

A First Look at Operational RAN Updates and Their Impact on Carrier Traffic Demands and Prediction

Antonio Boiano*, Nadezda Chukhno[†], Zbigniew Smoreda[‡], Alessandro Enrico Cesare Redondi*, Marco Fiore[†]

*Politecnico di Milano, Italy, [†]IMDEA Networks Institute, Spain, [‡]Orange Innovation, France

{nadezda.chukhno, marco.fiore}@networks.imdea.org,

{antonio.boiano, alessandroenrico.redondi}@polimi.it, zbigniew.smoreda@orange.com

Abstract—Radio Access Networks (RANs) are critical infrastructures that mobile operators continuously upgrade to accommodate increasing data traffic demands, stricter performance requirements, and evolutions in radio technologies. RAN updates can affect carrier-level Key Performance Indicators (KPIs) that are the foundational input to data-driven models for network management. However, to date, no study has systematically examined the dynamics of RAN deployments, and little is known about the actual prevalence of RAN updates or their impact on Machine Learning (ML) models for network automation. This paper presents a first characterization of RAN updates in a nationwide operational infrastructure composed of over 500,000 carriers. A network-side vantage point lets us (i) investigate the type and frequency of RAN modifications, (ii) assess the impact of such changes on a primary KPI for network management, i.e., the traffic volume served by individual carriers, and (iii) verify the final effects on a classical downstream ML application, i.e., traffic prediction. Our results reveal that RAN updates take place with notable frequency, e.g., occurring every few days even in medium-sized cities. Also, they affect in a significant way the demands at a considerable fraction of pre-existing carriers, where they can curb the accuracy of ML traffic forecasting models.

I. INTRODUCTION AND MOTIVATION

Radio Access Networks (RANs) are the backbone of wireless connectivity, enabling enable mobile services that range from real-time navigation and ubiquitous video streaming to remote healthcare and immersive augmented reality. RANs are also the main target for investment by Mobile Network Operators (MNOs): market analyses indicate that RANs account for over 67% of the global 5G infrastructure spend in 2024 and the dominant expense segment in the mobile ecosystem [1].

The high cost of RANs stems from the need for MNOs to continuously optimize radio deployments in response to increasing user-generated traffic, stricter quality-of-service requirements from mobile applications, and the integration of new Radio Access Technologies (RATs) while guaranteeing the operation of older RAT generations. RAN upgrading is indeed a ceaseless process, as also indirectly demonstrated by longitudinal measurements that attest the enhancement of 5G performance across recent years [2], [3].

Despite the inherent time-varying nature of RAN deployments, our current understanding of the dynamics of radio access infrastructures is extremely limited: as later discussed in Section II, the literature on the subject is quasi-nonexistent. As a result, important questions remain unanswered, which we list next with reference to the representative example in Figure 1, based on real-world measurements from Section III.

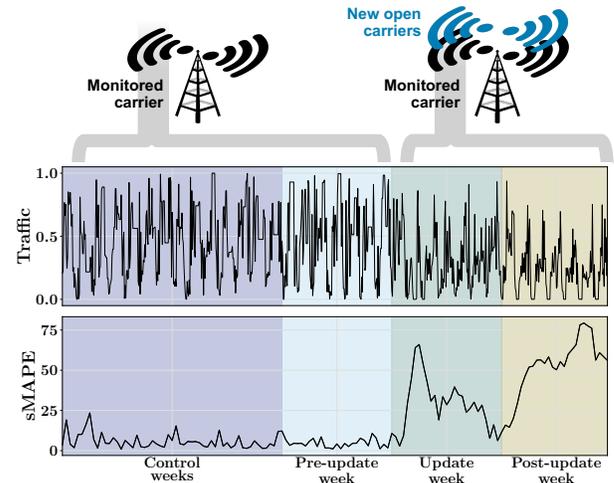


Fig. 1: Example of a RAN update event. The addition of two carriers operating in a new band at a site with two pre-existing 180° cells (top) impacts on (i) the real-world traffic at one pre-existing carrier (middle) and (ii) the prediction accuracy of a N-BEATS model for the traffic at such a carrier (bottom).

1. *How often do MNOs update their RANs and what type of modifications do they perform?* As shown in the top portion of Figure 1, an MNO may, e.g., add new carriers at an existing site or reconfigure the parameters of an in-place carrier. Yet, the frequency of such upgrades—regardless of type—remains unknown, making it impossible to gauge the practical relevance and implications of the phenomenon.
2. *How do RAN updates affect prior RAN Key Performance Indicators (KPIs)?* The middle plot of Figure 1 exhibits the time series of a specific KPI, i.e., the (normalized) traffic served by a pre-existing carrier, before and after a RAN update at its site. Here, the deployment of new carriers during the ‘Update week’ results in a noticeable decrease in the traffic volume handled by the original carrier.
3. *How do KPI deviations induced by RAN updates condition downstream Machine Learning (ML) tasks?* The bottom plot of Figure 1 portrays the prediction error of the popular N-BEATS model [4] in forecasting traffic at the pre-existent carrier. The ML model, trained on historical data, achieves high accuracy prior to the RAN update, but suffers from significant performance deterioration afterward.

This paper addresses the questions above from a network-side vantage point and makes the following contributions.

- We employ a longitudinal dataset describing the multi-year evolution of a nationwide operational RAN infrastructure composed of over 500,000 carriers and conduct the first characterization of RAN updates in production. We investigate the type and frequency of RAN modifications at different geographical scales, and reveal that RAN updates are prevalently linked to new carrier deployments and take place with notable frequency. Analyses are in Section IV.
- We leverage a two-year-long dataset of measurements collected in the same RAN deployment and assess the impact of RAN updates on a chief KPI for network management, *i.e.*, the carrier-level traffic. We find that RAN modifications can cause statistically significant traffic load and pattern changes in existing carriers. Analyses are in Section V.
- We investigate the effect of the observed KPI deviations on ML-based solutions informed by such indicators. We focus on a widely researched task in anticipatory network management, *i.e.*, traffic forecasting at carrier level. We show that RAN updates can curb the accuracy of state-of-the-art ML traffic predictors, dramatically inflating forecasting errors at affected carriers. Analyses are in Section VI.

Overall, our study unveils and quantifies the high prevalence of RAN updates in operational infrastructures—a phenomenon that has been largely overlooked to date—and the potential strong implications on data-driven network management. Thus, our results call into question common research practices adopted to validate ML models for RANs, which often assume completely static carrier deployments. Designing improved ML models that can cope with the effect of RAN updates is beyond the scope of this paper. Yet we release¹ a dataset of 1.800 traffic time series measured at carriers affected by RAN modifications, such as the sample in the middle plot of Figure 1, to support further research on the subject.

II. RELATED WORKS

Our work is related to prior studies on both measurements of RAN deployments and impact of non-stationary network traffic on the performance of data-driven traffic predictors.

RAN deployments. Numerous measurement studies have evaluated in recent years the performance and deployment of mobile networks. The vast majority of the literature adopts a client-side viewpoint, collecting data from mobile devices while they are associated to the target radio access infrastructure [5]–[18]. As a result, many of these studies consider sparse measurements from a limited set of devices [5]–[11], relatively short periods of time of days or weeks [12]–[16], or circumstantial scenarios encompassing a few RAN sites [17], [18], which do not allow appreciating RAN updates.

A few client-side investigations span one or more (possibly non-continuous) years and let observe evolutions in the RAN deployment. However, these works primarily focus on end-user performance, in terms of, *e.g.*, throughput, latency, or Quality of Service (QoS) [2], [3], [19]–[21]. Observations on the network infrastructure are instead hindered by the vantage

point and limited to, *e.g.*, the coverage of different RATs [3], [20] or implementation of carrier aggregation strategies [21].

