IEEE Open Journal of Vehicular Technology

Received 14 November 2023; accepted 20 November 2023. Date of publication 8 December 2023;
date of current version 2 April 2024. The review of this article was coordinated by Editor Edward Kwok Shum Au.

Digital Object Identifier 10.1109/OJVT.2023.3340993

3GPP-Compliant Datasets for xG Location-Aware Networks

ANDREA CONTI [1] (Fellow, IEEE), GIANLUCA TORSOLI [1] (Graduate Student Member, IEEE), CARLOS A. GÓMEZ-VEGA [1] (Graduate Student Member, IEEE), ALESSANDRO VACCARI [1] (Graduate Student Member, IEEE), GIANLUCA MAZZINI [1] (Fellow, IEEE), AND MOE Z. WIN [2] (Fellow, IEEE)

(Invited Paper)

[1]Department of Engineering and CNIT, University of Ferrara, 44122 Ferrara, Italy
[2]Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139 USA

CORRESPONDING AUTHOR: ANDREA CONTI (e-mail: a.conti@ieee.org).

The fundamental research described in this paper was supported, in part, by the Office of Naval Research under Grant N62909-22-1-2009, by the National Science Foundation under Grant CNS-2148251, and by federal agency and industry partners in the RINGS program.

ABSTRACT Location awareness is vital in next generation (xG) wireless networks to enable different use cases, including location-based services (LBSs) and efficient network management. However, achieving the service level requirements specified by the 3rd Generation Partnership Project (3GPP) is challenging. This calls for new localization algorithms as well as for 3GPP-standardized scenarios to support their systematic development and testing. In this context, the availability of public datasets with 3GPP-compliant configurations is essential to advance the evolution of xG networks. This paper introduces xG-Loc, the first open dataset for localization algorithms and services fully compliant with 3GPP technical reports and specifications. xG-Loc includes received localization signals, measurements, and analytics for different network and signal configurations in indoor and outdoor scenarios with center frequencies from micro-waves in frequency range 1 (FR1) to millimeter-waves in frequency range 2 (FR2). Position estimates obtained via soft information-based localization and wireless channel quality indicators via blockage intelligence are also provided. The rich set of data provided by xG-Loc enables the characterization of localization algorithms and services under common 3GPP-standardized scenarios in xG networks.

INDEX TERMS 3GPP, xG, localization, dataset, next-generation networks.

I. INTRODUCTION

Location awareness is vital for next generation (xG) wireless networks [1], [2], [3], [4], [5]. On the one hand, it is a key enabler for several use cases (UCs), including autonomous driving [6], [7], [8], assets tracking [9], [10], [11], Internet-of-Things (IoT) [12], [13], [14], virtual reality [15], [16], [17], and public safety [18], [19], [20]. On the other hand, the positional information can be exploited to enhance the network capabilites, for example via radio resource and beam management [21], [22], [23]. In this context, the 3rd Generation Partnership Project (3GPP) has defined seven positioning service levels (PSLs) for UCs enabled by localization: the key performance indicators (KPIs) requirements include horizontal and vertical localization accuracy, availability, and latency [24]. Fulfilling such specification requirements calls for new localization and location-aware algorithms as well as for a common ground to support their systematic development and testing. Therefore, to advance the evolution of xG networks, it is fundamental to make publicly available datasets compliant with 3GPP technical reports and specifications that serve as common ground for companies and research institutions for developing and testing new algorithms. In particular, the 3GPP has defined a set of reference settings in terms of scenarios,

signals, channels, and resource allocations [25], [26], [27]. However, implementing software to generate 3GPP-compliant data can be particularly challenging due to the complexity of the 3GPP specifications that make such task time-consuming and computationally intensive.

Starting from Release 16, the 3GPP has been dedicating significant efforts to extend the localization architecture and capabilities of mobile wireless networks for both downlink (DL) and uplink (UL). For example, the new enablers introduced by fifth generation (5G) New Radio, such as multiple-input multiple-output (MIMO) antennas and millimeter wave signals, can be leveraged to achieve unprecedented localization performance levels in cellular networks [1], [28]. Currently, few datasets for localization with 5G radio access technology (RAT)-dependent and RAT-independent measurements are publicly available [29], [30], [31], [32], [33], [34], [35]. In particular, fully 3GPP-compliant datasets for localization in 5G and beyond 5G wireless networks are still missing.

The goal of this paper is to introduce xG-Loc, an open dataset for localization and location-based services (LBSs) fully compliant with 3GPP technical reports and specifications [36]. xG-Loc includes received localization signals, measurements, and analytics for 3GPP-standardized indoor factory (InF)-dense high (DH), InF-sparse high (SH), indoor open office (IOO), and urban microcell (UMi) scenarios with central frequencies from micro-waves in frequency range 1 (FR1) to millimeter-waves in frequency range 2 (FR2) and bandwidths from 5 MHz to 400 MHz, and others as standardization will evolve. Specifically, the rich set of data provided by xG-Loc can be used to develop and test localization and location-aware algorithms. xG-Loc also includes position estimates obtained via soft information (SI)-based localization [37] and wireless channel quality indicators obtained via blockage intelligence (BI) [38] as performance benchmarks. In particular, xG-Loc is publicly available on IEEE DataPort [36]. The key contributions of this paper can be summarized as follows:

- introduction of xG-Loc, the first open dataset for the evaluation of xG localization algorithms and LBSs in 3GPP-standardized configurations;
- overview of the key characteristics of xG-Loc, including 3GPP-standardized reference signals (RSs) structure, measurements, and scenarios; and
- performance benchmarks for xG-Loc obtained via SI-based localization and BI.

The remainder of the paper is organized as follows. Section II provides an overview of the signals and architectures defined by the 3GPP for localization in beyond 5G wireless networks; Section III describes the setup used for the generation of xG-Loc; Section IV presents the structure of xG-Loc; Section V provides performance benchmarks on xG-Loc obtained via SI-based localization and BI. Finally, Section VI gives concluding remarks.

Notations: Random variables are displayed in sans serif, upright fonts; their realizations in serif, italic fonts. Vectors are denoted by bold lowercase letters. For example, a random variable and its realization are denoted by x and x, respectively; a random vector and its realization are denoted by $\mathbf{x}$ and $\boldsymbol{x}$, respectively; For a vector $\boldsymbol{x}$, its transpose is denoted by $\boldsymbol{x}^{\mathrm{T}}$. Sets are denoted by calligraphic fonts. For example, a set is denoted by $\mathcal{X}$. Symbol x^* denotes the complex conjugate of x. The function $f_{\mathbf{x}}(\boldsymbol{x})$ and, for brevity when possible, $f(\boldsymbol{x})$ denotes the probability density function (PDF) of a continuous random vector $\mathbf{x}$; $\mathbb{E}_{\mathsf{x}|\mathbf{y}}\{\cdot|\boldsymbol{y}\}$ denotes the expectation with respect to the random variable x conditioned on $\mathbf{y} = \boldsymbol{y}$.

II. LOCALIZATION IN CELLULAR NETWORKS

Localization in beyond 5G networks is performed via the transmission of specific RSs, namely the positioning reference signal (PRS) in DL and the sounding reference signal (SRS) in UL. Let N_{b} be the number of base stations (BSs), referred to as gNodeBs (gNBs) in 5G networks, and indexed by $j \in \mathcal{N}_{\mathrm{b}} = \{1, 2, \ldots, N_{\mathrm{b}}\}$. The position of each BS is known and denoted by $\boldsymbol{p}_j$. Considering a cellular network composed of N_{s} sites with N_{sec} sectors each, then the total number of BSs is given by $N_{\mathrm{b}} = N_{\mathrm{s}} N_{\mathrm{sec}}$ [39]. The estimation of the user equipment (UE) position can be performed by exchanging RSs with the BSs. According to current specifications [40], [41], the RSs can be transmitted with different carrier frequencies either in FR1 (carrier frequency between 410 MHz and 7.125 GHz) or FR2 (carrier frequency between 24.25 GHz and 52.6 GHz). A brief description of the structure of the PRS and of the SRS is presented next.

A. REFERENCE SIGNALS STRUCTURE

The PRS and the SRS are both based on orthogonal frequency division multiplexing (OFDM), with a specific organization in the time and frequency domains [42], [43], [44]. In the frequency domain, the signal is organized into N_{RB} resource blocks (RBs), with $N_{\mathrm{RB}} \in \{4, 8, \ldots, 272\}$, where each RB contains $N_{\mathrm{SC}} = 12$ subcarriers. Thus, the total number of subcarriers used for the RS transmission is given by $N_{\mathrm{F}} = N_{\mathrm{RB}} N_{\mathrm{SC}}$. The subcarrier allocation is organized into a comb structure, i.e., only one subcarrier out of K contains data, while the others are padded to zero. In the time domain, the RSs transmission is organized into multiple slots, with L consecutive symbols used. The k-th subcarrier of the l-th symbol is referred to as resource element (RE) and is indexed by (l, k). The slots are transmitted within radio frames with a fixed duration of 10 ms. The number of slots within a radio frame (and therefore their duration) depends on the numerology parameter $\mu \in \{0, 1, 2, 3, 4, 5, 6\}$. The parameter μ is used to configure the specific subcarrier spacing for the RS transmission, which is given by $\Delta_{\mathrm{f}} = 2^{\mu} \times 15\,\mathrm{kHz}$. Therefore, the RS bandwidth is given by $B = N_{\mathrm{F}} \Delta_{\mathrm{f}}$. The complete description of the possible configurations for allocating the time-frequency structure of the RSs can be found in [42].

