A Novel TOA-Based Mobile Localization Technique Under Mixed LOS/NLOS Conditions for Cellular Networks

Zohair Abu-Shaban, *Student Member, IEEE*, Xiangyun Zhou, *Member, IEEE*, and Thushara D. Abhayapala, *Senior Member, IEEE*

***Abstract*—The presence of a non-line-of-sight (NLOS) link between a base station (BS) and a mobile station (MS) in a cellular network is a major issue that limits the performance of the majority of time-of-arrival (TOA) localization methods. Due to blocking obstacles, a signal has to travel a longer distance to reach the other end of the communication link. Thus, the additional distance introduced by the presence of an NLOS link is usually modeled by a positive measurement bias. In contrast to most of relevant works that are either search based or iterative, in this paper, we propose a two-stage closed-form estimator to localize an MS by three BSs in cellular networks. We use a distance-dependent bias model to derive a range estimator as a first step. We then use trilateration to find an estimate of the MS position. To assess the performance of our technique, we derive the mean square error (MSE) of the estimator and numerically evaluate the Cramer–Rao lower bound (CRLB) as a benchmark. We investigate the performance of the proposed method under mixed line-of-sight/NLOS scenarios in four environments, ranging from a bad urban environment to a rural environment. The provided Monte Carlo simulations show that our technique performs, on average, closely with the CRLB and provides localization capability with an average error of approximately 21 m in the worst environment among the four environments.**

***Index Terms*—Cellular networks, mobile location estimation, non-line-of-sight (NLOS) bias, time-of-arrival (TOA).**

I. Introduction

OVER the past two decades, the estimation of a mobile user location has attracted considerable research focus due to its vital role in different wireless network applications, such as cellular networks, wireless local area networks, and wireless sensor networks. Location-aware services are on a growing demand in the cellular network field, particularly with the introduction of E-911 [1], which allows the authorities to locate the caller and provide emergency services efficiently. This would require a very accurate location estimation capability.

Manuscript received March 6, 2015; revised November 11, 2015; accepted December 16, 2015. Date of publication January 12, 2016; date of current version November 10, 2016. The review of this paper was coordinated by Dr. Y. Ji.

The authors are with the Research School of Engineering (RSEng), Australian National University (ANU), Canberra, ACT 0200, Australia (e-mail: zohair.abushaban@anu.edu.au; xiangyun.zhou@anu.edu.au; thushara.abhayapala@anu.edu.au).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2016.2517151

Mobile station (MS) localization has been extensively studied under line-of-sight (LOS) conditions, over the last few decades (see, for example, [2]–[5]). However, one major issue that limits the performance of many available methods is the presence of a non-line-of-sight (NLOS) link between the MS and the base station (BS), i.e., when an obstacle or more interrupt the direct path between an MS and a BS. One of the most popular localization methods is estimating the range between the two ends by multiplying the signal speed by the time-of-arrival (TOA) [6]. In the LOS case, the measured location is contaminated by only Gaussian measurement noise, whereas in the NLOS case, a measurement *bias* is added on the measured range and its corresponding noise. Since the presence of obstacles between the MS and the BS will force the signal from one end to be reflected on these obstacles before finding its way to the other end, the measured range will always be greater than the actual distance. Therefore, the distance bias under NLOS is always *positive*.

Researchers in this field considered the problem in different approaches. These approaches can be broadly classified into five categories.

1) Identify-and-localize [7]–[11]: In this category, the link status is identified as being LOS or NLOS. Based on this identification, localization is performed by either incorporating the NLOS links or discarding them.
2) Mathematical programming: The idea is to formulate the MS localization problem as a constraint optimization problem and solve it by techniques such as linear programming [12], linear quadratic programming [13], the interior-point method [3], and sequential quadratic programming [14].
3) Least squares (LS) solution: These include LS [15] and weighted LS (WLS) [16], [17] search-based techniques.
4) Robust estimation techniques: These methods try to suppress the effect of NLOS outliers on the measured ranges. To do this, they use estimators, such as the Huber estimator [18], [19] and least median of squares [20]. On the other hand, in [21] and [22], robust methods are implemented to mitigate the NLOS effect by estimating the measurement bias probability distributions and the position, iteratively.
5) Hybrid methods: Techniques in this category mix TOA with other localization methods, such as angle-of-arrival [23] and received signal strength [24].

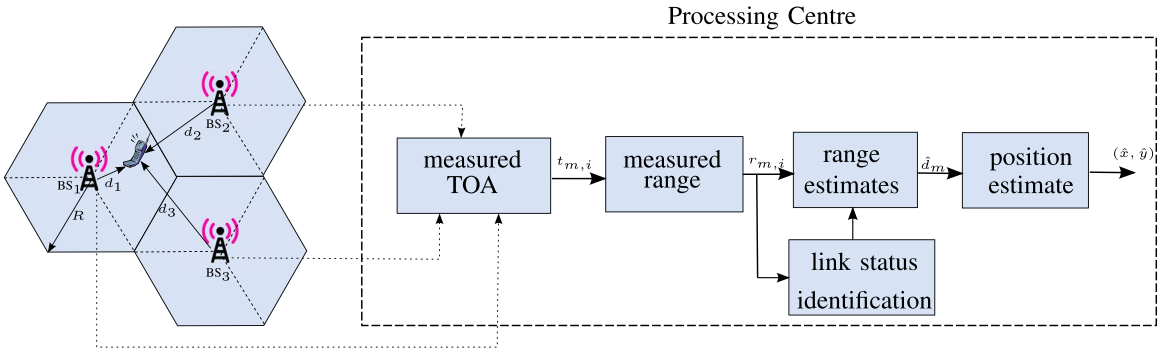


Fig. 1. Proposed localization technique block diagram. BS_1 is the hosting BS, whereas BS_2 and BS_3 are the neighboring BSs. Note that every cell is divided into three diamond-shaped areas called sectors. R is the cell radius.