Measurement studies from a network or provider perspective have explored traffic dynamics [22]–[24] and application performance [25] or consumption [26]–[28] but overlooked the RAN deployments. To the best of our knowledge, only one study has examined to date the evolution of an MNO commercial infrastructure [29]. However, the focus of that work is on the heterogeneity of 5G user devices and on the traffic they induce on the network; from the viewpoint of RAN dynamics, the study just provides high-level trends of the incidence of different RATs over the 2020–2022 period.

In summary, no prior work has systematically characterized radio access deployment updates or their effect on traffic loads at RAN elements as we do in this paper.

Data-driven traffic forecasting. The literature on mobile network traffic prediction is extensive and, over the past decade, has witnessed a shift from traditional auto-regressive approaches [30]–[33] to ML-driven models. In the ML space, a plethora of architectures have been proposed that leverage recurrent [34], convolutional [35], or graph [36] neural networks, as well as more advanced techniques such as three-dimensional convolution [37], transfer learning [38], original input operators [39], or meta-learned loss functions [40]. As the body of works is too vast to be reported here, we refer the reader to recent comprehensive surveys [41]–[44].

However, the established norm in the works above is to consider static network settings. When real-world measurements are used to train and validate the ML models, KPI data is collected from a portion of the RAN that is stable over several continuous weeks [36], [40] or the measurements are cleaned or aggregated to ensure invariant traffic settings [34], [35], [37], [38]. Hence, ML models are tested under traffic conditions that closely match those seen in the training phase.

The prediction of non-stationary mobile traffic is instead largely unexplored. We provide first evidence that RAN updates can hinder the performance of ML traffic predictors. Given the magnitude of the phenomenon, which affects a non-negligible portion of the network with a notable frequency, our results represent an important reality check for the community.

III. MEASUREMENT DATA

Our work builds on measurements data acquired in the nationwide commercial mobile network of Orange, the incumbent MNO in France. As illustrated in Figure 2, the target infrastructure covers 2G, 3G, 4G, and 5G Non-Standalone (NSA) RATs. We monitor the network from multiple vantage points, through passive probes deployed in the RAN, Core Network (CN), and internal management system of the MNO. We next detail the collection procedures and resulting datasets.

A. RAN Configuration

We collect Radio Site Configuration (RSC) information from the curated databases maintained by the radio planning department of the MNO. The databases contain detailed records of the configuration states of all carriers in the Orange

¹Data available at <https://github.com/nds-group/ran-updates>.

RAN infrastructure serving the whole France, including sites planned for future deployment or recently decommissioned.

We extract weekly snapshots of the geographical location, status (*i.e.*, under development, active, or decommissioned), and relevant configuration parameters (*i.e.*, antenna type, operating frequency, azimuth) of all carriers in each RAT, covering a period from the commercial opening of 5G in late 2020 to the end of 2024. Overall, we collect longitudinal data about the status of 520,000 carriers, with a rough split of 9%, 19%, 65%, and 7% across 2G, 3G, 4G, and 5G, respectively. It is worth noting that Orange launched commercial 5G Standalone (SA) services only in early 2025, hence all our measurements pertain to 5G NSA; however, the absence of SA data is inconsequential for our study, as SA involves CN-related upgrades while our focus is on RAN modifications.

B. Mobile Data Traffic

We measure the mobile network traffic that flows through the whole infrastructure using passive probes monitoring the SGi interface of the Packet Gateways (PGWs) that is serving both 4G and 5G users in 5G NSA deployments. The recording of network traffic from 2G and 3G equipment is performed over the Gi and Gn interfaces that connect Gateway GPRS Support Nodes (GGSNs) to external Public Data Networks (PDN). It is important to note that 5G NSA lacks a dedicated core, hence 5G gNodeBs (gNBs) leverage the 4G Mobility Management Entity (MME) via the X2 interface with co-located 4G eNodeBs (eNBs) for control-plane communication, and use a modified S1-U interface to connect the 4G Service Gateways (SGWs) and PGWs for data plane transmissions, as illustrated in Figure 2. To separate 4G and 5G NSA traffic at PGWs, we leverage information collected by the RAN probes at the MME, namely (*i*) Packet Data Protocol (PDP) context session data that include a flag for sessions with 5G NSA support and (*ii*) a *secondary* traffic indicator that is used by 4G eNBs to label traffic received through the X2 interface.

The resulting mobile network traffic dataset captures the total volume of exchanged bytes (in uplink and downlink) at every active carrier in the target nationwide infrastructure. The traffic data is aggregated with a temporal resolution of 15 minutes and is recorded for 15 consecutive months, from October 2023 through December 2024.

C. Ethical Considerations

We analyze user-generated data consisting of transport-layer sessions in the Orange CN. Spatiotemporal aggregations at carrier level and over 15-minute intervals render individual user sessions not identifiable, ensuring compliance with General Data Protection Regulation (GDPR) and protecting data subject privacy. The aggregation process is conducted within the secure infrastructure of the MNO by authorized personnel and under the guidance of the Data Protection Officer (DPO).

IV. RAN EVOLUTION CHARACTERIZATION

By leveraging the RAN configuration dataset presented above, we conduct the very first analysis of the dynamics of a large-scale mobile network infrastructure in production.

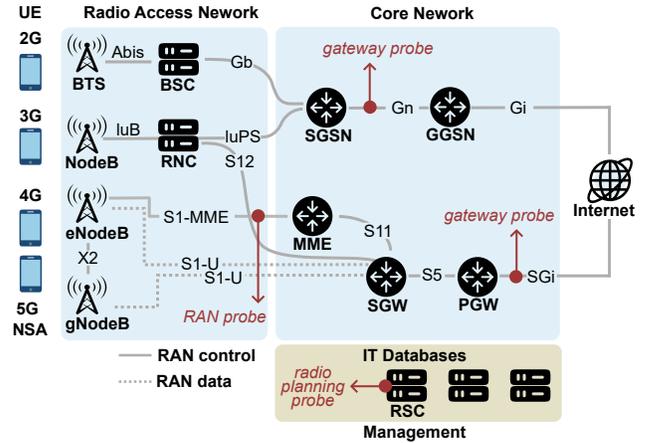


Fig. 2: High-level diagram of the packet-switched portion of the target multi-RAT 2G, 3G, 4G, and 5G NSA network, illustrating the location of the passive monitoring probes.

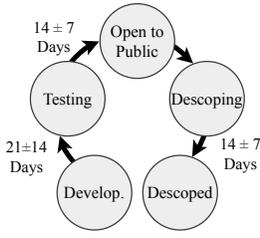
A. Carrier Life Cycle and RAN Update Types

We start by outlining the typical life cycle of a carrier as observed in the weekly snapshots from the Orange RAN deployment. Each carrier follows a rather established timeline that is illustrated in Figure 3a. A new carrier first appears as *developing* in the configuration dataset when it is planned for deployment. Within three weeks on average, the carrier transitions to a *testing* phase following its physical deployment, and undergoes a series of evaluations that validate the correct antenna installation and its expected impact on the RAN coverage and capacity. During this stage, which generally lasts another couple of weeks, the carrier is operational and capable of serving User Equipments (UEs).

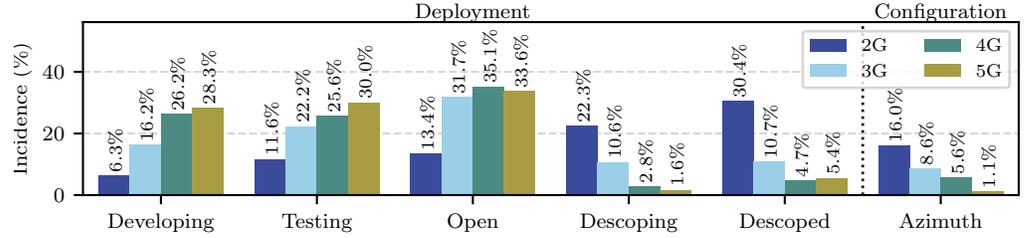
Once tested, the carrier becomes operational and *open* to public. This state is the longest in the life cycle, usually spanning years, until the hardware malfunctions, the MNO renovates its radio planning, or the RAT technology becomes obsolete. When one of these conditions occurs, the carrier enters its end-of-life: first, it undergoes a *descope* phase where the antenna is deactivated and withdrawn from the commercial network; then, after a couple of weeks, the antenna is *descope*d and physically removed from the site.

