, Mohammad Rajiullah, Giuseppe Caso, Ozgu Alay, Anna Brunstrom, Luca De Nardis, Marco Neri, Usman Ali, and Maria-Gabriella Di Benedetto. 2022. Implications of handover events in commercial 5G nonstandalone deployments in Rome. In *Proceedings of the ACM SIGCOMM Workshop on 5G and Beyond Network Measurements, Modeling, and Use Cases* (Amsterdam, Netherlands) *(5G-MeMU '22)*. Association for Computing Machinery, New York, NY, USA, 22–27. https://doi.org/10.1145/3538394.3546041

[25] William H. Kruskal and W. Allen Wallis. 1952. Use of Ranks in One-Criterion Variance Analysis. *J. Amer. Statist. Assoc.* 47, 260 (1952), 583–621. https://doi.org/10.2307/2280779

[26] Yuanjie Li, Qianru Li, Zhehui Zhang, Ghufran Baig, Lili Qiu, and Songwu Lu. 2020. Beyond 5G: Reliable Extreme Mobility Management. 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* (Virtual Event, USA) *(SIGCOMM '20)*. Association for Computing Machinery, New York, NY, USA, 344–358. https://doi.org/10.1145/3387514.3405873

[27] Yang Li, Hao Lin, Zhenhua Li, Yunhao Liu, Feng Qian, Liangyi Gong, Xianlong Xin, and Tianyin Xu. 2021. A nationwide study on cellular reliability: measurement, analysis, and enhancements. In *Proceedings of the 2021 ACM SIGCOMM 2021 Conference* (Virtual Event, USA) *(SIGCOMM '21)*. Association for Computing Machinery, New York, NY, USA, 597–609. https://doi.org/10.1145/3452296.3472908

[28] Yuanjie Li, Chunyi Peng, Zengwen Yuan, Jiayao Li, Haotian Deng, and Tao Wang. 2016. Mobileinsight: extracting and analyzing cellular network information on smartphones. In *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking* (New York City, New York) *(MobiCom '16)*. Association for Computing Machinery, New York, NY, USA, 202–215. https://doi.org/10.1145/2973750.2973751

[29] Guangyi Liu, Yuhong Huang, Zhuo Chen, Liang Liu, Qixing Wang, and Na Li. 2020. 5G Deployment: Standalone vs. Non-Standalone from the Operator Perspective. *IEEE Communications Magazine* 58, 11 (2020), 83–89. https://doi.org/10.1109/MCOM.001.2000230

[30] Zhihong Luo, Silvery Fu, Mark Theis, Shaddi Hasan, Sylvia Ratnasamy, and Scott Shenker. 2021. Democratizing Cellular Access with CellBricks. In *Proceedings of the 2021 ACM SIGCOMM 2021 Conference* (Virtual Event, USA) *(SIGCOMM '21)*. Association for Computing Machinery, New York, NY, USA, 626–640. https://doi.org/10.1145/3452296.3473336

[31] Andra Lutu, Byungjin Jun, Alessandro Finamore, Fabián E. Bustamante, and Diego Perino. 2020. Where Things Roam: Uncovering Cellular IoT/M2M Connectivity. In *Proceedings of the ACM Internet Measurement Conference* (Virtual Event, USA) *(IMC '20)*. Association for Computing Machinery, New York, NY, USA, 147–161. https://doi.org/10.1145/3419394.3423661

[32] Ajay Mahimkar, Ashiwan Sivakumar, Zihui Ge, Shomik Pathak, and Karunasish Biswas. 2021. Auric: using data-driven recommendation to automatically generate cellular configuration. In *Proceedings of the 2021 ACM SIGCOMM 2021 Conference* (Virtual Event, USA) *(SIGCOMM '21)*. Association for Computing Machinery, New York, NY, USA, 807–820. https://doi.org/10.1145/3452296.3472906

[33] Michael S Mollel, Attai Ibrahim Abubakar, Metin Ozturk, Shubi Felix Kaijage, Michael Kisangiri, Sajjad Hussain, Muhammad Ali Imran, and Qammer H Abbasi. 2021. A Survey of Machine Learning Applications to Handover Management in 5G and Beyond. *IEEE Access* 9 (2021), 45770–45802.

