In-depth Study of RNTI Management in Mobile Networks: Allocation Strategies and Implications on Data Trace Analysis

Giulia Attanasio, Claudio Fiandrino, Marco Fiore, Joerg Widmer, Norbert Ludant, Bastian Bloessl, Konstantinos Kousias, Özgü Alay, Lise Jacquot, Razvan Stanica

In-depth Study of RNTI Management in Mobile Networks: Allocation Strategies and Implications on Data Trace Analysis

Giulia Attanasio[a], Claudio Fiandrino[a], Marco Fiore[a], Joerg Widmer[a], Norbert Ludant[b], Bastian Bloessl[c], Konstantinos Kousias[d], Özgü Alay[e], Lise Jacquot[f], Razvan Stanica[f]

[a] *IMDEA Networks Institute, Madrid, Spain*
[b] *Khoury College of Computer Sciences, Northeastern University, Boston, MA, USA*
[c] *Technischen Universität Darmstadt, Darmstadt, Germany*
[d] *Simula Research Laboratory, Oslo, Norway*
[e] *University of Oslo, Norway*
[f] *Univ Lyon, INSA Lyon, Inria, CITI, Villeurbanne, France*

Abstract

The advance of mobile network technologies and components heavily relies on data-driven techniques. This is especially true for fifth generation (5G) and the upcoming sixth generation (6G) networks, as the optimization of network components and protocols is expected to be fueled by artificial intelligence (AI) based solutions. When using real-world radio access measurement traces, the identity of individual users is not directly accessible because at runtime operation Base Stations (BSs) assign Radio Network Temporary Identifiers (RNTIs) to users. RNTIs are not bound to a user but are reused upon expiration of an inactivity timer, whose duration is operator dependent. This implies that, over time, multiple users are mapped to the same RNTI. In fact, the allocation of RNTIs to users is implemented in diverse and proprietary ways by operators and equipment vendors. Distinguishing individual users within the RNTI space is a non-trivial task and key to analyze traffic traces properly. In this paper, we make the following contributions: *i)* we propose and validate two complementary methodologies to identify the RNTI inactivity threshold, and we characterize *ii)* the RNTI allocation process of network operators, and *iii)* the user traffic patterns given the specific RNTI allocation process. Our study is based on a large dataset we collected from production BSs of several mobile network operators across five different countries. We find that there exist heterogeneous strategies for RNTI allocation that BSs dynamically use depending on the traffic load and daytime. We further observe that the RNTI expiration threshold is in the order of minutes, and demonstrate how using thresholds around 10 seconds, as in the vast majority of the literature, can bias subsequent analyses. Overall, our work provides an important step towards dependable mobile network trace analysis, and lays solid foundations to research relying on

Email addresses: giulia.attanasio@imdea.org (Giulia Attanasio), claudio.fiandrino@imdea.org (Claudio Fiandrino), marco.fiore@imdea.org (Marco Fiore), joerg.widmer@imdea.org (Joerg Widmer), ludant.n@northeastern.edu (Norbert Ludant), bbloessl@seemoo.tu-darmstadt.de (Bastian Bloessl), kostas@simula.no (Konstantinos Kousias), ozgua@ifi.uio.no (Özgü Alay), lise.jacquot@insa-lyon.fr (Lise Jacquot), razvan.stanica@insa-lyon.fr (Razvan Stanica)

traffic traces for data-driven analysis.

Keywords: Mobile networks, network measurements, RNTI expiration threshold, RNTI allocation.

1. Introduction

The widespread availability of fifth generation (5G) networks is nowadays a reality, and the identification of the advanced features that will shape the evolution of 5G into the sixth generation (6G) of mobile network systems has already started [1]. 5G and 6G networks in particular are expected to support a diverse set of services, each more demanding than those of previous generation networks (*e.g.*, holographic communications, unmanned mobility, etc.), and to serve unprecedented traffic demands. According to the Ericsson mobility report [2], during the first quarter of 2022 the number of 5G subscriptions increased by 70 million, reaching 620 million overall, and is projected to surpass the 1 billion barrier by end of this year. Flexibility is key for 5G and 6G mobile networks to quickly adapt to changes in the network state and traffic [3]. However, it is very hard to correctly model the behavior of these networks that are becoming increasingly complex, heterogeneous, dynamic and dense [4]. Therefore, there are strong expectations that model-free solutions based on artificial intelligence (AI) will be a core concept to shape the architectural evolution from 5G to 6G networks.

Data-driven approaches have become more and more popular as techniques to advance mobile network technologies and components. Relevant examples encompass deployment planning [5] and automatic parameter configuration of newly deployed base stations [6], and optimizations at all the layers of the mobile network protocol stack, such as resource allocation in cloud Radio Access Networks (RAN) [7] and resource scaling of 5G core virtualized network functions [8], network slicing [9], energy savings with intelligent BS sleeping strategies [10], improved mobility management [11] and self-configuration of handover parameters [12], down to the physical layer to optimize beam management of radios operating at millimeter-wave frequencies [13]. The characterization of the user behavior is a key aspect to properly analyze traffic traces that are used in data-driven network optimization.

In mobile networks, the operator identifies both the users and their equipment [14]. The International Mobile Equipment Identity (IMEI) and the International Mobile Subscriber Identity (IMSI) are permanent and identify respectively the device and the mobile subscriber. Since these identifiers are permanent, their knowledge by an attacker would allow, among others, user tracking. Therefore they are considered sensitive and are transmitted, whenever possible, on encrypted channels (there are some exceptions where the IMSI is transmitted in plain text in pre-5G networks). To protect the IMSI, a Temporary Mobile Subscriber Identity (TMSI) is allocated to the user and changed frequently (every few hours). However, IMSI/TMSI and IMEI are core network identifiers and they are not used by the Base Station (BS) (*i.e.*, the gNB in 5G NR or the eNB in LTE). This is because core network identifiers

need to be unique in the entire mobile network domain, while the BS only needs identifiers to distinguish between User Equipments (UEs)[1] within the radio cell. Therefore, in the RAN, shorter identifiers known as Radio Network Temporary Identifiers (RNTI) are used. RNTIs themselves do not reveal the user identity and expire after the user is inactive for a certain period of time. Nevertheless, several works have shown that RNTIs can be mapped to other identifiers like TMSI, allowing an adversary to track and localize users within a cell [15, 16], and even to eavesdrop and decipher a user's signaling and data traffic [17]. Most of the protocol vulnerabilities of 5G NR and LTE assume knowledge of the RNTIs [18]. In practice, a timer monitors the user inactivity: when it exceeds a given threshold, the user-RNTI association is released and the same RNTI can be re-assigned to another user. The 3GPP standards define the process, but leave the specific implementation of the timer, threshold, and RNTI value selection to the operators. This makes identifying single users from RNTI sequences in measurement data a complex task. At the same time, the incorrect identification of user sessions has clear consequences on traffic analysis: counting multiple users as one, or vice versa, results in incorrect estimates of traffic distribution, user activity, and resource allocation, which affects modeling and simulations and potentially biases the conclusions. Traffic traces containing RNTI information have been proven useful to identify radio resource utilization of mobile traffic patterns [19, 20], to fingerprint applications [21, 22] and possibly reveal the user identity [23], to proactively identify user behavior for resource optimization [24] and to perform classification of Downlink Control Information (DCI) messages [25]. While previous studies on mobile data traffic were only able to differentiate traffic profiles according to time usage patterns [26, 27], the use of BS traffic traces allows more sophisticated studies, for example to identify application classes [19] and standalone applications [21, 22].

In this paper, we extend our previous preliminary work [28] and provide an in-depth study of the RNTI allocation process. We use the well-known FALCON [29] LTE passive monitoring tool, and collect traffic traces from 20 different BSs of 12 major network operators across the world (see Fig. 1 for an example of the measurement setup). Overall, our dataset contains more than 6 months of measurement data throughout 2020-21. Our analysis of the collected traces sheds light on a number of aspects related to RNTI allocation and management, including insights on the actual values of inactivity thresholds and on the allocation schemes adopted by different operators. We demonstrate that different RNTI allocation schemes may be used at a BS, and investigate the correlation between the user behavior and the employed RNTI allocation scheme. Overall, our results help to better understand and analyze BS measurement data, and pave the way for more dependable research on data-driven modeling and simulations of 5G/6G mobile network components.

The major contributions of this work are as follows.

[1] In this paper, we use the terms UE and *user* interchangeably.

- We propose a new methodology for the estimation of the RNTI expiration threshold. This methodology complements the one originally proposed in [28]. The two methodologies allow to correctly set and validate the RNTI inactivity threshold for trace analysis and to properly identify the life time of user connections to a BS. We verify that these methodologies hold for operators in five different countries (Spain, USA, Germany, Norway and France). In contrast to past research that commonly assumes inactivity thresholds of ∼10 seconds, our analysis shows that actual values for the threshold range between 30 seconds and 10 minutes in the monitored BSs.

- We characterize the RNTI allocation process adopted by multiple network operators across different countries, and find that different operators implement different RNTI allocation processes. We observe that the adopted allocation scheme may change over time at the same BS or according to the environment, *i.e.*, urban and rural areas.

- We characterize user traffic patterns given the specific RNTI allocation process. We find a correlation between the given RNTI allocation scheme and frame occupancy in terms of Physical Resource Block (PRB) usage.

- For all the monitored mobile network operators, we show that an incorrect setting of the expiration threshold affects the distributions of extracted users and per-user load and, for the scope of trace analyses, these errors may bias the conclusions that would be drawn from such traces.

- We intend to release the functional artifacts of our study upon acceptance.

The rest of the paper is organized as follows. Sec. 2 introduces the reader to user identification in mobile networks. Sec. 3 presents our real-world LTE traffic dataset and Sec. 4 delves into the RNTI analysis. Sec. 5 highlights specific behaviors and Sec. 6 evaluates the impact of incorrect inactivity thresholds on user life time and per-user load (we report detailed results in Appendix A). Sec. 7 outlines the research areas that may benefit from our results and provides further discussion. Finally, Sec. 8 concludes the work.

2. Temporary User Identifiers in Mobile Networks

To obtain service from the network, a mobile phone that is just turned-on has to perform an Initial Access (IA) procedure. IA consists of three phases: cell search, system information extraction and random access (RA) procedure. Once these phases are completed, the user switches from IDLE to CONNECTED mode and can communicate with the BS over scheduled channels from that moment on. Specifically, after the random access procedure is completed, the UE is mapped to a Cell Radio Network Temporary Identifier value (C-RNTI). If the user moves between BSs, *i.e.*, during handovers,

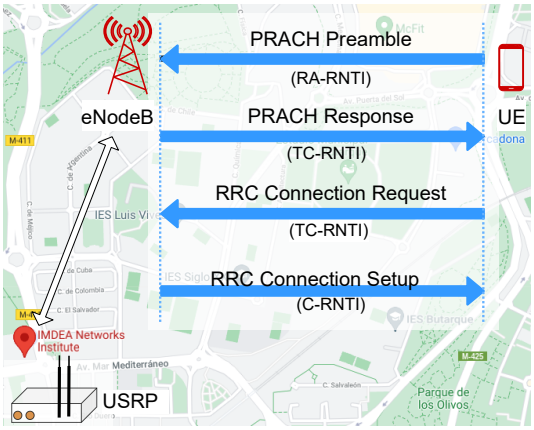


Figure 1: Example of the suburban area measurement setup and RNTI allocation procedure during RA

or if the user remains inactive within the same cell for a certain time (note that this parameter is not explicitly defined by the 3GPP standards), a new RA procedure is required to re-establish the CONNECTED mode.

Fig. 1 outlines the RA procedure through which a C-RNTI is assigned to an incoming user, going through several intermediary RNTIs in the process. The UE sends a Physical Random Access Channel (PRACH) preamble signal to the BS to initiate the RA. The preamble is selected from a subset of predefined waveforms. A so-called RA-RNTI is computed by both the UE and the BS based on the selected waveform and the random access slot used by the UE for the preamble transmission.

The BS replies to the PRACH Preamble by sending a PRACH Response. This response contains the Temporary C-RNTI (TC-RNTI) that will be used by the UE for the RRC connection request. If the connection is successfully set up, the TC-RNTI is promoted to C-RNTI. This value will be then used to uniquely identify UE traffic within the cell. In the rest of this paper, we will refer to C-RNTI simply as RNTI.