Updates associated with the life cycle of a carrier are classified as RAN *deployment* updates. They are not the sole type of modifications, since the MNO may adjust parameters of an operational carrier while it is in an open state, *e.g.*, varying the azimuth, tilt, or frequency of the corresponding antenna. These are referred to as RAN *configuration* updates.

Figure 3b shows different types of RAN updates, along with the fraction of total modifications they represent in each RAT, as derived from our longitudinal configuration data. The vast majority of RAN updates across all RATs are deployment-related. Configuration updates are almost exclusively azimuth adjustments and only contribute 5.6% and 1.1% of total RAN changes in 4G and 5G, respectively. It is also important to note that the total amount of modifications is very different across technologies: as shown in Figure 4a, 4G updates account for almost 70% of the total changes, which is expected as

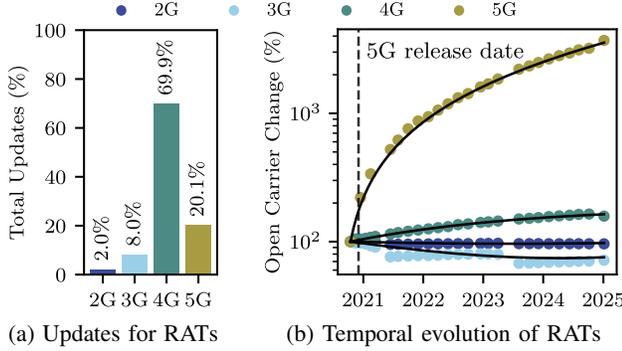


(a) Carrier life cycle



(b) Types of RAN updates

Fig. 3: (a) Typical carrier life cycle observed in the RAN configuration data with transition times reported as median \pm Median Absolute Deviation (MAD). (b) Types of RAN updates and their percent incidence on all modifications recorded for each RATs.



(a) Updates for RATs

(b) Temporal evolution of RATs

Fig. 4: (a) Fraction of RAN updates associated to each RAT. (b) Percent change of open carriers per RAT over time.

they concern the largest RAT infrastructure; 5G updates are instead responsible for 20% of the overall modifications—a large fraction considering that 5G only represents 7% of the total RAN deployment as mentioned in Section III-A.

As far as RAN deployment updates are concerned, a clear correlation emerges between the RAT generation and the incidence of each type of update in Figure 3b: 2G has the lowest incidence of developing, testing, or open carriers, and the highest incidence of descoping and descoped carriers; for 5G, the opposite holds—a sign of older technologies being phased out as 4G is still expanding and 5G is being rolled out. As a confirmation, Figure 4b shows the percent change in the number of open carriers with respect to late 2020, *i.e.*, $100 \cdot c_g(t)/c_g(T)$, where $c_g(t)$ is the number of open carriers for RAT g at time t , and T is October 2020. The trends are consistent with those observed until 2022 [29] and reflect how 5G deployment is growing much faster than all other RATs, while 2G and 3G are both characterized by a contraction.

Overall, we conclude that *RAN updates on the technologies that are today relevant, i.e., 4G and 5G, represent 90% of the total access network modifications and are dominated by the deployment of new carriers*—with events linked to developing, testing or newly open carriers amounting to more than 86% and 91% of the total changes in 4G and 5G, respectively.

B. Strategy and Frequency of Open Carrier RAN Updates

Drawing on the previous results, we focus our attention on RAN updates associated to the deployment of new carriers and limit our analysis to 4G and 5G RATs. Since the developing,

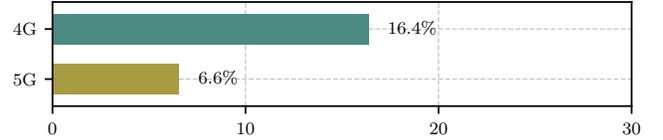


Fig. 5: Percentage of open carrier updates in new sites.

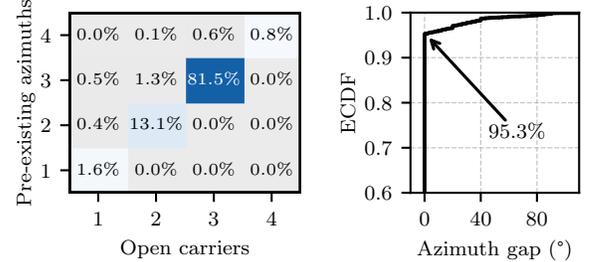


Fig. 6: Left: relationship between the number of new carriers open concurrently at one site (columns) and the number of pre-existing azimuth angles at the same site (rows). Right: distribution of the maximum gap in degrees between the azimuths of carriers deployed concurrently at one site and the azimuths of the pre-existing carriers at the same site.

testing and open states occur sequentially in the life cycle in Figure 3a, we consider their concatenation as an atomic event and only look at updates matching newly open carriers.

We first investigate the strategy adopted by the MNO when deploying new carriers at a target site. Figure 5 depicts the percentage of 4G and 5G carriers that are installed in brand new sites where no RAT infrastructure was present before. The capillarity of the 4G radio access is still expanding, around 16% new open carriers establishing a new site; conversely, 94% of 5G carriers leverage existing sites. When combining the two technologies, a significant majority of open carriers, around 86%, are introduced at already present sites. Focusing on this more frequent case, the left plot in Figure 6 shows that the number of carriers open concurrently at one site (columns) generally matches the number of azimuth angles that are already present at the same site (rows)—which most often is organized into three sectors separated by 120° . The right plot in Figure 6 confirms that the maximum difference between the azimuths of open and pre-existing carriers is zero degrees in over 95% of cases. Jointly, these results let us conclude that the MNO nearly always deploys new carriers

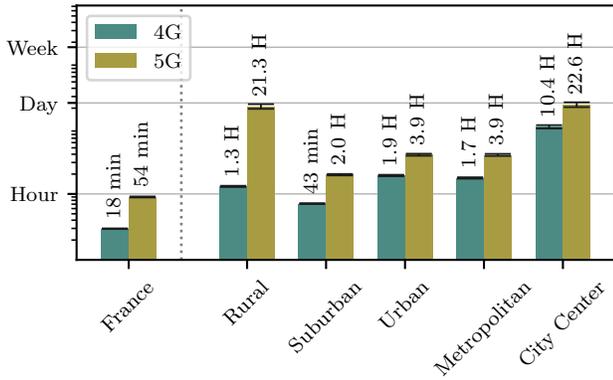


Fig. 7: Average time between open RAN carrier updates for 4G and 5G in the whole France and across urbanization levels.

that exactly match the former sectorization at the target site.

We then analyse how frequent is the deployment of new carriers in the RAN. In order to quantify such a frequency, we count the number of RAN updates one week to the next and assume that they are uniformly distributed in time during that week; while this is a rough approximation, the weekly resolution of our RAN configuration dataset does not allow to work under a better premise and we believe that this estimation still provides a relevant measure of the magnitude of the phenomenon we intend to capture.

The first pair of bars on the left of Figure 7 show that the average time between the opening of any two subsequent RAN carriers is just 18 and 54 minutes in the 4G and 5G nationwide infrastructures, respectively. These figures reveal that *production-grade RANs are surprisingly dynamic—to the point that, on average, country-scale deployments receive upgrades multiple times per hour*. It is also apparent that 4G open carrier updates are more frequent than 5G ones, which is aligned with the incidence of each RAT in Figure 4a.