[34] Arvind Narayanan, Eman Ramadan, Jason Carpenter, Qingxu Liu, Yu Liu, Feng Qian, and Zhi-Li Zhang. 2020. A First Look at Commercial 5G Performance on Smartphones. In *Proceedings of The Web Conference 2020* (Taipei, Taiwan) *(WWW '20)*. Association for Computing Machinery, New York, NY, USA, 894–905. https://doi.org/10.1145/3366423.3380169

[35] Arvind Narayanan, Eman Ramadan, Rishabh Mehta, Xinyue Hu, Qingxu Liu, Rostand A. K. Fezeu, Udhaya Kumar Dayalan, Saurabh Verma, Peiqi Ji, Tao Li, Feng Qian, and Zhi-Li Zhang. 2020. Lumos5G: Mapping and Predicting Commercial mmWave 5G Throughput. In *Proceedings of the ACM Internet Measurement Conference* (Virtual Event, USA) *(IMC '20)*. Association for Computing Machinery, New York, NY, USA, 176–193. https://doi.org/10.1145/3419394.3423629

[36] Arvind Narayanan, Muhammad Iqbal Rochman, Ahmad Hassan, Bariq S. Firmansyah, Vanlin Sathya, Monisha Ghosh, Feng Qian, and Zhi-Li Zhang. 2022. A Comparative Measurement Study of Commercial 5G mmWave Deployments. In *IEEE INFOCOM 2022 - IEEE Conference on Computer Communications*. IEEE, Virtual Conference, 800–809. https://doi.org/10.1109/INFOCOM48880.2022.9796693

[37] Arvind Narayanan, Xumiao Zhang, Ruiyang Zhu, Ahmad Hassan, Shuowei Jin, Xiao Zhu, Xiaoxuan Zhang, Denis Rybkin, Zhengxuan Yang, Zhuoqing Morley Mao, Feng Qian, and Zhi-Li Zhang. 2021. A Variegated Look at 5G in the Wild: Performance, Power, and QoE Implications. In *Proceedings of the 2021 ACM SIGCOMM 2021 Conference* (Virtual Event, USA) *(SIGCOMM '21)*. Association for Computing Machinery, New York, NY, USA, 610–625. https://doi.org/10.1145/3452296.3472923

[38] NGMN 5G Initiative. 2020. *5G White Paper, Version 2.* White Paper. NGMN.

[39] Yunzhe Ni, Feng Qian, Taide Liu, Yihua Cheng, Zhiyao Ma, Jing Wang, Zhongfeng Wang, Gang Huang, Xuanzhe Liu, and Chenren Xu. 2023. POLYCORN: Data-driven Cross-layer Multipath Networking for High-speed Railway through Composable Schedulerlets. In *20th USENIX Symposium on Networked Systems Design and Implementation (NSDI 23)*. USENIX Association, Boston, MA, 1325–1340. https://www.usenix.org/conference/nsdi23/presentation/ni

[40] Santi Phithakkitnukoon, Zbigniew Smoreda, and Patrick Olivier. 2012. Socio-Geography of Human Mobility: A Study Using Longitudinal Mobile Phone Data. *PLoS One* 7, 6 (June 2012), e39253.

[41] Michele Polese, Mischa Dohler, Falko Dressler, Melike Erol-Kantarci, Rittwik Jana, Raymond Knopp, and Tommaso Melodia. 2024. Empowering the 6G Cellular Architecture With Open RAN. *IEEE Journal on Selected Areas in Communications* 42, 2 (2024), 245–262. https://doi.org/10.1109/JSAC.2023.3334610

[42] Darijo Raca, Dylan Leahy, Cormac J. Sreenan, and Jason J. Quinlan. 2020. Beyond throughput, the next generation: a 5G dataset with channel and context metrics. In *Proceedings of the 11th ACM Multimedia Systems Conference* (Istanbul, Turkey) *(MMSys '20)*. Association for Computing Machinery, New York, NY, USA, 303–308. https://doi.org/10.1145/3339825.3394938

[43] Darijo Raca, Jason J. Quinlan, Ahmed H. Zahran, and Cormac J. Sreenan. 2018. Beyond Throughput: A 4G LTE Dataset with Channel and Context Metrics. In *Proceedings of the 9th ACM Multimedia Systems Conference* (Amsterdam, Netherlands) *(MMSys '18)*. Association for Computing Machinery, New York, NY, USA, 460–465. https://doi.org/10.1145/3204949.3208123

[44] 3GPP. 2024. *Universal Mobile Telecommunications System (UMTS); LTE; 5G; NR; Multi-connectivity; Overall description; Stage-2.* Technical Specification (TS) 37.340. 3rd Generation Partnership Project (3GPP).