Chen et al. [30] illustrate the details of the procedures to obtain RNTIs in both LTE and 5G NR. The main difference in the procedures for the two technologies is that the space of the control region occupies the whole channel bandwidth in LTE, but not in 5G NR. To support bandwidth-intensive applications, 5G NR extends the maximum channel bandwidth from the 20 MHz of LTE to 100 MHz (in low and mid bands, known as FR1 – at sub-6 GHz) and 400 MHz (at high bands, known as FR2 – at millimeter wave frequencies, *i.e.*, above 24 GHz). The continuous scan of all such bands would be power-costly for modems, therefore 5G NR features the so-called *bandwidth parts* (BWP), *i.e.*, smaller portions of the entire channel bandwidth that are flexibly assigned to the users according to application requirements. Using the same procedure of LTE for acquiring C-RNTI would require scanning the whole bandwidth and thus would be costly. For the sake of efficiency, the random access procedure and the distribution of TC-RNTIs is configured to be completed only within the initial BWP. Chen et

Table 1: RNTI Assignment

LTE		5G NR	
Value (Hexa-Decimal)	Type of RNTI	Value (Hexa-Decimal)	Type of RNTI
0000	N/A	0000	N/A
0001-0960	RA-RNTI, C-RNTI, Semi-Persistent Scheduling C-RNTI, TC-RNTI, eIMTA-RNTI, TPC-PUCCH-RNTI, TPC-PUSCH-RNTI, SL-RNTI, G-RNTI, SL-V-RNTI, UL Semi-Persistent Scheduling V-RNTI, SL Semi-Persistent Scheduling V-RNTI, SRS-TPC-RNTI, and AUL C-RNTI	0001-FFEF	RA-RNTI, TC-RNTI, C-RNTI, MCS-C-RNTI, CS-RNTI, TPC-PUCCH-RNTI, TPC-PUSCH-RNTI, TPC-SRS-RNTI, INT-RNTI, SFI-RNTI, and SP-CSI-RNTI
0961-FFF3	C-RNTI, Semi-Persistent Scheduling C-RNTI, eIMTA-RNTI, TC-RNTI, TPC-PUCCH-RNTI, TPC-PUSCH-RNTI, SL RNTI, G-RNTI, SL-V-RNTI, UL Semi-Persistent Scheduling V-RNTI, SL Semi-Persistent Scheduling V-RNTI, SRS-TPC-RNTI and AUL C-RNTI	FFF0-FFFD	Reserved for future use
FFF4-FFF8	Reserved for future use		
FFF9	SI-RNTI		
FFFA	SC-N-RNTI		
FFFB	SC-RNTI		
FFFC	CC-RNTI		
FFFD	M-RNTI		
FFFE	P-RNTI	FFFE	P-RNTI
FFFF	SI-RNTI	FFFE	P-RNTI

al. [30] observe that monitoring the PRACH channel looking for successful random access is a proxy to obtain in a computationally efficient manner those TC-RNTI that evolve into C-RNTI.

The 3GPP standards [31] define a limited pool of available RNTIs that ranges from 0x1 to 0xFFFF. A subset of them is not allocated to any UE explicitly, but reserved for specific messages such as paging (P-RNTI), system information (SI-RNTI) and power control purposes (TPC-RNTI). Tab. 1 provides the complete list of LTE RNTI values. For 5G NR, 3GPP specifications [32] introduce new types of RNTI to support novel uses such as the SFI-RNTI and the MCS-C-RNTI (see Tab. 1 for a complete list). Specifically, the first one is generally assigned to a group of users and used for the notification of slot format information. The second enables for both downlink and uplink the usage of an alternative Modulation and Coding Scheme (MCS) table for the scheduling of packets with high reliability. Our analysis covers LTE. As we are interested in the RNTI associated to actual user data transmissions, we consider RNTIs in the range [0xB - 0xFFF3] and the RA-RNTI associated to the PRACH channel, *i.e.*, from 0x1 to 0xA.

Once an RNTI is assigned to a UE, it is maintained until the user remains inactive for a certain amount of time. In that case, the corresponding RNTI value is released and it can be reassigned to a new user. Thus, if a UE whose RNTI was released wants to reconnect to the BS, it needs to perform a new access procedure to obtain a new RNTI. Existing works that focused on RAN metrology used short inactivity timers to estimate the expiration of an RNTI. Specifically, we found that the inactivity timer is set to 10 seconds [33], 10.15 seconds [34], 10.24 seconds [35] and 11.57 seconds [36] in previous studies. These adopted values are close to the System Frame Number cycle. This counter, incrementally

assigned to each LTE frame, repeats itself after 1024 frames, *i.e.*, 10.24 seconds. Another work [37] observed that an LTE smartphone connected to a major US operator maintained the same RNTI for over 4 hours but does not mention whether the smartphone was in fact exchanging traffic with the BS or not. A parallel field of study is the reverse engineering of the RRC state machine, where previous works inferred values for the inactivity timers in the order of tens of seconds [38, 39, 40].

The few studies that investigated RNTI allocation schemes mainly focused on the blind decoding mechanism performed by the UE at the receiver side. In order to decode the DCI message transmitted through the PDCCH, the UE will use its RNTI to descramble the Cyclic Redundancy Check part of the radio channel messages over a set of possible control channel elements. Within such set, the PDCCH carries information about the resource allocation for each user. In order to increase the control channel element utilization, several allocation schemes have been proposed, among them a table-based scheme in [41] and preamble power estimation based approach in [42].

No previous work analyzes the RNTI allocation process adopted by different network operators. Also, no characterization of traffic patterns according to the used allocation scheme is presently available. Our study fills these important gaps.

3. Measurement Dataset

This section presents the data collection methodology, the dataset of raw DCI information from different BSs in all the countries and operators and the pre-processing required for our study. For our analysis, we consider real LTE traffic information gathered from the PDCCH through FALCON [29]. Our initial dataset was collected for over 6 months from 8 different BSs of 3 major Spanish operators. During this period we monitored up to two different BSs at the same time. All the considered BSs are located on different sites and measurements were collected in the city center, in the suburban area and in the inner city. We extended the initial dataset with targeted measurements in several other countries using the same methodology for collection and analysis. We remark that one of the operators is present in both Spain and Germany, while it also partially owned and strongly collaborates with one of the operators in France. Tab. 2 summarizes the characteristics of our complete dataset.

The gathered PDCCH data contains per user scheduling information in both uplink and downlink directions with a resolution of 1 ms. Within each frame and transmission time interval, FALCON reports per-user resource allocation in terms of PRB, Transport Block Size (TBS) and MCS. The collected raw traces from each monitored BS are then grouped per unique registered RNTI. For our study, we collect traffic traces running FALCON with two setups, an Intel Core CPU at 3.6 GHz connected via 10 Gigabit Ethernet to an X310 USRP and an Intel Core CPU at 3.4 GHz connected via USB 3.0 to a B210 USRP (located at the USRP site in Fig. 1). Both PCs run Linux Ubuntu 20.04. We

Table 2: The dataset collected in the countries of the measurement campaign. We identify each country with its ISO 3166-1 alpha-2 code

Country	BS	Operator	Location	Collected Time
ES	1	Operator A	inner city	2 months
ES	2	Operator A	inner city	2 months
ES	3	Operator A	suburban area	1 month
ES	4	Operator B	inner city	2 months
ES	5	Operator B	inner city	2 months
ES	6	Operator B	suburban area	1 month
ES	7	Operator C	suburban area	1 month
ES	8	Operator C	city center	1 month
DE	9	Operator D	city center	6 days
DE	10	Operator D	rural scenario	6 days
DE	11	Operator E	city center	6 days
DE	12	Operator E	rural scenario	6 days
DE	13	Operator F	city center	6 days
DE	14	Operator F	rural scenario	6 days
NO	15	Operator G	city center	2 days
NO	16	Operator H	city center	2 days
FR	17	Operator I	city center	2 days
FR	18	Operator J	city center	2 days
US	19	Operator K	city center	4 days
US	20	Operator L	city center	4 days

do not observe any difference in terms of DCI decoding success rate with the two setups.

Since we are only interested in user traffic, we discard RNTI values that are not UE-related. Specifically, we remove the ones associated to paging and system information (0xFFF4 to 0xFFFF), and only include the RNTI values associated to actual UE transmissions, *i.e.*, from 0xB to 0xFFF3 and to the PRACH channel, *i.e.*, from 0x1 to 0xA in frequency division duplex (FDD), see Tab. 1.

4. RNTI Analysis

Our RNTI analysis has two objectives. First, we investigate how RNTIs are allocated and reassigned to identify the RNTI timer expiration thresholds. Second, we characterize the RNTI behavior across different BSs and operators to analyze how the RNTI allocation varies over daily and weekly cycles.

4.1. Detecting RNTI Expiration Thresholds

The data traffic associated to an RNTI is generated by different UEs that take the same identifier in different time periods (see Sec. 1). Since the set of available RNTI is limited, the BS reassigns identifiers that have been inactive for a certain period to new users. We denote as *user life time* the time elapsed between the first and the last transmission of a given user. At the time of the user life time end, the RNTI expiration timer starts and upon expiration, the BS reassigns the RNTI. Our goal is to identify which is the timer duration. The traces collected provide as information only the times of RNTI activity, hence we need to estimate the timer. For this, we derive two methodologies that are hereafter explained. The first one (Sec. 4.1.1, [28]) builds upon the observation that the period of inactivity should be set so that it is larger than the what is commonly observed as silent periods between subsequent user data transmissions. The second methodology is new and builds upon the observation that the identification of successful RA procedure allows to map RA-RNTI to C-RNTI (Sec. 4.1.2). The joint use of the two methodologies in pipeline allows to increase the confidence of the estimation.

4.1.1. Inter-transmission time analysis

We define the inter-transmission time as the time elapsed among two consecutive transmissions for the same RNTI. Then, if the inter-transmission time is smaller than the RNTI expiration threshold T, we consider that the consecutive transmissions belong to the same UE; if it is instead greater than T, we consider the transmissions to be from different UEs, as shown in Fig. 2.

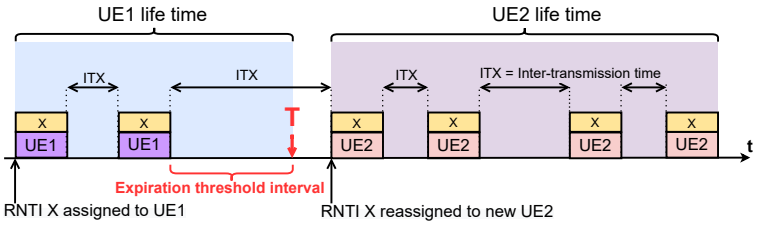
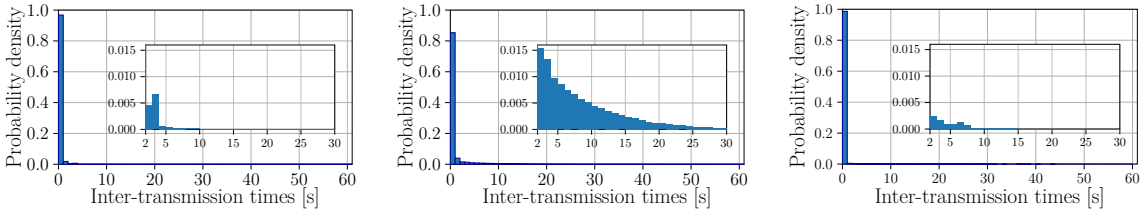


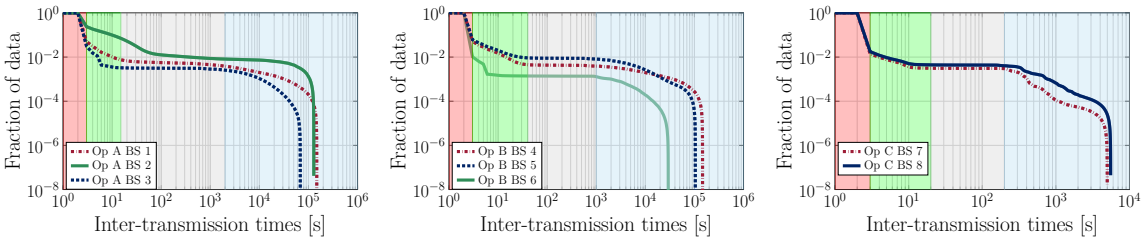
Figure 2: Diagram illustrating inter-transmission times (ITX) for the same RNTI (X). If the inter-transmission time exceeds the expiration threshold T, the RNTI can be reassigned to a new user.

