IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received June 7, 2020, accepted July 2, 2020, date of publication July 14, 2020, date of current version July 24, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3009127

Service-Based Network Dimensioning for 5G Networks Assisted by Real Data

M. UMAR KHAN, A. GARCÍA-ARMADA, (Senior Member, IEEE), AND J. J. ESCUDERO-GARZÁS, (Member, IEEE)

Department of Signal Theory and Communications, Universidad Carlos III de Madrid, 28911 Leganés, Spain

Corresponding author: M. Umar Khan (mumar@tsc.uc3m.es)

This work was supported by the Spanish National Project TERESA-ADA (MINECO/AEI/R, UE), under Grant TEC2017-90093-C3-2-R.

ABSTRACT The fifth-generation (5G) of cellular communications is expected to be deployed in the next years to support a wide range of services with different demands of peak data rates, latency and quality of experience (QoE). In this work, we propose a novel approach for radio network dimensioning (RND), named as Heuristic RND (HRND), which uses real open data in the network dimensioning process. This procedure, named as *NetDataDrilling*, provides the dimensioning target area by means of network data selection and visualization from the existing infrastructure. Moreover, the proposed *NetDimensioning* heuristic considers the necessary parameters of numerology and bandwidth parts (BWP) supported by New Radio (NR) to provide a balanced network design mediating among the requirements of coverage, capacity, QoE and cost. The proposed HRND is based on the novel quality of experience (QoE) parameter ζ by probabilistically characterizing the 5G radio resource control (*rrc*) states to ensure the availability of peak data rates for the MNO's required percentage of the time. The simulation results show the fulfilment of QoE and load balancing parameters with significant cost savings compared to the conventional RND methodology.

INDEX TERMS 5G, new radio, cellular network dimensioning, network data, capacity model.

ABBREVIATIONS

3GPP	Third-Generation Partnership Project
5GDA	5G Deployment Area
5G	Fifth Generation
BWP	Bandwidth Part
CET	Cell Edge Throughput
CID	Cell tower ID
CBS	Confidence Benchmark Samples
eNB	Evolved Node-B
eMBB	Enhanced Mobile Broadband
FR1	Frequency Range One
FR2	Frequency Range Two
GPS	Global Positioning System
gNB	Next Generation Node-B
HTTAC	Highest Traffic TAC
IID	Independent & Identically Distributed
mMTC	Massive Machine Type Communication
MCC	Mobile Country Code
MNC	Mobile Network Code
MNO	Mobile Network Operators
NR	New Radio
PRB	Physical Resource Block
PLMN	Public Land Mobile Network
RAT	Radio Access Technology
RRC	Radio Resource Control
RND	Radio Network Dimensioning
TA	Tracking Area
TAC	Tracking Area Code
TAI	Tracking Area Identity
URLLC	Ultra-Reliable Low Latency Communication
UMi	3GPP Urban Micro channel model

The associate editor coordinating the review of this manuscript and approving it for publication was Olabisi Falowo.

I. INTRODUCTION

The new generation of mobile networks, or 5G, is envisioned to offer new services defined by 3GPP, from ultra-reliable low latency (URLLC) services to enhanced mobile broadband (eMBB) and massive machine type communications (mMTC) [1]. These services should be able to support the stringent requirements demanded by applications such as vehicle-to-all (V2X) communications, emergency communications and augmented reality. Consequently, infrastructures supporting 5G communications are being deployed. In this context, cellular radio network planning is of paramount importance as the process of selecting the minimum number of radio sites, including the adequate setting of network

parameters, to guarantee a satisfactory quality of service to all mobile subscribers. Network planning, on one hand, makes it possible to provide adequate coverage for subscribers in the intended area [2], [3]. On the other hand, it is an effective way to reduce capital and operational expenditure (CAPEX and OPEX, respectively) [4], [5].

Network planning has originated a significant body of research (see [6], [7] and references therein) for LTE/ LTE-A networks, though it faces a number of challenges regarding the deployment of 5G networks, including the self-organizing capability of 5G networks [8], broadcast and multicast in smart grids [9] and ultra-dense topologies [10], [11].

Traditionally, cell planning has followed a human-centric approach [12]–[14], leading to network deployments where the number and location of base stations satisfy the traffic demanded by applications such as video streaming, Voice over IP (VoIP) and social networks. However, the needs of new machine-centric applications and scenarios (e.g., smart cities and emergency alerts) bring to 5G networks very stringent requirements in the form of strict delays or ultra-reliable links. Therefore, the proposals for 5G radio network dimensioning (RND) should explore how to deal with both types of services simultaneously. The use of numerology and bandwidth parts (BWP) endows 5G New Radio (NR) [15] with a degree of flexibility so that the radio interface is configured to accommodate such diversity of services.

An important aspect of telecommunication networks is the quality of service (QoS) [16], which has been taken into consideration in mobile networks under very different forms, from guaranteeing a minimum data rate [17] to throughput and delay constraints [18]. However, 5G services and applications require more than ensuring a given value of a conventional communications metric, leading to a broad adoption of the concept of Quality of Experience (QoE) introduced in LTE networks [19]–[21].

The majority of studies on radio network planning make use of forecast data, usually estimated traffic demand or traffic modeling, to design the network deployment (see for instance [3], [22]–[24]). As an alternative, there exist organizations that collect real mobile network data and make these data available on line, both in raw and processed form, making a more informed interpretation and decision-making possible. In this study, we focus on utilizing the available network data for 5G NR network planning, as these data can help us achieve a more efficient network planning design, especially in terms of topology and cost. The current 4G cellular networks are generating a huge amount of network data which are stored and processed by different entities for commercial usage [25], [26]. However, these databases are expensive and the network data are sensitive information of the MNOs that cannot be made available publicly. Consequently, we resource to crowdsourced open data to leverage the existing infrastructure from OpenCelliD by Unwired Labs, a collaborative community project that collects GPS positions of cellular towers and their corresponding location area identity [27], available for academic research.

In this work, we propose a new dimensioning approach whose objective is to minimize the network cost. This proposal contains significant contributions that differentiate it from existing works. In the first place, real open data are used to estimate traffic and identify the area of highest user density for the deployment of new services. Second, this work makes use of flexible numerology and BWPs in the design of network dimensioning, conforming a framework that is able to accommodate both human- and machine-centric services. Third, the quality of experience (QoE) is explicitly considered in the dimensioning process, consideration that is absent in the previous works in the literature.

The manuscript is organized as follows. In Section II, we provide the state of the art for different planning strategies adopted by the research community and our contributions. Section III describes our proposed framework. In Section IV, we discuss the data acquisition process to visualize subscriber density patterns for 5G gNB (Next Generation Node-B) deployments. In Section V, we describe the transmission model and the proposed radio configuration for the 5G services. Section VI details the network dimensioning approach, while the heuristic algorithm is developed in section VII. In section VIII, we introduce the conventional RND approach. We present our results and analysis in Section IX. Finally, we provide our conclusions and future research lines.

II. RELATED WORK AND CONTRIBUTIONS

Cellular network data has received attention even in the earlier studies to analyze the spatial and temporal variations of the subscribers [28], to identify the network traffic patterns [29] and in determining the network parameter settings for enhanced performance and QoE to the end user [30]. These studies being carried out using network data collected by the corresponding MNO. However, thanks to the existence of crowdsourced database, researchers have access to data that make it possible to analyze and compare the performance of different networks and operators without resourcing to data owned by MNOs. For instance, in [31] the authors estimate the current and future traffic demands and in [32] the state of the legacy infrastructure deployed by multiple MNOs is inferred. In our study, a crowdsourced database is used to identify the densest area in terms of subscribers, which is the target area for network planning in this work.

One significant challenge in wireless networks is to utilize network resources efficiently while delivering guaranteed QoE to end users, which can also be used to maximize operators' profit. In [33], the authors propose a multi-constraint optimization model for selecting the best possible network in a heterogeneous wireless environment to maximize the user experience, which eventually increases the revenue stream for MNOs. On the other hand, in [34], a novel QoE-based radio resource allocation for OFDMA networks is designed by using a utility function to characterize fairness in terms

of QoE and power consumption. Another approach to optimize the overall system QoE is proposed in [35] for LTE, which is based on an iterative algorithm to dynamically change the priorities of the different services. The self-tuning approach of [36] improves the QoE in multi-tier LTE networks based on parameters from network data traces. Another self-tuning approach [37] computes the priority score in a multi-service environment and allocates the resource blocks of the highest priority score to enhance QoE. Though QoE parameters have gained significant importance in live/running cellular networks [38], considering QoE in network planning is a crucial task as user expectations are changing the way MNOs manage and plan their networks. Therefore, we find it relevant to feed the QoE factor to the 5G RND exercise as proposed in this work.

LTE network planning has been studied in several directions. In [39], Hurley conceives network planning by proposing an optimization framework for radio site selection and determining the suitable configuration of antennas, power, azimuth and tilt, for a range of scenarios. The authors of [40] provide a theoretical analysis based on numerical models for developing and optimizing LTE network planning. In [10], genetic- and graph-based approaches are proposed to address the cell planning problem, improving the throughput by using interference mitigation. In [5], a geographic division approach divides the target service regions into multiple subregions of identical traffic loads, which is discussed in terms of coverage, capacity, QoS and reduction of CAPEX/OPEX. In the context of ultra-dense networks (UDN), González *et al.* present an optimization framework for planning, optimization and deployment of UDNs taking devices scalability, fairness and the cost of densifying the network into consideration [41]. In [42], network planning is posed as a Voronoi mapping between radio sites and service demands, discussing also power allocation. A different approach is developed in [43] to consolidate existing LTE sites so that the requirements of coverage, capacity and cost are accomplished.

In the 5G deployment race, various studies investigate different aspects of network planning. In [44], the energy efficiency prospects of 5G wireless networks are analyzed by proposing novel solutions for reducing energy consumption in different parts of the networks. The work in [45] introduces an analytical model for planning and dimensioning of 5G cloud RAN in the context of energy efficiency. Other works study 5G network planning and deployments in different perspectives. For instance, in [46], a model for choosing an optimal place of the base station to maximize the coverage area for different 5G frequencies in an urban scenario is presented. At the same time, [47] introduces an in-house open-source techno-economic assessment tool for 5G deployments in macrocell scenarios to minimize the cost through infrastructure sharing strategies [48], [49]. In [50], the authors studied the coverage and capacity dimensioning procedures to optimize the 4G LTE networks for the upcoming 5G deployments in the non-standalone scenario, while considering capacity and coverage planning. In [51], a detailed technical and economic analysis to minimize the 5G deployment cost in the context of different architectural scenarios is presented; however, it lacks the discussion on coverage, capacity and QoE. A recent work [52] studies the 5G network planning as a profit maximization model for the MNO with minimal deployment cost.

The above revision of the existing literature suggests that, to the best of our knowledge, our work is the first focused on 5G RND that incorporates NR numerology and bandwidth parts (BWP), so bringing QoE to network design. Besides, this perspective incorporates the requirements of use cases to planning. We also define a QoE parameter to deliver peak data rates for a required percentage of the time. Finally, our proposed network dimensioning framework provides a balanced network design with minimum deployment cost.

A. CONTRIBUTIONS

This study presents a complete radio network dimensioning approach in 5G NR based on real network data and pays particular attention to guaranteeing quality of experience for different type of services. In brief, we bring the following contributions:

1) We propose the utilization of real network data to estimate network traffic and the target deployment area where the subscribers concentration is maximum, reflecting the most congested area. As a result of deploying the 5G gNBs in this area, the MNOs increase network capacity and possibly reduce the cost per bit.
2) To the best of our knowledge, this is the first study that considers the pragmatic aspects of numerology and bandwidth parts for different 5G use cases offered as services. We adapt the conventional LTE RND as the foundation of our work to propose a heuristic dimensioning procedure for 5G, which provides a reliable technique for MNOs, infrastructure providers (InP) and 5G tool developers to address RND.
3) We introduce a QoE parameter in our heuristic dimensioning to ensure the peak data rates provided to subscribers for the percentage of time specified by the MNOs. This is accomplished by probabilistically characterizing the 5G *rrc* states. This QoE parameter is particularly relevant to mission critical 5G services.
4) We have evaluated our data analysis procedure and dimensioning heuristic based on data from the top three Spanish MNOs.

III. THE NETWORK DIMENSIONING FRAMEWORK

Network dimensioning is the core of our study as it refers to the design of the 5G radio network by performing coverage and capacity dimensioning, to be discussed in Sections VI-A and VI-B respectively (see Fig.1). This process crystallizes in the network dimensioning heuristic presented in Section VII. Different to legacy networks, where radio dimensioning is focused on providing high-speed data rates, in 5G there are multiple services with different data rates and latency

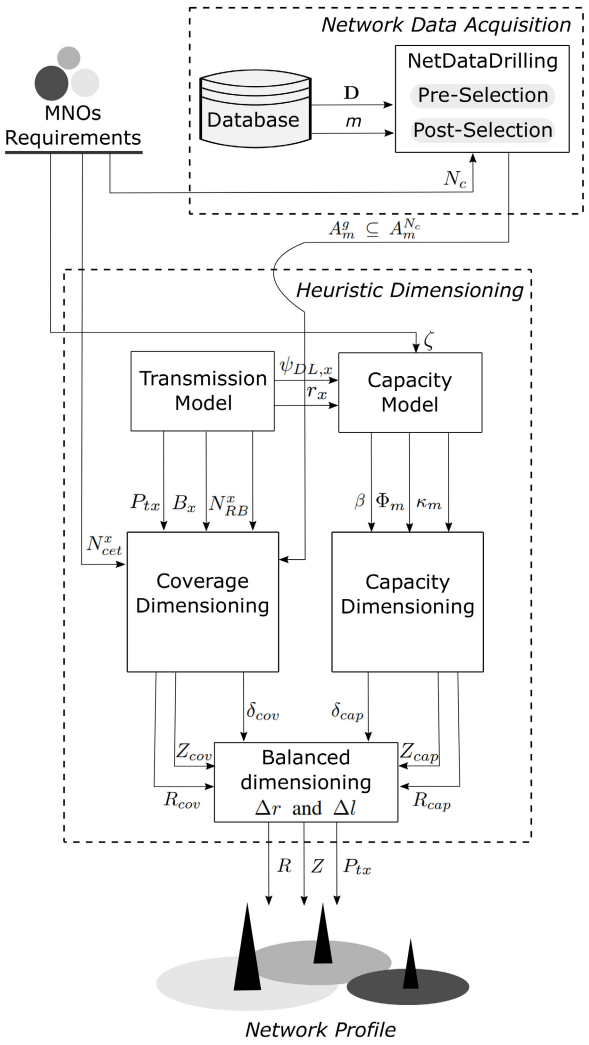

FIGURE 1. Network dimensioning framework.

requirements which should be an integral part of the network dimensioning phase. Our proposal takes these requirements into consideration to find the network parameters that mediate among the considered objectives: coverage, capacity, cost and QoE.