The right portion of Figure 7 breaks down the result per urbanization level. Specifically, we tell apart five levels of development, from *rural* to *city center*, based on the population density information provided by the French national institute for statistics. The results highlight that the RANs covering each of the suburban, urban and metropolitan regions of the country undergo new carrier updates at every few hours or less. The values are higher in the case of rural areas (where the infrastructure is sparse) and in city centers (which cover a much smaller surface and set of carriers than other levels).

In order to make the results more interpretable, we compute the same frequency metric over geographical areas of comparable size. First, we consider the IRIS zoning, which is the finest-grained statistical tessellation available for the French territory, and draw the map in Figure 8. The average number of days between RAN updates is clearly higher than in Figure 7, since IRIS zones are as small as a neighborhood in a major city hence the number of locally deployed carriers is generally low. However, the map highlights how higher update frequency pertains to larger cities, some of which are labeled in the figure: in IRIS zones located within these dense urban

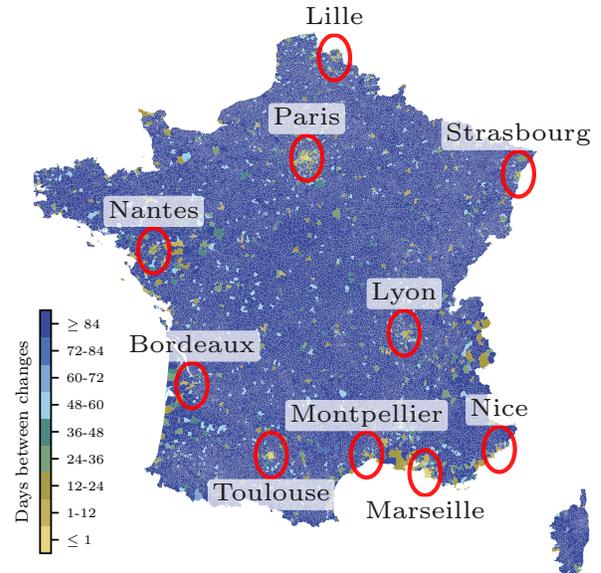


Fig. 8: Map of the average time between open 4G or 5G RAN carrier updates in all IRIS zones of metropolitan France.

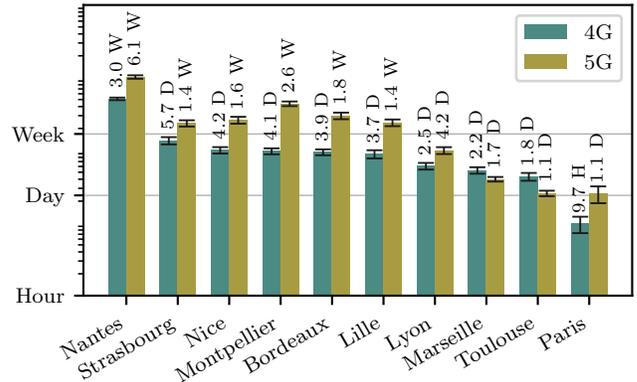


Fig. 9: Average time between open RAN carrier updates for 4G and 5G within each of the ten major French cities.

areas, new carriers may be open every week or less on average.

The result is corroborated by Figure 9, which portrays the frequency of updates in the RANs serving individual major cities in France. New carriers are opened in each city at every few days; *in Paris, the combined 4G and 5G deployment is upgraded somewhere within the urban area around four times per day on average*. We also run a separate analysis (omitted here due to space limitations) for the metropolitan region of Paris during the period immediately preceding the 2024 Olympics, which were held in that city in July and August; results reveal that the frequency of RAN carrier updates grew in the months preceding the event, with an average of a new 5G carrier open at every three hours during June.

In summary, our analysis above quantifies for the first time the mutability of a commercial RAN, proving that these infrastructures are far from static. Based on our insights, in metropolitan-scale month-long scenarios that are commonly used for networking research [45], [46], the RAN may undergo

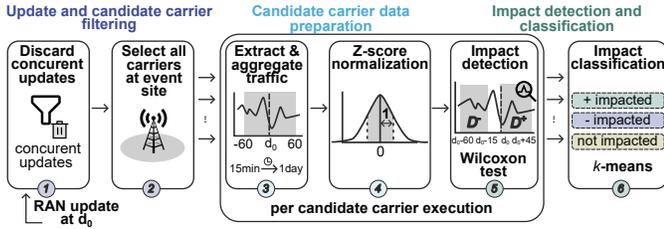


Fig. 10: Workflow of the methodology adopted in our study to quantify the effect of new open carriers on pre-existing ones.

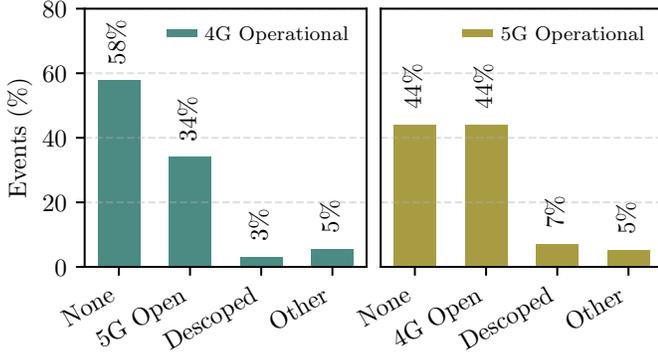


Fig. 11: Proportion of 4G (left) and 5G (right) open carrier updates that occur in isolation (first bar in each plot) or concurrently with (i) the augmentation of the other RAT at the same site (second bar), (ii) the descoping of local carriers, or (iii) any other deployment or configuration RAN change.

many hundreds of updates during the traffic measurement period—possibly inducing important but currently overlooked side effects on the results of studies based on such data.

V. IMPACT OF RAN UPDATES ON CARRIER TRAFFIC

Having brought to light the significant prevalence of RAN updates associated with open carriers, we investigate the impact that such infrastructure modifications yield on the traffic volumes served by pre-existing carriers. We focus on carrier-level traffic load as it is a prominent KPI collected at the RAN and used to inform a wide range of network management tasks, including baseband function placement [47], slice orchestration [48], power control [49], or O-Cloud management [50]. We acknowledge that other KPIs, such as latency, packet loss rates, or radio resource utilization, are also relevant, but, due to space limitations, we circumscribe the scope of this work to traffic load and leave other KPIs for future investigation.

A. Methodology

We propose a methodology to quantify the impact of open carrier updates on the load of already present RAN carriers. The workflow is outlined in Figure 10 and detailed below.

1) *Update and candidate carrier filtering*: Given an input RAN update, we verify in stage ① if the update occurs in isolation or concurrently with other modifications at the same site and azimuth. We only advance in the workflow in the former case, while concurrent updates are discarded. The rationale is that joint updates of a same sector yield

overlapping impacts that cannot be separated and quantified individually a posteriori. Figure 11 shows the fraction of open 4G and 5G carrier updates that coincide with other deployment of configuration updates at the same site and azimuth: 58% and 44% of such updates do not concur with other co-located RAN changes and are thus retained for subsequent analysis.

Next, in stage ②, we identify a set *candidate* carriers that are already present in the network and may be impacted by the update. For this, we adopt a strategy of selecting all existing carriers co-located at the same site where the update occurs. While simple, this approach is substantiated by recent analyses of operational RANs where network configuration updates—namely, carrier switch-off/on events—were demonstrated to only affect same-site carriers in the majority of cases [51].