In order to verify if users associated to the same RNTI can be correctly discriminated through the commonly adopted expiration threshold of 10 seconds, we plot the probability density function (PDF) of the inter-transmission times for each registered RNTI. Fig. 3 shows the results for all the monitored operators and BSs. For simplicity, we report only downlink traffic given that is predominant in mobile networks, but our approach is also valid for uplink traffic. In fact, we confirm that the fraction of uplink traffic is tiny with respect to the downlink traffic. The results show inter-transmission times from 0 to 60 seconds and, for each plot, we show an inset in the range from 2 to 30 seconds.



(a) Operator A. (b) Operator B. (c) Operator C.

Figure 3: PDF of inter-transmission times for different operators for values smaller than 60 seconds



(a) Operator A BSs. (b) Operator B BSs. (c) Operator C BSs.

Figure 4: Survival function of the inter-transmission times with a rapid decay phase (red), a smoothly decreasing stage (green), the leveling out period (gray) and distribution tail (blue). Each phase is visually defined from the survival function

Our aim is to identify the proper threshold T that determines the end of the user life time. As soon as the user remains inactive for a period of time that exceeds T, its RNTI becomes available for reassignment. Assume that within its life time, a user transmits $k \in \mathbb{N}$ times. For this RNTI reassignment, we expect to see $k-1$ inter-transmission time values smaller than T and one inter-transmission time value that exceeds T. Considering all active users, the probability of observing inter-transmission times greater than T (which requires that a new user arrives at that time and happens to be assigned the same RNTI) should be much smaller than the probability of observing values smaller than T (where the same user is still active). Thus, we expect that the number of registered occurrences to drop sharply for values larger than T.

Fig. 3 shows that the number of occurrences within the interval 2-30 seconds is two orders of magnitude smaller with respect to the one registered within the interval 0-1 seconds. However, no significant drop can be observed in the interval 10-11 seconds, where according to existing literature, the threshold should be set. Therefore, according to the results presented in Fig. 3, there is no indication that the commonly adopted threshold of 10 seconds is correct. In fact, an indication of the correct threshold setting can be derived from a significant variation of the inter-transmission time values. A change in behavior can be appreciated by looking at the survival function (complementary CDF) of consecutive RNTI transmissions. We evaluate the inter-transmission times considering one week of continuous traffic. Fig. 4 shows the results for each of the monitored BSs.

A suitable threshold for all BSs and operators falls within the range from 30 to 100 seconds. Indeed,

Table 3: Comparison for the different countries. We report the start inter-transmission time for each of the phases

Operator	Country	RD	SD	LO	TA
Operator D	DE	1-3 s	3-40 s	40-2,020 s	2,020-74,780 s
Operator E	DE	1-3 s	3-40 s	40-1,200 s	1,200-8,488 s
Operator F	DE	1-3 s	3-40 s	40-160 s	160-15,330 s
Operator G	NO	1-3 s	3-70 s	70-1,657 s	1,657-52,620 s
Operator H	NO	1-3 s	3-60 s	60-995 s	995-60,890 s
Operator I	FR	1-3 s	3-50 s	50-1,000 s	1,000-17,420 s
Operator J	FR	1-3 s	3-40 s	40-200 s	200-30,770 s
Operator K	US	1-3 s	3-30 s	30-409 s	409-21,880 s
Operator L	US	1-3 s	3-30 s	30-1,079 s	1,079-29,350 s

from Fig. 4(b), we observe four main stages. First, there is a rapid decay phase (RD - denoted with a red background) that ranges from 1 to 3 seconds for all the considered BSs. Most inter-transmission times are in this range, *i.e.*, the majority of consecutive user transmissions are within 3 seconds. The rapid decay is followed by a smoothly decreasing stage (SD - denoted with a green background) that spans over a larger inter-transmission time interval. After this stage, we can identify a leveling out period (LO - denoted with a gray background), followed by the tail of the distribution (TA - denoted with a blue background) for which the number of observed values drops significantly. For brevity, Tab. 3 shows the timings of all the phases for all the operators of the countries in the dataset.

For the threshold setting we are interested in the change of behavior that occurs between the second and the third phase, *i.e.*, the smoothly decreasing stage and the leveling out period. In fact, from the beginning of the leveling out period up to the tail of the distribution few inter-transmission times values can be observed as the survival function remains almost constant. Hence, smaller inter-transmission times within the rapid decay and the smoothly decreasing stage identify sessions of the same user. Thus, the most likely threshold setting is right after the knee point between the smoothly decreasing stage and the leveling out period. We do not consider the tail of the distribution since the number of registered observations is very sparse.

Furthermore, the significant change that occurs between the rapid decay and the smoothly decreasing stage would not represent a reasonable interval for the threshold setting. In this case the threshold would be 3 seconds. However, it would not make sense for the BS to set such a small value since also data continuity should be guaranteed to connected users. In fact, such a small value would trigger too many re-establishment procedures leading the system to collapse, a phenomenon known as signalling storm [43]. Thus, we set the expiration threshold to 60 seconds for BS 1, 2, 4, 5, and to 100 seconds for

BS 3, 6, 7 and 8. This result is in line with the studies that focus on the analysis of RRC state transitions [38, 39, 40]. We adjust the estimated expiration threshold for all the considered BSs, according to the methodology described in Sec 4.4.

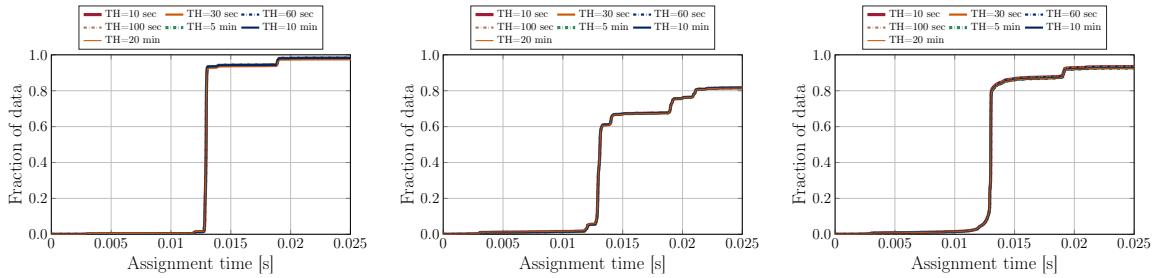
4.1.2. RA-RNTI to C-RNTI mapping

According to the procedure described in Sec. 2 and summarized in Fig. 1, a new incoming user initiates the random access procedure by computing the RA-RNTI based on the selected waveform and the random access slot used for the preamble transmission. If the connection is successfully set up, a C-RNTI is then assigned to the new incoming user. Since FALCON allows to decode both RA-RNTI and C-RNTI, we aim at identifying, within the collected traffic traces, the RA-RNTI and C-RNTI value pair associated to each new successful random access procedure. In order to identify RA-RNTI to C-RNTI pairs within traffic traces, we consider different possible expiration thresholds values *i.e.*, 10 sec, 30 sec, 60 sec, 100 sec, 5 min, 10 min and 20 min. By setting the expiration threshold, we filter out C-RNTI transmissions of the same user while preserving allocations of new incoming ones. This procedure allows to associate with higher probability RA-RNTI to new C-RNTI transmissions only.

Thus, for each of the identified pairs of RA-RNTI and C-RNTI, we measure the assignment time that represents the elapsed time between the RA-RNTI and the associated C-RNTI. Finally, we set the assignment time to the value that in the CDF corresponds to the inflection point of the sigmoid, *e.g.*, 13 ms for operators A and B and 12 ms for operators C, see Fig. 5. The estimation of the assignment time is next used to further refine the number of identified RA-RNTI to C-RNTI pairs. If within traffic traces, consecutive RA-RNTI transmissions are registered, we map RA-RNTIs to C-RNTIs transmissions that are within the assignment time interval. A correct setting of the expiration threshold would maximize the number of RA-RNTIs to C-RNTI pairs and minimize the number of unallocated RA-RNTI. Fig. 6 presents the results. For operators A, B and C we show both the number of RA-RNTI to C-RNTI pairs mapping (blue) and the unallocated RA-RNTI (orange) according to different thresholds values. A reasonable value for the threshold should be set where there exists a good trade-off between the number of identified RA-RNTI to C-RNTI pairs and the number of unallocated RA-RNTI. For both operators A and B a reasonable threshold value ranges between 60 sec to 10 min while for Operator C should range between 30 to 100 sec. This is in line with values obtained from the inter-transmission time analysis.

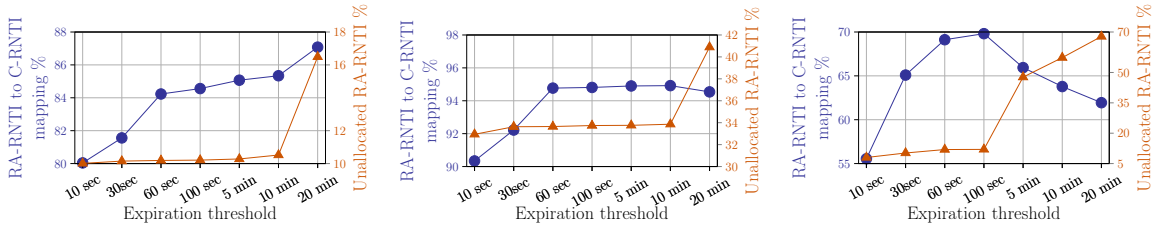
Thus, we propose to properly identify the setting of the expiration threshold T in two steps with the two methodologies. First, we identify the knee point between the SD and the LO period in the survival function of the inter-transmission times as a possible candidate for T. Next, we validate our choice with the RA-RNTI to C-RNTI mapping analysis.

Takeaway message. The threshold for the RNTI expiration should be set to much higher values than



(a) Operator A. (b) Operator B. (c) Operator C.

Figure 5: RA-RNTI to C-RNTI assignement time for different operators



(a) Operator A. (b) Operator B. (c) Operator C.

Figure 6: RA-RNTI to C-RNTI mapping analysis for different operators

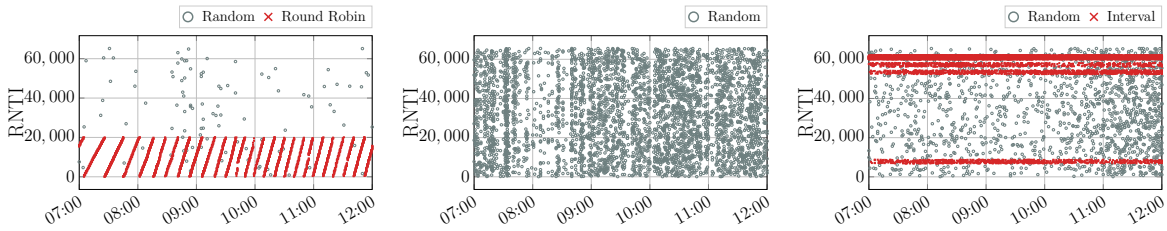
those around 10 seconds assumed in the literature. Based on our measurements, a reasonable threshold value matches the start of the leveling out period in the survival function of the inter-transmission times analysis (see Fig. 4 and Tab. 3). For the RA-RNTI to C-RNTI mapping, a reasonable threshold lies in the proximity of high one-to-one mappings and low unallocated RNTIs (see Fig. 6). This leads to threshold values that are in the range of 30-60 seconds to 10 minutes for the observed BSs.

4.2. RNTI Allocation Schemes

We now investigate how the monitored BSs assign RNTIs to the users. For this purpose, we set the inactivity expiration threshold to 60 s for BS 1, 2, 4 and 5 and to 30 s for BS 3 and 6 according to the results presented in Sec. 4.1. By considering the same procedure we set the inactivity expiration threshold to 50 seconds for Operators K, L, to 60 seconds for Operators D, E, F and J, to 70 seconds for Operator I, to 80 seconds for Operator H and to 90 seconds for Operator G. Thus, we plot for each RNTI, the inter-transmission times that are greater than the selected inactivity expiration threshold. In this way, each RNTI transmission will actually represent a new incoming user. From this analysis we visually observe that operators allocate RNTIs according to different strategies.