Given a time-frequency allocation, the digital OFDM-based signal for the l-th symbol is obtained via the inverse fast

Fourier transform as

$$s_l[n] = \frac{1}{\sqrt{N_\mathrm{F}}} \sum_{k=0}^{N_\mathrm{F}} a_{k,l} \exp\left\{\frac{2\pi n k}{N_\mathrm{F}}\right\} \tag{1}$$

where

$$a_{k,l} = \begin{cases} \beta\, q[m] & \text{if } m \text{ is mapped to } k \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

in which β denotes a scaling coefficient and $q[m]$ is a sequence to be allocated in the subcarriers (different in the cases of PRS and SRS).

B. POSITIONING REFERENCE SIGNAL (PRS)

The PRS has been introduced by the 3GPP specifically for DL localization [42]. The PRS leverages the OFDM-based structure described in Section II-A, where $K \in \{2, 4, 6, 12\}$ and $L \in \{2, 4, 6, 12\}$. The sequence $q[m]$ is obtained via quadrature phase-shift keying (QPSK) modulation as

$$q[m] = \frac{1}{\sqrt{2}}\left(1 - 2c[m]\right) + \jmath \frac{1}{\sqrt{2}}\left(1 - 2c[m+1]\right) \tag{3}$$

where $c[m]$ denotes a 31-bit long Gold sequence initialized based on the physical cell identity (PCI) according to [42].

C. SOUNDING REFERENCE SIGNAL (SRS)

The SRS was originally introduced by the 3GPP to perform UL channel sounding to support communication but its use has been extended also to localization purposes. The SRS exploits the OFDM-based structure described in Section II-A, where $K \in \{2, 4, 8\}$ and $L \in \{1, 2, 4, 8, 12\}$. The sequence $q[m]$ is a low peak-to-average power ratio Zadoff-Chu sequence, generated according to [42].

D. 3GPP MEASUREMENTS FOR LOCALIZATION

According to the 3GPP specifications, the received RSs can be processed to extract relevant measurements for localization. These measurements, denoted by $\{\hat{\theta}_j\}_{j\in\mathcal{N}_\mathrm{b}}$ include time-based measurements, such as DL-time difference-of-arrival (TDOA), UL-TDOA, and round-trip time (RTT), as well as angle-based measurements, such as angle-of-departure (AOD) [21], [45]. The set of measurements is related to a set of positional features $\{\theta_j(\boldsymbol{p}_j, \boldsymbol{p})\}_{j\in\mathcal{N}_\mathrm{b}}$ which are a function of both the position $\boldsymbol{p}_j$ of the BSs and on the unknown position $\boldsymbol{p}$ of the UE. To improve the accuracy of the measurements, several retransmissions, also referred to as occasions, of the RS can be performed in different radio frames and coherently accumulated to obtain a waveform that presents a higher signal-to-noise ratio. Moreover, given the 3GPP specifications for the symbol allocations, portions of the received signals corresponding to K OFDM symbols in the same slot can be coherently accumulated if $\mathrm{mod}(L, K) = 0$ [25].

Let $\bar{r}[n]$ and $s[n]$ denote the sampled versions of the received RS after accumulation and of the transmitted RS, respectively, then their cross-correlation is given by

$$R[n] = \sum_{k=0}^{N_\mathrm{s}-1} \bar{r}[n] s^*[n-k] \tag{4}$$

where $n = 0, 1, \ldots, N_\mathrm{c} - 1$. Such cross-correlation can be intended as an estimation of the channel impulse response (CIR) and can be processed to obtain time-based measurements for localization [46], [47]. Several approaches can be exploited to estimate the time-of-arrival (TOA), which is necessary to obtain both the TDOA and the RTT [45]. In particular, TOA is typically determined from (4) based on the estimation of the delay associated with the first peak.

Given a set of TOA measurements $\{\hat{\tau}_j\}_{j\in\mathcal{N}_\mathrm{b}}$ obtained from the RSs, the TDOA measurements are obtained as

$$\hat{\theta}_j = \hat{\tau}_{\hat{k}} - \hat{\tau}_j \tag{5}$$

where the BS indexed by $j = \hat{k}$ is the reference BS (possibly selected under an optimality criterion, e.g., see [48]). If the TOA estimate is obtained via the PRS or the SRS transmission, the measurements are referred to as DL-TDOA or UL-TDOA, respectively.

Differently, given a set of TOA estimates obtained via the PRS transmission, denoted by $\{\hat{\tau}_{\mathrm{d}j}\}_{j\in\mathcal{N}_\mathrm{b}}$, and a set of TOA estimates obtained via the SRS transmission, denoted by $\{\hat{\tau}_{\mathrm{u}j}\}_{j\in\mathcal{N}_\mathrm{b}}$, the RTT measurements for localization are obtained as

$$\hat{\theta}_j = \frac{1}{2}\left(\hat{\tau}_{\mathrm{d}j} + \hat{\tau}_{\mathrm{u}j}\right). \tag{6}$$

In addition to time-based measurements, angle-based measurements can be exploited for localization purposes [45]. This is performed via a beam sweeping procedure that involves multiple PRS transmissions with different beam steering angles [49]. The AOD is estimated as the steering angle that determined the PRS transmission with the highest reference signal received power (RSRP) [21]. Specifically, let $\mathcal{A}_j = \{\alpha_{j0}, \alpha_{j1}, \ldots, \alpha_{jN_\mathrm{A}}\}$ be a set of N_A angles to evaluate via beam sweeping, and let $\rho(\alpha_{jl})$ denote an RSRP measurement obtained via beam steering toward the angle α_{jl}. Then, the AOD is obtained as

$$\hat{\theta}_j = \underset{\alpha_{jl}\in\mathcal{A}_j}{\arg\max}\, \rho\left(\alpha_{jl}\right). \tag{7}$$

E. SOFT INFORMATION-BASED LOCALIZATION

SI-based localization has been recently proposed to overcome the limitations of classical localization algorithms based on single-value estimates (SVEs) [37], [50]. In particular, SI-based localization combines soft feature information (SFI), i.e., the information on features obtained from measurements, and soft context information (SCI), i.e., the information on the wireless environment, to localize the UE. The SFI can be classified into soft range information (SRI) and soft angle information (SAI), respectively, depending on whether time or angle measurements are considered. For any measurement

$\boldsymbol{y}$ in a 5G network, the SFI is given by

$$\mathcal{L}_{\boldsymbol{y}}(\boldsymbol{\theta}) \propto f_{\mathbf{y}}(\boldsymbol{y};\boldsymbol{\theta})\,. \tag{8}$$

Denoting by $\{\hat{\theta}_j\}_{j\in\mathcal{N}_\mathrm{b}}$ a collection of independent measurements obtained from different BSs, the position $\boldsymbol{p}$ of a UE can be inferred via maximum likelihood estimation as

$$\hat{\boldsymbol{p}} = \arg\max_{\tilde{\boldsymbol{p}}} \prod_{j\in\mathcal{N}_\mathrm{b}} \mathcal{L}_{\boldsymbol{y}_j}(\boldsymbol{\theta}_j(\boldsymbol{p}_j,\tilde{\boldsymbol{p}})). \tag{9}$$

The SFI is obtained as proportional to a generative model, i.e., an approximation of the joint probability distribution of measurements and positional features. Given the wireless propagation complexity of 5G wireless scenarios, it cannot be determined a priori and requires to be obtained via density estimation. In particular, the use of Gaussian mixture models (GMMs) with N_M components as generative models has been demonstrated to be effective for localization in 5G and beyond wireless networks [50], [51].

F. BLOCKAGE INTELLIGENCE

The concept of BI has been recently proposed to overcome the limitations of conventional non-line-of-sight (NLOS) identification approaches. It has been demonstrated that BI provides an indicator of the wireless channel quality that can be effectively employed to improve the performance of both conventional and SI-based localization algorithms [38]. Specifically, this can be obtained by leveraging the information on wireless propagation encapsulated in a set of statistical features $\boldsymbol{\nu}$ extracted from (4).

This problem can be cast to a two-class supervised classification problem. Let $\boldsymbol{\gamma} \in \{+1,-1\}$ be a binary random variable which takes value $+1$ and -1 for NLOS and line-of-sight (LOS) propagation conditions, respectively. By considering an exponential loss function, a model $c(\boldsymbol{\nu})$ for classification can be obtained by solving [52]

$$c(\boldsymbol{\nu}) = \arg\min_{\check{c}:\mathbb{R}^d\to\mathbb{R}} \mathbb{E}_{\boldsymbol{\gamma}|\boldsymbol{\nu}}\{e^{-\boldsymbol{\gamma}\check{c}(\boldsymbol{\nu})} \mid \boldsymbol{\nu}\} \tag{10}$$

which has a closed-form solution given by

$$\psi(\boldsymbol{\nu}) = \mathbb{P}\{\boldsymbol{\gamma}=+1 \mid \boldsymbol{\nu}\} = \frac{e^{c(\boldsymbol{\nu})}}{e^{-c(\boldsymbol{\nu})}+e^{c(\boldsymbol{\nu})}} \tag{11}$$

representing the probability of being in NLOS given a vector of features. However, solving (10) requires the joint probability distribution of $\boldsymbol{\gamma}$ and $\boldsymbol{\nu}$, which is not known a priori and calls for machine learning (ML)-based techniques to be obtained [38], [53], [54].