The complete framework is provided in Fig.1. We distinguish two main blocks, *Network Data Acquisition* and *Heuristic Dimensioning*. Network Data Acquisition processes the variables from an external database to identify the target area for deployment, referred to as 5G deployment area (5GDA) and denoted as A_m^g. The 5GDA is supplied to Heuristic Dimensioning, which processes this information and other inputs to provide the cell range, power distribution and the number of radio sites required for deployment.

The acquisition of network data is performed on variables from the database **D** for m MNOs to estimate network traffic and identify the 5GDA A_m^g, which is a sub-area of $A_m^{N_c}$ that covers the N_c cells of interest of the MNOs requirement. Specifically, we focus on the variables radio access technology (RAT) r_E, mobile network code (MNC) m_E, tracking area code (TAC) t_E, cell tower ID (CID) n_E, geographic location of a cell tower (LOC) l_E, and the number of samples recorded per cell tower s_E, where E represents the number of rows of the database. The NetDataDrilling algorithm described in Section IV-C processes these variables to identify A_m^g through two phases, i.e., Pre-Selection and Post-Selection. While the Pre-Selection phase extracts the data corresponding to the MNOs, the Post-Selection phase guarantees reliability of the acquired data.

The Heuristic Dimensioning entity in Fig.1 makes use of two models, namely the transmission model and the capacity model. The transmission model specifies the new services being offered by the MNO based on 5G numerology. Each service x is allocated a BWP B_x with its corresponding number of PRBs N_{RB}^x and the gNB transmit power level P_{tx}^x, which are provided as inputs to the Coverage Dimensioning phase. The maximum transmission data rate $\psi_{DL,x}$ per service and the peak data rate r_x per user are provided as inputs to the capacity model. This capacity model is based on the MNO's requirement to guarantee QoE by providing the peak data rate r_x for a required percentage of time ζ. Based on ζ and r_x, the capacity model decides the gNB site capacity β by considering the subscriber density Φ_m devices/km^2 and the market share κ_m of MNO m.

The proposed heuristic dimensioning of Section VII is based on the coverage and capacity dimensioning to get a balanced network design and the minimization of the deployment cost. Along with NR transmission model parameters, the coverage dimensioning of service x is based on the concept of cell edge throughput (CET) by allocating N_{cet}^x number of PRBs per service [53]. The coverage cell range R_{cov} is then provided along with the required number of radio sites Z_{cov} for coverage of the deployment area. At the same time, capacity dimensioning provides capacity cell range R_{cap} and the required number of radio sites Z_{cap} to meet the capacity based on the gNB site capacity β, Φ_m and κ_m per MNO. The balanced dimensioning approach then determines if the difference in cell ranges $|R_{cov} - R_{cap}|$ and cell loads $|\delta_{cov} - \delta_{cap}|$ are under the given thresholds Δr and Δl, respectively. Finally, the network profile is obtained, which consists of the suitable transmit power level P_{tx}, cell range R and the minimum number of required radio sites Z.

The different parts of this framework are developed in the subsequent sections.

IV. NETWORK DATA ACQUISITION

This section describes the procedure to extract the network information from the database, labelled as Network Data Acquisition in Fig.1. The objective of this procedure is to identify the area with maximum density of subscribers of each MNO, as this area would represent the region with the highest traffic with high probability. As 5G networks must offer new services with a lower cost per bit compared to LTE networks, the higher number of subscriptions increases network utilization and help MNOs to reduce the cost per bit. For instance, consider the two 7×7 km^2 areas given

in Fig.2 representing densely populated (left) and sparsely populated (right) regions. The same number of base stations would cover both identical areas. However, the subscriber densities are different, which possibly leads to under utilization of the resources in the sparsely populated area (Fig.2, right). Assuming that the CAPEX and OPEX of base station deployment are constant for both areas, the lower network utilization causes a higher cost per bit for the operator. On the other hand, in the case of a densely populated area (Fig.2, left) network utilization will be higher, which leads to lower cost per bit as compared with the former case. Therefore, we propose to utilize network data to identify the area with highest subscriber density per MNO as the target area of deployment to minimize the cost per bit.

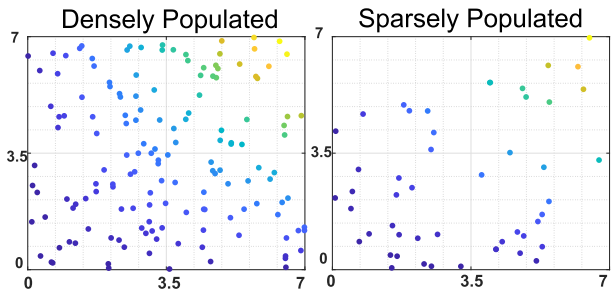

FIGURE 2. Example of subscriber density in 7 × 7 km^2 identical areas.

The following sections describe the complete process for network data acquisition. We first describe the open database in detail and then discuss the procedure.

A. THE OpenCellID DATABASE

OpenCellID is an open-access database/platform to which mobile devices send data related to their connection (e.g. position and different identifiers) through a crowdsourcing application [27]. This application records the GPS coordinates (latitude, longitude) of both device and base station (BS), the radio access technology (RAT), mobile country code (MCC), mobile network code (MNC), tracking area code (TAC), cell tower ID (CID), approximate area in meters where the cell possibly exists (range), and the number of samples recorded to prepare these data. When a device establishes a connection with the base station, it scans all the information of the above variables and each scan, denoted as *sample*, is recorded by the device and submitted to the database. The database also provides information about the BS location, whether it is obtained from the samples or provided to the database by the MNO itself.

The recorded variables in the database are the LTE identifiers as defined by 3GPP, which serve to provide unique and global identification throughout the network. For instance, consider the logical topology of the LTE cellular network inside the public land mobile network (PLMN) in Fig.3, with TACs (TAC-1 and TAC-2). Each TAC is defined by a unique code which represents the local tracking

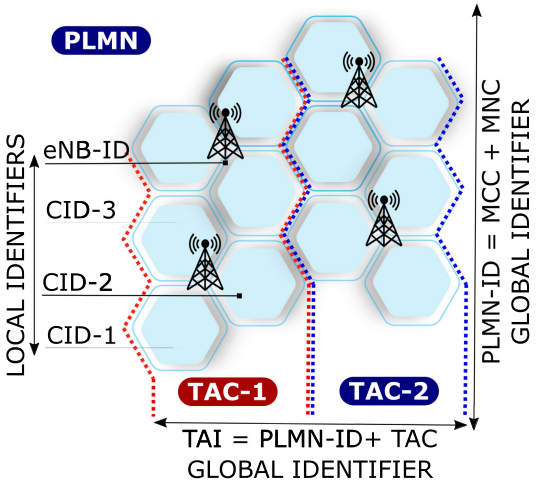

FIGURE 3. The standard LTE identifiers for tracking area identity (TAI) used to identify and process different variables/vectors in the considered database.

area inside the PLMN. In contrast, PLMN-ID represents a globally unique identifier of the network with the combination of MCC + MNC. The tracking area ID (TAI) is defined as a globally unique identifier in the network for a specific area corresponding to the combination PLMN-ID + TAC. Now, inside every tracking area (TA), each MNO has unique local IDs for each cell (CID) and each eNB (eNB-ID).

Nevertheless, eNB-IDs are not provided by the database to preserve the privacy of MNOs. Therefore, there exists the need for a selection procedure that extracts the information to uniquely identify the cells. The proposed procedure is introduced in Fig.1 (Network Data Acquisition) and consists of two phases, namely Pre-Selection and Post-Selection. In the first phase, we identify the top three MNOs and check that the RATs provided in the database are correct by comparing with a external data source. In the second phase, we first define the confidence benchmark sample (CBS) approach to ensure the quality of our analysis. Next, we combine two LTE identifiers to uniquely determine the cells, referred to as TAC + CID combination in the Post-Selection phase. Once the cells are uniquely identified, we obtain the region with highest traffic areas, denoted as HTTAC. From these tracking areas, we identify the LTE cells where MNOs select a feasible sub-area, denoted as 5G deployment area (5GDA). The criterion to select the 5GDA is to choose the region where cells are highly concentrated. A detailed discussion is provided in the next paragraphs, and the 5GDA selection through data visualization is described in Section IX-A.

Network data from OpenCellID database is used to obtain information about the state of the legacy network of MNOs related to traffic and subscriber density. The network database matrix $\mathbf{D}$ for Spain (MCC = 214) contains the raw data, being E the number of samples in millions, and each row $\mathbf{d}_e$ of $\mathbf{D}$ represents a data point $\mathbf{d}_e = (r_e, m_e, t_e, n_e, l_e, s_e)$. Hence, $\mathbf{D} := \{\mathbf{d_e}\}_{e=1}^{E}\ \forall\ \mathbf{d_e} \in \mathbb{R}^v$, where $v = 6$ represents the number

of data variables:

$$\mathbf{D} = \begin{pmatrix} r_1 & m_1 & t_1 & n_1 & l_1 & s_1 \\ r_2 & m_2 & t_2 & n_2 & l_2 & s_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r_e & m_e & t_e & n_e & l_e & s_e \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r_E & m_E & t_E & n_E & l_E & s_E \end{pmatrix},$$

- r_e: radio access technology (RAT) of row e.
- m_e: mobile network code (MNC) of row e.
- t_e: tracking area code (TAC) of row e.
- n_e: cell ID (CID) of row e.
- l_e: bi-dimensional location of the cell with latitude and longitude $l_e = (l_e^1, l_e^2)$.
- s_e: number of samples of row e.

For each MNO m we have a set of TACs $\mathcal{T}_m = \{t_1, \ldots, t_j, \ldots, t_J\}$, where t_j represents an individual TAC; a set of cells $\mathcal{N}_m = \{n_1, \ldots, n_k, \ldots, n_K\}$, where n_k represents the CID and $N_m = |\mathcal{N}_m|$ is the total number of cells for MNO m; and $s_m^{j,k}$ stands for the number of samples from cell k inside the j-th TAC.[1] We assume that the highest number of samples per cell represents the highest traffic density in that area. Therefore, the ID of the highest traffic TAC H_m, named as HTTACA, is obtained from the maximum aggregated number of samples of all existing cells inside the tracking area TAC as:

$$H_m = \arg_j \max H_m^j, \; H_m^j = \sum_{k=1}^{K} s_m^{j,k}, \quad \forall j = 1, \ldots, J. \quad (1)$$

Each MNO may have its own roll-out strategy concerning the deployment cost of the 5G infrastructure. Therefore, it is pertinent to identify the highest traffic area. We proceed by selecting N_c legacy cells inside H_m such that $N_c \subseteq \mathcal{N}_m$, being N_c those cells supporting the highest traffic density. The geographical area covering the N_c cells is represented by $A_m^{N_c}$. Besides, MNOs may have financial constraints for the deployment of 5G services in that area. Therefore, we define A_m^g as the area of high traffic concentration which is determined through data visualization where the gNBs will be deployed for 5G services such that $A_m^g \subseteq A_m^{N_c}$. These cells provide an estimate of the highest traffic areas where most of their subscribers are localized so that an MNO may start deploying new 5G services in order to minimize cost.

The dataset contains the coordinates of existing GSM, UMTS and LTE cells in Spain. In this study, we only consider the legacy LTE cells as the source to estimate current data utilization in the cellular networks, while GSM and UMTS traffic make a very little contribution. From these data, we find the highest density area, as we discuss next.

B. DATASET SELECTION

To guarantee the reliability of the dataset, we propose the following data selection procedure with a set of procedural checks. Two phases, named as Pre-Selection and Post-Selection, compose this procedure as follows (see Fig.1(Network Data Acquisition)).

In the Pre-Selection phase, we extract the relevant data related to the country of interest (MCC) and its networks (MNC) as described below:

- First, we extract the subset of network data belonging to Spain (country code MCC = 214).
- We select the top three MNOs through their respective market shares in Spain [54].
- We confirm that the LTE RATs provided by the OpenCellID database have been actually deployed by the MNOs by contrasting this information with [55].

The Post-selection phase is the core of the data selection procedure, ensuring the quality of our analysis, avoiding data redundancy, and revealing patterns to recognize higher density subscriber areas. To ensure the quality of our analysis, we define the confidence benchmark samples (CBS) interval as the number of samples of legacy cells in the range $[a, b]$ so that the existence of the N_c cells is guaranteed. Note that in the database there are cells whose recorded number of samples is small (e.g., lower than 100), that may well represent cells that were installed and then removed, reconfigured or merged. There are also a few cells that have a very large number of samples compared to the majority of cells (more than 2000), which are considered outliers. Therefore, we do not include those cells in our analysis by defining a and b as the lower and upper bound, respectively. This phase also contemplates the following aspects:

- There exist multiple samples of a single cell in the database. A single CID may correspond to multiple recorded GPS coordinates. To avoid redundancy of multiple cells at the same true location, we define the combination of TAC and CID to take individual cells.
- To estimate the actual traffic of each MNO, we assume that the areas (TAC) with the highest number of samples support the highest traffic density in that area (TA), referred to as highest traffic TAC area (HTTACA).
- We then select the N_c cells inside the HTTACA such that these cells have the highest number of samples, i.e. the number of connections established by UEs.
- Finally, the 5GDA sub-area $A_m^g \subseteq A_m^{N_c}$ is selected, representing the highest subscriber's density area.

C. DATA SELECTION AND EXTRACTION

The procedure *NetDataDriling* for data selection and extraction is described in this subsection and its pseudocode is provided in Algorithm 1. It requires the following inputs:

- The network database $\mathbf{D}$.
- The constant N_c for the selection of legacy cells.
- The top three MNOs represented with MNCs m.
- The constants a and b as lower and upper bound of the CBS approach.