Fig. 7(a) presents, over a period of 24 hours, the allocation schemes adopted by Operator A BS 2. From the plots, it emerges that the BSs from Operator A and B use two RNTI allocation strategies combined, namely random (gray circles) and round robin (red crosses along slopes). Specifically, we found that the round robin scheme always uses RNTIs that are smaller than $32,060$ (hereafter defined as 0-headed, according to the first digit of their binary representation). Instead, the random scheme

can either use 0-headed type of RNTI or RNTIs whose value is greater than 32,060 (hereafter defined as 1-headed). Note that 1-headed RNTIs are always allocated in random fashion. In our dataset, we found that such allocation mechanism is also present for both the operators in the USA (with the very same threshold of 32,060) and for Operator E in Germany (we observe that in this case the threshold is set to 22,060 and not 32,060). It is worth highlighting the special case of Operator H in Norway: for BS operating around 1,500 MHz it operates a round robin and random allocation strategy (using like Spain and USA the threshold of 32,060) while for BS operating around 800 MHz it operates in random-only mode. Fig. 7(b) shows an example of a random-only allocation that is used by Operator F in Germany, by Operator J in France and by Operator G in Norway in addition to the special case observed for Operator H of the same country.



(a) Random (gray) & round robin (red). (b) Random-only (gray). (c) Random (gray) and interval (red).

Figure 7: Allocation strategies of different operators. Figure best viewed in colors

The BSs from Operator C in Spain use instead a random allocation strategy (gray circles) and an interval-based scheme (red crosses along horizontal lines) as show in Fig. 7(c). In particular, the interval-based scheme allocates randomly a subset of RNTI values. However, those values are allocated per hour 15 times more frequently than the other random allocated ones (IDs not belonging to the interval). For both Operator C BSs, we identified three intervals, *i.e.*, low, medium and high according to the range that they cover. With the sole exception of one interval of BS 7, all the other intervals cover a range of 1400 RNTI values. Specifically, the interval that ranges from RNTI 48600 to 50000, is common to both BSs. The other two are instead different: 22000-24000, 3000-4400 for BS 7 and 56200-57600, 60000-61400 for BS 8. This interesting interval RNTI allocation behavior has been observed as well by the authors in [44], where they identify that users with multiple bands assigned for downlink transmission (Carrier Aggregation), share the same RNTI across different bands. Due to this phenomenon, we believe operators use a coordinated interval RNTI allocation scheme across BSs to avoid RNTI collisions across bands when using Carrier Aggregation. Also Operator D in Germany and Operator I in France use a random plus interval-based RNTI allocation strategy, although we observe that the width of the interval bands are different. The BS 10 from Germany uses one interval only that ranges from 60,000 to 62,000 while the BS 18 from France uses four different intervals that range from 6,750-8,750, 52,500-54,500, 56,200-58,200 and from 60,000 to 62,000, respectively.

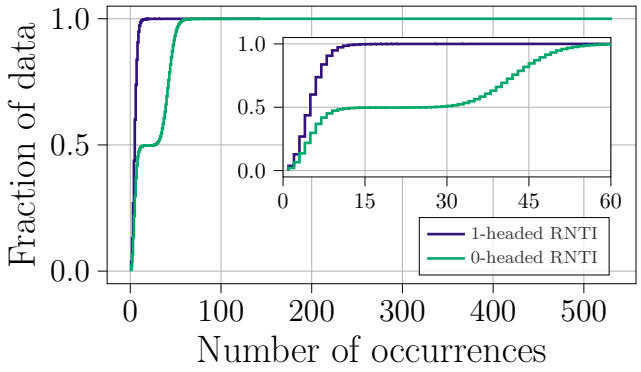


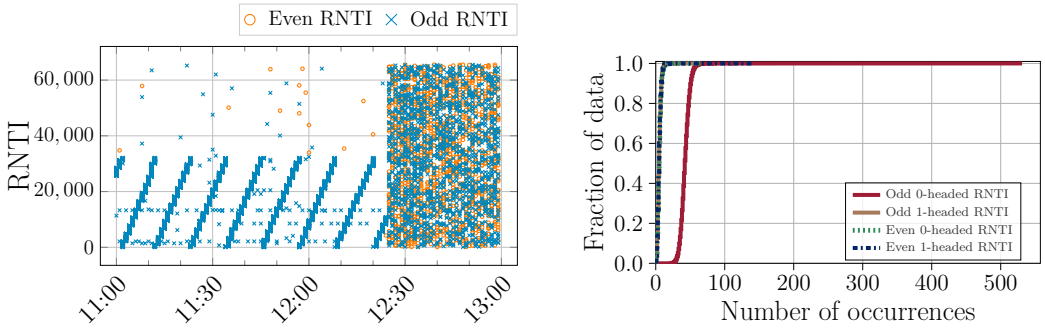
Figure 8: Occurrences of 0-headed and 1-headed RNTIs for Operator A BS 1

Motivated by the fact that operators can allocate RNTIs to incoming users following different strategies, we investigate the behavior of 0-headed and 1-headed RNTIs with respect to the random/round robin allocation scheme. To this end, we evaluate the CDF of the number of occurrences for each registered RNTI over one week. This analysis is performed on the traffic traces collected from Operator A and B of Spain, but the same considerations hold for the operators of the other countries that use random in combination with a round robin strategy (*i.e.*, operators of the USA, Operator E in Germany and Operator H in Norway). We do not perform this analysis for the remaining operators since none of them allocates RNTIs according to a round robin strategy. The results presented for Operator A BS 4 in Fig. 8 are consistent across all BSs from Operator A and B. We omit graphs for the other BSs for the sake of brevity. For both operators, 95 % of the 1-headed RNTIs, *i.e.*, the ones associated to random periods only, are concentrated within the interval that ranges from 1 to 10. From the same figure, the analysis of 0-headed RNTIs reveals the existence of two groups. The first group registers a significant smaller number of occurrences with respect to the second one. Specifically, the first one ranges from 1 to 15, while the second from 30 up to 531. By analyzing the RNTIs of both groups, we discover that the first one only comprises even IDs, while the second only comprises odd IDs values. Motivated by this finding, in the next section we separately characterize odd and even RNTIs.

Takeaway message. Operators adopt different RNTI allocation strategies. Specifically, they either use a random-only strategy, or combine random allocation with either round robin or interval-based schemes.

4.3. RNTI Characterization

We plot the inter-transmission times of odd and even RNTIs. Fig. 9(a) shows the results for Operator A BS 4 over a period of 2 hours. We know that the random allocation strategy can use either 0-headed or 1-headed RNTIs: we find that these can be either be even or odd. By contrast, the round robin mechanism, which we know it only uses 0-headed RNTIs, differs from the random strategy because all RNTIs are exclusively odd. This behavior holds across all monitored Operator A and



(a) Odd RNTIs can be assigned both randomly and in round robin (linear slopes), whereas even RNTIs are only assigned randomly.

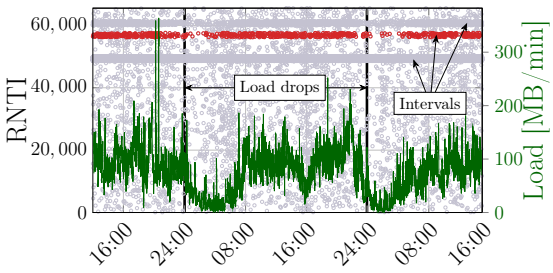
(b) Occurrences of odd and even 0-headed and 1-headed RNTIs.

Figure 9: Odd and even RNTI analysis. Figure best viewed in colors

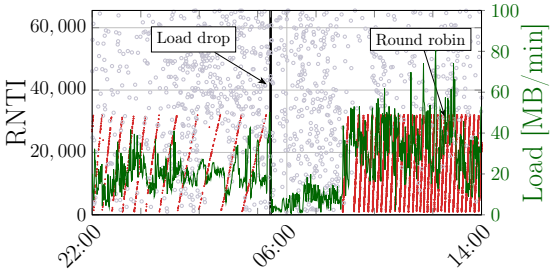
Operator B BSs. Fig. 9(b) illustrates the CDF of the number of occurrences of even/odd 0-headed and 1-headed RNTIs over one week of traffic data from Operator A BS 3. Given that round robin occurs more frequently than random and during high-load periods, it is not surprising that the occurrences of odd 0-headed RNTIs are much higher than odd 1-headed, 0-headed and 1-headed even RNTIs. The highest number of occurrences, *i.e.*, > 100 is registered for spurious 0-headed odd RNTIs (see Sec. 4.4). As stated in Sec. 4.2, Operator C combines a random and an interval scheme allocation. The same analysis reveals that Operator C does not allocate even and odd RNTIs differently within random and interval periods.

We now analyze whether RNTIs assigned according to different strategies (random, round robin, and interval-based) have different traffic characteristics. Furthermore, we investigate the relationship between the traffic load (that exhibits patterns, *e.g.*, day/night) registered at the BS and the RNTI allocation scheme that is used. From the analysis of Operator C traffic traces, we find that the medium interval for both BS is less used during night periods and early mornings, within a range that approximately lasts from 23:00 to 6:00. Fig. 10(a), shows that, as soon as the medium interval (red) is progressively less used, a drop in the BS load is registered too. The first load drop occurs at 24:00: we observe a load decrease of 89 % in the next minute. A second drop is registered around 24:00 of the next day and in this case we observe a load decrease of 85 % in the next minute. In both cases, as soon as the load grows over approximately 80 MB/min (*e.g.*, , around 07:00), RNTIs start to be assigned again to that interval. Our conjecture is that the BS changes the allocation to random for that interval whenever it detects a low level of traffic volume per period. The same considerations hold for Operator A BS 2 as shown in Fig. 10(b). In this case the BS swaps the allocation from round robin to random as soon as the load drops below 1 MB/min (05:20). In the same way, the allocation swaps from random to round robin as soon as the load grows above 25 MB/min (08:18).

Thus, we consider frame occupancy in terms of PRB utilization which is recommended as the

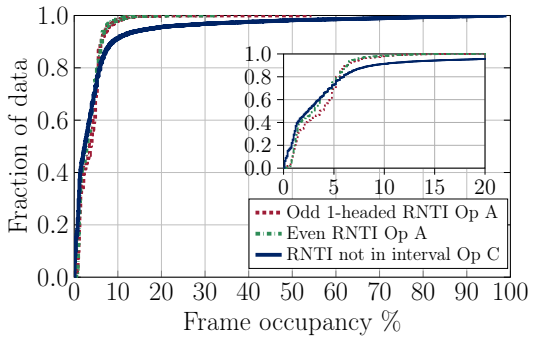


(a) Operator C BS 8 (48 hours). The interval RNTI allocation is highlighted in red.

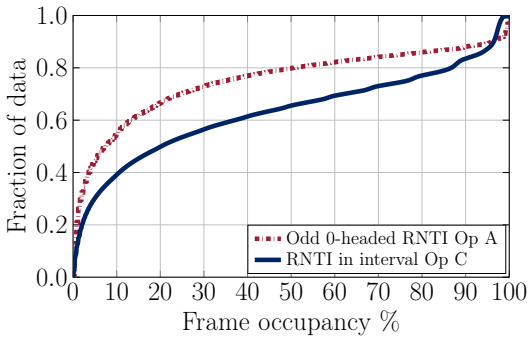


(b) Operator A BS 2 (16 hours). The round robin RNTI allocation is highlighted in red.

Figure 10: Traffic load over time for different RNTI allocation schemes. Figure best viewed in colors



(a) RNTI assigned by random allocation only. Inset over 0-20% interval.



(b) Round robin and interval.

Figure 11: Pairwise comparison of (a) random and (b) round robin vs interval for operators A and C with different RNTI allocation combinations

evaluation criteria of load [45] and depends on the application in use and on the amount of resources available. PRB utilization is defined as the ratio between the number of PRBs used for transmission and the total number of PRB of the BS, in each transmission time interval [46]. By considering one week of continuous traffic, we evaluate the CDF of PRB utilization. Fig. 11 presents the results for Operator A BS 3 and Operator C BS 8.

Fig. 11(a) shows similar results between odd 1-headed RNTIs, even RNTIs and Operator C RNTIs that are allocated randomly. In this case, the 90 % of the IDs registers a frame occupancy smaller than 10 %. Instead, when we look at odd 0-headed RNTIs and to the interval assigned ones, see Fig. 11(b), it emerges that a higher frame occupancy is achieved by both with respect to previous cases. From this result it appears that, when the random and round robin allocation schemes are used by the BS, a higher number of PRB resources is consumed by UEs with odd RNTI. The same result holds for Operator C. In this case, we register a higher frame occupancy for the interval-allocated RNTIs with respect to the ones that are not interval-allocated.