2) *Candidate carrier data preparation*: For each candidate carrier, in stage ③ we extract its recorded mobile data traffic within a 120-day window \mathcal{W} centered at the update date. The time series data is aggregated at daily resolution, which mitigates the natural burstiness of carrier-level traffic but retains enough granularity to appraise the impact of the RAN update in later stages. In stage ④, we standardize the daily traffic to remove bias in terms of absolute volume and variability, making the fluctuations of time series from different candidate carriers comparable and facilitating their processing in the remaining stages. We use the standard score normalization

$$t'(d) = \frac{t(d) - \bar{\mu}_t}{\bar{\sigma}_t}, \quad d \in \mathcal{W}, \quad (1)$$

where $t(d)$ is the traffic measured at the candidate carrier in day d , and $t'(d)$ is its standardized version. The mean $\bar{\mu}_t$ and standard deviation $\bar{\sigma}_t$ are computed from the daily traffic measured at the candidate carrier during a reference period R of 30 days, situated two months before the target update, as

$$\bar{\mu}_t = \frac{1}{|R|} \sum_{d \in R} t(d), \quad \bar{\sigma}_t = \sqrt{\frac{1}{|R| - 1} \sum_{d \in R} (t(d) - \bar{\mu}_t)^2}, \quad (2)$$

hence they are indicative of traffic statistics prior to the update.

3) *Impact detection and classification*: The standardized time series of each candidate carrier undergoes a statistical test in stage ⑤, with the aim of assessing if the traffic load has been actually impacted by the RAN update under consideration. To this end, let us define two time intervals $\mathcal{D}^- = [d_0 - 60, d_0 - 15]$ and $\mathcal{D}^+ = [d_0, d_0 + 45]$; here, d_0 denotes the last day of the week of the update in the RSC data, hence the intervals cover 45 days before² and after the update. We define two populations of normalized daily traffic samples, a pre-update $\mathbf{t}^- = \{t'(d) : d \in \mathcal{D}^-\}$ and a post-update $\mathbf{t}^+ = \{t'(d) : d \in \mathcal{D}^+\}$. Finally, we run a Wilcoxon signed-rank test [52] on the hypothesis that the update induced a statistically significant impact on the candidate carrier traffic:

$$H_0 : \mathbf{t}^- \stackrel{d}{=} \mathbf{t}^+, \quad \alpha = 0.01. \quad (3)$$

²We consider for \mathcal{D}^- a 15-day safety margin before the update so as to compensate for the weekly resolution of the RSC data in Section III-A and ensure that the samples in \mathcal{D}^- are representative of pre-update carrier traffic.

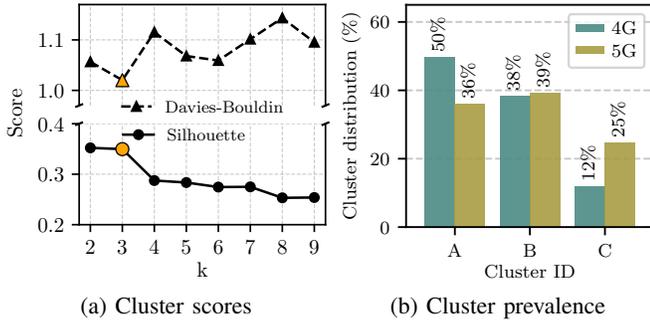


Fig. 12: (a) Silhouette (higher is better) and Davies-Bouldin (lower is better) scores of clusterings versus the parameter k . (b) Cluster distribution upon 4G and 5G open carrier updates.

We opt for this test since (i) it is non-parametric, (ii) it uses the ranks of the groups instead of mean values hence it is suitable for comparing samples drawn from non-Gaussian distributions such as those often encountered in mobile traffic, and (iii) it operates on paired observations of non-independent samples like those extracted from a same underlying time series.

In the final stage ⑥ of our workflow, all the standardized time series of the candidate carriers that pass the Wilcoxon test in (3) are grouped in a single dataset, independently of the open carrier RAN update that impacted them. We then run a clustering method on the dataset of time series, so as to explore the existence of different *types* of effects that the updates may produce on pre-existing carriers.

We define a feature vector $\mathbf{f} = [\bar{\mu}^-, \bar{\mu}^+, \mu, \sigma, \Delta]^\top$ for each time series, where: $\bar{\mu}^-$, $\bar{\mu}^+$, and μ are the mean values of $t'(d)$ samples computed in the intervals \mathcal{D}^- , \mathcal{D}^+ , and \mathcal{W} , respectively; σ is the standard deviation of $t'(d)$ samples over \mathcal{W} , and $\Delta = \bar{\mu}^+ - \bar{\mu}^-$. This compact representation lets us efficiently group the time series based on pre- and post-update statistics. Classic Silhouette and Davies-Bouldin scores are used to determine the optimal set of clusters among those returned by an exhaustive search of the k -means parameter k .

B. Analysis of Traffic Volumes and Patterns

In order to provide a macroscopic view of the impact of open RAN carrier updates on the existing infrastructure, we compute the total number of pre-existing carriers that undergo a statistically significant change in their served traffic, *i.e.*, pass the Wilcoxon test in stage ⑤ of our workflow. The result reveals that, *for the 4G RAT only, the traffic dynamics of over 20,000 carriers—corresponding to 7.1% of the nation-wide RAN deployment—are affected in a relevant manner by the installation of new open carriers*. It is worth noting that, while considerable and consistent with the high frequency of open carriers discussed in Section IV, this percentage is actually a lower bound, since we are considering only the most common class of updates and the total incidence of all RAN updates presented in Section IV-A will be inherently higher.

Having corroborated that a substantial proportion of carriers is affected by RAN updates, we now examine the different consequences that one such update can have on the traffic

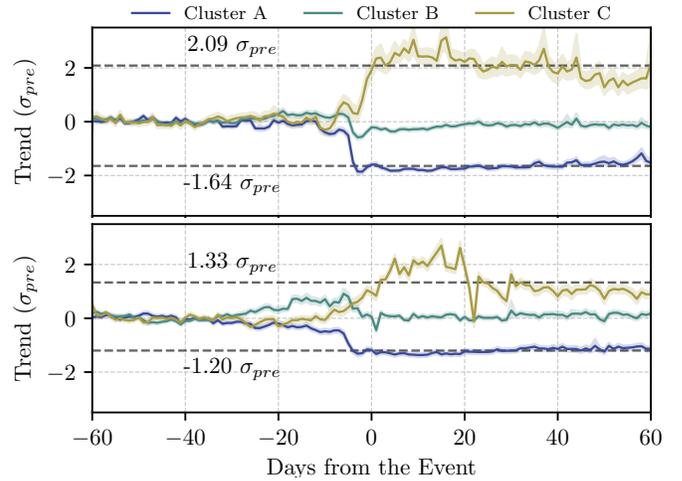


Fig. 13: Centroids of the standardized daily time series of the traffic served by pre-existing carriers that are impacted by 4G (top) and 5G (bottom) open carriers updates, for each cluster.

demands of impacted pre-existing carriers. To this end, we consider the outcome of the clustering in stage ⑥ of our workflow. Figure 12a shows the evolution of the scores versus k , which consistently indicates that three clusters retain the best grouping of traffic time series patterns upon an update. Figure 12b shows the fraction of impacted carriers in each of such three clusters, depending on whether the RAN update is induced by the opening of a new 4G or 5G carrier. Finally, Figure 13 provides a visual illustration of the typical impact of a new open 4G (top) and 5G (bottom) carrier on pre-existing carriers categorized in each of the three clusters, represented by the centroid of all standardized time series in each cluster.

- By joining Figures 12b and 13, three effects are apparent.
- In *cluster A*, the majority of carriers (50% upon a 4G update and 36% upon a 5G update) experience a significant drop in served traffic, by 1.64 or 1.20 standard deviations depending on the RAT of the newly installed carrier.
 - In *cluster B*, a second large fraction of carriers (38% upon a 4G update and 39% upon a 5G update) do not record significant variations in their traffic demands.
 - In *cluster C*, a minority of impacted carriers (12% upon a 4G update and 25% upon a 5G update) observe an important increase in their traffic volume.