The NetDataDriling's outcome is H_m, the HTTACA ID for each MNO m, and the set $\mathcal{N}_c$ with the N_c CIDs.

[1] Note that $s_m^{j,k} = 0$ if the cell is not part of the tracking area.

Algorithm 1 NetDataDrilling

1 **procedure** NetDataDrilling($\mathbf{D}, N_c, m, a, b$)
2 $S_{aux} \leftarrow \mathbf{0}$
3 $V_{aux} \leftarrow \mathbf{0}$
4 $W_{aux} \leftarrow \mathbf{0}$
5 **for** ***all*** m **do**
6 **for** ***all*** e **do**
7 **if** $(s_e < a \vee s_e > b)$ **then**
8 $\mathbf{D} \leftarrow \mathbf{D} \backslash \mathbf{d}_e$
9 **end**
10 **end**
11 $[\mathcal{T}_m, \mathcal{N}_m] \leftarrow$ UNIQUE($\mathbf{D}(t_e, n_e)$) $\forall e$
12 **for** ***all*** $t_e \in \mathcal{T}_m$
13 $k \leftarrow j : t_e = \mathcal{T}_m(j)$ **do**
14 **for** ***all*** e **do**
15 **if** $t_e = \mathbf{D}(t_e)$ *and* $\mathbf{D}(n_e) \in \mathcal{N}_m$ **then**
16 $S_{aux}(k) \leftarrow S_{aux}(k) + s_e$
17 **end**
18 **end**
19 **end**
20 $H_m = \max S_{aux}$
21 **for** ***all*** $e : \mathbf{D}(t_e) = H_m$ **do**
22 Find index $k : \mathcal{N}_m(k) = \mathbf{D}(n_e)$
23 $V_{aux}(k) \leftarrow V_{aux}(k) + s_e$
24 $W_{aux}(k) \leftarrow \mathcal{N}_m(k)$
25 **end**
26 $\mathcal{N}_c \leftarrow$ SORT(V_{aux}, W_{aux})
27 $\mathcal{N}_c \leftarrow \{n_{c,1}, \ldots, n_{c,k}, \ldots, n_{c,N_c}\}$
28 **end**

In the procedure, performed for each MNO (*for* loop in step 5), some auxiliary variables store intermediate results. S_{aux} represents the summation of the samples of those cells within each TAC, then having a dimension equal to the number of tracking areas. V_{aux} represents the summation of samples of each cell inside the HTTACA (with dimension equal to the number of cells forming the HTTACA) and the CIDs of these cells are stored in W_{aux}.

We first remove the data points $\mathbf{d}_e$ such that their number of samples is out of the CBS bounds, i.e. $a > s_e > b$ (steps 6-10). Next, we find the unique TAC and CIDs. Since there may exist multiple versions of the same cell (same CID) due to different recorded GPS coordinates, the combination of TAC and CID determines unique TACs and CIDs and eliminates redundancy. The process UNIQUE uses the pair (TAC, CID) to identify uniqueness and store the indices of TACs and CIDs in $\mathcal{T}_m$ and $\mathcal{N}_m$, respectively (step 11).

In steps 12-19, the total number of samples per TAC is stored in S_{aux} (step 16) by summing the samples of those CIDs associated with each TAC. We begin the loop in step 13, keeping in k the index of the TAC in $\mathcal{T}_m$ (given by j) to the TAC corresponding to the TAC (t_e) of that iteration e of the loop. For every TAC ID in $\mathcal{T}_m$, we match their IDs with those in the database $\mathbf{D}(t_e)$ and check whether the cell of the analyzed row e, denoted by $\mathbf{D}(n_e)$, is part of the MNO's network (step 15). Here, $\mathbf{D}(t_e)$ and $\mathbf{D}(n_e)$ from $\mathbf{D}$ represent respectively the TAC ID (3rd column) and the CID (4th column) in row e. In step 20, we find the tracking area with the highest traffic H_m for MNO m, corresponding to the TAC with the maximum number of recorded samples.

In steps (21-25), we calculate the total number of samples of the cells residing inside the target area H_m. So, once H_m is found in $\mathbf{D}$, the number of samples corresponding to the cells within H_m are stored in V_{aux} (step 23). The CIDs of the cells residing inside the target area H_m are then stored in W_{aux} (step 24). Finally, we need to identify the N_c cells supporting the highest traffic. In step 26, the cells corresponding to H_m are sorted according to the number of samples in decreasing order. Then, as shown in step 27, the selected N_c cells are the first N_c positions of the set $\mathcal{N}_c$, where $n_{c,k}$ denotes the CID of the cell with k-th highest traffic within H_m.

The complexity of Algorithm 1 is based on the size of the database $\mathbf{D}$ represented by the number of samples E. To process $\mathbf{D}$ a finite number of iterations is required given by the *for* loops. Therefore, the complexity of Algorithm 1 represented by the inner loops can be calculated as:

$$\mathcal{O}(E) + \mathcal{O}(E \times |\mathcal{T}_m|) + \mathcal{O}(|\mathbf{D}(t_e)|). \tag{2}$$

Note that (2) provides the complexity required for each MNO m and, in practice, this would be the total complexity as the samples from each MNO can be independently processed in parallel.

V. TRANSMISSION MODEL AND SERVICES

In this section, we present the transmission model, which is based on the bandwidth parts (BWPs), NR numerology and the definition of the three types of services. It is referred to as "Transmission Model" in Fig.1.

A. SERVICE DEFINITION

In 5G NR, the physical resources of the deployed gNBs can be allocated to several BWPs and configured via the numerology factor μ according to the data rate and latency requirements of the services. In this study, we consider three types of services as defined in the next paragraph, each service requiring different values of waveform parameters, namely sub-carrier spacing (Δf), number of slots per subframe (N^{μ}_{slot}) and slot duration (T_{slot}). Numerology in 5G is defined as the flexibility to use multiple values of Δf and is configured through parameter μ [56]. The values of Δf depends upon the selected frequency range. 5G NR supports two frequency ranges FR1<6GHz and FR2>6GHz. In this study, we consider FR1 for our scenario, which allows sub-carrier spacings of 15KHz, 30KHz and 60Khz for $\mu = 0$, $\mu = 1$ and $\mu = 2$ respectively, as summarized in Table 1.

B. TRANSMISSION MODEL

We consider a total transmission bandwidth of 50 MHz [57], where B_x denotes the BWP assigned to service x, with $x = \{1, 2, 3\}$. Service-1 serves eMBB devices requiring

TABLE 1. 5G transmission numerologies for FR1.

μ	Δf	N^{μ}_{slot}	$T_{slot} = 1/2^{\mu}$
0	15	1	1 ms
1	30	2	0.5 ms
2	60	4	0.25 ms

peak data rates of up to 100 Mbps and an acceptable latency of $T_{slot} = 0.5$ ms, with $\mu = 0$ corresponding to $\Delta f = 15$ Khz, $B_1 = 30$ MHz. Service-2 is intended to serve eMBB devices with peak rates of up to 15Mbps, with $\mu = 1$ corresponding to $\Delta f = 30$ Khz, $B_2 = 10$ MHz. Service-3 provides low data rates of up to 7 Mbps for URLLC/mMTC and lower latencies, for which we set $\mu = 2$ with $\Delta f = 60$ KHz, $T_{slot} = 0.25$ ms, $B_3 = 10$ MHz. The proposed multi-numerology configuration with three BWPs is shown in Fig. 4. For these three BWPs, the number of contiguous PRBs is $N_{RB} = \{106, 24, 11\}$, respectively. In particular, a large number of PRBs is suitable for eMBB traffic flows with moderate requirements of reliability and latency [58]. Therefore, eMBB services are configured with small sub-carrier spacings corresponding to numerologies $\mu = \{0, 1\}$ for $x = 1$ and $x = 2$, respectively. On the other hand, a larger sub-carrier spacing is required [59] to support ultra-reliability and very low latency in URLLC/mMTC. Consequently, service $x = 3$ is configured with $\Delta f = 60$KHz with $\mu = 2$ (see Table 1).

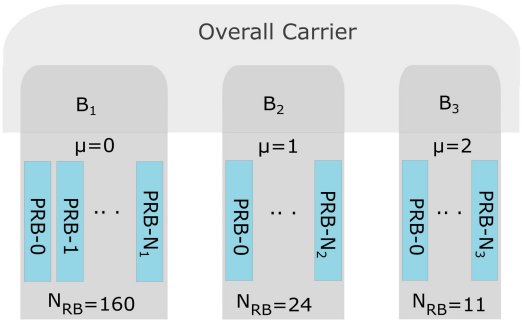

FIGURE 4. Considered BWPs with multi-numerology configuration of a gNB which is identical for each MNO.

Any 5G device supporting Multi-RAT Dual Connectivity can achieve a maximum transmission rate (ψ) given by [60]

$$\psi = 10^{-6} \cdot \left(\frac{V \cdot N_{RB} \cdot 12}{T_s^{\mu}} \cdot (1 - OH) \right), \qquad (3)$$

where $V = v_{layers} \cdot Q_m \cdot f \cdot R_{max}$, being v_{layers} the number of transmission streams (number of data layers) or maximum number of MIMO layers at the gNB, Q_m the modulation order, and f is the scaling factor and can take values 1, 0.8, 0.7 and 0.4. In our scenario, we consider $f = 1$ and $R_{max} = \frac{948}{1024}$ (maximum LPDC coding rate). The OFDM symbol period for each μ is given by $T_s^{\mu} = \frac{10^{-3}}{14*2^{\mu}}$ with 14 symbols per slot. The maximum resource block allocation per service x is N_{RB}. The overhead is represented by OH, with $OH = 0.14$ in for the DL and $OH = 0.08$ for the UL in FR1. In our scenario, Table 2 summarizes the values of v_{layers} and Q_m for the three services.

TABLE 2. Physical layer parameters of each BWP for supported capacity.

Parameters	B_1	B_2	B_3
v_{layers}	4	4	2
Q_m	8	8	6

VI. NETWORK DIMENSIONING

We now present our vision of network dimensioning, which will be fully developed in Sections VI-A (coverage dimensioning) and VI-B (capacity dimensioning), following the scheme of Fig.1. These two processes are integrated into the dimensioning heuristic algorithm presented in Section VII.

Coverage and capacity dimensioning are essential phases of the network planning process to determine the number of radio sites that satisfies the coverage and capacity requirements. Coverage dimensioning (Section VI-A) consists in computing the coverage cell range R_{cov} in the target area A^g_m to determine which service x provides the most restrictive coverage cell range and then we compute the required number of sites Z_{cov} based on this range. The calculation of R_{cov} assumes a certain cell load δ_{cov}. Capacity dimensioning (Section VI-B) determines the capacity cell range R_{cap} based on the effective capacity β of the gNB, subscriber's density Φ_m and the market share κ_m of MNO m. These two phases can ensure that the network design is balanced, i.e., the difference of cell ranges and cell loads from coverage and capacity dimensioning are under some given thresholds Δr and Δl, respectively. Therefore, the conditions $|R_{cov} - R_{cap}| \leqslant \Delta r$ and $|\delta_{cov} - \delta_{cap}| \leqslant \Delta l$ are imposed so that the two phases provide a similar solution in terms of radius and cell load. As a result, once the required number of sites Z_{cap} for capacity is obtained in the capacity dimensioning phase, the number of gNB sites is given by $Z = \min\{Z_{cov}, Z_{cap}\}$.

In this article, we study the DL RND scenario. The procedure can be straightforwardly adapted for the UL, given that the only difference is the value of a few parameters like UE transmit power P^{UE}_{tx}, antenna gain G_{UE} and interference margin I_L. The detailed discussion on coverage and capacity dimensioning is provided in the following sections.

A. COVERAGE DIMENSIONING

Coverage dimensioning refers to compute the required number of radio sites Z_{cov} and the coverage cell range R_{cov}, as shown in Fig.1, and follows the following steps. First, we decide the power scheme, which allocates transmit power to sub-carriers per service x following the CET approach [53]. Second, we determine the cell radius R_{cov} with particular emphasis on preventing coverage gap.

TABLE 3. List of variables and parameters.

Symbol	Description
A	Area coefficient 3-sectored site
$A_m^{N_c}$	Area of N_c number of cells of m-th MNO
A_m^g	High traffic concentration area
A_c	Area of active devices/user
B_x	BWP of service x
d_e	e-th data point in the database D
E	total number of samples (millions) in D
$E(r_x)$	Expected data rate per device
E_x	Effective power per sub-carrier
H_m	Highest traffic TAC ID of m-th MNO
L	Sum of all the losses during propagation
N_m	total number of cells of m-th MNO
N_x	total number of sub-carriers per x
N_{RB}^x	total number of physical resource blocks
N_{SC}	Number of sub-carriers in a resource block
N_{cet}	Number of PRBs to provide CET
N_{RB}^{thr}	Number of total resource blocks per
PL^{max}	Maximum allowed path loss
Q_m	Modulation order
r_x	Peak data rate per device
$\hat{r}_x$	Transmitted data rate per device at given time instant
r_x^{min}	Minimum data rate per subscriber
R_{max}	Maximum LPDC coding rate
R_{cov}^{DL}	DL coverage cell range
R_{cov}^{UL}	UL coverage cell range
R_{cov}	Coverage cell range from coverage dimensioning
R_{cap}	Capacity cell range from capacity dimensioning
$s_m^{j,k}$	Number of samples of k-th cell in j-th TAC
W_x^{max}	Maximum number of active devices per BWP
W_x	Number of active devices at given instant
Z_{cov}	Number of radio sites for coverage
Z_{cov}	Number of radio sites for capacity
Z	Minimum number of radio sites
v_{layers}	MIMO data streams
δ_{cov}	Assumed cell load
δ_{cap}	Actual cell load
Φ_m	Subscriber density
$\hat{\kappa}_m$	Market share
β	Effective capacity per gNB
ΔR	Difference in cell ranges
Δr	Thershold difference in cell ranges
Δl	Thershold difference in cell loads

1) POWER SCHEME

We first need to determine the distribution of PRBs among services. In 5G NR, the number of sub-carriers N_{sc} per PRB is 12 for all numerologies and services, as shown in Fig. 5, with N_{RB}^x denoting the number of PRBs per service x, each service x configured with different sub-carrier spacing Δf for $\mu = \{1, 2, 3\}$ and a total number of sub-carriers N_x.