Most of the relevant previous TOA-based works are either search based or iterative. On the contrary, in this paper, we contribute to the first category by proposing a closed-form, two-step localization technique for cellular networks. This technique first estimates the range between the MS and three BSs, i.e., the hosting BS and two neighboring cell BSs, by an asymptotically unbiased estimator. Subsequently, it estimates the user location using trilateration. In most of the identify-and-localize work, the measurement bias is considered either constant or a random process of a Gaussian [25], exponential [7], or uniform distribution [12] with *given* parameters. On the contrary, we consider the bias model proposed in [26] and adopted in the European standard COST 259 (TD) [27]. This model is more realistic in that it suggests that the bias follows a conditionally exponential distribution whose parameter is a function of the distance between the BS and the MS, the median root mean square (rms) delay spread, and shadowing. Although the model in [26] is used to generate simulation data in [17], [18], and [28], it was not incorporated in the respective localization algorithm. In our work herein, we use this bias model knowledge to derive an asymptotically unbiased estimator that finds approximate values of the range between the MS and all the three BSs. To achieve this, a processing center first collects TOA measurements and uses them to obtain range estimates and identify the link status, as shown in Fig. 1. Once all the three ranges are estimated, they are used to define three circles. The closest three intersection points of these circles are used to define a triangle whose centroid is taken as the user location estimate. In the case where two circles do not overlap, the center of the gap between them is taken as a triangle vertex. To assess the performance of our localization method, we investigate its performance in the four environments classified in [26] as follows: bad urban, urban, suburban, and rural.

The contributions and merits of this paper can be summarized as follows.

- In contrast to [17], [18], and [28], we use a range-dependent bias model [26] to derive a TOA-based closed-form range estimator that is asymptotically unbiased. Subsequently, we use trilateration to obtain an estimator of the MS location in closed form.
- The mean square error (MSE) of the range estimator is derived and compared with the Cramer–Rao lower bound (CRLB), which we evaluate numerically.[1] The range probability distribution function (pdf) for the three BSs is derived in this paper and is used to determine the average CRLB.
- Finally, extensive Monte Carlo simulations are carried out to assess the performance of the proposed localization technique in the four environments previously mentioned, using the performance measures defined in Section V. The results show that the proposed localization technique performs, on average, closely with the CRLB and provides localization capability with an average error of approximately 21 m in the worst environment.

The rest of this paper is organized as follows: Section II presents the problem formulation and sets the assumptions of this work. Subsequently, Section III describes in detail the range estimator and shows its unbiased behavior, whereas Section IV explains the trilateration procedure of estimating the user location. Section V discusses the CRLB and the difficulties of obtaining it analytically. It also lists the pdf of the range and defines the performance measure we use for our technique assessment. The numerical results of the Monte Carlo simulation are given in Section VI. A thorough discussion of the results is also included in that section. Finally, the conclusions are presented in Section VII, and the detailed derivation of the range pdf is given in the Appendix.

II. Problem Formulation

Here, we formulate the problem under consideration. First, we present the assumptions that our work is based on. Subsequently, the signal model is thoroughly discussed.

A. Assumptions

This work is based on the following assumptions.

- TOA measurements are readily available at the processing center and were obtained by the TOA method used by the air interface, i.e., wireless standard. The ℓth TOA measurement, i.e., t_ℓ, is used to calculate the ℓth range, i.e., $r_\ell = t_\ell c$, where $c = 0 \times 10^8$ m/s is the speed of the signal.

[1] We resorted to numerical computation due to the difficulties in evaluating it analytically. These difficulties are discussed in Section V.

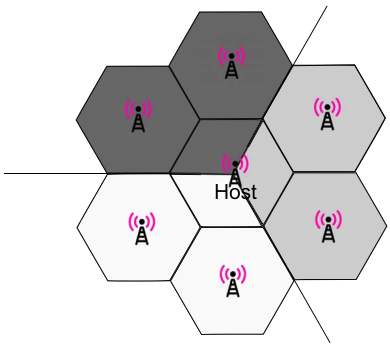


Fig. 2. BS selection is based on the sector of the host BS where the mobile is. Three scenarios are possible as illustrated: white, light gray, and dark gray.

- As shown in Fig. 2, we consider cellular networks with regular hexagonal cells. Each cell is divided into three sectors.[2] The user position is uniformly distributed in the cell, and the sector in which the user exists is assumed to be known. This can be identified using handover information available at the processing center, as used in [18]. The sector boundaries are used to select the three BSs to perform localization. In other words, considering the host cell in Fig. 2, when the MS is in a particular sector, the BSs in the cells shaded with the corresponding color are selected to perform the localization. As a result from the hexagonal pattern, BS locations are also assumed to be given. The backhaul link between BSs and the processing center is assumed to be error free.
- Similar to [17] and [18], the measurement bias due to NLOS is modeled as an exponential random process conditional to a zero-mean lognormal random process. These distributions are described by the nature of the environment, as detailed in [26]. Consecutive bias samples are considered independent and identically distributed (i.i.d.). On the other hand, the measurement noise is assumed to be an i.i.d. zero-mean Gaussian process and independent of the measurement bias.
- The parameters specifying the environment surrounding the MS are considered known and fixed during the measurement acquisition phase. The status of each MS–BS link whether LOS or NLOS is assumed to stay unchanged during the measurement acquisition phase.

B. Signal Model

Based on the given assumptions, the ith range measurement at the mth BS can be written as

$$r_{m,i} = d_m + n_{m,i} + b_{m,i}, \quad i = 1, 2, \ldots, N, \; m = 1, 2, 3 \quad (1)$$

where N is the sample size. The actual distance between the MS and the mth BS, positioned at $\mathbf{x} = (x, y)$ and $\mathbf{x}_m = (x_m, y_m)$, respectively, is given by

$$d_m = \sqrt{(x - x_m)^2 + (y - y_m)^2}. \quad (2)$$

[2]A sector is a partial area of a cell that is served by a directional antenna, usually with a beamwidth of 60° or 120° [29].