The first pattern is expected, as newly deployed carriers start serving part of the local demand and alleviate the load of already present carriers. The fact that 5G antennas induce smaller variations may be explained by the reduced number of 5G users compared to 4G ones (*e.g.*, in our reference mobile traffic dataset, 5G generates only one fifth of the traffic produced by 4G), which translates into a smaller potential to cause shifts in the pre-existing demand. The second behavior is an indicator that a decent fraction of open carriers tend to have only temporary or minimal effects on the traffic served by co-located RAN equipment. Finally, the third cluster gathers more rare cases where the traffic of existing carriers grows after the opening of a new local carrier; this effect is presumably due

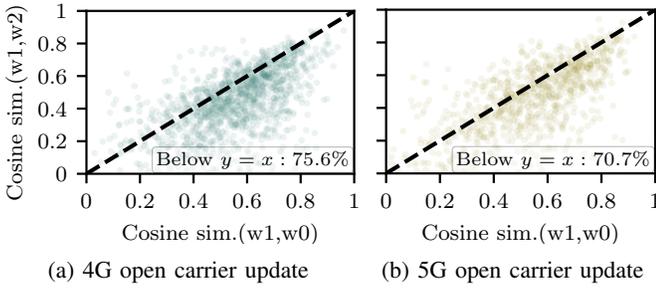


Fig. 14: Comparison among median-week traffic profiles at each impacted carrier before the update, (w_1, w_0) , and after it, (w_1, w_2) . Likelihood is measured using pairwise CS.

to external factors that out of our control, such as increases of the demand in the area or removal of capacity at nearby sites.

We remark that changes by more than one standard deviation are very large, and typically indicative of drops in peak traffic by tens of percent points. Therefore, the results of our analysis let us conclude that more than 45% of already present carriers impacted by open carrier RAN updates sustain significant drops in their served traffic volume.

It is also worth remarking that the deployment of new carriers induces not only reductions in the served load at pre-existing carriers, but also modifies short-timescale temporal fluctuations in the carrier-level traffic. To study this effect, for each impacted carrier retained after stage 5 of our workflow, we extract its original traffic time series, and compute median-week profiles [53] in three intervals: a month $w_0 = [d_0 - 63, d_0 - 35]$ well in advance of the update, the month $w_1 = [d_0 - 35, d_0 - 7]$ immediately preceding the update, and the month $w_2 = [d_0 + 7, d_0 + 35]$ after it. We then measure the pairwise likeness of these profiles via Cosine Similarity (CS), which, unlike the Wicoxon test, compares the weekly temporal fluctuations of the demands. Figure 14 juxtaposes the CS of pairs (w_1, w_0) (x axis) and (w_1, w_2) (y axis), revealing that the majority of the points are located below the diagonal: the traffic similarity is typically higher in the two months before the RAN update, which actually alters the traffic dynamics in over 70% of cases.

Overall, we find that a large fraction of open carrier RAN updates lead to important changes in the demands at existing carriers, in terms of both volume and temporal dynamics.

VI. IMPACT OF RAN UPDATES ON TRAFFIC FORECASTING

The findings in Sections IV and V undermine common assumptions in the literature about data-driven network management, *i.e.*, the immutability of the RAN infrastructure and the consequent stationary nature of carrier-level traffic demands. We thus study if the established practice of considering static network settings through ML model training, validation, and test periods may overlook important effects due to RAN updates that we showed to condition real-world deployments.

To this end, we consider one representative ML task, *i.e.*, carrier-level traffic forecasting, and study the repercussions that the opening of new carriers in the RAN has on the

inference performance of the model at pre-existing carriers. The choice of task is motivated by the pivotal role that forecasts play in anticipatory network management, as well as by the prevalence of the problem in the scientific literature. We stress that many other ML-driven tasks may be affected by RAN updates, yet it is not the purpose of this work to provide an extensive assessment over a full range of ML problems and we leave such a holistic investigation for future works.

A. Experimental Setup

The forecasting task consists in predicting traffic at 1,800 carriers in Cluster A returned by our workflow in Section V-B, *i.e.*, carriers that observe the most frequent effect of a significant drop in traffic following an update at their site. The traffic has hourly resolution and we consider a single-step prediction, *i.e.*, the model output is the demand foreseen for the next hour.

We select six state-of-the-art ML forecasting models: (i) a baseline Long Short-Term Memory (LSTM) recurrent neural network that is very popular for modeling sequential dependencies [54]; (ii) an encoder-decoder Transformer that captures long-range dependencies without recurrence [55]; (iii) a Temporal Convolutional Network (TCN) that uses dilated causal convolutions for efficient sequence modeling [56]; (iv) the Neural Basis Expansion Analysis Time Series Forecasting (N-BEATS) model that uses a deep residual architecture for interpretable and accurate predictions [4]; (v) the Time-Series Mixer (TSMixer) model, a feedforward architecture using multilayer perceptrons that addresses Transformers' computational demands while maintaining long-term forecasting performance [57]; (vi) the DLinear model, which captures trend components through simple linear layers and recently opened new trends in ML-driven time series prediction [58].

Each model is trained for up to 100 epochs, with early stopping triggered if the validation accuracy fails to improve for 10 consecutive epochs so as to avoid overfitting. We conduct a grid search over multiple input window sizes and, based on validation performance, select an input of 72 time steps (*i.e.*, three days) for all models. We adhere to the default settings proposed by Darts [59] for all remaining parameters.

We train the per-carrier models with traffic data from the three months preceding the update, specifically $[d_0 - 118, d_0 - 28]$; 80% of such data is used for training, and 20% for validation. Each trained model is then used in inference after $d_0 - 28$, measuring its performance in the four different time intervals that are illustrated in Figure 1: (i) a two-week *control* period $[d_0 - 27, d_0 - 14]$ where the traffic is expected to be stationary with respect to the training data; (ii) the immediate *pre-update* week $[d_0 - 13, d_0 - 7]$; (iii) the week during which the *update* takes place $[d_0 - 6, d_0]$; and, (iv) the *post-update* week $[d_0 + 1, d_0 + 7]$ during which the effect of the RAN modification shall be visible with certainty.

We evaluate the prediction error during the inference periods with Symmetric Mean Absolute Percentage Error (sMAPE) and Weighted Average Percentage Error (WAPE), *i.e.*,

$$sMAPE = \frac{200}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{|y_i| + |\hat{y}_i|}, \quad WAPE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{\sum_{i=1}^n |y_i|}, \quad (4)$$

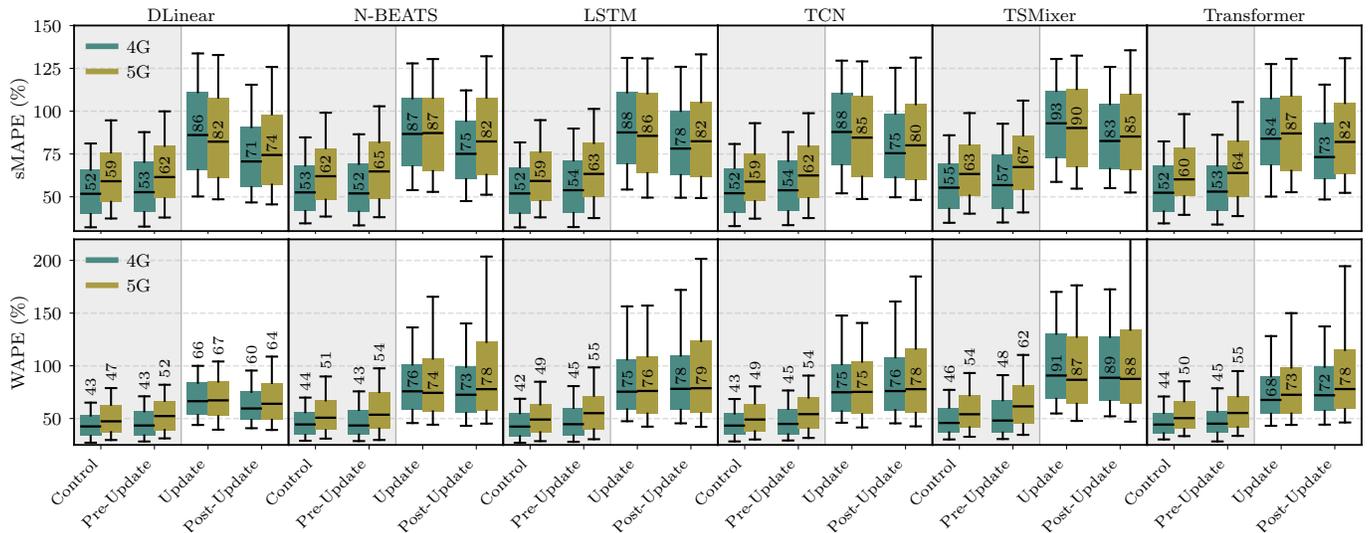


Fig. 15: Performance of ML models in terms of sMAPE and WAPE for control, pre-update, update, and post-update weeks.