To provide a quantification of the BI performance, it is possible to use the area under the curve (AUC)-receiver operating characteristic (ROC), a conventional metric used in ML and estimation theory for evaluating probabilistic classification algorithms [55], [56]. Specifically, the AUC-ROC can be intended as a measure of the probability that a randomly chosen wireless link in NLOS conditions is assigned a higher probability of being in NLOS compared to a randomly chosen wireless link in LOS conditions. In addition, the mean-square error (MSE) can be used to evaluate the BI NLOS probability estimation quality [57]. High values of ROC-AUC and low values of MSE indicate a high quality of the probability indicator provided by BI.

Note that the probabilistic value provided by BI can be transformed into a conventional binary NLOS indicator, hereafter referred to as discretized blockage intelligence (DBI), by applying a threshold η [38].

III. SETUP FOR xG-Loc GENERATION

xG-Loc [36] is generated in full compliance with 3GPP technical reports and specifications, including environments, channels, and signals [25], [26], [39], [58]. This enables providing data under a common setting used by companies and research institutions to evaluate their localization algorithms. In addition, the localization performance obtained via SI-based localization [37], [50], as well as the probabilistic indicator provided by BI [38], can be leveraged to evaluate the performance of LBSs that require accurate positional information under 3GPP-compliant settings (e.g., see [59]).

In the following sections, the main parameters used for generating xG-Loc are reported, including the wireless scenarios considered, the UE characteristics, and the RS time-frequency allocation parameters.

A. REFERENCE SIGNALS CHARACTERISTICS

The physical structure and the time-frequency allocation of the RSs are generated according to Section II-A, Section II-B, and Section II-C [42]. Specifically, the RSs are allocated in a single slot, which is retransmitted on 4 occasions for coherent accumulation. The PRS is generated with a number of symbols $L = 12$ and a comb size $K = 6$, while the SRS is generated with a number of symbols $L = 8$ and a comb size $K = 4$. These configurations enable an intra-slot coherent accumulation of the K symbols for the PRS and SRS [25]. The coherent accumulation of the RSs is performed over $N_\mathrm{OCC} = 4$ occasions. For the RS transmission and reception, no synchronization errors and ideal muting are considered in the network [60]. Moreover, the RSs are transmitted without the power boosting configuration [39].

B. WIRELESS CHANNELS

The wireless channels are generated according to the 3GPP specifications in [39]. The channel parameters are generated according to specific probability distributions, which depend on the wireless scenario and on the NLOS conditions [39]. In addition, the wireless channels, as well as the NLOS conditions, are generated using the 3GPP spatial consistency procedure via the QuaDRiGa wireless channel simulator [61].

C. 3GPP 38.901 INDOOR FACTORY-DENSE HIGH

The InF-DH scenario is a 3GPP-standardized scenario modeled to represent an indoor industrial environment characterized by high density of metallic objects with irregular structures [39]. Such characteristics determine heavy multipath propagation and high NLOS condition probabilities. The

scenario is modeled with correlation distances for the wireless channel parameters and for the NLOS conditions equal to 10 m and 1 m, respectively. The layout of the InF-DH scenario consists of an environment of dimension 120 m × 60 m, where 18 sites with an inter-site distance of 20 m and a height of 8 m are deployed (see Fig. 1(a)).

According to the specifications provided by 3GPP, the BSs are composed of 1 antenna sector in FR1 and of 3 antenna sectors in FR2, each of them covering 120°. The antenna patterns and characteristics are different in FR1 and in FR2, and are reported in [26]. The BSs transmitted power is equal to 24 dBm, and their noise figure is equal to 5 dB and 7 dB in FR1 and FR2, respectively. For determining the NLOS conditions, the effective clutter height is 2 m, the typical clutter size is 2 m, and the clutter density is 40% [26], [39]. Fig. 1(a) shows the layout of the InF-DH scenario, where the background represents an instantiation of the spatially consistent LOS map.

D. 3GPP 38.901 INDOOR FACTORY-SPARSE HIGH

The InF-SH scenario is a 3GPP-standardized scenario modeled to represent an indoor industrial environment, characterized by low density of metallic objects with regular metallic surfaces [39]. Such characteristics determine moderate multipath propagation and low NLOS condition probabilities. The scenario is modeled with correlation distances for the wireless channel parameters and for the NLOS conditions equal to 5 m and 1 m, respectively. The layout of the InF-DH scenario consists of an environment of dimension 300 m × 150 m, where 18 sites with an inter-site distance of 50 m and a height of 8 m are deployed (see Fig. 1(b)).

According to the specifications provided by 3GPP, the BSs are composed of 1 antenna sector in FR1 and of 3 antenna sectors in FR2, each of them covering 120°. The antenna patterns and characteristics are different in FR1 and in FR2, and are reported in [26]. The BSs transmitted power is equal to 24 dBm, and their noise figure is equal to 5 dB and 7 dB in FR1 and FR2, respectively. For determining the NLOS conditions, the effective clutter height is 2 m, the typical clutter size is 10 m, and the clutter density is 20% [26], [39]. Fig. 1(b) shows the layout of the InF-SH scenario, where the background represents an instantiation of the spatially consistent LOS map.

E. 3GPP 38.901 INDOOR OPEN OFFICE

The IOO scenario is a 3GPP-standardized scenario modeled to represent open cubicle areas typical of indoor environments, such as offices, shops, and shopping malls [39]. Accordingly, the IOO scenario exhibits scarce multipath propagation and low NLOS conditions probabilities. The scenario is modeled with correlation distances for the wireless channel parameters and for the NLOS conditions equal to 10 m. The layout of the IOO scenario consists of an environment of dimension 120 m × 50 m, where 12 sites with an inter-site distance of 20 m and a height of 3 m are deployed (see Fig. 1(c)).

According to the specifications provided by 3GPP, the BSs are composed of 1 antenna sector in FR1 and of 3 antenna sectors in FR2, each of them covering 120°. The antenna patterns and characteristics are different in FR1 and in FR2, and are reported in [25]. The BSs transmitted power is equal to 24 dBm, and their noise figure is equal to 5 dB and 7 dB in FR1 and FR2, respectively [25], [39]. Fig. 1(c) shows the layout of the IOO scenario, where the background represents an instantiation of the spatially consistent LOS map.

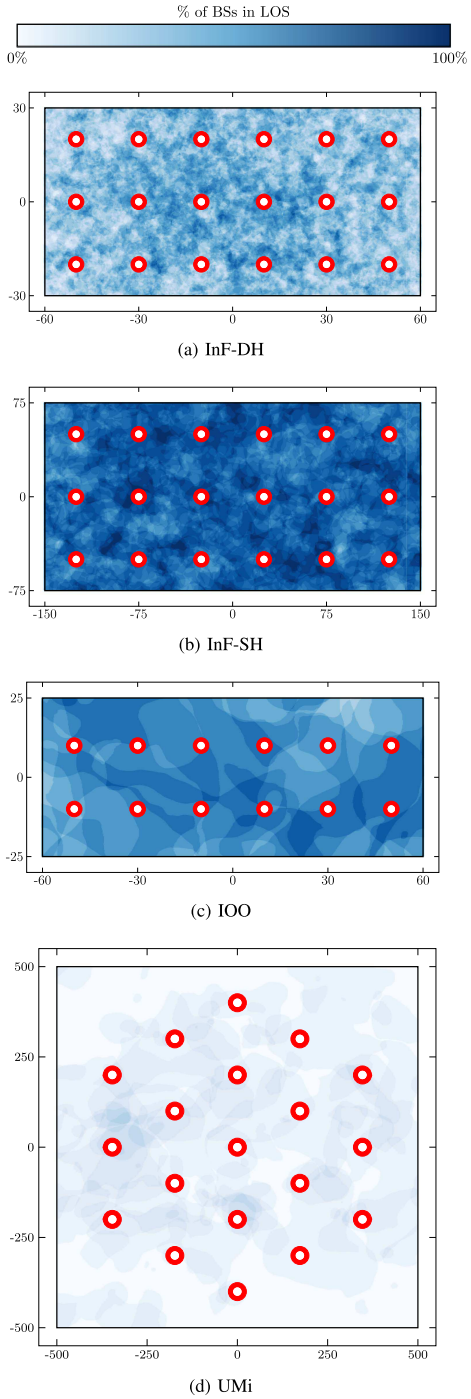

FIGURE 1. Layout of the 3GPP-standardized scenarios considered in xG-Loc. The red annuluses denote the BSs positions, while the background indicates a single random instantiation of the spatially consistent LOS map. The color map indicates the percentage of BSs in LOS conditions for each point of the map. The coordinates on the axes are in meters.