Takeaway message. Our analysis shows that specific ranges of RNTIs are reserved for different

allocation strategies, *i.e.*, random, round robin, or interval-based. In addition, round robin and interval-based allocations are only observed during time periods characterized by high traffic loads. Finally, users whose RNTIs are allocated via round robin or interval-based strategies generate traffic with a higher frame occupancy than that of users with randomly assigned RNTIs. This behavior may suggest the existence of a relationship between the RNTI assigned by the BS and the RA-RNTI used by the UE. Specifically, RACH preambles are divided in two categories, based on the data requirements of the UE. The UE selects the preamble randomly from a class of preambles related to its traffic needs. Thus, the BS may decide to allocate RNTI via round robin, interval or random to meet the traffic requirements of the UE.

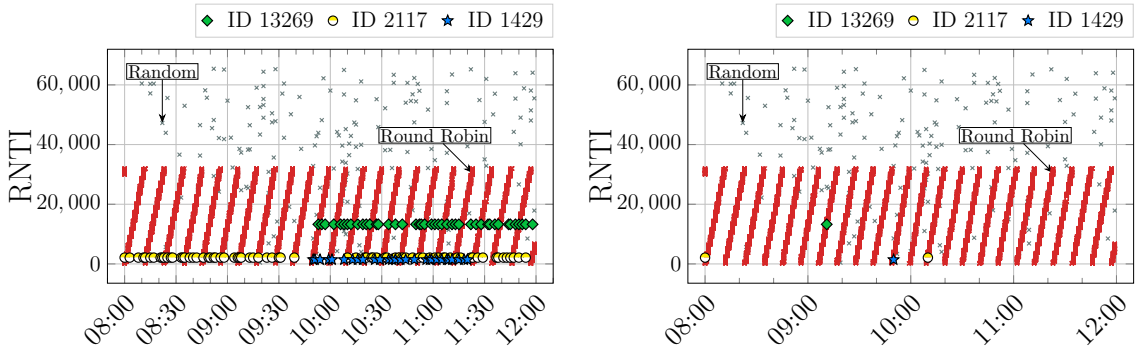
4.4. Threshold Tuning

We now perform an in-depth study on the micro-level behavior of RNTI allocation. This analysis reveals that the threshold T can be tuned further. By analyzing our traces, we notice that some RNTIs are first allocated in round robin fashion and then obtain subsequent allocations over random periods. We define these as spurious cases and Fig. 12(a) exemplifies this behavior for three RNTIs of BS 4 of Operator B. The frequency through which these allocations occur indicate that such transmissions actually belong to the same UE (since the probability of this occurring frequently at random is vanishingly small). Thus, the RNTI is preserved for an inactivity period which is larger than the one that we set in our first analysis, *i.e.*, 60 or 100 seconds according to the monitored BS. We refine the expiration threshold per BS by removing these random spurious RNTI allocations (see Fig. 12(b)). For this set of RNTIs, a real allocation to a new incoming user happens when the next round robin period occurs. For each spurious RNTI from the set, we evaluate the minimum threshold T' that allows to filter out all random RNTI allocations and preserve only the round robin ones. Finally, among all the T' obtained for each spurious RNTI, we take the minimum one. This analysis leads to $T' = 9$ minute.

Takeaway message. Whenever both random and round robin allocation schemes are used by the BS, it is possible to further refine T. We identify spurious RNTIs that are first allocated through round robin and then are reallocated by the random scheme. By eliminating random reallocations and preserving real round robin ones, we observe that the threshold can be extended up to tens of minutes.

5. Special Cases

The analysis of the traces collected in Germany, Norway, France and USA revealed special cases in RNTI management and in BS operation that we discuss hereafter.



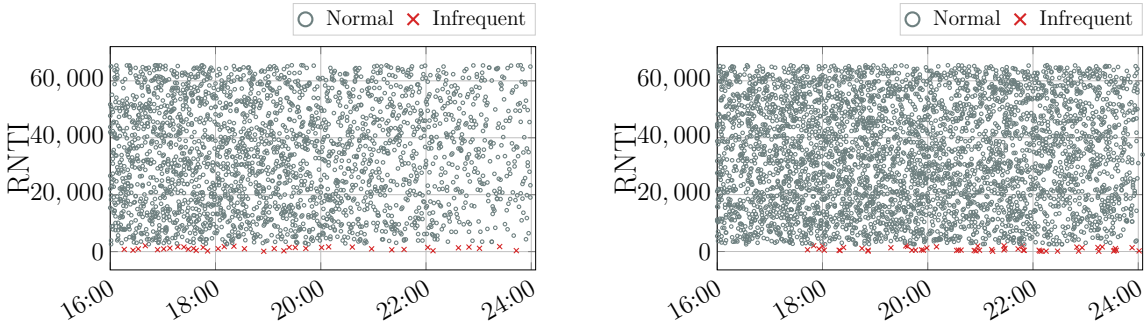
(a) RNTI expiration threshold of 60 seconds. Three spurious cases are highlighted, *i.e.*, RNTIs 2117, 1429, 13269.

(b) RNTI threshold of 9 minute. Spurious random allocations disappear while round robin allocations are preserved.

Figure 12: Threshold tuning procedure. Figure best viewed in colors

5.1. Infrequent Utilization of an Interval of RNTIs

Fig. 13 shows that for two operators, Operator F in Germany and Operator I in France, a whole interval of RNTIs that ranges between 72 and 2,500 is used seldom with respect to the full range of available RNTIs. We compared how often the RNTIs in the range $[72 - 2,500]$ are utilized with respect to RNTIs in other intervals of the same size (*i.e.*, $2,500 - 72$). The comparison is repeated for time periods of 1 hour each during the entire trace. The allocation frequency of RNTIs in the range $[72 - 2,500]$ is on average 12 times lower then RNTIs used in other intervals.



(a) Operator F, BS 13

(b) Operator I, BS 17

Figure 13: RNTI utilization. We highlight in red the range of infrequently used IDs.

5.2. Variation of Round Robin implementation in Norway and USA

Operator H of Norway and Operator K in the USA both share a different round robin implementation than that described in Sec. 4.2. Unlike the classical round robin implementation that has one slope, these operators implement three different round robin processes and each of them is characterized by a different slope of the increase. Fig. 14 highlights the behavior after filtering out the random RNTIs for Operator H of Norway. We can observe that RNTIs within slope 1 are allocated up to 32,060 while

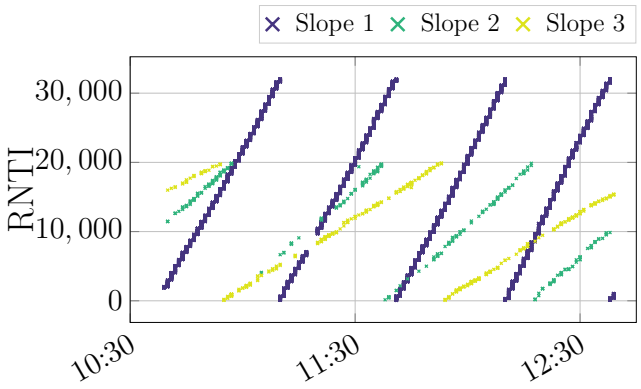
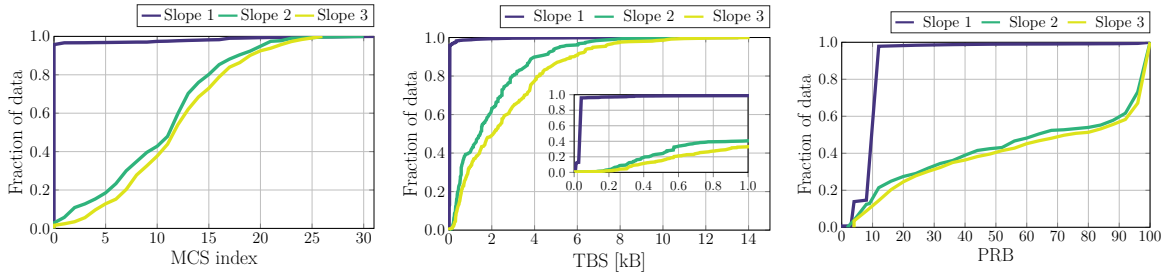


Figure 14: Round robin allocation with three different slopes for Operator H of Norway

RNTIs within slope 2 and 3 are allocated up to 22, 060. We investigate the characteristics of the RNTIs within each slope in terms of MCS, TBS and PRB. This analysis reveals the existence of two different clusters, Fig. 15 presents the results. Specifically, the first cluster consists of RNTIs allocated through slope 1. In this case, 98% of RNTIs show MCS $= 0$, TBS < 0.1 kB and PRB < 12. The second cluster consists of RNTIs allocated through slope 2 and 3 and shows that higher values for MCS, PRB and TBS are used. The characteristics of data transmission suggest that RNTIs within slope 1 most likely carry test/control messages or Narrowband-IoT traffic, while RNTIs within slope 2 and 3 carry actual data traffic.



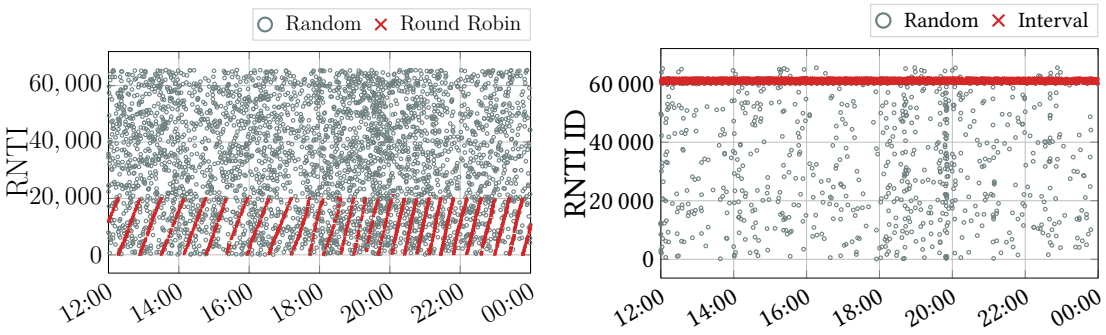
(a) MCS (b) TBS (c) PRB

Figure 15: Characteristics of the traffic carried by RNTIs allocated with the different round robin slopes for Operator H in Norway

5.3. Rural vs. Urban RNTI Allocation

The analysis of the traces of Operator D (Germany) unveiled an interesting change of the RNTI allocation strategy in urban (see Fig. 16(a)) and rural environments (see Fig. 16(b)). The urban measurements were conducted in Darmstadt, while the rural measurements were conducted in the country side, covering a village and sections of a highway and freeway. In the latter case, the BS is configured to assign RNTIs with an interval-based scheme while in the former case a random plus round robin strategy is enforced. The reason for such change of strategy is unclear given that there

are operators (*e.g.*, Operator C in Spain and Operator I in France) that in urban scenarios use an interval-based strategy, albeit with more interval bands than only one.



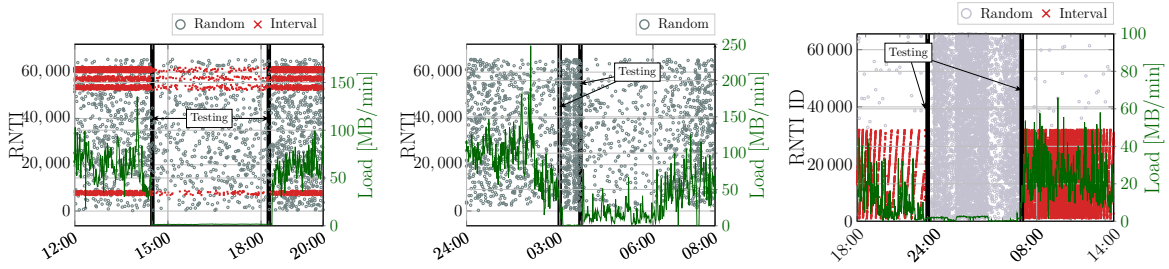
(a) Urban environment.

(b) Rural environment.

Figure 16: Operator 4 allocation scheme variation depending on the scenario. Figure best viewed in colors

5.4. Operator Tests

For both operators in France and for Operator K in the USA, we noted periods of time with sudden drops in traffic load from tens of MB/min down to 1 MB/min. Fig. 17 shows the behavior for the French operators: in Fig. 17(a) the time in which the load remains constantly low ranges from 14:40 to 18:10 while in Fig. 17(b) the time period is shorter, from 03:00 to 03:30.



(a) Operator H France.

(b) Operator I France.