In coverage dimensioning the role of the transmit power P_{tx} per gNB is essential when planning small cell deployment. Different cell types have different value parameters, as shown in Table 4 [61], so P_{tx} ranges from $P_{min} = 1$W to $P_{max} = 10$W dBm (micro-cell scenarios). The power available for service x is calculated as

$$P_{tx}^x = P_{tx} - 10\log_{10}\frac{N_{RB}^x}{N_{RB}^{thr}}, \tag{4}$$

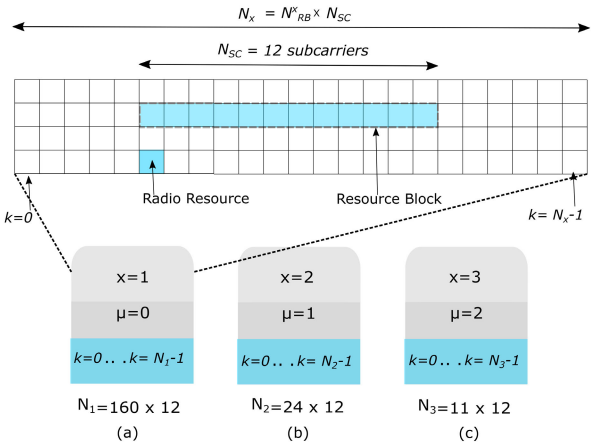

FIGURE 5. 5G resource grid for service x, PRBs per BWP.

TABLE 4. Cell classification.

Parameters	Femto cells	Pico cells	Micro cells	Macro cells
Transmit power	1-250mW	250mW-1W	1-10W	10-50+ W
Cell radius	10-100m	100-200m	0.2-2km	8-30km
Users	1-30	30-100	100-2000	2000+

being N_{RB}^{thr} the total number of PRBs per gNB and the effective power per sub-carrier for service x is $E_x = P_{tx}^x - 10\log_{10}(N_x)$[dBm]. Besides, we base our power allocation on the cell edge throughput (CET) criterion. To guarantee CET availability, N_{cet}^x PRBs are allocated to one user at the cell edge and the power associated with the sub-carriers of these PRBs can be computed as:

$$E_x^{cet} = E_x + 10\log_{10}(N_{cet}^x \times N_{SC}). \tag{5}$$

2) DETERMINATION OF THE CELL RADIUS R_{cov}

Determining R_{cov} implies evaluating the maximum path loss through link budget analysis, which includes all potential gains and the losses from the transmitter (Tx) through the channel to the receiver (Rx). We adopt the UMi (Urban Micro) 3GPP channel model that considers the path loss as the dominant factor in line-of-sight micro-cell outdoor scenarios [62]. The UMi path loss model is characterized by the 3D distance, which represents the coverage cell range R_{cov} and can be determined as [62], [63]

$$\begin{aligned} PL^{max} = {} & 32.4 + 40\log_{10}(R_{cov}) + 20\log_{10}(f_c) \\ & - 9.5\log_{10}((d'_{BP})^2 + (h_{gNB} - h_{UE})^2) + \sigma, \end{aligned} \tag{6}$$

where f_c is the carrier frequency, d'_{BP} is the break-point distance,[2] σ is the shadowing factor, h_{gNB} and h_{UE} are the effective antenna heights of gNB and UE, respectively. The maximum allowed path loss PL_x^{max} determines the maximum cell range of a gNB considering the transmit power level and

[2]The breakpoint distance is defined as the distance that separates two different trends of path loss that follows two different path loss exponents.

the propagation loses. We calculate PL_x^{max} for each service x with the CET criterion as:

$$PL_x^{max} = E_x^{cet} - L, \tag{7}$$

with $L = I_L + P_L + S_{UE} + \sigma$ representing the total channel loss that includes the interference margin I_L, P_L is the penetration loss, S_{UE} is the UE sensitivity, and σ represents the shadow fading. Note that I_L is the most relevant parameter in the coverage dimensioning as it accounts for the cell load δ_{cov} at which the link budget is performed, with $I_L = -10\log_{10}(1 - \delta_{cov})$. For an isolated cell we have $I_L = 0$, i.e. there are no neighbouring interfering cells. In our study we assume that the neighbouring cells are equally loaded with δ_{cov} in the busy hour, which corresponds to the interference margin I_L.

a: COVERAGE RANGE CALCULATION

Our proposal is to calculate the coverage cell range R_{cov}^x for each service x and select the most restrictive range among them as the coverage cell range denoted by R_{cov}. The first reason to adopt this criterion is that it provides a good trade-off between QoS and cost, as a larger coverage range reduces the number of gNBs at the cost of increasing inter-cell interference from neighbouring cells, which may deteriorate network performance. Second, a larger range would mean a higher number of devices, while a gNB can only satisfy user' demands if the number of devices is below a given threshold.

Once R_{cov} is determined, we can compute the required number of sites Z_{cov} for coverage. Considering three-sector cells design, the site area is calculated as $A \times R_{cov}^2$ [6], where $A = 1.95$ is the area coefficient, and the number of sites required for coverage is

$$Z_{cov} = \frac{A_m^g}{A \times R_{cov}^2}. \tag{8}$$

B. CAPACITY DIMENSIONING

Capacity dimensioning refers to determining the required number of radio sites Z_{cap} and capacity cell range R_{cap}. Conventional capacity dimensioning approaches find the cell range based on the maximum capacity assumption and the estimated traffic density in the area [6]. In our study, we propose a capacity model based on probabilistically characterizing the RRC (Radio Resource Control) states of 5G NR to determine the effective capacity of the gNB site.

Our model consists of three phases and involves the threshold number of allowed served devices W_x^{thr}, the probability p of the devices to be active, the requirement of MNO's peak data rate per user r_x for each service x, the novel parameter ζ that guarantees the availability of r_x for a required percentage of the time, and subscribers density Φ_m, as shown in Fig.1.

The first step is determining site capacity β based on the input parameters and requirements set by the MNO, including a capacity model that probabilistically characterizes the RRC states to determine W_x^{thr} under ζ and r_x requirements. Next, we find the gNB area A_c. Finally, we compute the capacity cell range R_{cap} to determine the number of sites Z_{cap}. The following subsections develop these three steps.

1) CAPACITY MODEL

In this section, we develop the capacity model based on the NR RRC states to determine the threshold number of devices that can be served by the gNB, represented by W_x^{thr}, such that the MNO requirement of guaranteeing the peak data rate r_x for at-least ζ percent of the time is fulfilled. To guarantee the peak data rate per user, the number of active devices must be limited as each BWP has a finite number of PRBs (N_{RB}).

Any device in the network can only be in one of the three RRC states, namely *rrc_idle*, *rrc_inactive* and *rrc_connected* (see Fig.6.). The idle state is the same as in LTE. Once an RRC connection is established there are two possible states, i.e., inactive and connected. The *rrc_connected* state acquires radio resources during the session according to the required data rate, and if there is no activity from the device, the session is suspended and the device passes to the *rrc_inactive* state. The *rrc_inactive* state is used to receive small payloads or short bursts. We characterize the RRC states with probabilities p_1, p_2, and p_3 for *rrc_idle*, *rrc_inactive* and *rrc_connected*, respectively such that $p_1 + p_2 + p_3 = 1$. Any user in the *rrc_idle* state is identified as non-active, while users either in the *rrc_inactive* or *rrc_connected* state are classified as active (see Fig.6).

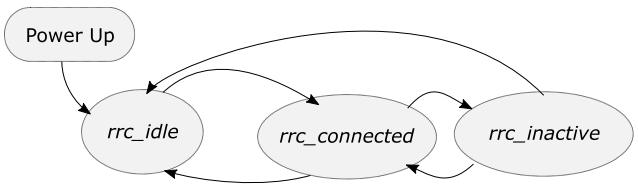

FIGURE 6. RRC state model for the 5G radio access network differentiating the *rrc* states.

Let r_x denote the peak data rate per user set by an MNO for service x. We define W_x as the random variable that represents the number of active devices at a given random instant. As any device in the network can only be non-active or active at a time, a 5G device's activity can be modelled as a sequence of independent and identically distributed (iid) random variables. This means that either the device is active (*rrc_inactive* or *rrc_connected*) with probability $p = p_2 + p_3$ or non-active with probability $1 - p = p_1$. With these assumptions, W_x follows a binomial distribution $W_x \sim B(W_x^{thr}, p)$ [64] with probability density function:

$$\begin{aligned} P(W_x = k) &= \binom{W_x^{thr}}{k} p^k (1-p)^{W_x^{thr}-k} \\ &= \binom{W_x^{thr}}{k} (p_2 + p_3)^k (p_1)^{W_x^{thr}-k}, \\ & k = 0, 1, \ldots, W_x^{thr}, \quad 0 \le W_x \le W_x^{thr}. \end{aligned} \tag{9}$$

The peak data rates r_x are usually optimistic and can be lower due to propagation conditions and the number of active

devices W_x. We define the random variable $\hat{r}_x$ representing the data rate transmitted per device for service x, which depends on W_x at a given time instant and is calculated as:

$$\hat{r}_x(W_x) = \min\left\{\frac{\psi_{DL,x}}{W_x}, r_x\right\}, \tag{10}$$

where $\hat{r}_x \leq r_x$ and $\psi_{DL,x}$ is the DL transmission rate defined by (3). We choose the DL as the representative case, compared to the UL, as significantly larger data rates are generaly required for DL applications. Note that the active devices may observe a data rate larger than $\frac{\psi_{DL,x}}{k}$ when the number of active devices is below a given k, and the probability associated to this event is

$$P\left(\hat{r}_x \geq \frac{\psi_{DL,x}}{k}\right) = P(W_x \leq k). \tag{11}$$

The two relevant metrics in the model are the average expected rate per active device $E(\hat{r}_x)$ for a given W_x^{thr} and the probability p. The value p also refers to the cell load, as it corresponds to the percentage of active devices. Hence, $E(\hat{r}_x)$ must consider the active devices W_x with their respective probabilities as

$$E(\hat{r}_x) = \sum_{k=0}^{W_x^{thr}} \hat{r}_x(W_x) P(W_x = k). \tag{12}$$

The expected number of active devices per service x can be calculated as:

$$E(W_x) = pW_x^{thr}. \tag{13}$$

For the sake of capacity dimensioning, our goal is to determine the threshold number of devices W_x^{thr} such that r_x is provided for at least ζ percentage of time. It is clear that the peak data rate r_x is only achievable when the random number of active devices W_x becomes $W_x = W_x^{max}$, where $W_x^{max} = \lfloor \frac{\psi_{DL,x}}{r_x} \rfloor$ is the maximum number of active devices to provide r_x in the given transmission bandwidth $\psi_{DL,x}$. Note that, due to the limited number of PRBs, W_x^{thr} takes a range of discrete values such that $W_x^{max} \leqslant W_x^{thr} \leqslant W$, where W is a seed value. From this discussion, we see that ζ gives the probability that W_x takes a value below or equal to W_x^{max} as:

$$\begin{aligned}\zeta &= P\left(W_x \leq W_x^{max}\right) \\ &= \sum_{k=0}^{W_x^{max}} \binom{W_x^{thr}}{k} (p_2 + p_3)^k \left(p_1\right)^{W_x^{thr}-k}.\end{aligned} \tag{14}$$

The calculation of W_x^{thr} from (14) is not straight but requires an iterative process such that the required value of ζ is fulfilled. Therefore, the proposed capacity model is presented as a procedure in Algorithm 2, which is part of the RND heuristic proposed in Section VII. The inputs are a seed value W, the probability of active devices p, the peak data rates r_x, and the maximum transmission capacity $\psi_{DL,x}$. We perform capacity modelling for each service x (*for* loop in step 4) to specifically determining the threshold number of served devices W_x^{thr}. We compute the PDF $P(W_x = k)$, expected data rate $E(\hat{r}_x)$, expected number of active devices $E(W_x)$ and ζ for the range (W_x^{max}, W_x^{thr}) in steps 5-10. Finally, if the computed value $P(W_x \leqslant W_x^{max})$ is greater or equal to the required ζ (steps 11-13), the result is updated.

Algorithm 2 Capacity Modelling

```
1  procedure CapModelling(W, ζ, ψ_{DL,x}, p, r_x)
2  result ← 0
3  W_x^max ← ⌊ψ_{DL,x} / r_x⌋
4  for all x do
5      for all W_x^max to W_x^thr do
6          calculate P(W_x = k) with (9)
7          calculate E(r̂_x) with (12)
8          calculate E(W_x) with (13)
9          calculate P(W_x ⩽ W_x^max) with (14)
10     end
11     if P(W_x ⩽ W_x^max) ⩾ ζ then
12         update result;
13     end
14 end
15 result ← [W_x^thr]
```

Next, we provide an example to illustrate how the value of W_x^{thr} is calculated by fulfilling the MNOs requirement of ζ.

2) NUMERICAL EXAMPLE

We consider $W_x^{thr} = 130$ as the threshold number of devices per gNB, BWPs of 25 MHz with maximum transmission capacity of $\psi_{DL,x} = 1138.5$ Mbps, and peak data rate $r_x = 30$ Mbps. To guarantee r_x, it is clear that the maximum number of active devices must be $W_x^{max} = 37$. For an MNO that wants to provide r_x for at least $\zeta = 40\%$ of the time, we need to find W_x^{thr} per gNB such that $P\left(W_x \leq W_x^{max}\right) \geqslant 40\%$. Also, from empirical observations reporting that only a small number of devices are active [65], [66], we consider $p_3 = 0.15$ and $p_2 = 0.20$. We compute the average number of active devices as:

$$E(W_x) = pW_x^{thr} = 45 \text{ devices},$$

and the average expected data rate as per (12) is $E(\hat{r}_x) = 25.2$ Mbps with probability $P(W_x = 45) = 2.1\%$.