F. 3GPP 38.901 URBAN MICROCELL

The UMi scenario is a 3GPP-standardized scenario modeled to represent outdoor street canyons, hence capturing environments such as cities and stations squares [39]. Such characteristics determine low multipath propagation and high NLOS conditions probabilities. The scenario is modeled with correlation distances for the wireless channel parameters and for the NLOS conditions equal to 10 m. The layout of the UMi scenario consists of an environment of dimension 1000 m $\times$ 1000 m, where 19 sites with an inter-site distance of 200 m and a height of 10 m are deployed (see Fig. 1(d)). The minimum distance between a site and the UEs is equal to 10 m.

According to the specifications provided by 3GPP, the BSs are composed of 3 antenna sectors, each of them covering 120°. The antenna patterns and characteristics are different in FR1 and in FR2, and are reported in [25]. In FR2, multiple antenna panels are employed according to [25]. The BSs transmitted power is equal to 44 dBm, in FR1, and to 37 dBm per panel in FR2. The BSs noise figure is equal to 5 dB and 7 dB in FR1 and FR2, respectively [25], [39]. Fig. 1(d) shows the layout of the UMi scenario, where the background represents an instantiation of the spatially consistent LOS map.

G. USER EQUIPMENTS CHARACTERISTICS

The 5G UEs are deployed in the aforementioned scenarios, with a height equal to 1.5 m. Their positions and orientations are generated randomly following the 3GPP drop-based procedure (i.e., the UEs positions are generated following a uniform distribution) [25]. The UEs are assumed to be static in the instant of the RSs transmission.

According to the specifications provided by 3GPP, the UE is equipped with an omnidirectional antenna in FR1, and with a directional antenna in FR2, whose specific pattern is reported in [25]. The UEs transmitted power is equal to 23 dBm, and their noise figure is equal to 9 dB and 13 dB in FR1 and FR2, respectively.

IV. xG-Loc CHARACTERISTICS

This section presents the characteristics of xG-Loc [36]. First, the configurations of central frequency and bandwidths considered for the generation of xG-Loc are reported. Then, the directory organization of xG-Loc is described. Finally, the performance of SI-based localization and BI are evaluated on xG-Loc.

A. xG-Loc GENERATION

Several conditions for the RS transmission are considered in xG-Loc, including different central frequencies (both FR1 and FR2) and bandwidths (from 5 MHz to 400 MHz) in the scenarios described in Section III. Moreover, different numerologies and numbers of RBs are taken into account. This enables to provide data that can be used to evaluate different operational conditions, from reduced capability devices to high-end 5G devices. The configurations of central frequency f_c and bandwidth B are selected according to [25], [26], [42] and depend on the scenarios and are reported in Table 1. In addition, a summary of the main parameters of the 3GPP scenarios considered for generating xG-Loc and described in Section III are reported in Table 2.

TABLE 1 xG-Loc Configurations Characteristics

	Dataset Name	f_c [GHz]	B [MHz]	μ	N_{RB}
InF-DH	infdh_5_3.5	3.5	5	0	24
	infdh_20_3.5	3.5	20	0	104
	infdh_50_3.5	3.5	50	1	136
	infdh_100_3.5	3.5	100	1	272
	infdh_100_28	28	100	2	136
	infdh_200_28	28	200	2	272
	infdh_400_28	28	400	3	272
InF-SH	infsh_5_3.5	3.5	5	0	24
	infsh_20_3.5	3.5	20	0	104
	infsh_50_3.5	3.5	50	1	136
	infsh_100_3.5	3.5	100	1	272
	infsh_100_28	28	100	2	136
	infsh_200_28	28	200	2	272
	infsh_400_28	28	400	3	272
IOO	ioo_5_2	2	5	0	24
	ioo_20_2	2	20	0	104
	ioo_50_4	4	50	1	136
	ioo_100_4	4	100	1	272
	ioo_100_30	30	100	2	136
	ioo_200_30	30	200	2	272
	ioo_400_30	30	400	3	272
UMi	umi_5_2	2	5	0	24
	umi_20_2	2	20	0	104
	umi_50_4	4	50	1	136
	umi_100_4	4	100	1	272
	umi_100_30	30	100	2	136
	umi_200_30	30	200	2	272
	umi_400_30	30	400	3	272

B. xG-Loc STRUCTURE

xG-Loc is structured in different directories, where each of them contains all the data related to a certain configuration of scenario, bandwidth B, and central frequency f_c. The name of these directories for the different configurations is reported in Table 1. For each configuration, $N_{RUN} = 100$ random instantiations of the scenarios, indexed by $n \in \{0, 1, \ldots, N_{RUN} - 1\}$ are generated with spatially consistent wireless channels and NLOS maps. For each n, a corresponding folder, named RUN_n, is generated to contain the information related to the RS transmission involving $N_{UE} = 10$ UEs, indexed by $m \in \{0, 1, \ldots, N_{UE} - 1\}$. For each UE, a corresponding folder named UE_m is used to store information related to the m-th UE. Such folders contain the RSs exchanged between the m-th UEs and all the BS in

TABLE 2 Main Parameters of the 3GPP Scenarios

Parameter	InF-DH		InF-SH		IOO		UMi	
	FR1	FR2	FR1	FR2	FR1	FR2	FR1	FR2
Environment size	120 m × 60 m		300 m × 150 m		120 m × 50 m		1000 m × 1000 m	
Number of sites	18		18		12		19	
Inter-site distance	20 m		50 m		20 m		200 m	
Site height	8 m		8 m		3 m		10 m	
n° antenna sectors	1	3	1	3	1	3	3	
BS noise figure	5 dB	7 dB	5 dB	7 dB	5 dB	7 dB	5 dB	7 dB
BS transmitted power	24 dBm		24 dBm		24 dBm		44 dBm	37 dBm
Channel corr. dist.	10 m		5 m		10 m		10 m	
NLOS corr. dist.	1 m		1 m		10 m		10 m	
NLOS clutter height	2 m		2 m		–		–	
NLOS typical clutter size	2 m		10 m		–		–	
NLOS clutter density	40 %		20 %		–		–	

a compressed folder named `RX_PRS_RUN_`n`_UE_`m`.zip` and `RX_SRS_RUN_`n`_UE_`m`.zip` for the received PRS and SRS, respectively.

The sampled versions of the received RSs are saved in text files in the aforementioned directories and named as `RX_PRS_RUN_`n`_UE_`m`_BS_`j`.txt`, for $j \in \{0, 1, \ldots, N_\mathrm{b}\}$. The signals saved in the text files are first pre-processed to reduce their dimensionality and to provide data that can be more easily handled by localization algorithms as in the following. First, the baseband RSs are processed to remove the cyclic prefix introduced for the transmission via OFDM [42]. Then, the N_OCC occasions of the RSs are coherently accumulated, and the samples not corresponding to allocated symbols are discarded. Finally, given the time-frequency configuration of the RSs reported in Section III-A, groups of K symbols are accumulated.

In addition to samples of RSs, a set of synthetic data obtained from the RSs are reported in a JavaScript object notation (JSON) file named `INFO_RUN_`n`_UE_`m`.json`, whose keys and values are described in the following:

- `UE_REAL_POS` indicates the vector

$$\boldsymbol{p} = [x, y, z]^\mathrm{T} \tag{12}$$

  with the coordinates of the real 3D position in meters for the UE in the scenario. For each scenario, the coordinate system is as in Fig. 1.
- `UE_EST_POS_SI_RTT` indicates the vector

$$\hat{\boldsymbol{p}} = [\hat{x}, \hat{y}, z]^\mathrm{T} \tag{13}$$

  with the coordinates of the estimated position in meters for the UE in the scenario via SI-based localization. SI-based localization is performed in 2D, hence only $\hat{x}$ and $\hat{y}$ are estimated. The position is estimated using only the information obtained from RTT measurements and with $N_\mathrm{M} = 8$ components in the GMM. The UE height $z = 1.5$ m is fixed and is included in $\hat{\boldsymbol{p}}$ as known information.
- `DIST_DL_TOA` and `DIST_UL_TOA`, indicate the vectors

$$\boldsymbol{d}_\mathrm{d} = \left[c\hat{\tau}_{\mathrm{d}0}, c\hat{\tau}_{\mathrm{d}1}, \ldots, c\hat{\tau}_{\mathrm{d}N_\mathrm{b}}\right]^\mathrm{T} \tag{14a}$$

$$\boldsymbol{d}_\mathrm{u} = \left[c\hat{\tau}_{\mathrm{u}0}, c\hat{\tau}_{\mathrm{u}1}, \ldots, c\hat{\tau}_{\mathrm{u}N_\mathrm{b}}\right]^\mathrm{T} \tag{14b}$$

  respectively, containing the 3D distance estimates in meters between the N_b BSs and the UE. The estimates are based on the transmission of the PRS in (14a), and of the SRS in (15), respectively. The TOA estimates are obtained via the detection of the highest peak in the estimated CIR. For simplicity, the estimated TOA $\hat{\tau}$ is mapped into distances via the multiplication with the speed of light c.
- `EST_AOD` indicates the vector

$$\boldsymbol{\alpha} = [\alpha_0, \alpha_1, \ldots, \alpha_{N_\mathrm{s}}]^\mathrm{T} \tag{15}$$

  containing the AOD estimates between the N_s sites and the UE in degrees. Given the 3GPP specifications for the BSs antenna configurations, a sufficient number of antennas to enable the AOD estimation is available only in FR2 [25], [26]. The AOD estimation is accomplished leveraging a beam sweeping procedure with an angular resolution of 10° as briefly described in Section II-D. The complete algorithm used for the AOD estimation can be found in [50]. Fig. 2 shows the angular coordinate system used for the AOD estimation. Moreover, it shows the normalized antenna radiation pattern of a single antenna sector for some steering configurations. Note that for the configurations where the AOD is not estimated, this field is left empty.
- `NLOS_STATE` indicates the vector

$$\boldsymbol{\gamma} = \left[\gamma_0, \gamma_1, \ldots, \gamma_{N_\mathrm{b}}\right]^\mathrm{T} \tag{16}$$

  containing the binary indicators of NLOS conditions between the BSs and the UE. The value "1" indicates that the UE is in NLOS condition with respect to the BS,