[45] Gyokov Solutions. 2024. G-NetTrack Pro. https://gyokovsolutions.com/g-nettrack/. Accessed: 2024-05-15.

[46] Gordon L Stüber and Gordon L Steuber. 2017. *Principles of mobile communication* (4th ed.). Springer, Cham, Switzerland.

[47] Rumeng Tan, Ying Shi, Yingying Fan, Wentao Zhu, and Tong Wu. 2022. Energy saving technologies and best practices for 5G radio access network. *IEEE Access* 10 (2022), 51747–51756.

[48] Muhammad Tayyab, Xavier Gelabert, and Riku Jäntti. 2019. A Survey on Handover Management: From LTE to NR. *IEEE Access* 7 (2019), 118907–118930. https://doi.org/10.1109/ACCESS.2019.2937405

[49] Ingo Viering, Henrik Martikainen, Andreas Lobinger, and Bernhard Wegmann. 2018. Zero-Zero Mobility: Intra-Frequency Handovers with Zero Interruption and Zero Failures. *IEEE Network* 32, 2 (2018), 48–54. https://doi.org/10.1109/MNET.2018.1700223

[50] Jing Wang, Yufan Zheng, Yunzhe Ni, Chenren Xu, Feng Qian, Wangyang Li, Wantong Jiang, Yihua Cheng, Zhuo Cheng, Yuanjie Li, Xiufeng Xie, Yi Sun, and Zhongfeng Wang. 2019. An Active-Passive Measurement Study of TCP Performance over LTE on High-Speed Rails. In *The 25th Annual International Conference on Mobile Computing and Networking (MobiCom '19)*. Association for Computing Machinery, New York, NY, USA, Article 18, 16 pages. https://doi.org/10.1145/3300061.3300123

[51] Qingyang Xiao, Ke Xu, Dan Wang, Li Li, and Yifeng Zhong. 2014. TCP Performance over Mobile Networks in High-Speed Mobility Scenarios. In *Proceedings of the 2014 IEEE 22nd International Conference on Network Protocols (ICNP '14)*. IEEE Computer Society, USA, 281–286. https://doi.org/10.1109/ICNP.2014.49

[52] Dongzhu Xu, Anfu Zhou, Xinyu Zhang, Guixian Wang, Xi Liu, Congkai An, Yiming Shi, Liang Liu, and Huadong Ma. 2020. Understanding Operational 5G: A First Measurement Study on Its Coverage, Performance and Energy Consumption. In *Proceedings of the Annual Conference of the ACM Special Interest Group on Data Communication on the Applications, Technologies, Architectures, and Protocols for Computer Communication* (Virtual Event) *(SIGCOMM '20)*. Association for Computing Machinery, New York, NY, USA, 479–494. https://doi.org/10.1145/3387514.3405882

[53] Xinjie Yuan, Mingzhou Wu, Zhi Wang, Yifei Zhu, Ming Ma, Junjian Guo, Zhi-Li Zhang, and Wenwu Zhu. 2022. Understanding 5G performance for real-world services: a content provider's perspective. In *Proceedings of the ACM SIGCOMM 2022* (Amsterdam, Netherlands) *(SIGCOMM '22)*. Association for Computing Machinery, New York, NY, USA, 101–113. https://doi.org/10.1145/3544216.3544219

[54] Hongtao Zhang and Lingcheng Dai. 2018. Mobility Prediction: A Survey on State-of-the-Art Schemes and Future Applications. *IEEE access* 7 (2018), 802–822.