Here, d_m is an unknown variable that is considered constant over the N measurements. $n_{m,i}$ denotes the measurement noise and is modeled as a zero-mean Gaussian random variable, i.e., $n_{m,i} \sim N(0, \sigma_n^2)$. The measurement bias, which is always nonnegative, is denoted by $b_{m,i}$. In fact, $b_{m,i} = 0$ in the case of a LOS link between the MS and the BS. On the other hand, $b_{m,i}$ is a positive random variable under NLOS conditions. This is due to the fact that a signal reflected by obstacles will travel a longer distance compared with a LOS signal. According to [26], the stochastic distribution of $b_{m,i}$ is given by

$$f_{B_m}(b_{m,i}) = \begin{cases} 0, & \text{LOS} \\ \int_{0^+}^{\infty} f_{B_m|Z_m}(b_{m,i}|z_m) f_{Z_m}(z_m) dz_m, & \text{NLOS} \end{cases} \quad (3)$$

where

$$f_{Z_m}(z_m) = \frac{1}{z_m \sigma_z \sqrt{2\pi}} e^{-\frac{(\ln z_m)^2}{2\sigma_z^2}} \quad (4)$$

is the pdf of Z_m, the lognormal random variable representing the shadowing that affects the signal from the MS to the mth BS, σ_z is measured in nepers ($\sigma_z' = \sigma_z(20/\ln 10)$ dB), and

$$f_{B_m|Z_,}(b_{m,i}|z_m) = \frac{1}{\beta_m} e^{-\frac{b_{m,i}}{\beta_m}}. \quad (5)$$

The parameter β_m is given by

$$\beta_m = c\tau_{\text{rms}} = cT d_m^{\epsilon} z_m = cT\sqrt{d_m} z_m \quad (6)$$

where τ_{rms} is the rms delay spread within the MS environment, T is the median value of τ_{rms} obtained at a distance of 1 km from the BS, and $\epsilon = 1/2$.

Depending on the severity of T, we focus on four environments: bad urban, urban, suburban, and rural [26]. Furthermore, we investigate mixed LOS/NLOS scenarios, where the status of a link is detected at the processing center.

Given the previously discussed assumptions and system model, our objective is to estimate the MS position $\mathbf{x}$ in the vicinity of three BSs located at $\mathbf{x}_m \in \Re^2$ and covering three hexagonal cells of radius R.

III. Proposed Closed-Form Non-Light-of-Sight Range Estimation

Here, we propose a two-stage location estimator to address the problem described in Section II. First, we present an asymptotically unbiased range estimator to estimate $\hat{d}_m$. Second, we use these range estimates to find an estimate position of the MS, using trilateration.

A. Proposed Range Estimator

We aim at using the first-order statistics of r_m in the range estimation. The classical approach to compute the pdf of R_m would be to obtain the $f_{B_m}(b_{m,i})$ in (3) in closed form and convolve it with $f_{N_m}(n_{m,i})$. However, evaluating $f_{B_m}(b_{m,i})$ is extremely intractable when considering the integration of the product of (4) and (5). In addition, as far as the derivation of

our estimator is concerned, we need to know the mean of R_m, whereas there is no need to know the complete distribution of R_m. Thus, we resort to computing $\mathbb{E}_{R_m}[r_{m,i}]$, relying on the law of total expectation and the independence between N_m and B_m. Steps (7)–(10) are used to obtain the expectation of R_m under NLOS conditions, which then can be *estimated* by the sample mean in (12).

Taking the conditional expectation of (1) w.r.t. $B_m|Z_m$, we have

$$\begin{aligned}\mathbb{E}_{B_m|Z_m}[r_{m,i}] &= d_m + n_{m,i} + \mathbb{E}_{B_m|Z_m}[b_{m,i}|z_m] \\ &= d_m + n_{m,i} + cT\sqrt{d_m}z_m \end{aligned} \tag{7}$$

where (7) follows from the exponential pdf in (5). Taking the expectation of (7) w.r.t. Z_m and applying the law of total expectation [30], we have

$$\begin{aligned}\mathbb{E}_{Z_m}\left[\mathbb{E}_{B_m|Z_m}[r_{m,i}]\right] &= \mathbb{E}_{B_m}[r_{m,i}] \\ &= d_m + n_{m,i} + cT\sqrt{d_m}\mu_z \end{aligned} \tag{8}$$

where $\mu_z = e^{\sigma_z^2/2}$ is the mean of the lognormal distribution given in (4).

Taking the expectation of (8) w.r.t. N_m, we have

$$\mathbb{E}_{N_m}\left[\mathbb{E}_{B_m}[r_{m,i}]\right] = \mathbb{E}_{B_m,N_m}[r_{m,i}] \tag{9}$$

$$= \mathbb{E}_{R_m}[r_{m,i}] = d_m + cT\sqrt{d_m}\mu_z. \tag{10}$$

Note that (9) follows from the assumption that the measurement noise and the bias are independent, whereas (10) uses the assumption that the noise is of zero mean. It should be remarked that (10) represents the theoretical mean of the collected measurement, which requires an infinite number of samples to be computed. Thus, we use the sample mean, instead, as a suitable estimate to the mean in (10).

Assuming ergodicity, the theoretical mean in (10) can be *estimated* by the sample mean as

$$\hat{\mathbb{E}}_{R_m}[r_{m,i}] = \bar{r}_m \tag{11}$$

$$= \frac{1}{N}\sum_{i=1}^{N} r_{m,i}$$

$$= d_m + \frac{1}{N}\sum_{i=1}^{N} b_{m,i}. \tag{12}$$

The sample mean in (12) represents the estimate of the recorded measurement mean in (10) using N samples. Note that this estimator is the best mean estimator, as shown in [31].

Based on the sample mean of R_m, we now proceed to derive the closed-form range estimator. Using (10) and (11), we can write

$$\hat{d}_m + cTmu_z\sqrt{\hat{d}_m} - \bar{r}_m = 0. \tag{13}$$

Taking $u = \sqrt{\hat{d}_m} \geq 0$ and solving for u, we have

$$u = \frac{-cT\mu_z + \sqrt{(cT\mu_z)^2 + 4\bar{r}_m}}{2}. \tag{14}$$

Subsequently

$$\hat{d}_m = \begin{cases} \frac{1}{2}\left(D^2 - D\sqrt{D^2 + 4\bar{r}_m}\right) + \bar{r}_m, & \text{NLOS} \\ \bar{r}_m, & \text{LOS} \end{cases} \tag{15}$$

where $D = cT\mu_z$. Note that when $D = 0$, the NLOS estimator reduces to the LOS estimator. Moreover, note that since d_m and $b_{m,i}$ are positive, $\bar{r}_m$ is positive. As a result, $\sqrt{D^2 + 4\bar{r}_m}$ is always real.

Proposition 1: The proposed range estimator under NLOS in (15) is asymptotically unbiased.