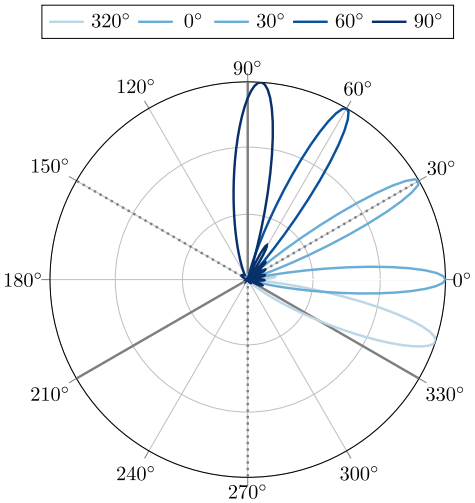

FIGURE 2. Polar plot representing the normalized antenna radiation patterns of a single BS sector considering different steering vectors. The grey solid lines delimit the antenna sectors, while the dotted lines indicate the angle to which the antenna is directed.

while the value "0" indicates that the UE and the BS are in LOS condition.

- `BI` indicates the vector

$$\boldsymbol{\psi} = \left[\psi(\bar{\boldsymbol{\nu}}_0), \psi(\bar{\boldsymbol{\nu}}_1), \ldots, \psi(\bar{\boldsymbol{\nu}}_{N_\mathrm{b}})\right]^\mathrm{T} \tag{17}$$

containing the probabilistic values provided by BI as described in Section II-F for each wireless link between the BSs and the UE. The vector of statistical indicators $\bar{\boldsymbol{\nu}}$ is obtained by averaging the feature vector $\boldsymbol{\nu}^{\mathrm{PRS}}$ obtained transmitting the PRS and the feature vector $\boldsymbol{\nu}^{\mathrm{SRS}}$ obtained transmitting the SRS. The selected features are as in [38].

In the outer directory, some other information necessary to perform localization is available. Specifically, there are two compressed folders, named `TX_PRS.zip` and `TX_SRS.zip`, which contain the transmitted PRS and SRS, respectively. To simplify the use of such data, the same preprocessing applied to the received RS is applied also to the transmitted RS.

In addition, a set of relevant information on the simulation settings is reported in a JSON file named `SIMULATION_INFO.json`, whose keys and values are described in the following:

- `SCENARIO_COORD` indicates the coordinates of the corners of the scenario as in [39] and as depicted in Fig. 1.
- `BS_COORD` indicates the coordinates of the BSs in the scenario; and
- `SAMPLING_RATE` indicates the sampling rate in samples/s used for saving the RSs.

The complete directory organization of a single configuration of xG-Loc is shown in Fig. 3.

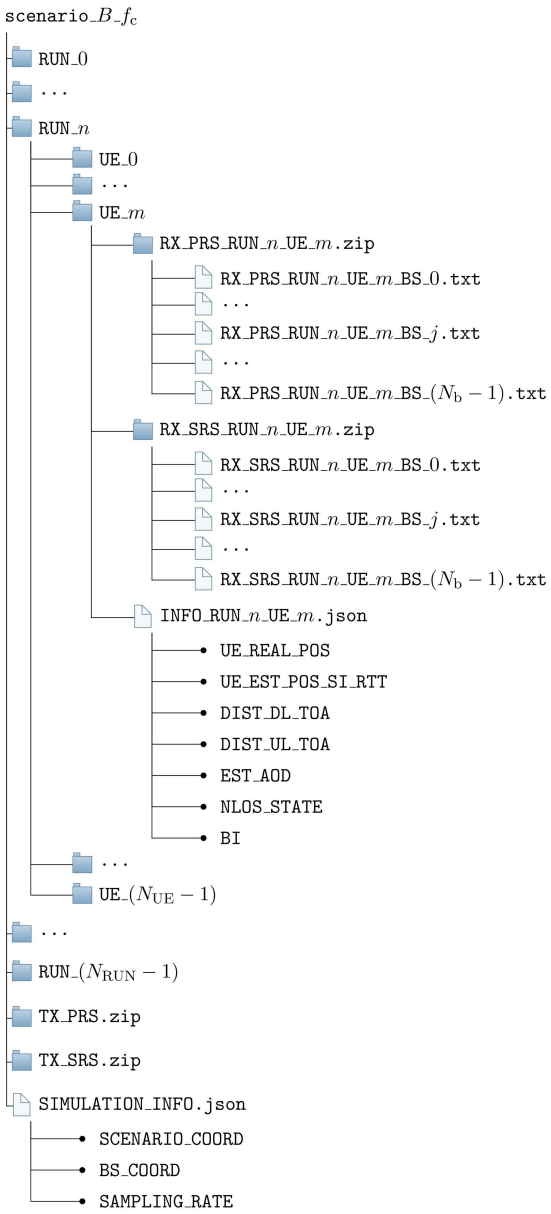

FIGURE 3. xG-Loc directory organization for a single configuration.

V. PERFORMANCE BENCHMARKS

This section presents performance benchmarks for the different datasets in xG-Loc. First, the performance of SI-based localization are reported for all the configurations considered. Then, the performance of BI are evaluated considering the metrics briefly described in Section II-F.

A. SI-BASED LOCALIZATION PERFORMANCE

Table 3 shows relevant percentiles of the horizontal localization error obtained via SI-based localization for the different configurations considered using only RTT measurements. Such statistics summarize and describe the error of the position estimates reported in the field `UE_EST_POS_SI_RTT` over all the runs and UEs for each configuration. Further