where y_i and \hat{y}_i are the actual value and its associated forecast, while n is the number of predicted samples in the period.

B. Impact of RAN Modifications on Traffic Forecasting

Figure 15 reports the outcome of our experiments, making it very clear that *the ML prediction accuracy is affected by RAN updates in a marked way that is consistent across all tested models and performance metrics*. Indeed, the performance of each model under both sMAPE and WAPE is very similar in the control and pre-update weeks, during which the traffic behavior at the carrier is consistent with that seen in training. Instead, both the update and post-update weeks record clear reductions of accuracy following the RAN modification, indicating a decline that is not just temporary but persistent in time. For sMAPE, the forecast error jumps by 41% and 27% on average upon the installation of a new 4G and 5G antenna, respectively; the same figures are even more pronounced under WAPE, at 66% and 40%, respectively. Model-wise, TSMixer and Transformer appear to be the architectures that most suffer the update, highlighting the vulnerability of multi-head attention to temporal drift and traffic pattern changes. DLinear is instead the only model to show a modest recovery in accuracy post-update, which confirms its superior generalization capabilities over Transformers [58].

VII. IMPLICATIONS OF OUR STUDY AND CONCLUSIONS

Our analysis and results unveil and quantify a phenomenon that is not well understood to date, *i.e.*, the continual evolution of production-grade RANs. We show that infrastructural updates are in fact frequent and induce significant variations in the traffic served by the carriers. In turn, such alterations in the demands may undermine the operation of ML models by introducing sudden changes to input distributions with respect to those seen during training with historical pre-update data.

The implications are manifold. First, our study calls for additional investigations of RAN deployment dynamics in order to shed further light on the upgrade or sunseting processes of

RATs, which we find to be complex and ceaseless. Second, the impact of RAN updates on the existing infrastructure is entangled and our work just scratches the surface of its characterization. Many aspects remain to be explored, such as the effects of KPIs beyond traffic volumes, at levels other than that of individual carriers, or for modifications that target the configuration of carrier parameters. Third, our conclusions are a reality check for common practices adopted in the design and evaluation of ML models for network management. Stationarity in the network KPIs cannot be given for granted when training or evaluating such models, or there is a significant risk of facing unexpected, disappointing real-world performance.

The last point lets us advocate in favor of ML models that remain robust in the presence of data distribution drifts. While a few recent studies have begun to address the challenge of detecting drifts in the KPIs used by ML-based forecasting models [60], [61], this area remains largely underexplored. Online re-training mechanisms based on such drift detection are also critical [62], and so are zero- or few-shot learners capable of generalizing to novel network configurations without requiring extensive labeled or historical data. While proposing a solution to the problems above is beyond the scope of our work, we release our labeled measurement data to support future collaborative research on the topic.

The authors have provided public access to sanitized data at <https://github.com/nds-group/ran-updates>.

ACKNOWLEDGMENT

Nadezda Chukhno's work is funded by the European Union under Grant Agreement No. 101206327 (6G-AI-TANGO). Zbigniew Smoreda's work was supported by CoCo5G (ANR-22-CE25-0016) funded by the French National Research Agency. Marco Fiore's work was supported by grant CNS2023-143870 (6G-IRONWARE) funded by MICIU/AEI /10.13039/501100011033 and by the European Union NextGenerationEU/PRTR.