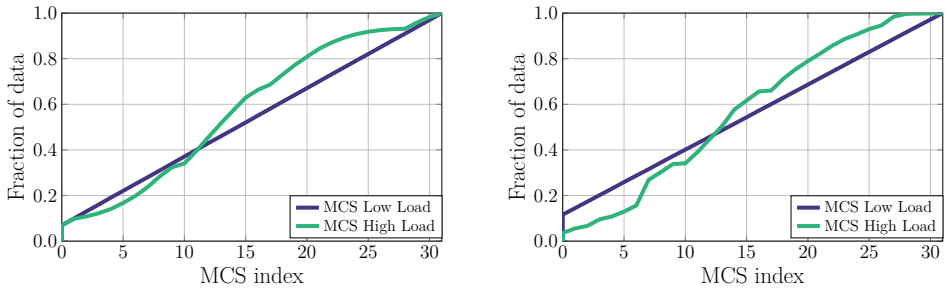
(c) Operator B Spain, BS 4.

Figure 17: Traffic load and RNTI allocation during the testing periods and normal operation. Figure best viewed in colors

To better understand such behavior, we analyzed the distribution of MCS usage, TBS allocation (number of transmissions and load). Surprisingly, during the *testing* periods the MCS distribution is different than during *normal activity* periods and the resulting CDF is a straight line for both operators (see Fig. 18). This means that MCS values are uniformly used across users. We tested the randomness of the MCS sequence of the traces with the ENT tool[2] and compared with a sequence of MCSs generated with `unix rand` function. ENT provides entropy, chi-square test and serial correlation coefficient among others metrics, and all indicate that the MCS sequence of the traces is indeed from a random number

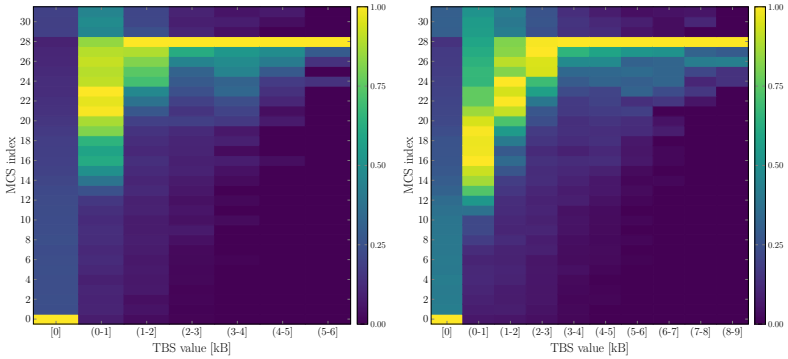
[2]Available at: `https://www.fourmilab.ch/random/`

generator (for example, for Operator J, the entropy are respectively 0.243550 and 0.323951 bits per bit while the serial correlation coefficient 0.374581 and 0.344085). The analysis of the TBS allocation, that we omit for brevity, shows that the per-user load in *testing* periods is significantly smaller than in *normal activity* periods and the number of individual user traffic allocations is significantly higher. All in all, this behavior suggests that some form of testing is ongoing and that these BSs are sending probes consisting of tiny packets with MCS values chosen uniformly at random. Fig. 19 shows the relation between MCS index and corresponding TBS with the occurrences normalized by the maximum of each column. For both operators, we observe that for small TBS values there exist a certain variation in the adopted MCS, this behavior suggests that tests with tiny packets over the full range of MCS are performed by the operator. Specifically, since MCS 29, 30 and 31 are reserved for retransmissions, the occurrences that we observe indicate that the testing packets that are lost get retransmitted by the BS.



(a) Operator H France CDF of MCS. (b) Operator I France CDF of MCS.

Figure 18: MCS testing periods



(a) Operator H. (b) Operator I.

Figure 19: MCS and TBS allocation during the testing periods

The same considerations hold for Operator K of the USA. Specifically, in this case we register one sudden drop in load from 05:24 to to 06:40. For the Spanish Operator B, BS 4, we observe testing periods during night times (see Fig. 17(c)). We performed the same study on the MCS sequence like for the operators of France and the USA and we find that they are not uniformly tested.

6. Evaluation

We now investigate the impact of the incorrect setting of the expiration threshold on the analysis of the traffic generated by individual users.

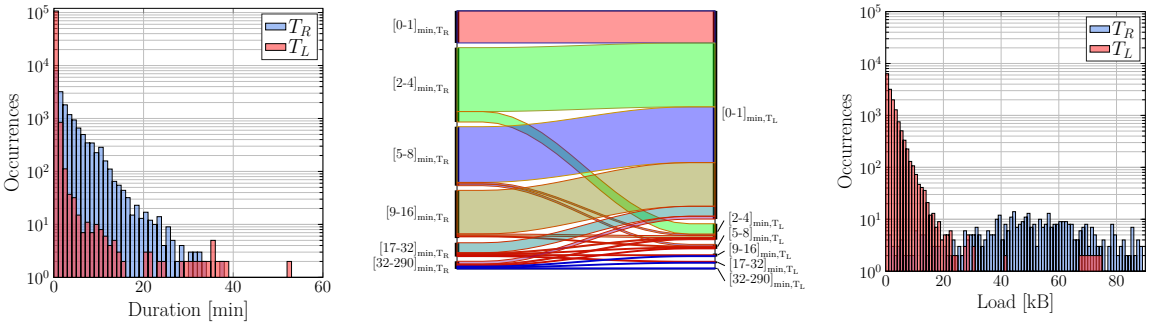
We present the results for Operator B BS 1 over a period of 24 hours. Specifically, we analyze how the number of extracted users and per-user load change when considering T_{R} and T_{L}. We denote as T_{R} our refined threshold, *i.e.*, 9 minutes (see Sec. 4.4) and as T_{L} the 10 seconds threshold used in the literature (see Sec. 2). Using T_{L} underestimates the user life time, *i.e.*, it assumes the RNTI was released while the user is in fact still active with the same RNTI, and the remaining time is assigned to a new user that in reality does not exist. This has two main implications i) a reduction of data attributed to the underestimated user and ii) the shift of such data to non-existent user. In fact, we observe that the number of extracted users is an order of magnitude higher for T_{L} than for T_{R}, *i.e.*, 107982 vs. 18170 respectively.

Fig. 20(a) presents occurrences of user life time duration extracted through T_{L} and T_{R} over 1-minute bins. The figure shows that for both thresholds the highest number of occurrences is registered for users whose duration is smaller than 1 minute. In the considered dataset, 40% of users have an estimated life time smaller than 1 minute, and this number grows to 98% when T_{L} is considered. The figure also shows that, as soon as the larger user durations when using T_{R} are considered, the corresponding number of occurrences in bin (0-1] minute drops by two orders of magnitude, since through T_{L} many fake users are created in precisely that range.

The Sankey diagram in Fig. 20(b) illustrates the user life time reallocation, where the width of the edge flow is proportional to the number of occurences that are shifted from larger to smaller intervals when using T_{L} instead of the correct value T_{R}. We can observe that all users' life time that ranges from 2 minutes up to the maximum are mostly allocated to non-existent users with a duration of (0-1] minute. Instead, only a small fraction of users preserve the same life time duration from T_{R} to T_{L}.

An incorrect choice of threshold also has an impact on the considered per-user load. Fig. 20(c) considers user durations that fall within the bin of (8-16] minutes and analyzes their characteristics in terms of per user-load. The average load of T_{R} extracted users is significantly higher than T_{L} extracted ones, *i.e.*, 830 kB vs. 32 kB. The figure shows a zoom over the interval 0 - 120 kB. We observe that when T_{L} is considered, 99% of the extracted users have a load which is smaller than 20 kB. Instead, the actual load when T_{R} is used as expiration threshold shows higher values ranging from 25 kB to 120 kB with a median value of 60 kB.

Takeaway message. An incorrect setting of the expiration threshold strongly affects the number of extracted users and the per-user load. When T_{L} instead of T_{R} is used, non-existent users are created, that have a smaller load value than the T_{R} extracted ones. These errors can have substantial impact



(a) Estimated user lifetime duration.

(b) Sankey diagram of user life time flows from users identified with T_{R} and with T_{L}.

(c) Load variation for user duration in the range 8-16 minute.

Figure 20: Implications of correct (T_{R}) and incorrect (T_{L}) RNTI expiration thresholds. Figure best viewed in colors

Table 4: Mean and standard deviation of user life time extracted according to correct and incorrect thresholds

Country	Mean		Stdev	
	$T_{\mathbf{L}}$	$T_{\mathbf{R}}$	$T_{\mathbf{L}}$	$T_{\mathbf{R}}$
FR	17 s	41 s	48 s	150 s
DE	9 s	84 s	31 s	203 s
NO	16 s	80 s	20 s	220 s
US	12 s	22 s	32 s	61 s

on the subsequent trace analyses, and may bias the conclusions.

7. Discussion

Our analysis serves a wide range of applications in resource management and traffic forecasting.

Resource allocation. *The resource allocation process at the BS side can benefit from the correct identification of the user life time over the network. Possible applications range from scheduling optimization for video streaming to allocation of pilot symbols.* For example, in [47], the number of active C-RNTIs is used to identify traffic variations and thus to detect urban anomalies. By correctly identifying the number of active UEs at a certain time period it is possible to discriminate normal and anomalous network behavior. Furthermore, by leveraging ms-level resource allocation patterns extracted from LTE, PERCEIVE in [48], aims at uplink throughput predictions to improve video streaming experience. TRADER [49] defines an efficient policy for aperiodic Sounding Reference Signal scheduling, leveraging per user traffic forecasts at the millisecond level.

Traffic forecasting. *The characterization of RNTIs* in Sec. 4.3 can *improve the predictive analysis of cellular networks.* For example in [34], the authors use the traffic information derived from the

DCI to predict mobile traffic up to the next 150 milliseconds. A neural network can benefit from the information about the adopted allocation scheme, *e.g.*, a change from one allocation scheme to another can be tied to a load variation (see Fig. 10a) over time in order to optimize the predictions. Furthermore, in [21], the RNTI is used to track the data flow of a certain user and then to learn a classifier on the applications executed at the connected mobile terminals. In particular, for some applications, *e.g.*, video streaming, a higher frame occupancy is desirable, and thus such type of traffic should be linked to certain RNTI allocation schemes, *i.e.*, round robin or interval allocation.

8. Conclusions

In this paper we provided an extensive study of the RNTI allocation mechanism by collecting and analyzing a real LTE traffic dataset from different BSs and operators in five different countries. By considering RNTI inter-transmission times and RA-RNTI to RNTI (C-RNTI) mapping, we proposed two methodologies for setting and validating the RNTI expiration threshold for trace analysis. Furthermore, according to the allocation scheme adopted by the BS, we characterized RNTIs in terms of frame occupancy and traffic load. This is an important step towards improving analysis and use of traffic traces gathered from operational mobile networks.

Our results reveal that the operators use or combine different allocation strategies (*e.g.*, random-only, random plus round robin and random plus interval) and which strategy is used varies according to the traffic load over night/day periods and rural/urban scenarios. We also show different types of round robin implementations and settings. Finally, we show how setting an incorrect RNTI expiration threshold affects the number of extracted users and the per-user load which highlights the importance of this parameter for proper analysis of traffic traces.

Upon acceptance, we intend to release the functional artifacts of our study to the scientific community. Future work will be devoted to compare our analysis for 5G NR which requires to extend the capabilities of the existing DCI sniffers.

Acknowledgment

This work is partially supported by a Juan de la Cierva grant from the Spanish Ministry of Science and Innovation (IJC2019-039885-I), by the Atracción de Talento Investigador grant number 2019-T1/TIC-16037 NetSense, funded by the Comunidad de Madrid, and by the Madrid Regional Government through the TAPIR-CM program (S2018/TCS-4496).

References

[1] C. D. Alwis, A. Kalla, Q.-V. Pham, P. Kumar, K. Dev, W.-J. Hwang, M. Liyanage, Survey on 6G frontiers: Trends, applications, requirements, technologies and future research, IEEE Open Journal of the Communications Society 2 (2021) 836–886. doi:10.1109/OJCOMS.2021.3071496.

[2] Ericsson, Mobility report, june 2022. technical report. (2021).

[3] F. Malandrino, C.-F. Chiasserini, 5G traffic forecasting: If verticals and mobile operators cooperate, in: Proc. of IEEE WONS, 2019, pp. 79–82.

[4] C. Fiandrino, C. Zhang, P. Patras, A. Banchs, J. Widmer, A machine learning-based framework for optimizing the operation of future networks, IEEE Communications Magazine 58 (6) (2020) 20–25. doi:10.1109/MCOM.001.1900601.