TABLE 3 Error Percentiles for SI-Based Localization With RTT Measurements

Dataset	10th	20th	30th	40th	50th	60th	70th	80th	90th	95th	99th
infdh_5_3.5	3.95 m	6.05 m	7.97 m	9.82 m	11.50 m	13.48 m	16.05 m	19.53 m	26.72 m	36.72 m	55.65 m
infdh_20_3.5	1.21 m	1.74 m	2.27 m	2.61 m	3.19 m	3.68 m	4.40 m	5.47 m	8.47 m	12.53 m	22.83 m
infdh_50_3.5	0.48 m	0.75 m	1.02 m	1.23 m	1.50 m	1.82 m	2.15 m	2.57 m	3.92 m	6.70 m	18.34 m
infdh_100_3.5	0.36 m	0.56 m	0.72 m	0.84 m	0.96 m	1.12 m	1.32 m	1.52 m	2.04 m	2.76 m	13.69 m
infdh_100_28	0.45 m	0.63 m	0.82 m	0.97 m	1.10 m	1.29 m	1.50 m	1.76 m	2.21 m	2.76 m	11.75 m
infdh_200_28	0.54 m	0.68 m	0.79 m	0.88 m	0.99 m	1.08 m	1.22 m	1.44 m	1.85 m	2.96 m	13.72 m
infdh_400_28	0.56 m	0.65 m	0.74 m	0.83 m	0.90 m	1.01 m	1.17 m	1.49 m	2.95 m	7.53 m	18.30 m
infsh_5_3.5	2.95 m	4.20 m	5.22 m	6.47 m	7.60 m	9.19 m	10.55 m	12.26 m	15.38 m	18.38 m	27.29 m
infsh_20_3.5	0.60 m	0.86 m	1.10 m	1.32 m	1.56 m	1.86 m	2.17 m	2.51 m	3.03 m	3.49 m	4.74 m
infsh_50_3.5	0.30 m	0.43 m	0.55 m	0.67 m	0.78 m	0.90 m	1.00 m	1.18 m	1.40 m	1.59 m	2.10 m
infsh_100_3.5	0.13 m	0.20 m	0.24 m	0.29 m	0.34 m	0.40 m	0.46 m	0.55 m	0.66 m	0.75 m	1.05 m
infsh_100_28	0.15 m	0.21 m	0.27 m	0.33 m	0.39 m	0.45 m	0.52 m	0.61 m	0.76 m	0.90 m	1.25 m
infsh_200_28	0.07 m	0.11 m	0.13 m	0.16 m	0.19 m	0.22 m	0.26 m	0.31 m	0.38 m	0.44 m	0.60 m
infsh_400_28	0.04 m	0.06 m	0.07 m	0.09 m	0.10 m	0.12 m	0.14 m	0.16 m	0.21 m	0.26 m	0.38 m
ioo_5_2	2.35 m	3.45 m	4.29 m	5.02 m	5.81 m	6.44 m	7.64 m	9.26 m	12.45 m	16.43 m	36.52 m
ioo_20_2	0.74 m	1.07 m	1.38 m	1.78 m	2.12 m	2.58 m	3.14 m	3.95 m	5.28 m	6.77 m	8.62 m
ioo_50_4	0.47 m	0.67 m	0.84 m	1.01 m	1.17 m	1.35 m	1.55 m	1.79 m	2.29 m	3.09 m	6.32 m
ioo_100_4	0.20 m	0.30 m	0.38 m	0.45 m	0.53 m	0.62 m	0.72 m	0.87 m	1.10 m	1.40 m	2.13 m
ioo_100_30	0.27 m	0.40 m	0.50 m	0.63 m	0.73 m	0.84 m	1.03 m	1.30 m	1.66 m	2.20 m	4.09 m
ioo_200_30	0.16 m	0.24 m	0.30 m	0.36 m	0.43 m	0.50 m	0.59 m	0.76 m	1.11 m	1.50 m	2.73 m
ioo_400_30	0.06 m	0.09 m	0.12 m	0.14 m	0.17 m	0.21 m	0.25 m	0.32 m	0.51 m	0.70 m	1.36 m
umi_5_2	2.86 m	4.08 m	5.11 m	5.98 m	7.10 m	8.37 m	9.66 m	10.98 m	13.62 m	15.61 m	20.07 m
umi_20_2	1.64 m	2.20 m	2.77 m	3.27 m	3.79 m	4.45 m	5.13 m	5.84 m	6.93 m	7.99 m	10.68 m
umi_50_4	1.14 m	1.69 m	2.16 m	2.68 m	3.11 m	3.60 m	4.16 m	4.88 m	5.86 m	7.08 m	8.64 m
umi_100_4	1.00 m	1.53 m	2.01 m	2.44 m	2.87 m	3.28 m	3.76 m	4.35 m	5.09 m	5.91 m	7.33 m
umi_100_30	1.50 m	2.11 m	2.65 m	3.19 m	3.79 m	4.45 m	5.11 m	6.07 m	7.34 m	8.96 m	14.92 m
umi_200_30	1.27 m	1.88 m	2.44 m	2.93 m	3.32 m	3.76 m	4.48 m	5.31 m	6.37 m	7.64 m	12.29 m
umi_400_30	1.19 m	1.90 m	2.49 m	2.97 m	3.56 m	4.04 m	4.75 m	5.82 m	7.60 m	10.45 m	40.27 m

localization performance benchmarks with different types of measurements can be found in [25], [26], [38], [50].

First, it can be observed that for all the configurations considered, the localization performance tends to improve by increasing the bandwidth. This is because of the higher sampling rate which determines a higher resolution in the TOA estimation. Moreover, it can be observed that the performance in the UMi scenario are worse with respect to the ones in indoor scenarios.

In the InF-DH scenario, it can be observed that an accuracy of around 1 m at the 60th percentile and of around 2 m at the 90th percentile (i.e., 90% of the times the error is below 2 m) can be obtained with a bandwidth higher than 100 MHz. Moreover, it can be observed that the use of a bandwidth of 5 MHz is not sufficient for providing localization in the scenario.

In the InF-SH scenario, it can be observed that an accuracy of around 3 m at the 90th percentile can be achieved in FR1 with a bandwidth of 20 MHz. It can also be observed that for every configuration with a bandwidth higher than 100 MHz, it is possible to achieve sub-meter localization accuracy at the 90th percentile with SI-based approach. Moreover, a sub-meter accuracy can be achieved even at the 99th in FR2 if the bandwidth is higher than 200 MHz. Specifically, the errors are 0.6 m and 0.38 m at the 99th percentile with bandwidths of 200 MHz and 400 MHz, respectively.

In the IOO scenario, it can be observed that accuracy of around 1 m at the 90th percentile can be achieved in FR1 with a bandwidth of 100 MHz. It can also be observed that the best localization performance is obtained in FR2 at 30 GHz with a bandwidth of 400 MHz. Specifically, with this configuration, it is possible to achieve a localization accuracy of 0.51 m and 1.36 m at the 90th and at the 99th percentile, respectively.

In the UMi scenario, it can be observed that the localization performance is worse with respect to other scenarios. This is due to the characteristics of such scenario, which is characterized by a high NLOS probability, as shown in Fig. 1(d). In particular, it can be observed that the localization performance is better in FR1 than in FR2. Given the size of the scenario, this is due to the degradation of the RSs caused by the path

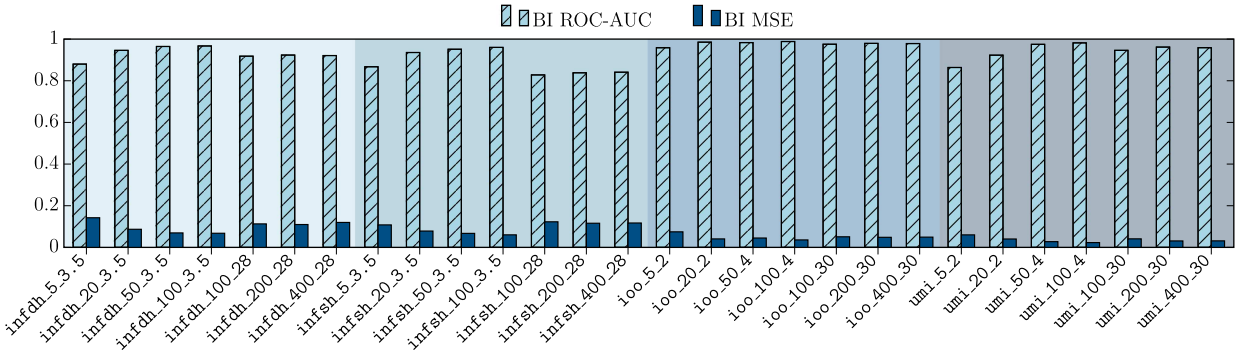

FIGURE 4. Histogram of the BI ROC-AUC and of the BI MSE for the different configurations in xG-Loc.

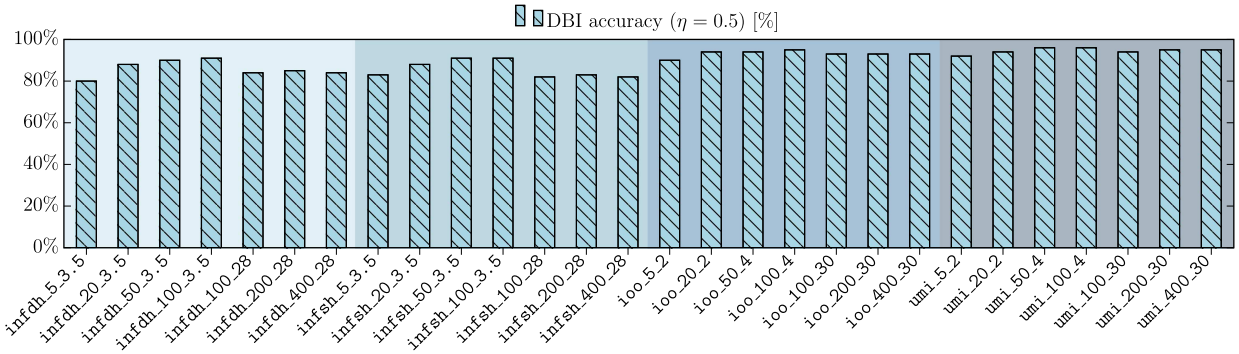

FIGURE 5. Histogram of the DBI identification accuracy for the different configurations in xG-Loc.

loss which highly impacts the quality of the RTT measurements used for localization.

B. BI PERFORMANCE

Fig. 4 shows the BI ROC-AUC and the BI MSE. These metrics are evaluated on the fields `BI` over all the runs and UEs for each configuration. It can be observed that the BI ROC-AUC increases with the bandwidth within the same frequency range, while the BI MSE decreases. This is because, similarly to localization performance, higher bandwidths correspond to higher sampling frequencies, which determines a better estimation of the CIR via the RSs cross-correlation. The high values of ROC-AUC and the low values of MSE demonstrate the quality of the probability estimation provided via BI. Fig. 5 shows the DBI NLOS identification accuracy assuming $\eta = 0.5$. It can be observed that the identification accuracy of DBI is between the 80% and the 90% for the configurations in the InF-DH and in the InF-SH scenarios, while it is over the 90% for all the configurations in the IOO and UMi scenarios.