REFERENCES

- [1] IMARC Group, “5G Infrastructure Market Size, Share, Trends and Forecast by Communication Infrastructure, Network Technology, Network Architecture, Frequency, End User, and Region, 2025-2033,” Tech. Rep., 2025.
- [2] G. Caso *et al.*, “An Initial Look into the Performance Evolution of 5G Non-Standalone Networks,” in *IFIP/IEEE TMA*, 2023.
- [3] O. Basit *et al.*, “5G Metamorphosis: A Longitudinal Study of 5G Performance from the Beginning,” in *ACM IMC*, 2025.
- [4] B. N. Oreshkin *et al.*, “N-BEATS: Neural basis expansion analysis for interpretable time series forecasting,” in *ICLR*, 2020.
- [5] A. Narayanan *et al.*, “A First Look at Commercial 5G Performance on Smartphones,” in *ACM WWW*, 2020.
- [6] A. Narayanan, E. Ramadan *et al.*, “Lumos5G: Mapping and Predicting Commercial mmWave 5G Throughput,” in *ACM IMC*, 2020.
- [7] A. Narayanan, X. Zhang *et al.*, “A Variegated Look at 5G in the Wild: Performance, Power, and QoE Implications,” in *ACM SIGCOMM*, 2021.
- [8] C. Fiandrino *et al.*, “Uncovering 5G Performance on Public Transit Systems with an App-based Measurement Study,” in *ACM MSWiM*, 2022.
- [9] R. A. Fezeu, J. Carpenter *et al.*, “Mid-band 5G: A Measurement Study in Europe and US,” in *ACM SIGCOMM*, 2024.
- [10] R. A. Fezeu *et al.*, “Roaming across the European Union in the 5G era: performance, challenges, and opportunities,” in *IEEE INFOCOM*, 2024.
- [11] R. A. Fezeu, C. Fiandrino *et al.*, “Unveiling the 5G Mid-Band Landscape: From Network Deployment to Performance and Application QoE,” in *ACM SIGCOMM*, 2024.
- [12] K. Kousias *et al.*, “Coverage and Performance Analysis of 5G Non-Standalone Deployments,” in *ACM WiNTECH*, 2022.
- [13] A. Narayanan *et al.*, “A Comparative Measurement Study of Commercial 5G mmWave Deployments,” in *IEEE INFOCOM*, 2022.
- [14] A. Hassan *et al.*, “Vivisecting Mobility Management in 5G Cellular Networks,” in *ACM SIGCOMM*, 2022.
- [15] M. Ghoshal *et al.*, “Performance of Cellular Networks on the Wheels,” in *ACM IMC*, 2023.
- [16] M. I. Rochman *et al.*, “A Comprehensive Real-World Evaluation of 5G Improvements over 4G in Low-and Mid-Bands,” *IEEE Trans. on Cognitive Communications and Networking*, 2025.
- [17] D. Xu *et al.*, “Understanding Operational 5G: A First Measurement Study on Its Coverage, Performance and Energy Consumption,” in *ACM SIGCOMM*, 2020.
- [18] R. A. Fezeu *et al.*, “A Peek into 5G NSA vs. SA Control Plane Performance,” in *ACM HotMobile*, 2025.
- [19] X. Yang *et al.*, “Mobile Access Bandwidth in Practice: Measurement, Analysis, and Implications,” in *ACM SIGCOMM*, 2022.
- [20] I. Khan *et al.*, “How Mature is 5G Deployment? A Cross-Sectional, Year-Long Study of 5G Uplink Performance,” *Computer Communications*, p. 108153, 2025.
- [21] W. Ye *et al.*, “Dissecting Carrier Aggregation in 5G Networks: Measurement, QoE Implications and Prediction,” in *ACM SIGCOMM*, 2024.
- [22] U. Paul *et al.*, “Understanding Traffic Dynamics in Cellular Data Networks,” in *IEEE INFOCOM*, 2011.
- [23] M. Z. Shafiq *et al.*, “Characterizing and Modeling Internet Traffic Dynamics of Cellular Devices,” *ACM SIGMETRICS Performance Evaluation Review*, vol. 39, no. 1, pp. 265–276, 2011.
- [24] A. Furno *et al.*, “Joint Spatial and Temporal Classification of Mobile Traffic Demands,” in *IEEE INFOCOM*, 2017.
- [25] X. Yuan *et al.*, “Understanding 5G Performance For Real-World Services: A Content Provider’s Perspective,” in *ACM SIGCOMM*, 2022.
- [26] M. Z. Shafiq *et al.*, “Characterizing Geospatial Dynamics of Application Usage in a 3G Cellular Data Network,” in *IEEE INFOCOM*, 2012.
- [27] S. Mishra *et al.*, “Characterizing 5G Adoption and its Impact on Network Traffic and Mobile Service Consumption,” in *IEEE INFOCOM*, 2024.
- [28] S. Mishra, D. Madariaga *et al.*, “An Urban Geography of Mobile Application Usage: Connecting Demand Dynamics and Urban Fabrics,” in *IEEE INFOCOM*, 2025.
- [29] P. Parastar *et al.*, “Spotlight on 5G: Performance, Device Evolution and Challenges from a Mobile Operator Perspective,” in *IEEE INFOCOM*, 2023.
- [30] F. Xu *et al.*, “Big Data Driven Mobile Traffic Understanding and Forecasting: A Time Series Approach,” *IEEE Trans. on Services Computing*, vol. 9, no. 5, pp. 796–805, 2016.
- [31] M. Zhang *et al.*, “Understanding Urban Dynamics From Massive Mobile Traffic Data,” *IEEE Trans. on Big Data*, vol. 3, no. 1, pp. 10–22, 2017.
- [32] S. Ntalampiras *et al.*, “Forecasting Mobile Service Demands for Anticipatory MEC,” in *IEEE WoWMoM*, 2018.
- [33] J. X. Salvat *et al.*, “Overbooking Network Slices through Yield-Driven End-to-End Orchestration,” in *ACM CoNEXT*, 2018.
- [34] J. Wang *et al.*, “Spatiotemporal Modeling and Prediction in Cellular Networks: A Big Data Enabled Deep Learning Approach,” in *IEEE INFOCOM*, 2017.
- [35] D. Bega *et al.*, “DeepCog: Optimizing Resource Provisioning in Network Slicing with AI-based Capacity Forecasting,” *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 2, pp. 361–376, 2019.
- [36] F. Sun *et al.*, “Mobile Data Traffic Prediction by Exploiting Time-Evolving User Mobility Patterns,” *IEEE Trans. on Mobile Computing*, vol. 21, no. 12, pp. 4456–4470, 2021.
- [37] C. Zhang *et al.*, “Long-Term Mobile Traffic Forecasting Using Deep Spatio-Temporal Neural Networks,” in *ACM MobiHoc*, 2018.
- [38] Q. Wu *et al.*, “Deep Transfer Learning across Cities for Mobile Traffic Prediction,” *IEEE/ACM Trans. on Networking*, vol. 30, no. 3, 2021.
- [39] C. Zhang *et al.*, “Cloudlstm: A Recurrent Neural Model for Spatiotemporal Point-Cloud Stream Forecasting,” in *AAAI*, 2021.
- [40] A. Collet *et al.*, “Automanager: A Meta-Learning Model for Network Management from Intertwined Forecasts,” in *IEEE INFOCOM*, 2023.
- [41] C. Zhang *et al.*, “Deep Learning in Mobile and Wireless Networking: A Survey,” *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2224–2287, 2019.
- [42] G. O. Ferreira *et al.*, “Forecasting Network Traffic: A Survey and Tutorial with Open-Source Comparative Evaluation,” *IEEE Access*, vol. 11, pp. 6018–6044, 2023.
- [43] A. Nascita *et al.*, “A Survey on Explainable Artificial Intelligence for Internet Traffic Classification and Prediction, and Intrusion Detection,” *IEEE Communications Surveys & Tutorials*, pp. 1–1, 2024.
- [44] O. Aouedi *et al.*, “Deep Learning on Network Traffic Prediction: Recent Advances, Analysis, and Future Directions,” *ACM Comput. Surv.*, vol. 57, no. 6, Feb. 2025.
- [45] G. Barlacchi *et al.*, “A Multi-Source Dataset of Urban Life in the City of Milan and the Province of Trentino,” *Scientific Data*, vol. 2, no. 1, p. 150055, 2015. [Online]. Available: <https://doi.org/10.1038/sdata.2015.55>
- [46] O. E. Martínez-Durive *et al.*, “The NetMob23 Dataset: A High-Resolution Multi-Region Service-Level Mobile Data Traffic Cartography,” 2023.
- [47] L. M. M. Zorello *et al.*, “Black-Box Optimization for Anticipated Baseband-Function Placement in 5G Networks,” *Computer Networks*, vol. 245, p. 110384, 2024.
- [48] V. Sciancalepore *et al.*, “Mobile Traffic Forecasting for Maximizing 5G Network Slicing Resource Utilization,” in *IEEE INFOCOM*, 2017.
- [49] J. X. Salvat Lozano *et al.*, “Kairos: Energy-Efficient Radio Unit Control for O-RAN via Advanced Sleep Modes,” in *IEEE INFOCOM*, 2025.
- [50] X. Foukas *et al.*, “Concordia: Teaching the 5G vRAN to Share Compute,” in *ACM SIGCOMM*, 2021.
- [51] O. E. Martínez-Durive *et al.*, “An Evaluation of RAN Sustainability Strategies in Production Networks,” in *IEEE INFOCOM*, 2025.
- [52] F. Wilcoxon, “Individual Comparisons by Ranking Methods,” *Biometrics Bulletin*, vol. 1, no. 6, pp. 80–83, 1945.
- [53] A. Furno *et al.*, “A Tale of Ten Cities: Characterizing Signatures of Mobile Traffic in Urban Areas,” *IEEE Transactions on Mobile Computing*, vol. 16, no. 10, pp. 2682–2696, 2017.
- [54] S. Hochreiter *et al.*, “Long Short-Term Memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [55] A. Vaswani *et al.*, “Attention is All You Need,” *NeurIPS*, 2017.
- [56] S. Bai *et al.*, “An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling,” *arXiv preprint arXiv:1803.01271*, 2018.
- [57] S.-A. Chen *et al.*, “TSMixer: An All-MLP Architecture for Time Series Forecasting,” *arXiv preprint arXiv:2303.06053*, 2023.
- [58] A. Zeng *et al.*, “Are Transformers Effective for Time Series Forecasting?” in *AAAI*, 2023.
- [59] J. Herzen *et al.*, “Darts: User-Friendly Modern Machine Learning for Time Series,” *Journal of Machine Learning Research*, vol. 23, no. 124, pp. 1–6, 2022.
- [60] S. Liu *et al.*, “LEAF: Navigating Concept Drift in Cellular Networks,” *Proc. ACM Netw.*, vol. 1, no. CoNEXT2, Sep. 2023.
- [61] A. P. Jagadeesan *et al.*, “First Steps in Concept Drift Management for Resilient Wireless Networks,” in *European Wireless*, 2024.
- [62] Y. Zeng *et al.*, “Online Scheduling of Edge Multiple-Model Inference with DAG Structure and Retraining,” in *IEEE INFOCOM*, 2025.