We can now compute ζ to check whether the MNO's requirement $\zeta \geqslant 40\%$ is met or not. With $W_x^{thr} = 130$, the calculated valued of ζ is 6.9%, i.e., the peak data rate r_x is only guaranteed for 6.9% of the time, which is very less than the required 40%. This means that for $W_x^{thr} = 130$ the given requirement of $\zeta \geqslant 40\%$ is not fulfilled. Therefore, we need a smaller value of W_x^{thr} to ensure $\zeta \geqslant 40\%$. By iterating (14), we find that $W_x^{thr} = 108$ ensures r_x for $\zeta = 48\%$. Given that $W_x^{thr} = 108$, the average expected data rate becomes $E(\hat{r}_x) = 28.7$ Mbps.

3) CELL RANGE

The capacity cell range R_{cap} is computed based on the target area and the supported cell load δ_{cap}, which represents the

percentage of W_x^{thr} active devices. These variables are related to the active area A_c as [6]:

$$A_c = \frac{\beta \times \delta_{cap}}{\Phi_m}, \tag{15}$$

where A_c is the area corresponding to a gNB site serving $\beta \times \delta_{cap}$ subscribers with subscriber density Φ_m per km^2, being β the site capacity of a s-sectored gNB site given by

$$\beta = s \sum_x W_x^{thr}, \quad \forall x. \tag{16}$$

The site capacity is based on W_x^{thr} ensuring the threshold number of devices per gNB per service x. As A_c can be calculated for three-sectored sites as $A_c = A * R_{cap}^2$ [6], with $A = 1.95$, the capacity cell range R_{cap} is given by

$$R_{cap} = \sqrt{\frac{A_c}{A}}. \tag{17}$$

C. DIMENSIONING EXAMPLE

To illustrate the linkage between coverage and capacity dimensioning, we present a numerical example. We consider that a service x can support $W_x^{thr} = 38$ devices for $\zeta \geqslant 50\%$, based on the capacity model of VI-B1. In this case, a maximum three-sectored capacity $\beta = 114$ subscribers is obtained with (16). We assume that the mth MNO has a subscriber density $\Phi_m = 200$ user/km^2. Given threshold values for Δr and Δl are set to 10 meters and 1%, respectively, to achieve balanced network design.

We first perform coverage dimensioning according to VI-A on the target area A_m^g with $P_{tx} = 40$ dBm and cell load $\delta_{cov} = 40$, obtaining a coverage radius $R_{cov} = 0.350$ km for a certain service x. On the other hand, following the capacity dimensioning of VI-B, we obtain $\delta_{cap} = 44.7$ and a single gNB site can achieve $\beta \times \delta_{cap} = 50$ active subscribers with $A_c = 0.248$ from (15), and from (17) we have $R_{cap} = 0.357$ km. It turns out that R_{cap} is very close to $R_{cov} = 0.350$ km, giving a difference between cell ranges $\Delta R = 7$ meters which is under $\Delta r = 10$ meters. However, the network design is not balanced because the difference between the coverage and capacity cell load is $\Delta L = 4.7$, which exceeds the given threshold $\Delta l = 1$. In this case the design is under-dimensioned as $\delta_{cov} < \delta_{cap}$, and the network would have over-utilized gNBs.

In the next section, we propose a network dimensioning heuristic to achieve balanced network design.

VII. HEURISTIC DIMENSIONING ALGORITHM

This section presents the network dimensioning heuristic algorithm referred to as Heuristic Dimensioning in Fig.1 and whose pseudocode is given in Algorithm 3. While there exist procedures and algorithms for network dimensioning of legacy radio networks (see for instance [6], [67]), we recall that, to the best of our knowledge, this is the first network dimensioning proposal taking into account multi-numerology radio configuration. The algorithm is based on the following inputs and outputs:

Input

- Deployment area A_m^g.
- Transmit power level P_{tx}.
- Subscriber density Φ_m per MNO m.
- Design parameters Δr and Δl.
- Number of sub-carriers N_x per service x.
- Number of PRBs N_{RB}^x and N_{cet} per BWP B_x.

Output

- The cell range R.
- The minimun number of radio sites Z.
- Transmit power level P_{tx}.

NetDimensioning calls the CapModelling (Algorithm 2) procedure to get the threshold number of devices W_x^{thr} (step 3). Based on W_x^{thr}, the gNB capacity β is then calculated in step 4 from (16). Next, network dimensioning is performed for each MNO m (*for* loop in step 5). Steps 6-12 implement coverage dimensioning for each service x to find the power distribution scheme E_x^{cet} from (5), represented by the process PowerScheme in Algorithm 3, and the coverage radius R_{cov}^x according to (6) and (7), represented by the process UMiModel in Algorithm 3. being ΔP in step 10 the transmit power granularity to increment P_{tx} for the next iteration. We then find which service x provides the most restrictive coverage cell range R_{cov} (steps 13-14).

The capacity dimensioning process consists of steps 15-25 and iterates over the cell load δ_{cap}. With the corresponding value of δ_{cap}, we compute the active area A_c from (15) and the capacity cell range R_{cap} from (17) (steps 16 and 17). Next, the difference between coverage and capacity cell range $\Delta R = |R_{cap} - R_{cov}|$ and the difference $\Delta L = |\delta_{cap} - \delta_{cov}|$ are updated (steps 18 and 19).

The criteria to get a balanced network design is applied in step 20. First, we need that ΔL is under the established threshold Δl. Second, ΔR must be under the established threshold Δr. If these two conditions are not fulfilled the dimensioning is imbalanced, potentially causing coverage gaps or overlaps with over or under-utilized base stations (see [6] for details). If the conditions in step 20 are true, we compute the number of radio sites Z_{cov} and Z_{cap} for coverage and capacity, respectively (step 21-22). In step 24, ΔQ represents the capacity cell load granularity to increment δ_{cap} for the next iteration. Finally, we find the index z which represents the minimum number of required radio sites $Z(z)$ (step 26), the corresponding cell range $R(z)$ and the transmit power level $P_{tx}(z)$ (step 27).

The complexity of Algorithm 3, for each MNO, can be calculated similarly to that for Algorithm 1 as

$$\mathcal{O}(x \times |\Delta P|) + \mathcal{O}(|\Delta Q|). \tag{18}$$

Again, (18) would be the total complexity in practice as the procedures for the MNOs can be independently executed in parallel.

Algorithm 3 NetDimensioning

procedure NetDimensioning(A_m^g, P_{tx}, Φ_m, N_x, ... N_{cet}, Δr, Δl)
  $result \leftarrow \mathbf{0}$
  $W_x^{thr} \leftarrow$ CapModelling(W, ζ, $\psi_{DL,x}$, p, r_x)
  calculate β from (16)
  for ***all*** m **do**
    for ***all*** x **do**
      for $P_{tx} \leftarrow P_{min}$ *to* P_{max} **do**
        $E_x^{cet} \leftarrow$ PowerScheme(P_{tx}, N_x, N_{cet})
        $R_{cov}^x \leftarrow$ UMiModel(E_x^{cet}, L, δ_{cov})
        $P_{tx} \leftarrow P_{tx} + \Delta P$
      end
    end
    Find index k : min(R_{cov}^x)
    $R_{cov} \leftarrow R_{cov}^x(k)$
    for $\delta_{cap} \leftarrow \delta_{min}$ *to* δ_{max} **do**
      $A_c \leftarrow \beta \times \delta_{cap}/\Phi_m$
      $R_{cap} \leftarrow \sqrt{A_c/A}$
      $\Delta R \leftarrow |R_{cov} - R_{cap}|$
      $\Delta L \leftarrow |\delta_{cov} - \delta_{cap}|$
      if $(\Delta L \leqslant \Delta l) \wedge (\Delta R \leqslant \Delta r)$ **then**
        $Z_{cov} \leftarrow A_m^g/A \times R_{cov}^2$
        $Z_{cap} \leftarrow A_m^g/A_c$
      end
      $\delta_{cap} \leftarrow \delta_{cap} + \Delta Q$
    end
    Find index z : min(Z_{cov}, Z_{cap})
    $result \leftarrow [Z(z), R(z), P_{tx}(z)]$
end

VIII. CONVENTIONAL RND

As a benchmark, we use the LTE conventional radio network dimensioning (CRND) of [6], [67]. The main differences between our approach and CRND are outlined next.

Regarding subcarrier spacing, the CRND approach does not consider the use of multiple numerologies, being based on a fixed sub-carrier spacing Δf irrespective of the data rate and latency required by different services. Therefore, the bandwidth can only be configured with the fixed sub-carrier spacing Δf to serve devices with a peak data rate r_x for a single service x. With respect to transmit power, it is uniformly distributed on all the sub-carriers and PRBs in CRND, and the EIRP per sub-carrier can be computed as:

$$E_x = P_{tx} - 10\log_{10}(N_{RB}^{thr} \times N_{SC}), \tag{19}$$

where N_{RB}^{thr} is the total available PRBs and N_{SC} is number of sub-carriers per resource block. To fulfil the CET criterion, the EIRP E_x^{cet} per sub-carrier can be computed with (5). In CRND, the key factor in determining site capacity is the burst traffic, which is based on the peak to average rate ratio. The site capacity β is then computed as [67]:

$$\beta = \frac{s \times (\psi_{DL,x} \times \delta_{cap})}{(1+\lambda) \times r_x^{min}}, \tag{20}$$

where s is the number of sectors and the term $(\psi_{DL,x} \times \delta_{cap})$ represents the designed DL cell capacity which depends on the transmission capacity $\psi_{DL,x}$ and the capacity cell load δ_{cap}. The term λ is the factor to accommodate burst traffic, known as peak-to-average margin. In LTE, a typical value is $\lambda = 10 - 20\%$ for capacity dimensioning [67].

The conventional capacity model does not provide QoE, in the case of this paper what percentage of time the peak data rates r_x are guaranteed. On the contrary, our capacity model guarantees r_x for at-least ζ percent of the time, as (14) shows, by estimating the number of served devices W_x^{thr}, which is afterward used to compute the site capacity β from (16). Note that QoE has become critical given the nature of next-generation cellular networks to provide mission-critical services [68].

IX. RESULTS AND ANALYSIS

This section presents the results and the corresponding analysis. In the first section, we discuss the results on network data acquisition and valuable insights for network dimensioning. The second part includes the evaluation of our proposed heuristic for the top three Spanish MNOs in terms of coverage, capacity, QoE and cost. We also compare our proposed heuristic with the conventional RND approach.

A. NETWORK DATA PROCESSING

The first step is the creation of the database **D**, used by the NetDataDrilling algorithm, from OpenCellID data. The accumulative samples for Spain corresponding to GSM + UMTS + LTE recorded per MNO are 2.7, 2.6 and 3.1 millions for MNO1, MNO2 and MNO3, respectively, which are significant enough for our analysis. We plotted the samples corresponding to MNO1 around Madrid region as shown on the map in Fig.7(a) by plotting the cell base stations. MNO2 and MNO3 have similar plots. The NetDataDrilling algorithm imposes the CBS approach by setting $a = 100$ and $b = 1000$. The LTE cells fulfilling the CBS criterion are shown in Fig.7(b). These cells are in thousands (3938, 3404, 3665) referring to MNO1, MNO2 and MNO3, respectively.

Unique TACs and CIDs are now identified. Checking of individual LTE cells is applied to the region plotted in Fig.7(b) with the combination of unique TAC ($\mathcal{T}_m$) and CIDs ($\mathcal{N}_m$) to obtain the number of unique LTE cells. The output of the combination shows that MNO1 has 3895, while MNO2 and MNO3 have 3296 and 3408 cells, with market shares $\kappa_1 = 23.55\%$, $\kappa_2 = 25.79\%$ and $\kappa_3 = 30.14\%$, respectively [54].

Finally, the 5G Deployment Area Selection (A_m^g) is determined. The network data analytics guide MNOs in the planning process to select those highly concentrated areas known as 5GDAs for new deployments. The HTTAC H_1 of MNO1 is shown in Fig.8(a). We found that the areas covered by H_1 and H_2 both lies in the centre of Madrid, while H_3 is located in the centre of Barcelona. To locate the 5GDA, we set $N_c = 100$ to find the highest traffic cells residing within each H_m, as Fig.8(b) shows for MNO1. The HTTAC of MNO1 exhibits

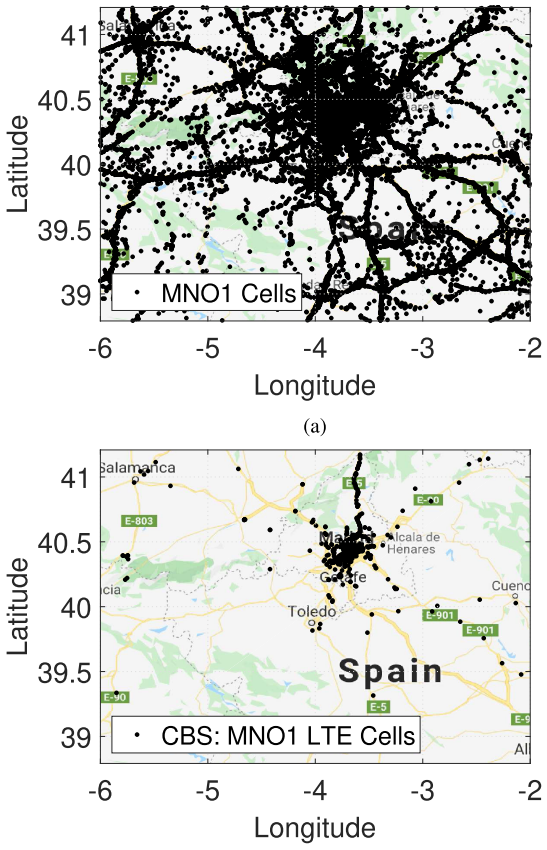

FIGURE 7. (a) Accumulative samples (GSM + UMTS + LTE) (b) LTE samples as a result of CBS for MNO1.

a head/tail pattern, in which a high density of cells represents the head of the HTTAC and the tail represents the lower concentration of cells. These type of patterns help the network planning engineer to locate the high-concentration sub-area $A_m^g \subseteq A_m^{N_c}$ for 5G deployments and perform the demarcation of the area as shown in Fig.8(c) for MNO1. The demarcated area A_m^g is an essential input to the NetDimensioning (Algorithm 3) from the *Network Data Acquisition* entity. The other demographic inputs, namely subscriber densities Φ_m and market shares κ_m per MNO, are shown in Table 5, which are required by Algorithm 3.