VI. FINAL REMARK

This paper introduces xG-Loc, the first open dataset for localization algorithms and services that is fully compliant with 3GPP technical reports and specifications. xG-Loc includes localization signals, measurements, and analytics for different network and signal configurations in indoor and outdoor scenarios with center frequencies from micro-waves in FR1 to millimeter-waves in FR2. The rich data provided by xG-Loc enables the systematic development and testing of localization algorithms and services in scenarios compliant with 3GPP specifications. The public availability of xG-Loc paves the way to advance the evolution of xG networks.

REFERENCES

[1] "Technical Specification Group Radio Access Network; Technical Specification Group Services and System Aspects; Study on Positioning Use Cases," 3GPP, Sophia Antipolis, France, Tech. Rep. 22.872 V16.1.0, 2018.
[2] M. Z. Win, Y. Shen, and W. Dai, "A theoretical foundation of network localization and navigation," *Proc. IEEE*, vol. 106, no. 7, pp. 1136–1165, Jul. 2018.
[3] W. Chen et al., "5G-advanced toward 6G: Past, present, and future," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 6, pp. 1592–1619, Jun. 2023.
[4] M. Z. Win et al., "Network localization and navigation via cooperation," *IEEE Commun. Mag.*, vol. 49, no. 5, pp. 56–62, May 2011.
[5] A. Conti et al., "Location awareness in beyond 5G networks," *IEEE Commun. Mag.*, vol. 59, no. 11, pp. 22–27, Nov. 2021.
[6] N. Decarli, A. Guerra, C. Giovannetti, F. Guidi, and B. M. Masini, "V2X sidelink localization of connected automated vehicles," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 1, pp. 120–133, Jan. 2024.
[7] J. Ji, A. Khajepour, W. W. Melek, and Y. Huang, "Path planning and tracking for vehicle collision avoidance based on model predictive control with multiconstraints," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 952–964, Feb. 2017.
[8] M. Girdhar, J. Hong, and J. Moore, "Cybersecurity of autonomous vehicles: A systematic literature review of adversarial attacks and defense models," *IEEE Open J. Veh. Technol.*, vol. 4, pp. 417–437, 2023.
[9] J. Thomas, J. Welde, G. Loianno, K. Daniilidis, and V. Kumar, "Autonomous flight for detection, localization, and tracking of moving targets with a small quadrotor," *IEEE Robot. Autom. Lett.*, vol. 2, no. 3, pp. 1762–1769, Jul. 2017.

[10] R. Karlsson and F. Gustafsson, "The future of automotive localization algorithms: Available, reliable, and scalable localization: Anywhere and anytime," *IEEE Signal Process. Mag.*, vol. 34, no. 2, pp. 60–69, Mar. 2017.

[11] W. Yuan, S. Li, L. Xiang, and D. W. K. Ng, "Distributed estimation framework for beyond 5G intelligent vehicular networks," *IEEE Open J. Veh. Technol.*, vol. 1, pp. 190–214, 2020.

[12] K. Lin, M. Chen, J. Deng, M. M. Hassan, and G. Fortino, "Enhanced fingerprinting and trajectory prediction for IoT localization in smart buildings," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 3, pp. 1294–1307, Jul. 2016.

[13] L. Catarinucci et al., "An IoT-aware architecture for smart healthcare systems," *IEEE Internet Things J.*, vol. 2, no. 6, pp. 515–526, Dec. 2015.

[14] Y. Li et al., "Toward location-enabled IoT (LE-IoT): IoT positioning techniques, error sources, and error mitigation," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4035–4062, Mar. 2021.

[15] J. Struye, F. Lemic, and J. Famaey, "CoVRage: Millimeter-wave beamforming for mobile interactive virtual reality," *IEEE Trans. Wireless Commun.*, vol. 22, no. 7, pp. 4828–4842, Jul. 2023.

[16] D. G. Morín, P. Pérez, and A. G. Armada, "Toward the distributed implementation of immersive augmented reality architectures on 5G networks," *IEEE Commun. Mag.*, vol. 60, no. 2, pp. 46–52, Feb. 2022.

[17] S. K. Sharma, I. Woungang, A. Anpalagan, and S. Chatzinotas, "Toward tactile internet in beyond 5G era: Recent advances, current issues, and future directions," *IEEE Access*, vol. 8, pp. 56948–56991, 2020.

[18] A. F. G. Gonçalves Ferreira, D. M. A. Fernandes, A. P. Catarino, and J. L. Monteiro, "Localization and positioning systems for emergency responders: A survey," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2836–2870, Fourthquarter 2017.

[19] J. Li et al., "5G New Radio for public safety mission critical communications," *IEEE Commun. Stand. Mag.*, vol. 6, no. 4, pp. 48–55, Dec. 2022.

[20] J. Bravo-Arrabal, P. Zambrana, J. J. Fernandez-Lozano, J. A. Gomez-Ruiz, J. S. Barba, and A. García-Cerezo, "Realistic deployment of hybrid wireless sensor networks based on ZigBee and LoRa for search and rescue applications," *IEEE Access*, vol. 10, pp. 64618–64637, 2022.

[21] "Technical Specification Group Radio Access Network; NG Radio Access Network (NG-RAN); Stage 2 functional specification of User Equipment (UE) positioning in NG-RAN," 3GPP, Sophia Antipolis, France, Tech. Specification 38.305 V17.6.0, 2023.

[22] M. Koivisto, A. Hakkarainen, M. Costa, P. Kela, K. Leppanen, and M. Valkama, "High-efficiency device positioning and location-aware communications in dense 5G networks," *IEEE Commun. Mag.*, vol. 55, no. 8, pp. 188–195, Aug. 2017.

[23] G. Kwon, A. Conti, H. Park, and M. Z. Win, "Joint communication and localization in millimeter wave networks," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 6, pp. 1439–1454, Nov. 2021.

[24] "Technical Specification Group Services and System Aspects; Service Requirements for the 5G System; Stage 1," 3GPP, Sophia Antipolis, France, Tech. Specification, 22.261 V19.4.0, 2023.

[25] "Technical Specification Group Radio Access Network; Study on NR Positioning Support," 3GPP, Sophia Antipolis, France, Tech. Rep. 38.855 V16.0.0, 2019.

[26] "Technical Specification Group Radio Access Network; Study on NR Positioning Enhancements," 3GPP, Sophia Antipolis, France, Tech. Rep. 38.857 V17.0.0, 2021.

[27] "Technical Specification Group Radio Access Network; Study on Expanded and Improved NR Positioning," 3GPP, Sophia Antipolis, France, Tech. Rep. 38.859 V18.0.0, 2023.

[28] S. Dwivedi et al., "Positioning in 5G networks," *IEEE Commun. Mag.*, vol. 59, no. 11, pp. 38–44, Nov. 2021.

[29] M. Pan et al., "5G CFR/CSI dataset for wireless channel parameter estimation, array calibration, and indoor positioning," *IEEE Dataport*, 2022, doi: 10.21227/k2f0-k132.

[30] Y. Zhang, J. Sun, G. Gui, H. Gacanin, and H. Sari, "Channel state information dataset for 5G New Radio system based on ray-tracing," *IEEE Dataport*, 2021, doi: 10.21227/1vss-kc05.

[31] J. Zhang, "WiFi fingerprinting dataset for IP-TAG," *IEEE Dataport*, 2023, doi: 10.21227/fp3p-bf42.

[32] G. Maus et al., "Bluetooth 5.1 angle of arrival based indoor localization," *IEEE Dataport*, 2021, doi: 10.21227/2j4h-3w77.

[33] Q. Wang, "NLOS occlusion UWB TOA ranging dataset in indoor environment," *IEEE Dataport*, 2023, doi: 10.21227/fx7b-7p46.

[34] K. Bregar, "UWB positioning and tracking data set," *Zenodo*, 2023, doi: 10.5281/zenodo.5556170.

[35] S. Angarano, V. Mazzia, F. Salvetti, G. Fantin, and M. Chiaberge, "DeepUWB," *Zenodo*, 2020, doi: 10.5281/zenodo.4290068.

[36] A. Conti, G. Torsoli, C. A. Gómez-Vega, A. Vaccari, and M. Z. Win, "xG-Loc: 3GPP-compliant datasets for xG location-aware networks," *IEEE Dataport*, 2023, doi: 10.21227/rper-vc03.

[37] A. Conti, S. Mazuelas, S. Bartoletti, W. C. Lindsey, and M. Z. Win, "Soft information for localization-of-things," *Proc. IEEE*, vol. 107, no. 11, pp. 2240–2264, Nov. 2019.

[38] G. Torsoli, M. Z. Win, and A. Conti, "Blockage intelligence in complex environments for beyond 5G localization," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 6, pp. 1688–1701, Jun. 2023.

[39] "Technical Specification Group Radio Access Network; Study on Channel Model for Frequencies From 0.5 to 100 GHz," 3GPP, Sophia Antipolis, France, Tech. Rep. 38.901 V17.0.0, 2022.

[40] "Technical Specification Group Radio Access Network; NR; User Equipment (UE) Transmission and Reception," 3GPP, Sophia Antipolis, France, Tech. Specification 38.101-1 V18.3.0, 2023.

[41] "Technical Specification Group Radio Access Network; NR; User Equipment (UE) Transmission and Reception; Part 2: Range 2 Standalone," 3GPP, Sophia Antipolis, France, Tech. Specification 38.101-2 V18.3.0, 2023.