[5] P. D. Francesco, F. Malandrino, L. A. DaSilva, Assembling and using a cellular dataset for mobile network analysis and planning, IEEE Transactions on Big Data 4 (4) (2018) 614–620. doi:10.1109/TBDATA.2017.2734100.

[6] A. Mahimkar, A. Sivakumar, Z. Ge, S. Pathak, K. Biswas, Auric: Using data-driven recommendation to automatically generate cellular configuration, in: Proc. of the ACM SIGCOMM, 2021, p. 807–820. doi:10.1145/3452296.3472906.

[7] L. Chen, T.-M.-T. Nguyen, D. Yang, M. Nogueira, C. Wang, D. Zhang, Data-driven C-RAN optimization exploiting traffic and mobility dynamics of mobile users, IEEE Transactions on Mobile Computing 20 (5) (2021) 1773–1788. doi:10.1109/TMC.2020.2971470.

[8] I. Alawe, A. Ksentini, Y. Hadjadj-Aoul, P. Bertin, Improving traffic forecasting for 5G core network scalability: A machine learning approach, IEEE Network 32 (6) (2018) 42–49. doi:10.1109/MNET.2018.1800104.

[9] D. Bega, M. Gramaglia, M. Fiore, A. Banchs, X. Costa-Perez, DeepCog: Optimizing resource provisioning in network slicing with AI-based capacity forecasting, IEEE JSAC 38 (2) (2020) 361–376.

[10] J. Lin, Y. Chen, H. Zheng, M. Ding, P. Cheng, L. Hanzo, A data-driven base station sleeping strategy based on traffic prediction, IEEE Transactions on Network Science and Engineering (2021) 1–1doi:10.1109/TNSE.2021.3109614.

[11] S. Zhao, X. Jiang, G. Jacobson, R. Jana, W.-L. Hsu, R. Rustamov, M. Talasila, S. A. Aftab, Y. Chen, C. Borcea, Cellular network traffic prediction incorporating handover: A graph convolutional approach, in: Proc. of IEEE SECON, 2020, pp. 1–9.

[12] C. Gijón, M. Toril, S. Luna-Ramírez, M. L. Marí-Altozano, A data-driven traffic steering algorithm for optimizing user experience in multi-tier LTE networks, IEEE Transactions on Vehicular Technology 68 (10) (2019) 9414–9424. doi:10.1109/TVT.2019.2933068.

[13] M. Polese, F. Restuccia, T. Melodia, DeepBeam: Deep Waveform Learning for Coordination-Free Beam Management in mmWave Networks, Proc. of ACM MobiHoc (2021).

[14] S. Ahmadi, Mobile WiMAX: A systems approach to understanding IEEE 802.16 m radio access technology, Academic Press, 2010.

[15] D. Rupprecht, K. Kohls, T. Holz, C. Pöpper, Breaking LTE on layer two, in: Proc. of IEEE SP, 2019, pp. 1121–1136.

[16] S. Kumar, E. Hamed, D. Katabi, L. Erran Li, LTE radio analytics made easy and accessible, Proc. of ACM SIGCOMM 44 (4) (2014) 211–222.

[17] J. Zhao, B. Ding, Y. Guo, Z. Tan, S. Lu, Securesim: Rethinking authentication and access control for sim/esim, in: Proceedings of the ACM International Conference on Mobile Computing and Networking, 2021, p. 451–464. doi:10.1145/3447993.3483254.

[18] S. R. Hussain, M. Echeverria, I. Karim, O. Chowdhury, E. Bertino, 5GReasoner: A property-directed security and privacy analysis framework for 5g cellular network protocol, in: Proceedings ACM SIGSAC Conference on Computer and Communications Security, 2019, p. 669–684. doi:10.1145/3319535.3354263.

[19] A. Rago, G. Piro, G. Boggia, P. Dini, Multi-task learning at the mobile edge: An effective way to combine traffic classification and prediction, IEEE Transactions on Vehicular Technology 69 (9) (2020) 10362–10374. doi:10.1109/TVT.2020.3005724.

[20] A. Rago, G. Piro, H. D. Trinh, G. Boggia, P. Dini, Unveiling radio resource utilization dynamics of mobile traffic through unsupervised learning, in: Network Traffic Measurement and Analysis Conference (TMA), 2019, pp. 209–214. doi:10.23919/TMA.2019.8784692.

[21] H. D. Trinh, A. F. Gambin, L. Giupponi, M. Rossi, P. Dini, Mobile traffic classification through physical control channel fingerprinting: A deep learning approach, IEEE Transactions on Network and Service Management 18 (2) (2021) 1946–1961. doi:10.1109/TNSM.2020.3028197.

[22] L. Zhai, Z. Qiao, Z. Wang, D. Wei, Identify what you are doing: Smartphone apps fingerprinting on cellular network traffic, in: IEEE Symposium on Computers and Communications (ISCC), 2021, pp. 1–7. doi:10.1109/ISCC53001.2021.9631415.

[23] F. Meneghello, M. Rossi, N. Bui, Smartphone identification via passive traffic fingerprinting: A sequence-to-sequence learning approach, IEEE Network 34 (2) (2020) 112–120.

[24] T. A. Nguyen, P. Martins, Cellular traffic type recognition and prediction, in: IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), 2021, pp. 1167–1172. doi:10.1109/PIMRC50174.2021.9569524.

[25] J.-W. Son, S. Lee, M.-h. Han, Supervised service classification using downlink control indicator in LTE physical downlink control channel, in: International Conference on Information and Communication Technology Convergence (ICTC), 2021, pp. 1533–1536. doi:10.1109/ICTC52510.2021.9621192.

[26] J. Wu, M. Zeng, X. Chen, Y. Li, D. Jin, Characterizing and predicting individual traffic usage of mobile application in cellular network, in: Proceedings of the ACM International Joint Conference and International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers, 2018, p. 852–861. doi:10.1145/3267305.3274173.

[27] E. Mucelli Rezende Oliveira, A. Carneiro Viana, K. Naveen, C. Sarraute, Mobile data traffic modeling: Revealing temporal facets, Computer Networks 112 (2017) 176–193. doi:https://doi.org/10.1016/j.comnet.2016.10.016.

[28] G. Attanasio, C. Fiandrino, M. Fiore, J. Widmer, Characterizing rnti allocation and management in mobile networks, in: Proceedings of the 24th International ACM Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems, Association for Computing Machinery, New York, NY, USA, 2021, p. 189–197.
URL https://doi.org/10.1145/3479239.3485685

[29] R. Falkenberg, C. Wietfeld, Falcon: An accurate real-time monitor for client-based mobile network data analytics, in: Proc. of IEEE GLOBECOM, 2019, pp. 1–7.

[30] J. Chen, R. Shi, K. Deng, Acquisition and separation of mobile communication cell users code stream through communication reconnaissance, in: International Conference on Communication Software and Networks (ICCSN), 2022, pp. 6–12. doi:10.1109/ICCSN55126.2022.9817609.

[31] G. T. 36.321, LTE Medium Access Control (MAC) protocol specification, Tech. Rep. 36.321, 3GPP, version 14.4.0 (10 2017).

[32] G. T. 38.321, 5G NR Medium Access Control (MAC) protocol specification, Tech. Rep. 38.321, 3GPP, version 15.6.0 (07 2019).

[33] S. Hailu, M. Säily, Hybrid paging and location tracking scheme for inactive 5G ues, in: Proc. of IEEE EuCNC, 2017, pp. 1–6. doi:10.1109/EuCNC.2017.7980730.

[34] H. D. Trinh, L. Giupponi, P. Dini, Mobile traffic prediction from raw data using LSTM networks, in: Proc. of IEEE PIMRC, 2018, pp. 1827–1832. doi:10.1109/PIMRC.2018.8581000.

[35] N. Bui, J. Widmer, OWL: A reliable online watcher for lte control channel measurements, in: Proc. of ACM Workshop on All Things Cellular: Operations, Applications and Challenges, 2016, pp. 25–30. doi:10.1145/2980055.2980057.

[36] J. Huang, F. Qian, A. Gerber, Z. M. Mao, S. Sen, O. Spatscheck, A close examination of performance and power characteristics of 4G LTE networks, in: Proc. of ACM MobiSys, 2012, pp. 225–238.

[37] R. P. Jover, LTE security, protocol exploits and location tracking experimentation with low-cost software radio, arXiv:1607.05171 (2016).

[38] F. Qian, Z. Wang, A. Gerber, Z. M. Mao, S. Sen, O. Spatscheck, Characterizing radio resource allocation for 3G networks, in: Proc. of ACM IMC, 2010, pp. 137–150.

[39] M. Z. Shafiq, L. Ji, A. X. Liu, J. Pang, S. Venkataraman, J. Wang, Characterizing and optimizing cellular network performance during crowded events, IEEE/ACM Transactions On Networking 24 (3) (2016) 1308–1321.

[40] A. Narayanan, X. Zhang, R. Zhu, A. Hassan, S. Jin, X. Zhu, X. Zhang, D. Rybkin, Z. Yang, Z. M. Mao, et al., A variegated look at 5G in the wild: performance, power, and QoE implications, in: Proc. of ACM SIGCOMM, 2021, pp. 610–625.

[41] X. Chen, W. Yang, C. Xu, Y.-I. Kim, RNTI allocation schemes for user equipments in LTE system, in: Proc. of IEEE WiCOM, 2012, pp. 1–4.

[42] C. Ramesh, A. Jafar, G. Prasad, An effective RNTI allocation scheme for user equipment in LTE systems, in: Proc. of WiSPNET, 2016, pp. 1737–1740. doi:10.1109/WiSPNET.2016.7566436.

[43] M. T. Raza, D. Kim, K.-H. Kim, S. Lu, M. Gerla, Rethinking LTE network functions virtualization, in: Proc. of IEEE ICNP, 2017, pp. 1–10.

[44] N. Ludant, N. Bui, A. García Armada, J. Widmer, Data-driven performance evaluation of carrier aggregation in LTE-advanced, in: Proc. of IEEE PIMRC, 2017, pp. 1–6.

[45] H. Zhang, X. Qiu, L. Meng, X. Zhang, Design of distributed and autonomic load balancing for self-organization LTE, in: Proc. of IEEE VTC Fall, 2010, pp. 1–5.

[46] G. T. 38.314, LTE Evolved Universal Terrestrial Radio Access Network (E-UTRAN);, Tech. rep., 3GPP, version 8.1.0 (04 2009).

[47] H. D. Trinh, L. Giupponi, P. Dini, Urban anomaly detection by processing mobile traffic traces with LSTM neural networks, in: Proc. of IEEE SECON, 2019, pp. 1–8.

[48] J. Lee, S. Lee, J. Lee, S. D. Sathyanarayana, H. Lim, J. Lee, X. Zhu, S. Ramakrishnan, D. Grunwald, K. Lee, et al., PERCEIVE: deep learning-based cellular uplink prediction using real-time scheduling patterns, in: Proc. of ACM MobiSys, 2020, pp. 377–390.

[49] C. Fiandrino, G. Attanasio, M. Fiore, J. Widmer, Traffic-driven sounding reference signal resourceallocation in (beyond) 5g networks, in: Proc. of IEEE SECON, 2021, pp. 1–9. doi:10.1109/SECON52354.2021.9491611.