[55] Dinko Zidic, Toni Mastelic, Ivana Nizetic Kosovic, Mario Cagalj, and Josip Lorincz. 2023. Analyses of Ping-Pong Handovers in Real 4G Telecommunication Networks. *Comput. Netw.* 227, C (5 2023), 12 pages. https://doi.org/10.1016/j.comnet.2023.109699

A Ethics

The datasets we leverage in our research are protected under Non-Disclosure Agreements (NDAs) that explicitly forbid the dissemination of information to unauthorized parties and public repositories. The procedures for data collection and storage within the network's infrastructure strictly follow the guidelines set forth by the MNO, and are in full compliance with local regulations. Moreover, while some metrics are computed on the user-level, our data-handling processes strictly focus on generating aggregated, anonymized insights, without access to the exact locations/trajectories of the users. No personal and/or contract information was available for this study and none of the authors of this paper participated in the extraction and/or encryption of the raw data. Ultimately, our datasets and research do not involve risks for the mobile subscribers, while they provide new knowledge about the dynamics of mobility management and handovers.

B Regression Analysis Details

Here, we complement the main regression models presented in §6.3 with additional models, which have comparable performance in terms of Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) with Random Forest (RF) [10]. The results are aligned and support the reported findings. We remind the reader that the analysis is performed on a dataset that records the daily percentage of failed HOs (i.e., HOF rate) during the studied 4-week interval.

We start by plotting the main statistics (boxplots with mean and median values) for the effect of HO type, antenna vendor, and sector area on the HOF rates. We also plot the ECDFs for the first two cases in Figure 16, while the summary statistics can be seen in Table 6. Performing a one-way ANOVA test we find that the effect of HO type on HOF rate is statistically significant and large $(F(2, 3857071) = 8.01 \cdot 10^6, p < .001; \eta^2 = 0.81, 95\%CI\ [0.82, 1.00])$, and Post-hoc pairwise comparisons (Tukey's HSD) verify that this effect is significant for all HO types. A Kruskal-Wallis test also supports this hypothesis ($p = 0$).

Next, we turn our attention to the vendor of the source sector (i.e., antenna vendor). Due to confidentiality issues, we refer to the 4 vendors with the codes V1, V2, V3, and V4, instead of using their actual names. First, we note that different vendors are used in sectors in different regions (North, South, West, Capital area), Figure 17 (top); while all but one vendors are involved in similar proportions in intra 4G/5G-NSA HOs and HOs to 3G, Figure 17 (bottom). In Figure 18 (top), we present the boxplots for the effect of the antenna vendor on HOF rates. In this case, we create one plot for each type of RAT and focus on HOF rates < 1% for 4G/5G-NSA, since the values are concentrated in the low-end of the spectrum. ANOVA tests for each HO type and for all HO types concurrently verify this effect is statistically significant but very small $((F(3, 4911927) = 30524.85, p < .001; \eta^2 = 0.02, 95\%CI[0.02, 1.00]))$. Finally, Figure 18 (bottom) studies the effect of the area type, where this feature takes two values: rural and urban. We observe a small effect of the area type and indeed performing an ANOVA test, we find it statistically significant but small $(F(2, 4664505) = 18559.77, p < .001, \eta^2 = 7.90 \cdot 10^{-3}, 95\%CI[7.76 \cdot 10^{-3}, 1.00])$, even when we subset per HO type.

Table 6: Summary Stats of Dataset.

| Feature	Min	1st Qu	Median	Mean	3rd Qu	Max
Daily HOs	1	76	1989	6431	8591	953287
HOF rate	0.0	0.0	0.069	6.131	4.191	100.0
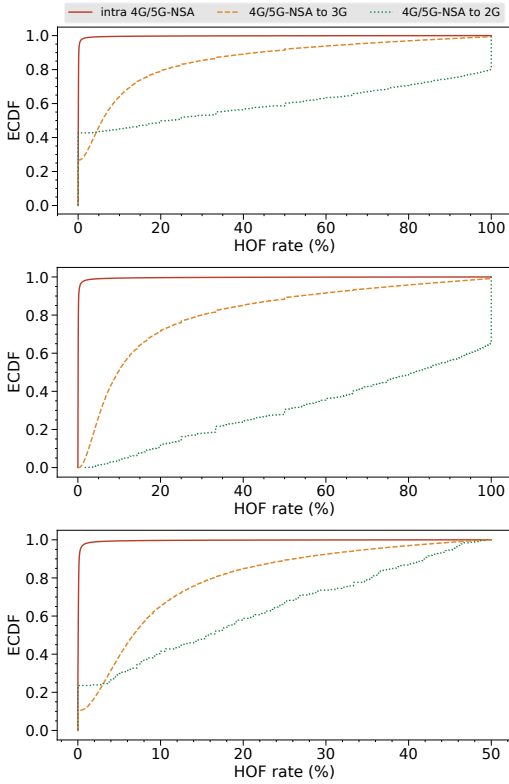
Figure 16: ECDF of HOF rates for HO type: all HOF rates (top); non-zero HOF rates (middle), HOF rates without outliers (bottom).