TABLE 5. Demographic inputs for network dimensioning.

MNOs	Φ_m (/km^2)	κ_m(%)	A_m^g (km^2)
MNO1	183	23.55	23.69
MNO2	200	25.79	25.20
MNO3	256	30.14	23.39

B. NETWORK DIMENSIONING

In this section, we discuss the most relevant parameters of our method. First, we present the results analyzing the effect of the threshold probability ζ on QoE, peak data

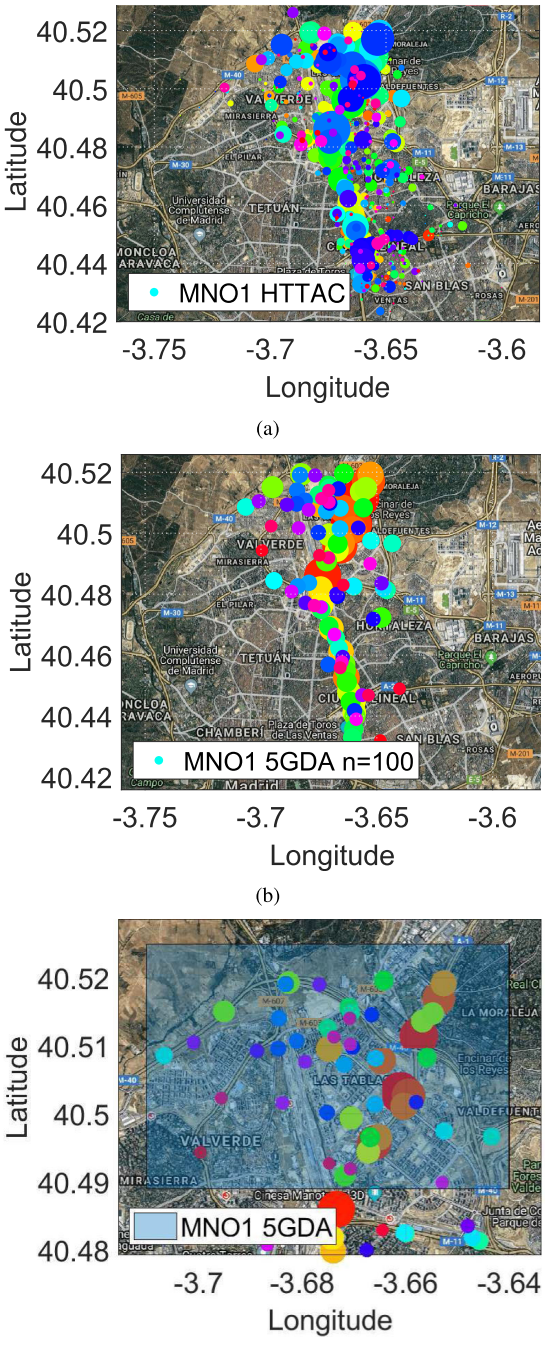

FIGURE 8. Coloured circles with largest radius represents the LTE cells with highest traffic (a) Highest traffic area represented by HTTACs (b) HTTAC of MNO1 showing only n = 100 LTE cells (c) Demarcation of the 5GDA inside the highest traffic area.

rate r_x, transmission bandwidth B_x and RRC state probabilities p, p_1, p_2, p_3. Next, we discuss coverage, capacity and cost minimization of the new gNB deployments concerning balance dimensioning.

The simulation covers a rectangular deployment area A_m^g given in Table 5 where users are distributed uniformly and randomly across the area with density Φ_m(/km^2) in a Line-of-Sight urban microcell scenario [62]. The wireless link

parameters are provided in Table 6. The curves have been generated by averaging the results from 100 executions of the NetDimensioning algorithm, each corresponding to one independent users distribution.

TABLE 6. Network dimensioning parameters.

Symbol	Description	Value	Unit
d'_{BP}	Breakpoint distance	$4h'_{gNB}h'_{UE}f_c/c$	meters
f_c	Carrier frequency	3500	MHz
G_{gNB}	gNB antenna gain	3	dB
G_{UE}	UE antenna gain	0	dB
$hgNB$	gNB antenna height	10	meters
h_{UE}	UE antenna height	1.5	meters
h'_{gNB}	Effective antenna height of gNB	$h_{gNB} - h_E$	meters
h'_{UE}	Effective antenna height of UE	$h_{UE} - h_E$	meters
h_E	Effective environment height	1.0	meters
I_L (a)	Interference margin	1.8(DL), 1(UL)	dB
L	Total losses	(a+b+c+d)	dB
N^{μ}_{RB}	Resource blocks	106, 24, 11	–
N^{μ}_{x}	Sub-carriers	$N^{\mu}_{RB} \times N^{RB}_{sc}$	–
P^{UE}_{tx}	UE transmit power	24	dBm
P_L (b)	Penetration loss	20	dB
S_{UE} (c)	UE sensitivity	-120	dB
σ (d)	Shadow fading	$\sim \mathcal{N}(0, 7)$	dB
W	Seed value for capacity model	500	–
Δf	Sub-carrier spacing	15, 30, 60	KHz
δ_{cov}	Assumed cell load	20	%
Δr	Threshold cell range difference	10	meters
Δl	Threshold cell load difference	1	%

1) PROPOSED RND

We first show the difference in cell range R^x_{cov} obtained for the three services as a function of gNB transmit power (see Fig.9). We observe that service $x = 1$ has the most restrictive cell range as it demands the highest data rate. Next, we check the effect of the transmission bandwidth B_x and peak data rate r_x, particularized on service $x = 1$, with B_x and r_x fixed for the other two services as Table 7 shows.

We fix $\zeta \geqslant 50\%$ and the probabilities are set to $p_2 = 0.15$ and $p_3 = 0.25$ for *rrc_inactive* and *rrc_connected*, respectively, for simulation purposes. We observe that the largest B_x, the highest the site capacity β, as shown in Fig.10, as a consequence of the increase in the available PRBs for $x = 1$. The higher number of PRBs enhances the transmission rates ($\psi_{DL,x} = \{222, 453, 684\}$) and allows to accommodate a higher number of served devices W^{thr}_x given that the condition $\zeta \geqslant 50\%$ is fulfilled. At the same time, Fig.10 captures the effect of the peak data rate r_x on β, which lessens with higher rates because it decreases the value of W^{max}_x required to guarantee the provision of r_x for $\zeta \geqslant 50\%$.

In Fig.11 we show the cell ranges of the most restrictive service $x = 1$ with the varying peak data rates r_x and transmission bandwidths $B_1 = \{10, 20, 30\}$. We find that higher peak data rates decrease the coverage and capacity cell ranges R_{cov} and R_{cap}, respectively. The fluctuations of R_{cov} and R_{cap} are also linked with radio sites Z_{cov} and Z_{cap} as given in (8) and (17) for coverage and capacity, respectively. We also see, from Fig.12, that the number of radio sites responds as a non-decreasing function of r_x. Therefore, a large number of sites is required for a cell with low capacity β to provide higher peak data rates r_x, as Fig.12 shows. Hence, the trade-off in providing higher peak data rates results in lower site capacity β, shorter cell ranges (R_{cov},R_{cap}) and deploying more radio sites to cover the same area. For instance, for $r_x = 100$ Mbps and 50 MHz bandwidth we need $Z_{cap} = 95$ compared to the case of $r_x = 500$ Mbps where the number of sites is larger, $Z_{cap} = 108$, so increasing deployment costs.

Varying requirements of r_x implies a variation in the active area of the subscribers A_c, as shown in Fig.13. We recall that A_c in (15) is the area corresponding to a gNB site serving the subscribers represented by $\beta \times \delta_{cap}$ with density Φ_m. A higher r_x decreases the site capacity β, leading to smaller A_c according to (15). Therefore, the non-increasing behaviour of A_c with respect to r_x explains the fact that the coverage is provided to a less number of subscribers.

In Fig.14, we see that the two conditions of balanced dimensioning are fulfilled since ΔR and ΔL are below the

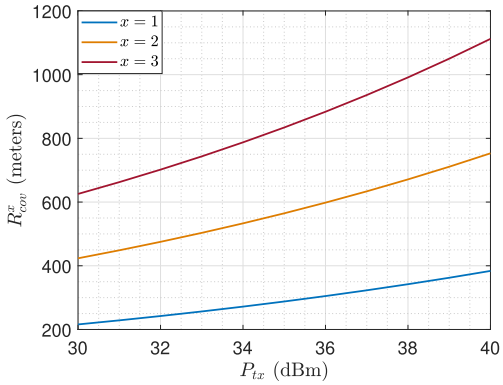

FIGURE 9. Effect of P_{tx} on coverage cell range R^x_{cov} of x services in a micro cell scenario.

TABLE 7. Peak data rates and transmission rates per BWP for x services.

Services	r_x (Mbps)	$\psi_{DL,x}$ (Mbps)	B_x (MHz)
$x = 1$	100-500	$\{222, 453, 684\}$	$\{10, 20, 30\}$
$x = 2$	15	205	10
$x = 3$	7	70	10

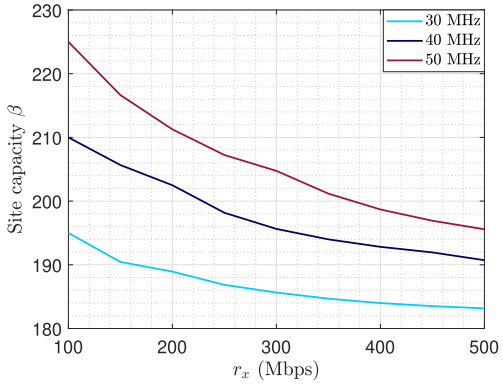

FIGURE 10. Site capacity β versus r_x is obtained by varying $B_x = \{10, 20, 30\}$ for $x = 1$ which increases the overall carrier bandwidths to $\{30, 40, 50\}$ MHz.

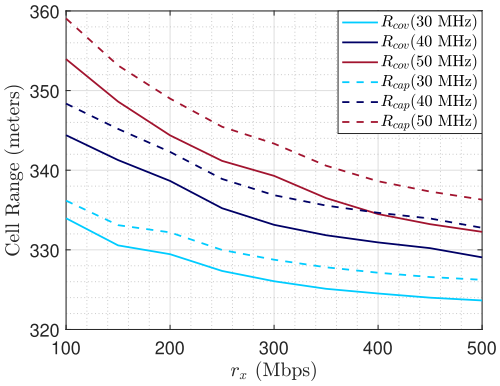

FIGURE 11. Coverage and capacity cell range for service $x = 1$.

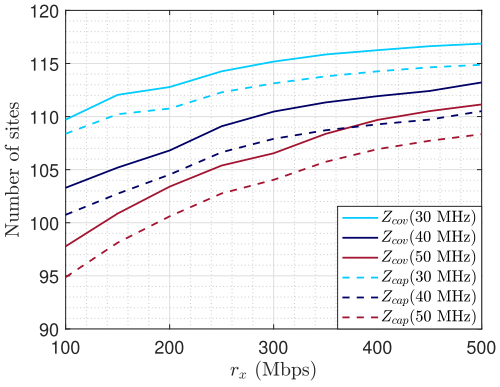

FIGURE 12. Required radio sites as a function of r_x and total carrier bandwidths to cover $A_m^g = 23.69\text{km}^2$ of MNO1.

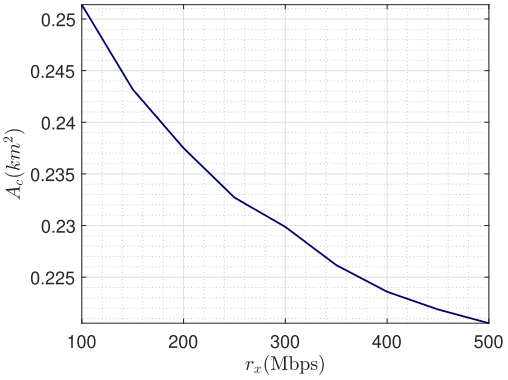

FIGURE 13. Area of active subscribers A_c as a function of r_x.

defined thresholds $\Delta r = 10$ m and $\Delta l = 1\%$, respectively. As ΔL is $|\delta_{cap} - \delta_{cov}|$ and we have that $\delta_{cap} < \delta_{cov}$, the network is less than 1% over-dimensioned, i.e. under-utilized gNBs with possible coverage overlaps. On contrary, if the network is under-dimensioned ($\delta_{cap} > \delta_{cov}$), the gNBs are over-utilized and the resulting network has lower QoE and creates coverage gaps.

Next, we evaluate the site capacity β as a function of p, the probability of the devices to be active, while keeping ζ constant. The site capacity reaches the peak

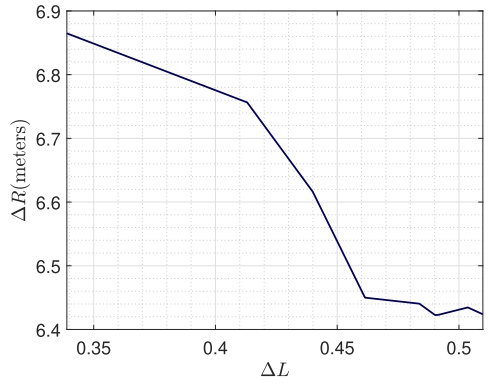

FIGURE 14. Fulfilment of the balanced RND conditions in ΔR and ΔL.

at p around 0.1 for the three analyzed values $\zeta(\%) = \{50, 70, 90\}$ (see Fig.15). We also observe that higher values of ζ (higher QoE) imply lower site capacity β, ranging from peaks of 958 ($\zeta = 50$) to 716 ($\zeta = 90$). After the peak, β is a non-increasing function of p, meaning that as the value of p increases the capacity of the gNB (served devices) decreases. Note also that the required number of radio sites increases linearly with p (see Fig.16) as lower capacity leads to smaller area A_c and smaller cell range.