Proof: Using the assumption of zero-mean noise and substituting (12) in (15), we have

$$\hat{d}_m - d_m = \frac{1}{2}\left(D^2 - D\sqrt{D^2 + 4d_m + \frac{4}{N}\sum_{i=1}^{N} b_{m,i}}\right) + \frac{1}{N}\sum_{i=1}^{N} b_{m,i}. \tag{16}$$

Taking the expectation[3] of (16), we obtain

$$\mathbb{E}[\hat{d}_m] - d_m = \frac{1}{2}\mathbb{E}\left[D^2 - D\sqrt{D^2 + 4d_m + \frac{4}{N}\sum_{i=1}^{N} b_{m,i}}\right] + \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}[b_{m,i}]. \tag{17}$$

If N is large enough, then

$$\frac{1}{N}\sum_{i=1}^{N} b_{m,i} \to \mathbb{E}[b_{m,i}] = \mathbb{E}_z[\beta_m] = D\sqrt{d_m} \tag{18}$$

which implies that

$$\mathbb{E}[\hat{d}_m] - d_m \to \frac{1}{2}\left(D^2 - D\sqrt{D^2 + 4d_m + 4D\sqrt{d_m}}\right) + D\sqrt{d_m}. \tag{19}$$

By completing the square, we get

$$\mathbb{E}[\hat{d}_m] - d_m \to 0 \tag{20}$$

which leads to the conclusion that the proposed range estimator is asymptotically unbiased. ■

It should be stressed that the notions of *measurement bias* with the *estimator bias* are completely distinct from each other. The measurement bias is inherent to the environment and is caused by the signal traveling a further distance in the case of NLOS. However, the estimator bias is the difference between the average of the estimated parameter and the actual value of this parameter. This is an estimator property and is not related to the environment. For an unbiased estimator, this difference is zero [31].

[3]Dropping the subscript from the expectation operator means that the expectation is taken with respect to the random variable between brackets.

B. MSE of the Range Estimator

Since the proposed range estimator in (15) was shown in Section III-A to be asymptotically unbiased, the estimator error variance, i.e., $\sigma^2_{d_m}$, and the MSE, i.e., $\rho^2_{d_m}$, are equal [31]. In other words

$$\begin{aligned}\sigma^2_{d_m} &= \mathbb{E}[(\hat{d}_m - d_m)^2] \\ &= \mathbb{E}\left[\hat{d}^2_m\right] - d^2_m = \rho^2_{d_m}. \end{aligned} \tag{21}$$

Substituting (15) in (21) and simplifying the result, we obtain

$$\begin{aligned}\rho^2_{d_m} = \frac{D^4}{2} + 2D^2\mathbb{E}[\bar{r}_m] + \mathbb{E}[\bar{r}^2_m] \\ - \frac{D}{2}\mathbb{E}\left[(D^2+2\bar{r})\sqrt{D^2+4\bar{r}_m}\right] - d^2_m. \end{aligned} \tag{22}$$

From (12) and (18), we have

$$2D^2\mathbb{E}[\bar{r}_m] = 2d_mD^2 + 2D^3\sqrt{d_m} \tag{23}$$

$$\mathbb{E}\left[\bar{r}^2_m\right] = \mathbb{E}\left[\left(d_m + \frac{1}{N}\sum_{i=1}^{N} n_{m,i} + \frac{1}{N}\sum_{i=1}^{N} b_{m,i}\right)^2\right] \tag{24}$$

$$\begin{aligned}&= d^2_m + \frac{1}{N^2}\mathbb{E}\left[\left(\sum_{i=1}^{N} n_{m,i}\right)^2\right] \\ &\quad + \frac{1}{N^2}\mathbb{E}\left[\left(\sum_{i=1}^{N} b_{m,i}\right)^2\right] + \frac{2d_m}{N}\sum_{i=1}^{N}\mathbb{E}[b_{m,i}] \end{aligned} \tag{25}$$

$$\begin{aligned}&= d^2_m + \frac{1}{N^2}\mathbb{E}\left[\sum_{i=1}^{N} n^2_{m,i}\right] + \frac{1}{N^2}\mathbb{E} \\ &\quad \times\left[\sum_{i=1}^{N} b^2_{m,i} + 2\sum_{i=1}^{N}\sum_{j=i+1}^{N} b_{m,i}b_{m,j}\right] + 2d_mD\sqrt{d_m} \\ &= d^2_m + \frac{\sigma^2_n}{N} + \frac{\mathbb{E}\left[b^2_{m,i}\right]}{N} + \frac{(N-1)\mathbb{E}^2[b_{m,i}]}{N} \\ &\quad + 2d_mD\sqrt{d_m} \\ &= d^2_m + 2d_mD\sqrt{d_m} + \frac{\sigma^2_n + \sigma^2_b}{N} + \mathbb{E}^2[b_{m,i}] \\ &= d^2_m + D^2d_m + 2d_mD\sqrt{d_m} + \frac{\sigma^2_n + \sigma^2_b}{N}. \end{aligned} \tag{26}$$

Note that the terms $(2d_m/N)\sum_{i=1}^{N}\mathbb{E}[n_{m,i}]$ and $(1/N^2)\sum_{i=1}^{N}\mathbb{E}[n_{m,i}]\sum_{i=1}^{N}\mathbb{E}[b_{m,i}]$ vanish in (25) since the noise is of zero mean. Here, σ^2_b is the measurement bias variance calculated by

$$\begin{aligned}\sigma^2_b &= \mathbb{E}\left[b^2_{m,i}\right] - \mathbb{E}^2[b_{m,i}] \\ &= \mathbb{E}_Z\left[\mathbb{E}_{B|Z}\left[b^2_{m,i}|z_m\right]\right] - \mathbb{E}^2_z\left[\mathbb{E}_{B|Z}[b_{m,i}|z_m]\right] \\ &= \mathbb{E}_Z\left[\sigma^2_{B_m|Z_m} + \mathbb{E}^2_{B|Z}[b_{m,i}|z_m]\right] - \mathbb{E}^2_z\left[\mathbb{E}_{B|Z}[b_{m,i}|z_m]\right] \\ &= \mathbb{E}_Z\left[c^2T^2d_mz^2_m + c^2T^2d_mz^2_m\right] - (cT\sqrt{d_m}\mu_z)^2 \\ &= D^2d_m\left(2\mu^2_z - 1\right). \end{aligned} \tag{27}$$