[42] "Technical Specification Group Radio Access Network; NR; Physical Channels and Modulation," 3GPP, Sophia Antipolis, France, Tech. Specification, 38.211 V17.5.0, 2023.

[43] "Technical Specification Group Radio Access Network; NG-RAN; Physical Layer Procedures for Data," 3GPP, Sophia Antipolis, France, Tech. Specification 38.214 V17.4.0, 2023.

[44] A. F. Molisch, *Wireless Communications: From Fundamentals to Beyond 5G*, 3rd ed. U.K.: IEEE Press, Wiley, 2022.

[45] "Technical Specification Group Radio Access Network; NR; Physical Layer Measurements," 3GPP, Sophia Antipolis, France, Tech. Specification, 38.215 V17.2.0, 2022.

[46] H. Ryden, A. A. Zaidi, S. M. Razavi, F. Gunnarsson, and I. Siomina, "Enhanced time of arrival estimation and quantization for positioning in LTE networks," in *Proc. IEEE 27th Int. Symp. Pers., Indoor, Mobile Radio Commun.*, 2016, pp. 1–6.

[47] I. Guvenc and C.-C. Chong, "A survey on TOA based wireless localization and NLOS mitigation techniques," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 3, pp. 107–124, Third Quarter 2009.

[48] G. Torsoli, M. Z. Win, and A. Conti, "Selection of reference base station for TDOA-based localization in 5G and beyond IIoT," in *Proc. IEEE Glob. Telecommun. Conf.*, 2022, pp. 317–322.

[49] M. Giordani, M. Polese, A. Roy, D. Castor, and M. Zorzi, "A tutorial on beam management for 3GPP NR at mmWave frequencies," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 173–196, Firstquarter 2019.

[50] F. Morselli, S. M. Razavi, M. Z. Win, and A. Conti, "Soft information based localization for 5G networks and beyond," *IEEE Trans. Wireless Commun.*, vol. 22, no. 12, pp. 9923–9938, Dec. 2023.

[51] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.

[52] V. N. Vapnik, *Statistical Learning Theory*. New York, NY, USA: Wiley, 1998.

[53] J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: A statistical view of boosting," *Ann. Statist.*, vol. 28, pp. 337–407, Apr. 2000.

[54] L. Breiman, J. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. Belmont, CA, USA: Wadsworth Publishing Company, 1984.

[55] H. L. Van Trees*Detection, Estimation, and Modulation Theory*, 1st ed. New York, NY, USA: Wiley, 1968.

[56] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. New York, NY, USA: Springer, 2009.

[57] N. Japkowicz and M. Shah, *Evaluating Learning Algorithms: A Classification Perspective*. New York, NY, USA: Cambridge Univ. Press, 2011.

[58] "Technical Specification Group Radio Access Network; Study on New Radio Access Technology Physical Layer Aspects," 3GPP, Sophia Antipolis, France, Tech. Rep. 38.802 V14.2.0, 2017.

[59] G. Torsoli, S. Stracca, M. Puleri, M. Z. Win, and A. Conti, "Beyond 5G-assisted automatic guided vehicles in seaport environments," in *Proc. IEEE Int. Conf. Commun.*, 2023, pp. 1038–1043.

[60] "Technical Specification Group Radio Access Network; Base Station (BS) Radio Transmission and Reception," 3GPP, Sophia Antipolis, France, Tech. Rep. 38.104 V18.3.0, 2023.
[61] S. Jaeckel, L. Raschkowski, K. Börner, and L. Thiele, "QuaDRiGa: A 3-D multi-cell channel model with time evolution for enabling virtual field trials," *IEEE Trans. Antennas Propag.*, vol. 62, no. 6, pp. 3242–3256, Jun. 2014.

ANDREA CONTI (Fellow, IEEE) is a Professor and founding director of the Wireless Communication and Localization Networks Laboratory at the University of Ferrara, Italy. Prior to joining the University of Ferrara, he was with CNIT and with IEIIT-CNR.

In Summer 2001, he was with the Wireless Systems Research Department at AT&T Research Laboratories. Since 2003, he has been a frequent visitor to the Wireless Information and Network Sciences Laboratory at the Massachusetts Institute of Technology, where he presently holds the Research Affiliate appointment. His research interests involve theory and experimentation of wireless communication and localization systems. His current research topics include network localization and navigation, distributed sensing, adaptive diversity communications, and quantum information science.

Dr. Conti has served as editor for IEEE journals and chaired international conferences. He was elected Chair of the IEEE Communications Society's Radio Communications Technical Committee and is Co-founder of the IEEE Quantum Communications & Information Technology Emerging Technical Subcommittee. He received the HTE Puskás Tivadar Medal, the IEEE Communications Society's Fred W. Ellersick Prize, and the IEEE Communications Society's Stephen O. Rice Prize in the field of Communications Theory. He is an elected Fellow of the IEEE and of the IET, and a member of Sigma Xi. He has been selected as an IEEE Distinguished Lecturer.

GIANLUCA TORSOLI (Graduate Student Member, IEEE) received the Laurea degree (*summa cum laude*) in Electronics for ICT engineering in 2021 from the University of Ferrara, Italy. He is currently pursuing the Ph.D. degree with the Wireless Communication and Localization Networks Laboratory, University of Ferrara, Italy.

Since 2021, he has been with the Wireless Communication and Localization Networks Laboratory at the University of Ferrara. His research interests include network localization and navigation, and joint sensing and communication.

Mr. Torsoli serves as a reviewer for various IEEE journals and international conferences.

CARLOS A. GÓMEZ-VEGA (Graduate Student Member, IEEE) received the M.S. degree in electronic engineering from the Universidad Autónoma de San Luis Potosí, San Luis Potosí, Mexico, in 2020. He is currently pursuing the Ph.D. degree with the Wireless Communication and Localization Networks Laboratory, University of Ferrara, Italy.

Since 2020, he has been a Research Assistant with the Wireless Communication and Localization Networks Laboratory, University of Ferrara. In 2019, he was a Visiting Student with the Wireless Information and Network Sciences Laboratory, Massachusetts Institute of Technology, USA. His current research interests include network localization and navigation, and wireless resource optimization.

Mr. Gómez-Vega received the Best Paper Award from the IEEE LATINCOM in 2019. He also serves as a reviewer for various IEEE journals and international conferences.

ALESSANDRO VACCARI (Graduate Student Member, IEEE) received the Laurea degree (*summa cum laude*) in Electronics for ICT engineering in 2022 from the University of Ferrara, Italy. He is currently pursuing the Ph.D. degree with the Wireless Communication and Localization Networks Laboratory, University of Ferrara, Italy.

Since 2022, he has been with the Wireless Communication and Localization Networks Laboratory at the University of Ferrara. His research interests include multi-target tracking and classification.

Mr. Vaccari serves as a reviewer for various IEEE journals and international conferences.

GIANLUCA MAZZINI (Fellow, IEEE) is an Associate Professor with the University of Ferrara, Italy. Prior to joining the University of Ferrara, he was with the University of Bologna. He achieved milestones in both academic and professional journeys.

His research encompasses diverse areas, including spread spectrum communications, applications of chaos in telecommunications, efficient radio network architectures, peer-to-peer networks, and information security.

Dr. Mazzini has served as editor for IEEE journals and chaired international conferences. He leads several national and European projects, and sits on various boards of directors or management. He also served as the CEO of Lepida, an Italian in-house public company.

MOE Z. WIN (Fellow, IEEE) is a Professor at the Massachusetts Institute of Technology (MIT) and the founding director of the Wireless Information and Network Sciences Laboratory. Prior to joining MIT, he was with AT&T Research Laboratories and with NASA Jet Propulsion Laboratory.

His research encompasses fundamental theories, algorithm design, and network experimentation for a broad range of real-world problems. His current research topics include ultra-wideband systems, network localization and navigation, network interference exploitation, and quantum information science. He has served the IEEE Communications Society as an elected Member-at-Large on the Board of Governors, as elected Chair of the Radio Communications Committee, and as an IEEE Distinguished Lecturer. Over the last two decades, he held various editorial positions for IEEE journals and organized numerous international conferences. Recently, he has served on the SIAM Diversity Advisory Committee.

Dr. Win is an elected Fellow of the AAAS, the EURASIP, the IEEE, and the IET. He was honored with two IEEE Technical Field Awards: the IEEE Kiyo Tomiyasu Award (2011) and the IEEE Eric E. Sumner Award (2006, jointly with R. A. Scholtz). His publications, co-authored with students and colleagues, have received several awards. Other recognitions include the MIT Everett Moore Baker Award (2022), the IEEE Vehicular Technology Society James Evans Avant Garde Award (2022), the IEEE Communications Society Edwin H. Armstrong Achievement Award (2016), the Cristoforo Colombo International Prize for Communications (2013), the Copernicus Fellowship (2011) and the *Laurea Honoris Causa* (2008) from the Universitá degli Studi di Ferrara, and the U.S. Presidential Early Career Award for Scientists and Engineers (2004).

Open Access funding provided by 'Università degli Studi di Ferrara' within the CRUI CARE Agreement