**Table 7: Regression Summary: Linear Model w/o 2G HOs.**

Feature	Coeff.	Std Err	t value	Pr(> |t|)
(Intercept)	-3.64	0.0185	-196	0
HO type: 4G/5G-NSA→3G	5.23	0.00120	4348	0
Number of daily HOs	$-1.02 \cdot 10^{-5}$	0	-215	0
Area Type: Rural	0.416	0.00273	153	0
Area Type: Urban	0.365	0.00259	141	0
Antenna Vendor: V2	0.0241	0.00166	14.5	0
Antenna Vendor: V3	1.00	0.0183	54.6	0
Antenna Vendor: V4	0.227	0.0199	11.4	0
Sector Region: North	-0.107	0.0184	-5.81	$6.14 \cdot 10^{-9}$
Sector Region: South	-0.0527	0.00160	-32.9	0
Sector Region: West	0.577	0.0184	31.5	0
District population	$-1.52 \cdot 10^{-7}$	0	-54.7	0

$N = 4892154$,   RMSE=1.072901, $R^2 = 0.8502$,   AIC=14571839

After this first level of analysis, we proceed with regression models that complement those presented in §6.3. Table 5 reports the
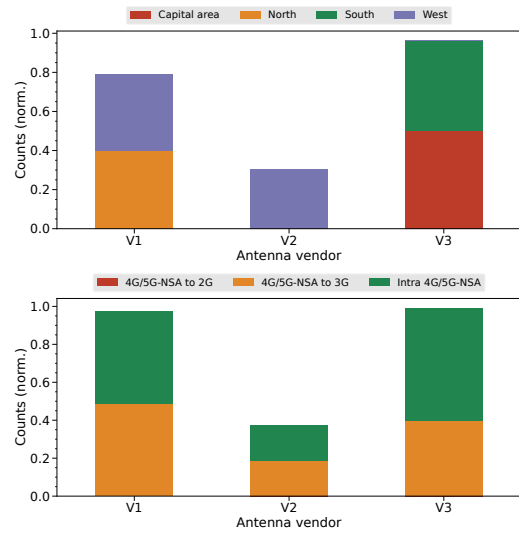
Figure 17: Antenna vendor per region (top); per HO type (bottom).
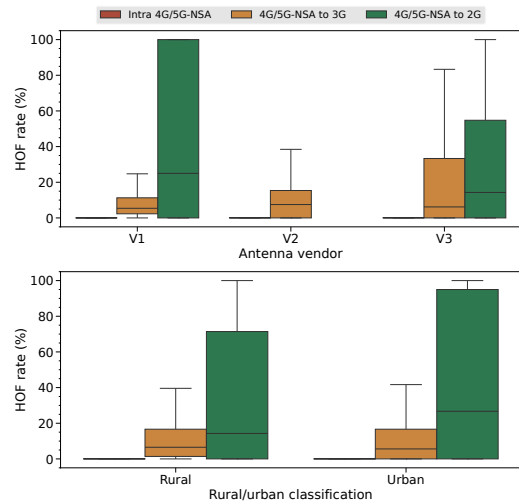
Figure 18: Boxplots of HOF rates vs antenna vendor (top); vs urban/rural areas (bottom).

results for a linear regression model, after log-transforming the dependent variable and excluding outliers (i.e., removing entries with HOF rates exceeding 50%, less than 10 HOs per day or more than 30k HOs per day) that includes all main features of the dataset. In line with the simpler univariate model in §6.3, we see that the HO type remains the main contributing factor on HOF, even when accounting for all other covariates. On the other hand, the rest of the features are significant, yet have a much smaller, often negligible, effect. To further delineate the effect of the other covariates, we repeat the analysis after excluding HOs to 2G since they represent only 0.04% of dataset entries and are skewed towards much higher HOFs (see boxplots). The results are summarized in Table 7 where

Table 8: Quantile Regression w/o Outliers.