Appendix A. Detailed Evaluation Results

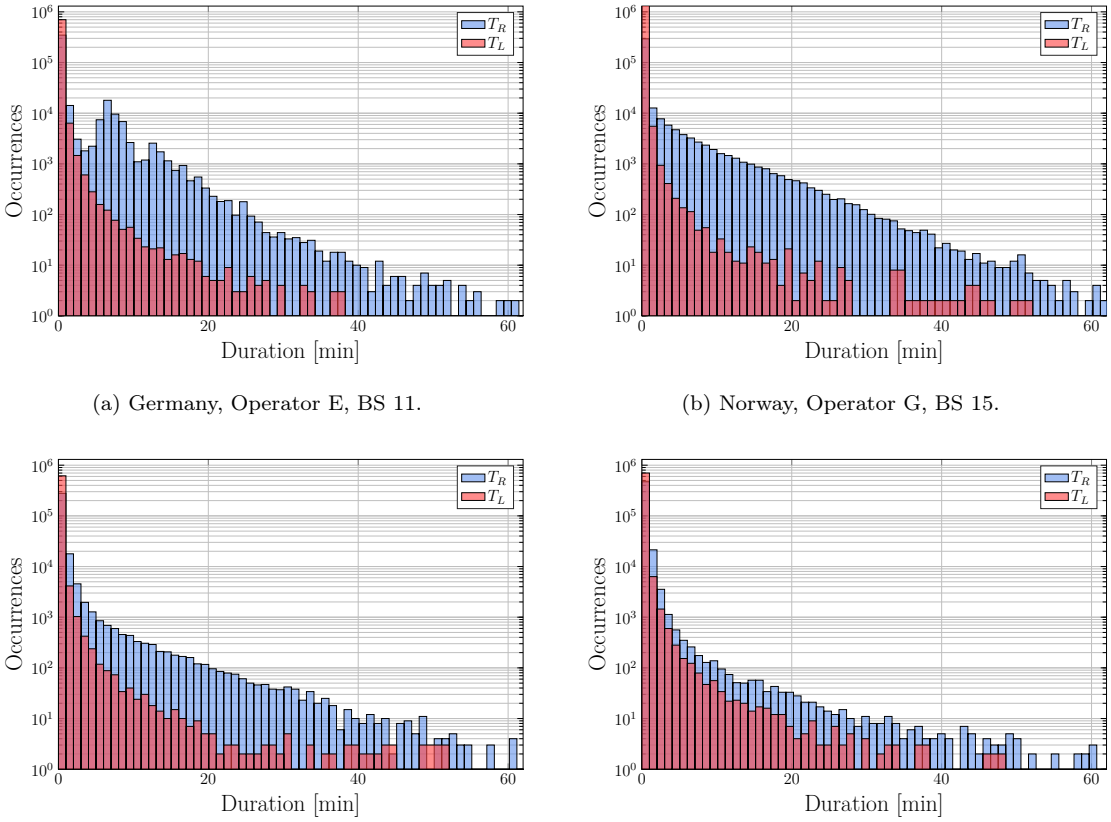
This section extends the results presented in Sec. 6 and shows the detailed results per-country of user life time computed with the correct (T_{R}) and incorrect (T_{L}) RNTI expiration threshold.

Specifically, Fig. A.21 shows the distribution of user life time with the two thresholds. We observe that the main result holds for all the countries: the use of T_{L} leads to a biased distribution where the vast majority of the user life time duration is at most 1 minute. Using the T_{R} threshold instead reduces by two orders of magnitudes the occurrences of users with life time duration less than 1 minute. The consequence is that the distribution shows more occurrences for user life times of longer duration: this becomes evident looking at the tail of the distributions where T_{R} shows occurrences not present with T_{L}. Depending on the traffic conditions, *i.e.*, number of connected users at the BS and their session duration, the difference between the distributions obtained with T_{R} and T_{L} can vary. For example, in Fig. A.21(d) the two distributions are closer than in Fig. A.21(b) and this indicates that the mistake of selecting an incorrect threshold is more severe in the second case.

Fig. A.23 highlights in more details for the specific countries how user life time changes with the use of T_{R} (left inside of the plots) with respect to the use of T_{L} (right inside of the plots). With such a graphical representation it becomes easier to observe how close are the distributions of user life time with correct and incorrect threshold. As mentioned above, in Fig. A.23(d) the two distributions are closer than in Fig. A.23(b) because the fraction of users that remains in the interval $[0, 1]$ minute is larger.

Biographies

Giulia Attanasio is a Ph.D. student in Telematics at the University of Carlos III. Her research interest lies between machine learning and mobile networks communications. Previously, she completed an MSc



(a) Germany, Operator E, BS 11.

(b) Norway, Operator G, BS 15.

(c) France, Operator J, BS 18.

(d) USA, Operator L, BS 20.

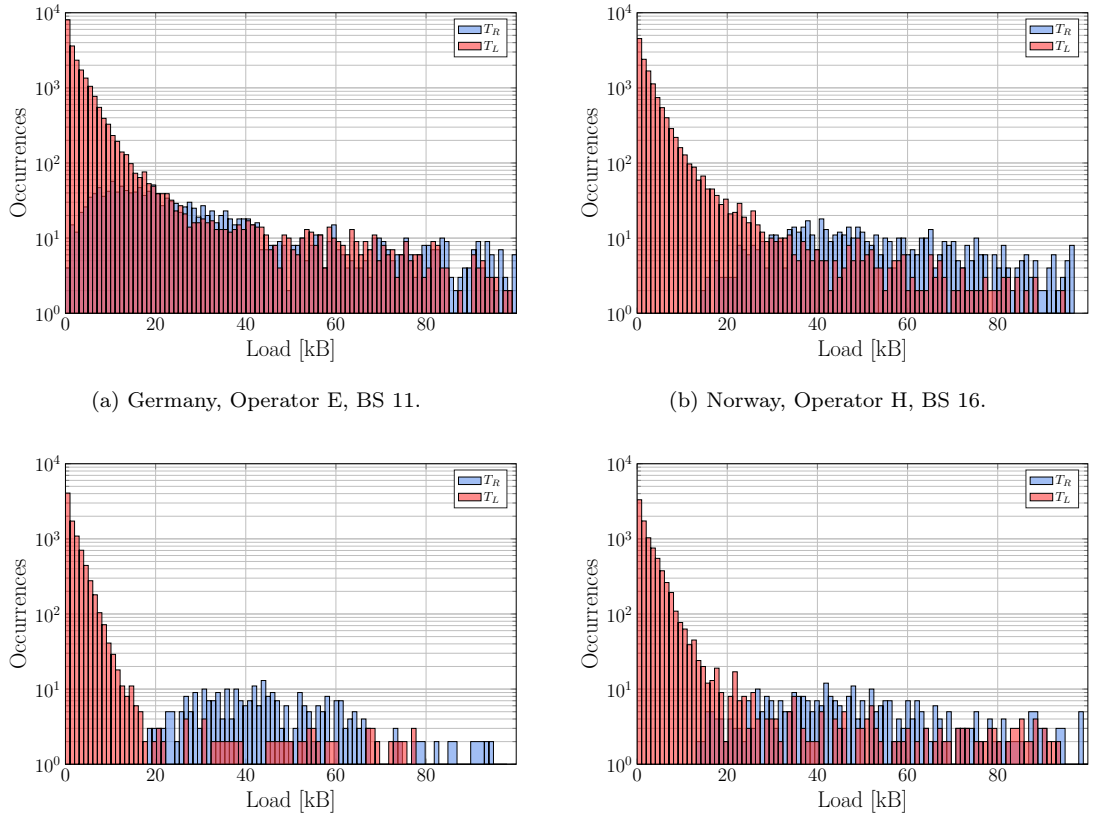
Figure A.21: Implication of setting correct (T_{R}) and incorrect (T_{L}) RNTI expiration thresholds on the estimated user lifetime

in Communications and Computer Networks Engineering and a BS in Telecommunications Engineering at the Polytechnic University of Turin.

Claudio Fiandrino is a postdoctoral researcher at IMDEA Networks Institute. He obtained his Ph.D. degree from the University of Luxembourg in 2016. His research focuses on mobile networks in the areas of AI-driven network optimization and edge computing. Claudio was awarded with two Spanish Juan de la Cierva grants and 3 Best Paper Awards for his research.

Marco Fiore is a Research Associate Professor at IMDEA Networks Institute, Spain, where he leads the Networks Data Science group on research activities at the interface of mobile networking and applied data science. He previously held tenured positions in France and Italy, and was a Marie Curie fellow and a Royal Society visiting research fellow.

Joerg Widmer is Research Professor and Research Director of IMDEA Networks in Madrid, Spain. His research focuses on wireless networks, in particular millimeter-wave communications. Joerg Widmer is an IEEE Fellow and Distinguished Member of the ACM, and was awarded an ERC consolidator



(a) Germany, Operator E, BS 11.

(b) Norway, Operator H, BS 16.

(c) France, Operator J, BS 18.

(d) USA, Operator L, BS 20.

Figure A.22: Implication of setting correct (T_{R}) and incorrect (T_{L}) RNTI expiration thresholds on per-user load

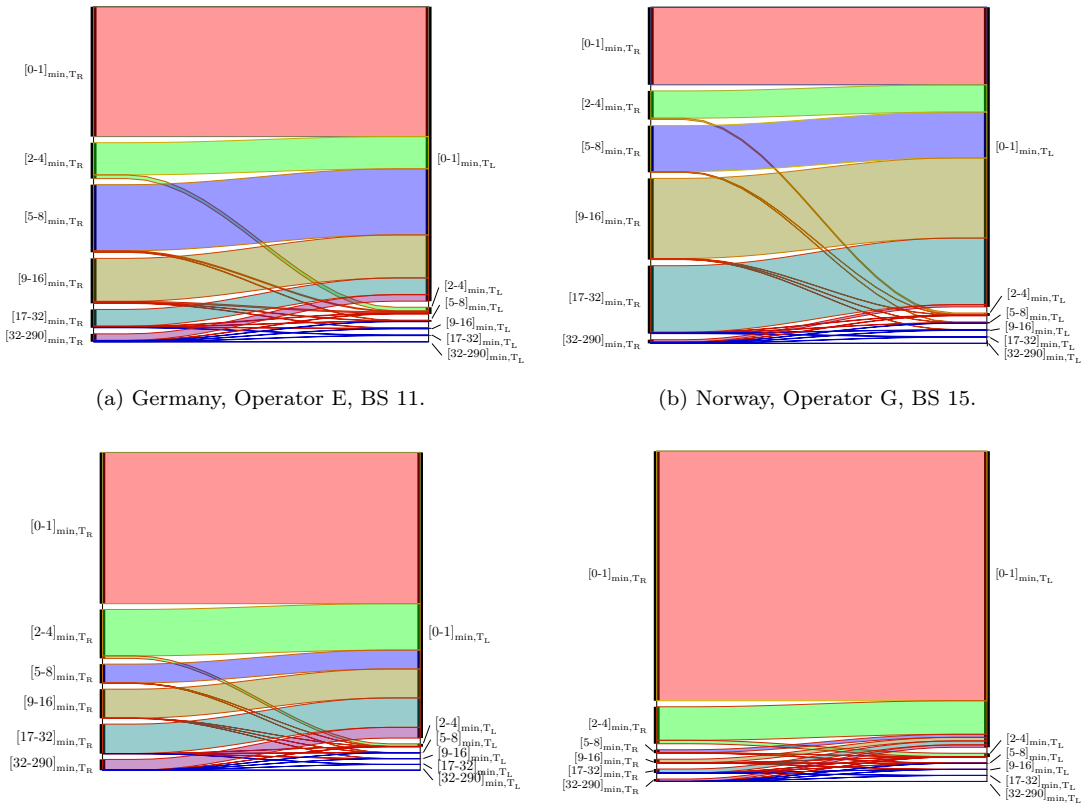
grant, the Friedrich Wilhelm Bessel Research Award of the Alexander von Humboldt Foundation, a Mercator Fellowship of the German Research Foundation, and a Spanish Ramon y Cajal grant.

Bastian Bloessl is a postdoctoral researcher at the Secure Mobile Networking Lab, TU Darmstadt, Germany. He previously was a research fellow at the CONNECT Center, Trinity College Dublin, Ireland's Research Center for Future Networks and Communications, where he was funded through a Marie Curie fellowship. His research focuses on software-defined wireless communication systems.

Norbert Ludant is a PhD student in Cybersecurity at Northeastern University. He received MS degrees in Telecommunications Engineering and Multimedia and Communications in 2017 from the University Carlos III de Madrid. His research focuses on mobile broadband wireless communications, signal processing, security and privacy.

Konstantinos Kousias is a postdoctoral researcher at Simula Research Laboratory, Norway. He received his Ph.D. from the University of Oslo in 2021. His research focuses on the empirical modeling and evaluation of mobile networks and Internet of Things (IoT) using AI-driven analytics.

Özgü Alay is an Associate Professor at the Department of Informatics and leads the Gemini IoT



(a) Germany, Operator E, BS 11.

(b) Norway, Operator G, BS 15.

(c) France, Operator J, BS 18.

(d) USA, Operator L, BS 20.

Figure A.23: Sankey diagram of user life time: flows from users identified with T_{L} and with T_{R}

Center at the University of Oslo. Her research focuses on mobile networks including 5G and B5G, IoT, low latency networking, multipath protocols and robust multimedia transmission over wireless networks.

Lise Jacquot is a MS student at INSA Lyon, France. Her research focus is on energy consumption modelling in cellular networks.

Razvan Stanica is an Associate Professor at INSA Lyon and a member of the Inria Agora team. He obtained his Ph.D. degree from the National Polytechnic Institute of Toulouse in 2011. His research focuses on wireless networks in dynamic environments, analytics of mobile traffic data and intelligent transportation systems.