Using (12) and (18), we can write

$$\begin{aligned}&-\frac{D}{2}\mathbb{E}\left[(D^2+2\bar{r})\sqrt{D^2+4\bar{r}_m}\right] \\ &= -\frac{D}{2}\mathbb{E}\left[\left(D^2 + 2d_m + 2\frac{1}{N}\sum_{i=1}^{N} b_{m,i}\right)\right. \\ &\qquad \left.\times\sqrt{D^2 + 4d_m + 4\frac{1}{N}\sum_{i=1}^{N} b_{m,i}}\right] \\ &= -\frac{D}{2}\left(D^2 + 2d_m + 2D\sqrt{d_m}\right)\sqrt{D^2 + 4d_m + 4D\sqrt{d_m}} \\ &= -\frac{D}{2}\left(D^2 + 2d_m + 2D\sqrt{d_m}\right)\left(D + 2\sqrt{d_m}\right) \\ &= -\frac{D^4}{2} - 2D^3\sqrt{d_m} - 3D^2d_m - 2Dd_m\sqrt{d_m}. \end{aligned} \tag{28}$$

Substituting (23), (26), and (28) into (22) gives

$$\begin{aligned}\rho^2_{d_m} &= \frac{1}{N}\left(\sigma^2_b + \sigma^2_n\right) \\ &= \frac{1}{N}D^2\left(2\mu^2_z - 1\right)d_m + \frac{1}{N}\sigma^2_n, \quad N \gg 1 \end{aligned} \tag{29}$$

$$\rho^2_{d_m} = \begin{cases} \frac{D^2\left(2\mu^2_z-1\right)d_m + \sigma^2_n}{N}, & N \gg 1, \quad \text{NLOS} \\ \frac{\sigma^2_n}{N}, & \text{LOS}. \end{cases} \tag{30}$$

Note that in the case of LOS, i.e., $D = 0$, $\rho^2_{d_m}$ reduces to the well-known CRLB of measurements contaminated by additive white Gaussian noise, i.e., σ^2_n/N [31].

IV. Location Estimation Based on Estimated Ranges

After estimating the range between every BS and the MS, we proceed to estimate the MS location by, first, drawing circles with the estimated ranges, i.e., $\hat{d}_m$, as radii. Then, we find the number of intersecting circles, which is denoted by $C \geq 0$, and the points of intersection of these circles. Two circles i, j are overlapping if

$$\hat{d}_i + \hat{d}_j > \|\mathbf{x}_i - \mathbf{x}_j\| \tag{31}$$

where $\|\cdot\|$ denotes the l_2 norm. Subsequently, we define a triangle and estimate the MS location as the centroid of this triangle. However, since our distance estimator is asymptotically unbiased, i.e., $\mathbb{E}[\hat{d}_m] \to d_m$, for some measurements, the distance will be underestimated $\hat{d}_m < d_m$, whereas for the other measurements, it will be overestimated, i.e., $\hat{d}_m > d_m$. However, in the long run, the average estimated distance, i.e., $\mathbb{E}[\hat{d}_m]$, approaches the actual distance, i.e., d_m. These two cases are addressed in Fig. 3, depending on which the triangle is defined.

1) $C = 3$: In this case, the triangle vertices are defined by the three intersection points bounding the intersection area of the three circles' overlap, as shown in Fig. 3(a). The three triangle vertices are given by

$$\mathbf{y}_1 = \arg\min_{\mathbf{u}\in\left\{\mathbf{c}^{(1)}_{12},\mathbf{c}^{(2)}_{12}\right\}} \|\mathbf{u} - \mathbf{x}_3\| \tag{32}$$

$$\mathbf{y}_2 = \arg\min_{\mathbf{u}\in\left\{\mathbf{c}^{(1)}_{23},\mathbf{c}^{(2)}_{23}\right\}} \|\mathbf{u} - \mathbf{x}_1\| \tag{33}$$

$$\mathbf{y}_3 = \arg\min_{\mathbf{u}\in\left\{\mathbf{c}^{(1)}_{31},\mathbf{c}^{(2)}_{31}\right\}} \|\mathbf{u} - \mathbf{x}_2\| \tag{34}$$

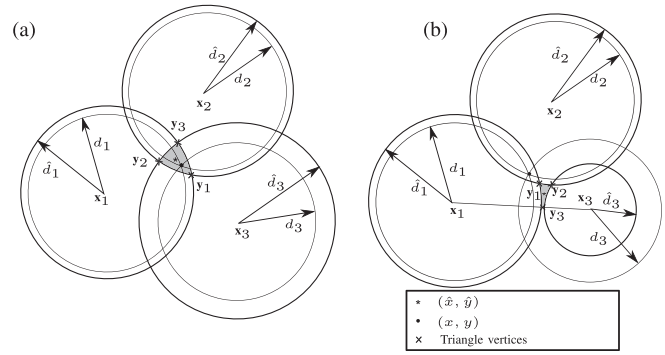


Fig. 3. Effect of range estimation error on location estimation (a) $C = 3$ and (b) $C = 2$.

where $\mathbf{c}_{ij}^{(1)}, \mathbf{c}_{ij}^{(2)} \in \Re^2$ are the two intersection points of the circles i and j.

2) $C < 3$: In this case, there are less than three circles overlapping, and hence, three intersection points are partially available. To decide on triangle vertices, the circles are investigated in pairs. Since any two intersecting circles will have two intersection points, we define a triangle vertex by the intersection point closer to the third circle (nonintersecting), similar to (32)–(34). For the nonoverlapping circles, we define the triangle vertex as the point in the gap center between the two circles. For the example shown in Fig. 3(b), two circles overlap, but a third point is needed for trilateration. Thus, we select the midgap point as a reasonable heuristic approach, i.e.,

$$\mathbf{y}_3 = \mathbf{x}_1 + (\hat{d}_1 + 0.5g)\frac{\mathbf{x}_3 - \mathbf{x}_1}{\|\mathbf{x}_3 - \mathbf{x}_1\|} \tag{35}$$

where $g = (\hat{d}_1 + \hat{d}_3) - \|\mathbf{x}_3 - \mathbf{x}_1\|$ is the gap between the two circles. To see why this approach is taken, consider Fig. 3(b). Circle 1 and circle 3 do not overlap. This means that $\hat{d}_1, \hat{d}_3$, or both of them are underestimated. However, we have no additional information to weigh among the three possibilities. For this reason, we choose the midgap point as a trilateration point.