Feature; Quantile	Coeff.	Std Err	t value	Pr(> $|t|$)
(Intercept); $\tau = 0.2$	-3.59	0.00072	-5000.50	0
HO type: 4G/5G-NSA→2G	5.80	0.07401	78.37	0
HO type: 4G/5G-NSA→3G	4.86	0.00113	4297.03	0
(Intercept); $\tau = 0.4$	-2.99	0.00077	-3865.27	0
HO type: 4G/5G-NSA→2G	5.880	0.07951	73.95	0
HO type: 4G/5G-NSA→3G	4.79	0.00122	3935.15	0
(Intercept); $\tau = 0.6$	-2.56	0.00066	-3874.20	0
HO type: 4G/5G-NSA→2G	5.84	0.06822	85.74	0
HO type: 4G/5G-NSA→3G	4.83	0.00104	4632.57	0
(Intercept) $\tau = 0.8$	-2.09	0.00092	-2281.89	0
HO type: 4G/5G-NSA→2G	5.72	0.09450	60.57	0
HO type: 4G/5G-NSA→3G	4.97	0.00145	3437.48	0

Table 9: Quantile Regression – All HOFs.

	$\tau = 0.2$	$\tau = 0.4$	$\tau = 0.6$	$\tau = 0.8$
(Intercept)	-3.62	-3.00	-2.58	-2.11
HO type: 4G/5G-NSA→2G	7.13	7.20	7.13	6.72
HO type: 4G/5G-NSA→3G	5.03	4.99	5.15	5.51

we see that the HO type (only related to 3G in this case) is pronounced, the rural/urban feature is significant but the two values have a similar effect, as well as a significant and large effect of the vendor and the region (West). We note these latter findings (effect of Vendor V2 and West) remain significant even if we exclude the HOs to 2G and 3G, and regress only over the intra 4G/5G-NSA HOs.

As a final robustness test and based on the (near) bimodal distribution of the log-transformed HOF rate variable, we perform quantile regression on 5 intervals ($\tau \in \{0.2, 0.4, 0.6, 0.8\}$), using the HO type as the only feature. Table 8 summarizes the results for the case we filter outliers as before, and Table 9 presents the coefficients for the entire dataset of non-zero HOF rates. These results reinforce the findings of the previous models, verifying the significant and large effect of the HO type on HOFs across the entire spectrum of observed values.

C List of Abbreviations

To facilitate the reading and understanding of this work, abbreviations and their full terms are provided in Table 10.

Table 10: List of Abbreviations (alphabetically).

Abbreviation	Full Term
5G-NR	5G New Radio
5G-NSA	5G-Non-Standalone
5G-SA	5G-Standalone
APN	Access Point Name
CN	Core Network
CS	Circuit Switched
DL	Downlink
ECDF	Empirical Cumulative Distribution Function
EPC	Evolved Packet Core
GSM	Global System for Mobile Communications
GSMA	Global System for Mobile Communications Association
HOF	Handover Failure
HO	Handovers
IMEI	International Mobile Equipment Identity
IMSI	International Mobile Subscriber Identity
IoT	Internet-of-Things
M2M	Machine-to-Machine
MME	Mobile Management Entity
MNO	Mobile Network Operator
MR	Measurement Report
MSC	Mobile Switching Center
PS	Packet Switched
QoE	Quality of Experience
QoS	Quality of Service
RACH	Random Access Channel
RAN	Radio Access Network
RAT	Radio Access Technology
RSRQ	Reference Signal Received Quality
RRC	Radio Resource Control
SGSN	Serving GPRS Support Node
SGW	Serving Gateway
TAC	Type Allocation Code
TAU	Tracking Area Update
UE	User Equipment
UL	Uplink
URLLC	Ultra Reliable Low Latency Communications