We now analyze the QoE parameter ζ as a function of the total number of devices W_x^{thr}, as shown in Fig.17. The non-increasing trend of ζ is captured with fixed p and W_x^{max} while W_x^{thr} satisfies $W_x^{max} \leqslant W_x^{thr} \leqslant 100$. Interestingly, we observe that there is a value of W_x^{thr} that can be established as a threshold and beyond that value the ζ requirement is not fulfilled. In the figure, this is labelled as the MNO's region of interest (ROI), in which each service x achieves the value of ζ greater than the MNO's requirement. We set $\zeta = 50\%$ for each x and all MNOs, i.e. r_x will be provided to subscribers for at least 50% of the time to provide QoE for mission critical 5G services. Fig.17 displays how to obtain the ROI for each service. For service $x = 1$, the threshold $W_x^{thr} = 16$ devices (floor value) ensures the availability of radio resources to provide $r_x = 100$ Mbps for $\zeta = 52.7\%$ of the time. In this

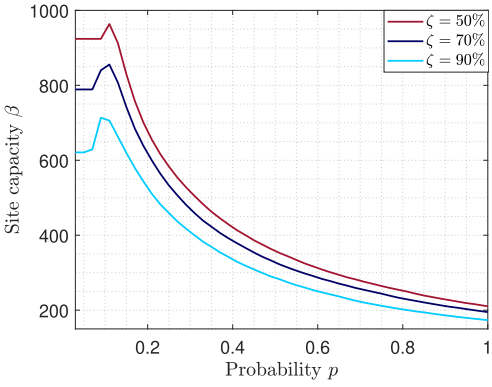

FIGURE 15. Site capacity β as a function of p guaranteeing $r_x = 100$Mbps for $\zeta(\%) = \{50, 70, 90\}$ of the time.

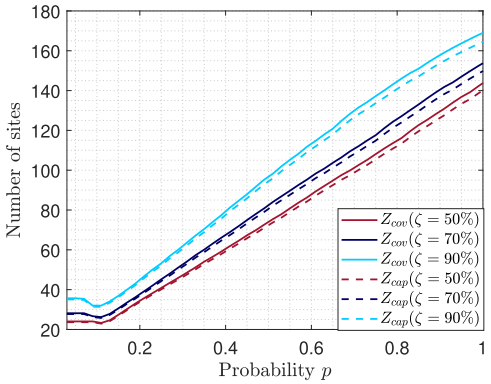

FIGURE 16. Required radio sites as a function of p referring to the probability of higher number of active subscribers.

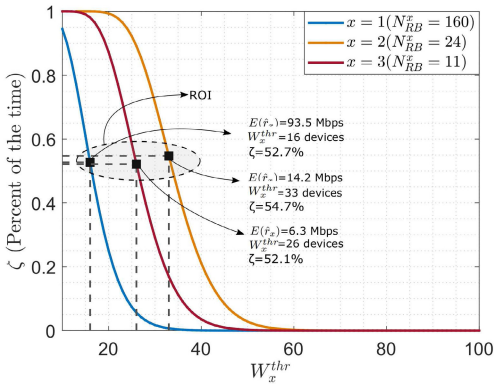

FIGURE 17. Effect on MNO's requirement of ζ as a function of W_x^{thr} for the offered services with $W_x^{max} = \{6, 13, 10\}$ respectively.

case, the average expected data rate is $E(\hat{r}_x) = 93.5$ Mbps. For the other two services ($x = 2, x = 3$), the threshold on W_x^{thr} is 33 and 26 devices respectively, providing peak data rates for $\zeta = 54.7\%$ and $\zeta = 52.1\%$ of the time. Note that service $x = 1$ has a large number of PRBs ($N_{RB} = 160$) to provide higher peak data rate. Consequently, $W_1^{thr} = 16$ is smaller compared with $W_2^{thr} = 33$, $W_3^{thr} = 26$ with $N_{RB} = 24$ and $N_{RB} = 11$ PRBs, respectively. The complete results are summarized in Table 8. The results from our capacity model achieve expected data rates $E(\hat{r}_x)$ close to the peak data rates r_x requirements. The peak data rates $r_x = \{100, 15, 7\}$ were set according to the service definition and transmission model presented in Section V.

TABLE 8. Results from proposed capacity model.

Service	W_x^{thr}	$E(\hat{r}_x)$	r_x^{min}
$x = 1$	16	93.5 Mbps	42.8 Mbps
$x = 2$	33	14.2 Mbps	6.2 Mbps
$x = 3$	26	6.3 Mbps	2.7 Mbps

C. COMPARATIVE ANALYSIS

This section provides a comparison between our heuristic approach, denoted as HRND, and the conventional RND methodology CRND, according to the parameters in Table 9. The total transmission bandwidth is set to be equal for both RND techniques, i.e. if we set 30 MHz for CRND we make the combination of the BWPs for the three services such that the total transmission bandwidth remains equal. In this study, we vary the bandwidth of the most restrictive service $x = 1$ as $B_1 = \{10, 20, 30\}$ MHz with fixed bandwidth of the other two services $B_2 = B_3 = 10$ MHz (see Table 9). In this way, the total transmission bandwidth for comparison is 30, 40 and 50 MHz. The sub-carrier spacing and number of PRBs are set accordingly. The minimum data rate r_x^{min} is also equal (see Table 9). The transmit power levels remain the same, though the power distribution schemes are different as described in (4) and (19), for HRND and CRND, respectively. The dimensioning results are generated by averaging the outputs from 100 executions of the NetDimensioning and conventional RND algorithm, each corresponding to one independent users distribution.

TABLE 9. Parameters for conventional and proposed RND comparison.

	Conventional RND (CRND)	Heuristic RND (HRND)
Bandwidth	30, 40, 50 MHz	$\{10, 20, 30\}$, 10, 10 MHz
Sub-carrier spacing	15 KHz ($\mu = 0$)	15, 30, 60 KHz ($\mu = \{0, 1, 2\}$)
Resource blocks	$\{160, 216, 270\}$	$\{52, 106, 160\}$, 24, 11
Minimum data rate	37, 41.2, 42.8 Mbps	[37, 41.2, 42.8], 6.2, 2.7 Mbps
Capacity model	Brust traffic($\lambda = 10\%$)	Mission critical ($\zeta = 50\%$)
Transmit power	30-40 dBm	30-40 dBm
Power distribution	Conventional	Proposed

We compare HRND and CRND in terms of the required number of radio sites Z and site capacity β as a function of transmission bandwidth, as presented in Table 10 for the three MNOs. Considering MNO1 and 30 MHz, there is a significant difference in terms of number of sites as we have $Z_{cap} = 204$ for CRND and $Z_{cap} = 109$ for HRND, respectively. In the case of 50 MHz, the number of radio sites largely decreases to $Z_{cap} = 141$ and $Z_{cap} = 95$ for CRND and HRND, respectively. This happens because of two reasons. First, the proposed HRND provides longer cell ranges $R = \{336, 349, 360\}$ compared to the ranges obtained for CRND $R = \{245, 267, 295\}$ for 30, 40 and 50 MHz, respectively. Second, the HRND capacity model covers more subscribers because of the larger site capacity $\beta = \{195, 210, 225\}$ compared to the site capacity of the conventional model $\beta = \{101, 123, 148\}$. The dimensioning results of the other two MNOs exhibit a similar pattern, as Table 10 shows. We also observe that the dimensioning results respond to the size of the deployment area A_m^g and to the subscriber density Φ_m. The deployment area for MNO1 is $A_1^g = 23.69$ km^2, larger than MNO2's $A_2^g = 25.20$ km^2, so MNO2 requires a larger number of radio sites. The area for MNO3 is very close to MNO1 ($A_3^g = 23.39$ km^2), though the subscriber densities are different ($\Phi_1 = 256$ user/km^2, $\Phi_1 = 183$ user/km^2), so MNO3 requires the deployment of more radio sites to meet the capacity demanded by the users.

Next, we present the cost saving obtained when our heuristic is applied, compared with the conventional approach

TABLE 10. Dimensioning results with balanced capacity cell load in both schemes is $\delta_{cap} = 21\%$ and cost savings of the MNOs with different transmission bandwidths.

	MNO1				MNO2				MNO3			
	30 MHz		50 MHz		30 MHz		50 MHz		30 MHz		50 MHz	
Parameters	CRND	HRND	CRND	HRND	CRND	HRND	CRND	HRND	CRND	HRND	CRND	HRND
β	101	195	148	225	101	195	148	225	101	195	148	225
Z	204	109	141	95	237	127	164	110	279	149	193	130
R	245	336	295	360	235	321	282	344	208	285	250	306
Cost Savings	46.5%		32.6%		46.4%		32.9%		46.5%		32.6%	

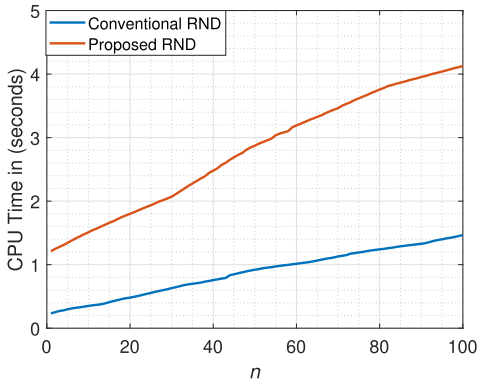

FIGURE 18. Conventional RND vs proposed RND computational complexity.

(see Table 10). The values of the cost components per gNB are provided in Table 11 [47]. The results show that dimensioning for 30MHz provides higher cost savings than for 50MHz. For instance, consider the case of MNO1, when HRND saves 46.5% of the cost with respect to CRND with 30MHz and this cost savings decrease to 32.6% when a higher transmission bandwidth is utilized. This means that HRND performs well for cost minimization for different bandwidths. However, the price paid for this cost is the computational complexity, as Fig.18 displays in terms of CPU time. We have presented the complexity curves by scaling the deployment area A_m^g of MNO1 by $n = 1$ up-to 100 such that the complexity is computed for $n \times A_m^g$. The CPU time spent by the heuristic is higher but still linear with n. Besides, the complexity in this case is far less important than the deployment cost, as HRND can be computed on a offline, non-real-time basis.

TABLE 11. Cost components for a single gNB site [47].

Components	Capex ($USD)	Quantity
Sector antenna	1,500	3
Remote radio unit	4,000	3
Baseband unit	10,000	1
Civil material	5,000	1
Transportation	10,000	1
Installation	5,000	1
Power system	5,000	1

X. CONCLUSION

Inferring about the existing network infrastructure through data can provide useful information to reduce the cost in mobile network dimensioning. In this paper, we show how the OpenCellID database has guided the demarcation of the deployment area before the dimensioning phase. The deployment area is provided to the proposed network dimensioning heuristic HRND along with requirements on ζ and peak data rates r_x to guarantee QoE. We show that our capacity model ensures the availability of the peak data rates for $\zeta \geqslant 50\%$ of the time for all services, as a use case. Moreover, dimensioning results show that we can attain up to 32% cost savings as compared with a conventional dimensioning approach. We have also demonstrated and discussed how our heuristic is capable of providing balanced dimensioning results by fulfilling QoE and load balancing parameters. We have evaluated our proposed heuristic with data regarding top three Spanish MNOs having different target areas and subscriber densities, and our heuristic provides consistent performance.

As future research lines, we are committed to exploring the machine learning applications in the context of network data combined with network dimensioning of mmWave.

REFERENCES

[1] *IMT Vision—Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond*, Standard ITU-R M.2083, 2015.
[2] D. Cao, S. Zhou, and Z. Niu, "Optimal base station density for energy-efficient heterogeneous cellular networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2012, pp. 4379–4383.
[3] Z. Niu, S. Zhou, Y. Hua, Q. Zhang, and D. Cao, "Energy-aware network planning for wireless cellular system with inter-cell cooperation," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1412–1423, Apr. 2012.
[4] M. Zheng, P. Pawelczak, S. Stanczak, and H. Yu, "Planning of cellular networks enhanced by energy harvesting," *IEEE Commun. Lett.*, vol. 17, no. 6, pp. 1092–1095, Jun. 2013.
[5] S. Wang and C. Ran, "Rethinking cellular network planning and optimization," *IEEE Wireless Commun.*, vol. 23, no. 2, pp. 118–125, Apr. 2016.
[6] M. Jaber, Z. Dawy, N. Akl, and E. Yaacoub, "Tutorial on LTE/LTE—A cellular network dimensioning using iterative statistical analysis," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1355–1383, 2nd Quart., 2016.
[7] A. Taufique, M. Jaber, A. Imran, Z. Dawy, and E. Yacoub, "Planning wireless cellular networks of future: Outlook, challenges and opportunities," *IEEE Access*, vol. 5, pp. 4821–4845, 2017.
[8] J. Perez-Romero, O. Sallent, R. Ferrus, and R. Agusti, "Knowledge-based 5G radio access network planning and optimization," in *Proc. Int. Symp. Wireless Commun. Syst. (ISWCS)*, Sep. 2016, pp. 359–365.
[9] N. Saxena, A. Roy, and H. Kim, "Efficient 5G small cell planning with eMBMS for optimal demand response in smart grids," *IEEE Trans. Ind. Informat.*, vol. 13, no. 3, pp. 1471–1481, Jun. 2017.

[10] F.-H. Tseng, L.-D. Chou, H.-C. Chao, and J. Wang, "Ultra-dense small cell planning using cognitive radio network toward 5G," *IEEE Wireless Commun.*, vol. 22, no. 6, pp. 76–83, Dec. 2015.

[11] A. L. Rezaabad, H. Beyranvand, J. A. Salehi, and M. Maier, "Ultra-dense 5G small cell deployment for fiber and wireless backhaul-aware infrastructures," *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 12231–12243, Dec. 2018.

[12] C. Y. Lee and H. G. Kang, "Cell planning with capacity expansion in mobile communications: A tabu search approach," *IEEE Trans. Veh. Technol.*, vol. 49, no. 5, pp. 1678–1691, Sep. 2000.

[13] T. Bauschert, C. Büsing, F. D'Andreagiovanni, A. C. A. Koster, M. Kutschka, and U. Steglich, "Network planning under demand uncertainty with robust optimization," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 178–185, Feb. 2014.

[14] H. Ghazzai, E. Yaacoub, M.-S. Alouini, Z. Dawy, and A. Abu-Dayya, "Optimized LTE cell planning with varying spatial and temporal user densities," *IEEE Trans. Veh. Technol.*, vol. 65, no. 3, pp. 1575–1589, Mar. 2016.