After obtaining the triangle vertices, i.e., $\mathbf{y}_1$, $\mathbf{y}_2$, and $\mathbf{y}_3$, the MS location is estimated by

$$\hat{\mathbf{x}} = (\hat{x}, \hat{y}) = \frac{\mathbf{y}_1 + \mathbf{y}_2 + \mathbf{y}_3}{3}. \tag{36}$$

V. Cramer–Rao Lower Bound

One of the most important performance benchmark for an estimator is the CRLB. In our problem, the CRLB of d_m in (1) is defined by [31]

$$\sigma^2_{\mathrm{CRLB}}(d_m) = \frac{-1}{\mathbb{E}\left[\frac{\partial^2 \ln f(\mathbf{r}_m; d_m)}{\partial^2 d_m}\right]} \tag{37}$$

where $\mathbf{r}_m = [r_{m,1}, r_{m,2}, \ldots, r_{m,N}]^T$. Note that in the case of LOS, this bound is given by σ_n^2/N [31]. Although there has been some works on CRLB for NLOS localization, e.g., [15] and [22], none of these works addressed the problem formulated in Section II. Therefore, here, we focus on NLOS CRLB for that problem.

To start, we need the joint probability distribution $f(\mathbf{r}_m; d_m)$. However, since $r_{m,1}, r_{m,2}, \ldots, r_{m,N}$ are i.i.d., we can write

$$f(\mathbf{r}_m; d_m) = \prod_{i=1}^{N} f(r_{m,i}; d_m) = f^N(r_{m,1}; d_m) \tag{38}$$

where the first equality follows from the independence assumption, and the second equality follows from the identical distribution assumption. In other words, since $r_{m,i}$ are i.i.d., $\prod_{i=1}^{N} f(r_{m,i}; d_m)$ can be written in terms of the pdf of a single sample, raised to the power of N. We choose this sample to the first sample, i.e., $r_{m,1}$.

Note that we have the following observations.

- $r_{m,1} = (d_m + n_{m,1}) + b_{m,1}$; hence, to evaluate $f(r_{m,1}; d_m)$, we would require the joint probability $f_{B,N+D_m}(b, n + d_m)$.
- From (1), $b_{m,1}$ and $d_m + n_{m,1}$ are not independent since $b_{m,1}$ is a function of d_m. Therefore, their joint distribution cannot be simply obtained by the convolution of the marginal distribution, $f_B(b)$ and d_m-shifted $f_N(n)$.
- Even if they were independent, the integral in (3) is very hard to evaluate from (4) and (5).

For these reasons, we proceed to evaluate the CRLB numerically as listed in Algorithm 1.

Algorithm 1 Numerical Evaluation of CRLB in (37), using MATLAB

Input $N, \sigma_n, \sigma_z, R, T, c, N_b, I_r, I_{\mathrm{crlb}}$.
initialize $\sigma^2_{\mathrm{CRLB}}(d_m) = 0,\ \forall d_m$
for $k = 1 : I_{\mathrm{crlb}}$ **do**
  for $d_m = 1 : 2R$ **do**
    initialize $\hat{f}(r; d_m) = 0$
    Generate z_m
    for $i = 1 : I_r$ **do**
      Generate N_b instances of $b_{m,i}$, and $n_{m,i}$,
      Calculate the corresponding N_b instances of $r_{m,i} = d_m + b_{m,i} + n_{m,i}$,
      Obtain the Kernel Density Estimate, $\hat{f}(r_i; d_m)$, using a normal kernel function ([32]).
      $\hat{f}(r; d_m) = \hat{f}(r; d_m) + \hat{f}(r_i; d_m)$
    end for
    $\hat{f}(r; d_m) = \hat{f}(r; d_m)/I_r$
  end for
  for $j = 1 : N_b$ **do**
    $g(j, d_m) = (\partial^2/\partial d_m^2) \ln \hat{f}(j; d_m)$, where the second derivative is obtained by the MATLAB function `diff(.,2)`
  end for
  $\sigma^2_{\mathrm{CRLB}}(d_m) = \sigma^2_{\mathrm{CRLB}}(d_m) + (-1/N \sum_j g(j, d_m) \hat{f}(j; d_m))$
end for
$\sigma^2_{\mathrm{CRLB}}(d_m) = \sigma^2_{\mathrm{CRLB}}(d_m)/I_{\mathrm{crlb}}$.

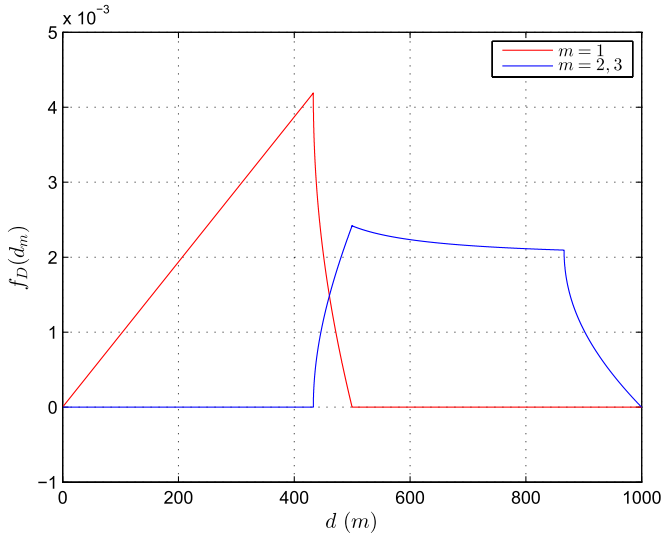


Fig. 4. PDF of the MS range from the mth BS, $m = 1$, 2 and 3. $R = 500$ m.

The CRLB procedure in Algorithm 1 gives a performance measure for individual range estimators, i.e., cells. However, to evaluate the overall localization performance, we need to define some performance measures that take into account the average performance over the three cells. Before that, we need to evaluate the pdf of the range, $f_{D_m}(d_m)$, for $m = 1$, 2, 3. As derived in the Appendix, for the hosting cell $(m = 1)$, this distribution is given by

$$f_D(d_m) = \begin{cases} \frac{4\pi d_m}{3\sqrt{3}R^2}, & 0 \le d_m < \frac{\sqrt{3}R}{2} \\ \frac{8d_m}{\sqrt{3}R^2}\left[\frac{\pi}{6} - \cos^{-1}\left(\frac{\sqrt{3}R}{2d_m}\right)\right], & \frac{\sqrt{3}}{2}R \le d_m < R. \end{cases} \tag{39}$$

For the two neighboring cells $(m = 2, 3)$, this distribution is given by

$$f_D(d_m) = \begin{cases} \frac{4d_m}{\sqrt{3}R^2}\cos^{-1}\left(\frac{\sqrt{3}R}{2d_m}\right), & \frac{\sqrt{3}R}{2} \le d_m < R \\ \frac{2d_m}{\sqrt{3}R^2}\sin^{-1}\left(\frac{\sqrt{3}R}{2d_m}\right), & R \le d_m < \sqrt{3}R \\ \frac{2d_m}{\sqrt{3}R^2}\left[\sin^{-1}\left(\frac{\sqrt{3}R}{d_m}\right) - \frac{\pi}{3}\right], & \sqrt{3}R \le d_m < 2R. \end{cases} \tag{40}$$

The full derivation of (39) and (40) is provided in the Appendix, whereas Fig. 4 shows their plots.

To measure the performance averaged over the three cells, we define the following measures.

- Average CRLB over three cells:

$$\tilde{\sigma}^2_{\text{CRLB}} = \tilde{\sigma}^2_{\text{CRLB|LOS}}\,\Pr(\text{LOS}) + \tilde{\sigma}^2_{\text{CRLB|NLOS}}\,\Pr(\text{NLOS}) \tag{41}$$

  where $\Pr(\text{LOS})$ and $\Pr(\text{NLOS})$ are the probabilities of the link being LOS or NLOS, respectively, and are assumed to be given *a priori*. The exact values of these two probabilities depend on several factors, such as the environment, the location, and being indoor or outdoor. $\tilde{\sigma}^2_{\text{CRLB|LOS}} = \sigma_n^2/N$, whereas $\tilde{\sigma}^2_{\text{CRLB|NLOS}} = (1/3)\sum_{m=1}^{3}[\int_{d_m}\sigma^2_{\text{CRLB}}(d_m)f_D(d_m)dd_m]$.

TABLE I
MEDIAN RMS DELAY SPREAD FOR THE CONSIDERED ENVIRONMENTS [26]

Environment	T (μs)
Bad Urban	1.0
Urban	0.4
Suburban	0.3
Rural	0.1

- Average range estimation MSE:

  When the link status is perfectly known at the processing center, the average range estimation MSE is denoted by

$$\tilde{\rho}^2_{de|k} = \left(\frac{\rho^2_{d_1} + \rho^2_{d_2} + \rho^2_{d_3}}{3}\middle|\ \text{link status known}\right). \tag{42}$$

  On the other hand, when the link status is estimated by a decision rule, the average range estimation MSE is denoted by

$$\tilde{\rho}^2_{de|i} = \left(\frac{\rho^2_{d_1} + \rho^2_{d_2} + \rho^2_{d_3}}{3}\middle|\ \text{link status identified}\right). \tag{43}$$

- Location estimation MSE:

  Depending on the link status knowledge, this MSE is denoted by

$$\tilde{\rho}^2_{xe|k} = \mathbb{E}\left[\|\mathbf{x} - \hat{\mathbf{x}}\|^2\middle|\ \text{link status known}\right]$$
$$\tilde{\rho}^2_{xe|i} = \mathbb{E}\left[\|\mathbf{x} - \hat{\mathbf{x}}\|^2\middle|\ \text{link status identified}\right].$$

VI. Numerical Results and Discussion

To investigate the performance of the proposed algorithm, we perform extensive Monte Carlo simulations. We also present the results for the numerical computation of the CRLB in this section, but before that, we start by describing the simulation setup.

A. Simulation Setup

The measured range samples were generated according to the model discussed in Section II. The model parameters we used here were chosen to match the recommended values in [26]. In this regard, the measurement noise is generated as a zero-mean Gaussian process with $\sigma_n = 60$ m. Moreover, the lognormal shadowing, i.e., Z_m, conditioning the measurement bias parameters is specified by $\sigma'_z = 4$ dB, whereas the values of T for the four environments under interest are given in Table I.

In our simulations, we consider seven regular hexagonal cells served by BSs located at (0, 0), $(1.5R, \sqrt{3}/2R)$, $(1.5R, -(\sqrt{3}/2)R)$, $(0, -\sqrt{3}R)$, $(-1.5R, -(\sqrt{3}/2)R)$, $(-1.5R, (\sqrt{3}/2)R)$, and $(0, \sqrt{3}R)$, where $R = 500$ m, as shown in Fig. 2. The user, which is located in the first cell, is localized by three BSs that are defined by the sector boundaries, as described in Section II-A. The model in [33] was used to generate user locations that are uniformly distributed over a regular hexagon.

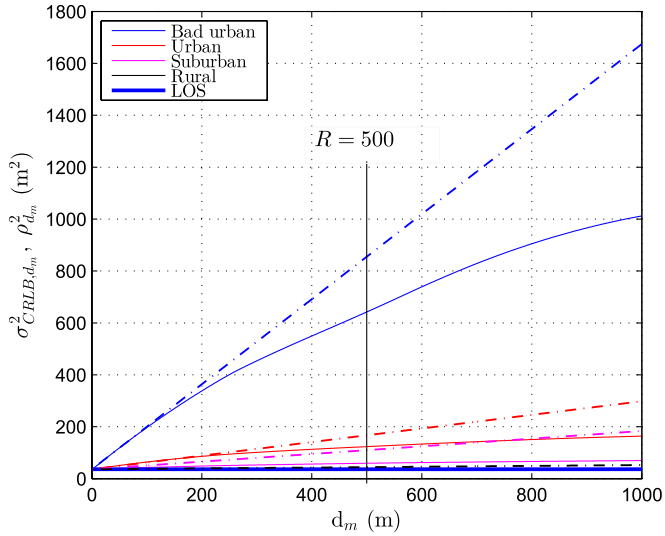


Fig. 5. CRLB of the range estimator performance in terms of (d_m) compared with the range estimation error variance $\rho^2_{d_m}$, for $N = 100$, $R = 500$ m. Solid lines represent $\sigma^2_{\text{CRLB}}(d_m)$, whereas dash-dotted lines represent $\rho^2_{d_m}$. $0 \leq d_1 < R$, and $(\sqrt{3}/2)R \leq d_2, d_3 < 2R$.