[15] *Technical Specification. NR and NG-RAN Overall Description; Stage 2 (Release 16)*, document TS 38.300 V16.0.0, 3GPP, 2019.

[16] T. Doumi, M. F. Dolan, S. Tatesh, A. Casati, G. Tsirtsis, K. Anchan, and D. Flore, "LTE for public safety networks," *IEEE Commun. Mag.*, vol. 51, no. 2, pp. 106–112, Feb. 2013.

[17] M. Kashef, M. Ismail, E. Serpedin, and K. Qaraqe, "Balanced dynamic planning in green heterogeneous cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3299–3312, Dec. 2016.

[18] M. Maule, D. Moltchanov, P. Kustarev, M. Komarov, S. Andreev, and Y. Koucheryavy, "Delivering fairness and QoS guarantees for LTE/Wi-Fi coexistence under LAA operation," *IEEE Access*, vol. 6, pp. 7359–7373, 2018.

[19] K. Zheng, X. Zhang, Q. Zheng, W. Xiang, and L. Hanzo, "Quality-of-experience assessment and its application to video services in LTE networks," *IEEE Wireless Commun.*, vol. 22, no. 1, pp. 70–78, Feb. 2015.

[20] A. E. Essaili, D. Schroeder, D. Staehle, M. Shehada, W. Kellerer, and E. Steinbach, "Quality-of-experience driven adaptive HTTP media delivery," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2013, pp. 2480–2485.

[21] L. Pierucci, "The quality of experience perspective toward 5G technology," *IEEE Wireless Commun.*, vol. 22, no. 4, pp. 10–16, Aug. 2015.

[22] Z.-H. Yang, M. Chen, Y.-P. Wen, L.-Q. Jia, and Y. Zhang, "Cell planning based on minimized power consumption for LTE networks," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Apr. 2016, pp. 1–6.

[23] D. G. González and J. Hämäläinen, "Looking at cellular networks through canonical domains and conformal mapping," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3703–3717, May 2016.

[24] P. Muñoz, O. Sallent, and J. Pérez-Romero, "Self-dimensioning and planning of small cell capacity in multitenant 5G networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4552–4564, May 2018.

[25] (2019). *Open Signal Mobile Analytics*. [Online]. Available: https://www.opensignal.com/

[26] (2019). *Telecommunication Coverage Maps, NPERF*. [Online]. Available: https://www.nperf.com/

[27] (2019). *Exisiting 4G Infrastructure From OpenCelliD by Uniwired Labs*. [Online]. Available: http://wiki.opencellid.org/wiki/

[28] D. Willkomm, S. Machiraju, J. Bolot, and A. Wolisz, "Primary users in cellular networks: A large-scale measurement study," in *Proc. 3rd IEEE Symp. New Frontiers Dyn. Spectr. Access Netw.*, Oct. 2008, pp. 1–11.

[29] F. Xu, Y. Li, H. Wang, P. Zhang, and D. Jin, "Understanding mobile traffic patterns of large scale cellular towers in urban environment," *IEEE/ACM Trans. Netw.*, vol. 25, no. 2, pp. 1147–1161, Apr. 2017.

[30] M. Z. Shafiq, L. Ji, A. X. Liu, J. Pang, and J. Wang, "Geospatial and temporal dynamics of application usage in cellular data networks," *IEEE Trans. Mobile Comput.*, vol. 14, no. 7, pp. 1369–1381, Jul. 2015.

[31] F. Malandrino, C.-F. Chiasserini, and S. Kirkpatrick, "Cellular network traces towards 5G: Usage, analysis and generation," *IEEE Trans. Mobile Comput.*, vol. 17, no. 3, pp. 529–542, Mar. 2018.

[32] Z. Frias, L. Mendo, and E. J. Oughton, "Evolving mobile networks towards 5G. a framework to infer the state of connectivity infrastructure in dense urban areas," in *Proc. 30th ITS Eur. Conf.*, Espoo, Finland, Jun. 2019, pp. 1–24.

[33] Q.-T. Nguyen-Vuong, N. Agoulmine, E. H. Cherkaoui, and L. Toni, "Multicriteria optimization of access selection to improve the quality of experience in heterogeneous wireless access networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 4, pp. 1785–1800, May 2013.

[34] H. Shao, W. Jing, X. Wen, Z. Lu, H. Zhang, Y. Chen, and D. Ling, "Joint optimization of quality of experience and power consumption in OFDMA multicell networks," *IEEE Commun. Lett.*, vol. 20, no. 2, pp. 380–383, Feb. 2016.

[35] P. Oliver-Balsalobre, M. Toril, S. Luna-Ramirez, and R. G. Garaluz, "Self-tuning of service priority parameters for optimizing quality of experience in LTE," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3534–3544, Apr. 2018.

[36] C. Gijón, M. Toril, S. Luna-Ramírez, and M. Luisa Marí-Altozano, "A data-driven traffic steering algorithm for optimizing user experience in multi-tier LTE networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, pp. 9414–9424, Oct. 2019.

[37] P. Uthansakul, P. Anchuen, M. Uthansakul, and A. A. Khan, "QoE-aware self-tuning of service priority factor for resource allocation optimization in LTE networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 887–900, Jan. 2020.

[38] K. Bouraqia, E. Sabir, M. Sadik, and L. Ladid, "Quality of experience for streaming services: Measurements, challenges and insights," *IEEE Access*, vol. 8, pp. 13341–13361, 2020.

[39] S. Hurley, "Planning effective cellular mobile radio networks," *IEEE Trans. Veh. Technol.*, vol. 51, no. 2, pp. 243–253, Mar. 2002.

[40] I. Siomina and D. Yuan, "Analysis of cell load coupling for LTE network planning and optimization," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2287–2297, Jun. 2012.

[41] D. G. González, E. Mutafungwa, B. Haile, J. Hämäläinen, and H. Poveda, "A planning and optimization framework for ultra dense cellular deployments," *Mobile Inf. Syst.*, vol. 2017, pp. 1–17, Mar. 2017.

[42] D. G. González, H. Hakula, A. Rasila, and J. Hämäläinen, "Spatial mappings for planning and optimization of cellular networks," *IEEE/ACM Trans. Netw.*, vol. 26, no. 1, pp. 175–188, Feb. 2018.

[43] G. V. Tsoulos, G. E. Athanasiadou, D. A. Zarbouti, and I. K. Valavanis, "Optimizing radio network planning evolution towards microcellular systems," *Wireless Pers. Commun.*, vol. 106, pp. 521–534, Feb. 2019.

[44] M. Masoudi *et al.*, "Green mobile networks for 5G and beyond," *IEEE Access*, vol. 7, pp. 107270–107299, Aug. 2019.

[45] R. Bassoli, M. Di Renzo, and F. Granelli, "Analytical energy-efficient planning of 5G cloud radio access network," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–4.

[46] A. A. Vaganova, N. N. Kisel, and A. I. Panychev, "Simulation model of 5G coverage in urban areas," in *Proc. Radiat. Scattering Electromagn. Waves (RSEMW)*, Jun. 2019, pp. 372–375.

[47] E. J. Oughton, K. Katsaros, F. Entezami, D. Kaleshi, and J. Crowcroft, "An open-source techno-economic assessment framework for 5G deployment," *IEEE Access*, vol. 7, pp. 155930–155940, Oct. 2019.

[48] E. J. Oughton, Z. Frias, S. van der Gaast, and R. van der Berg, "Assessing the capacity, coverage and cost of 5G infrastructure strategies: Analysis of The Netherlands," *Telematics Informat.*, vol. 37, pp. 50–69, Apr. 2019.

[49] E. Oughton, Z. Frias, T. Russell, D. Sicker, and D. D. Cleevely, "Towards 5G: Scenario-based assessment of the future supply and demand for mobile telecommunications infrastructure," *Technol. Forecasting Social Change*, vol. 133, pp. 141–155, Aug. 2018.

[50] S. Dadvandipour and M. Alsharif, "On the analysis and tool development of the long term evolution cell planning," in *Proc. 20th Int. Carpathian Control Conf. (ICCC)*, May 2019, pp. 1–6.

[51] J. R. Martín, R. Pérez-Leal, and J. Navío-Marco, "Towards 5G: Techno-economic analysis of suitable use cases," *NETNOMICS, Econ. Res. Electron. Netw.*, vol. 20, nos. 2–3, pp. 153–175, Dec. 2019.

[52] O. Bondarenko, D. Ageyev, and O. Mohammed, "Optimization model for 5G network planning," in *Proc. IEEE 15th Int. Conf. Exp. Designing Appl. CAD Syst. (CADSM)*, Feb. 2019, pp. 1–4.

[53] A. Elnashar and M. A. El-Saidny, "Looking at LTE in practice: A performance analysis of the LTE system based on field test results," *IEEE Veh. Technol. Mag.*, vol. 8, no. 3, pp. 81–92, Sep. 2013.

[54] Statista. (2019). *Mobile Telephony Market Share in Spain by Telecommunications Providers*. [Online]. Available: https://www.statista.com/statistics/745319/mobile-phone-provider-market-share-in-spain/

[55] National Mobile Network Operators. (2019). *Halber Bastion, Radio Frequency Technologies*. [Online]. Available: https://halberdbastion.com/intelligence/countries-nations/spain

[56] A. A. Zaidi, R. Baldemair, H. Tullberg, H. Bjorkegren, L. Sundstrom, J. Medbo, C. Kilinc, and I. Da Silva, "Waveform and numerology to support 5G services and requirements," *IEEE Commun. Mag.*, vol. 54, no. 11, pp. 90–98, Nov. 2016.

[57] *Technical Specification: 5G NR- User Equipment (UE) Radio Transmission and Reception, Version 15.2.0 Release 15*, document TS 38.101-1, 3GPP, 2018.
[58] J. J. Escudero-Garzás, C. Bousoño-Calzón, and A. García, "On the feasibility of 5G slice resource allocation with spectral efficiency: A probabilistic characterization," *IEEE Access*, vol. 7, pp. 151948–151961, Oct. 2019.
[59] V. N. Ha, T. T. Nguyen, L. B. Le, and J.-F. Frigon, "Admission control and network slicing for multi-numerology 5G wireless networks," *IEEE Netw. Lett.*, vol. 2, no. 1, pp. 5–9, Mar. 2020.
[60] *Technical Specification: 5G NR- User Equipment (UE) Radio Access Capabilities, Version 15.2.0 Release 15*, document TS 38.306, 3GPP, 2018.
[61] (2019). *Small Cells, Big Impact: Designing Power Solutions for 5G Applications*. [Online]. Available: http://www.ti.com/lit/wp/slyy166/slyy166.pdf
[62] *ETSI TR 138 901: 5G; Study on Channel Model for Frequencies From 0.5 to 100 GHz, Version 15.0.0 Release 15*, docment TR 38.901, 3GPP, 2018.
[63] J. T. J. Penttinen, *5G Network Planning and Optimization*. Hoboken, NJ, USA: Wiley, 2019, ch. 9, pp. 255–269.
[64] J. Fu and R. Sproule, "A generalization of the binomial distribution," *Commun. Statist.-Theory Methods*, vol. 24, no. 10, pp. 2645–2658, 1995.
[65] H. D. Trinh, N. Bui, J. Widmer, L. Giupponi, and P. Dini, "Analysis and modeling of mobile traffic using real traces," in *Proc. IEEE 28th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Oct. 2017, pp. 1–6.
[66] (2014). *Mobile Broadband With HSPA and LTE Capacity and Cost Aspects*. [Online]. Available: http://ec.europa.eu/information-society/newsroom/cf/dae/document
[67] A. Elnashar, *Coverage and Capacity Planning of 4G Networks*. Hoboken, NJ, USA: Wiley, 2014, ch. 6, pp. 349–444.
[68] N. A. Mohammed, A. M. Mansoor, and R. B. Ahmad, "Mission-critical machine-type communication: An overview and perspectives towards 5G," *IEEE Access*, vol. 7, pp. 127198–127216, 2019.

M. UMAR KHAN received the B.S. degree in telecommunication engineering from COMSATS University Islamabad, Pakistan, in 2012, and the master's degree in telematics engineering from Universidad Carlos III de Madrid (UC3M), in 2015, where he is currently pursuing the Ph.D. degree with the Department of Signal Theory and Communications. He was a Lecturer with the Center for Advanced Studies in Telecommunications (CAST), COMSATS University Islamabad, from 2015 to 2016. From 2012 to 2013, he was a Research Associate with CAST, COMSATS University Islamabad. His research interests include optimization and dimensioning of 5G networks.

A. GARCÍA-ARMADA (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the Polytechnical University of Madrid, in February 1998. She is currently a Professor with the University Carlos III of Madrid, Spain. She is leading the Communications Research Group, University Carlos III of Madrid. She has been a Visiting Scholar with Stanford University, Bell Labs, and University of Southampton. She has participated (and coordinated most of them) in more than 30 national and ten international research projects as well as 20 contracts with the industry, all of them related to wireless communications. She has published around 150 articles in international journals and conference proceedings, and she holds four patents. She has contributed to international standards organizations, such as ITU and ETSI, is a member of the expert group of the European 5G PPP and a member of the advisory committee 5JAC of the ESA as expert appointed by Spain on 5G. She was awarded the third place Bell Labs Prize 2014 for shaping the future of information and communications technology. She has served on the editorial board of several journals, such as the IEEE COMMUNICATIONS LETTERS and the IEEE TRANSACTIONS ON COMMUNICATIONS. She is the Secretary of the IEEE ComSoc SPCC committee. She was a Secretary of the IEEE ComSoc Women in Communications Engineering Standing Committee (2016–2017) and the Chair of this committee (2018–2019).

J. J. ESCUDERO-GARZÁS (Member, IEEE) received the Ph.D. degree in electrical engineering from the Universidad Carlos III de Madrid (UC3M). From 1997 to 2002, he was a Provisioning Engineer with Spanish telcos and the Head of the Communications Network Maintenance Department. He has been a Faculty Member with the Department of Signal Theory and Communications, UC3M. He was also a Postdoctoral Fellow with the Universitat Autònoma de Barcelona, from 2010 to 2012, and a Researcher with the University of Virginia, in 2012, the University of Florida, in 2017, and Texas A&M University, from 2018 to 2019. His research interests include optimization of wireless communication systems and resource management for 5G networks.