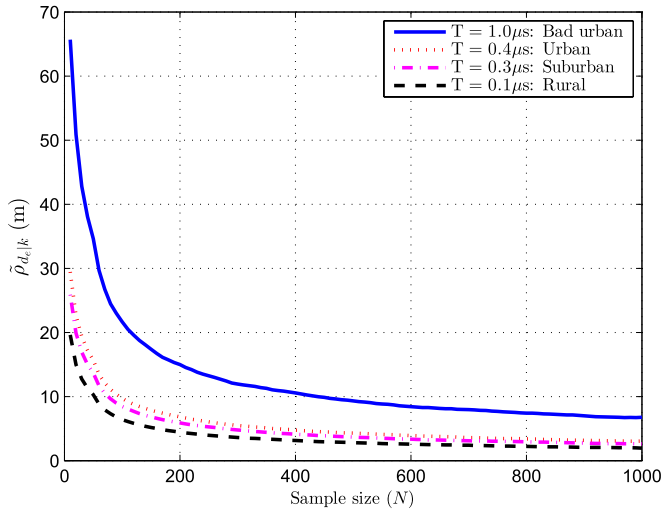


Fig. 6. Average range RMSE as a function of N when link statuses are known, obtained by averaging over 1000 user locations.

For the link status identification, we use the single-BS decision-theoretic method proposed in [25]. By recalling that $\sigma^2_{\bar{r}} = \sigma^2_n + \sigma^2_b = \sigma^2_n + D^2 d_m(2\mu^2_z - 1)$, a decision rule that can be used to identify the link status is

$$\sigma^2_{\bar{r}} \underset{\text{LOS}}{\overset{\text{NLOS}}{\gtrless}} \eta\sigma^2_n \tag{44}$$

where $\sigma^2_{\bar{r}}$ is the recorded sample variance, and $\eta > 1$ is a decision threshold that depends on the environment and the cell being either an MS host or a neighboring cell. In our simulations, a link status is modeled as an equally probable Bernoulli random variable.

For location estimation performance comparison, we use a WLS technique similar to [16]

$$\hat{\mathbf{x}}_{LS,i} = \arg\min_{\hat{\mathbf{x}}} \sum_{m=1}^{3} \frac{(r_{m,i} - \|\hat{\mathbf{x}} - \mathbf{x}_m\|)^2}{\alpha_m} \tag{45}$$

where

$$\alpha_m = \begin{cases} \sigma^2_n, & \text{LOS} \\ \sigma^2_n + D^2\left(2\mu^2_z - 1\right)\|\hat{\mathbf{x}} - \mathbf{x}_m\|, & \text{NLOS}. \end{cases} \tag{46}$$

Subsequently, the location estimates are taken as: $\hat{\mathbf{x}} = \text{Mean}(\hat{\mathbf{x}}_{LS,i}), i = 1, 2, \ldots, N$, where N is the sample size.

B. Range Estimation

Considering the range estimation stage only, the CRLB of the NLOS range estimator is given in Fig. 5. Since the CRLB increases with d_m, we can see that the farther the MS is from the BS the harder it is to get a lower error estimate. Moreover, we can infer that the CRLB is a function of the environment, in that it becomes worse for environments with higher delay spread. Recall that $0 \leq d_1 \leq R$, whereas $\sqrt{3}R/2 \leq d_2, d_3 \leq 2R$, which means that d_1 is generally better estimated than d_2 and d_3. Finally, note that the four CRLB curves approach $\sigma^2_n/N = 36$ m^2 when $d_m \to 0$. This is because the bias term in (1) vanishes and only noise is present, which leads the bound to become similar to that of the LOS case. Moreover, Fig. 5 shows that $\rho^2_{d_m}$ of the proposed distance estimator performs closely to the CRLB for lower d_m values but diverges as the distance between the MS and the BS increases.

We now present the results for the range estimation under mixed LOS/NLOS conditions with the status of the links assumed to be known. Fig. 6 shows how the selection of the data size N affects the range root mean square error (RMSE), which is measured by $\tilde{\rho}_{de|k}$, for the four user environments. The range RMSE decays as N increases, which is sensible because the range estimator requires averaging the recorded range sample, and this average is better approximated with larger history. We can also infer from this figure that the range estimation accuracy is generally inversely proportional to the value of T that specifies the environment. Note that, from (6), larger T implies higher delay spread, measurement bias mean, and measurement bias variance.

Fig. 7 shows the identification error effect on the range estimation. The identification error is the error made in the link status identification block. An identification error occurs when the actual link is LOS, whereas it was identified as NLOS, or *vice versa*, i.e., when the link is actually NLOS but was identified as LOS. The effect of this error is quantified by the difference between $\tilde{\rho}_{d_e|k}$ when the link statuses are perfectly known and $\tilde{\rho}_{d_e|i}$ when they are identified by using the decision rule in (44). In this figure, we can see that the identification error has a minor effect on the three environments with higher T, when $N \geq 100$. However, the identification error is substantial in rural areas. For example, at $N = 200$, the identification error is around 4.8 m. To see why, recall that the identification approach we are using from [25] and defined in (44) relies on the gap between the sample variance, i.e., $\sigma^2_{\bar{r}} = D^2 d_m(2\mu^2_z - 1) + \sigma^2_n$, and the noise variance, i.e., σ^2_n. Consequently, the larger this gap is, the more accurate the decision. Hence, we

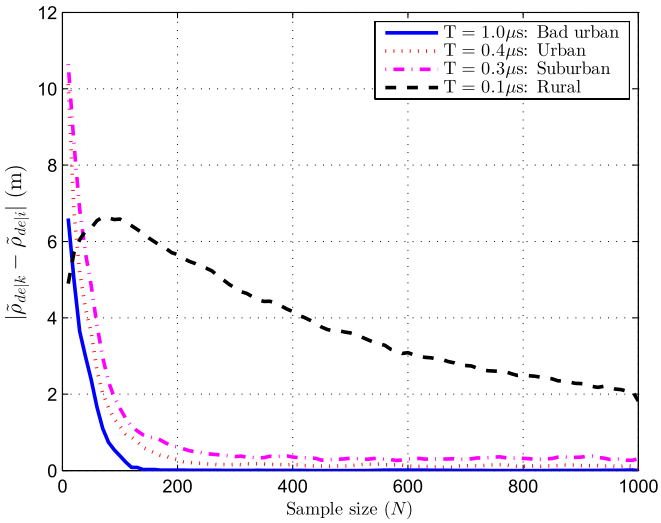


Fig. 7. Identification error effect on range estimation as a function of N obtained by averaging over 1000 user locations.

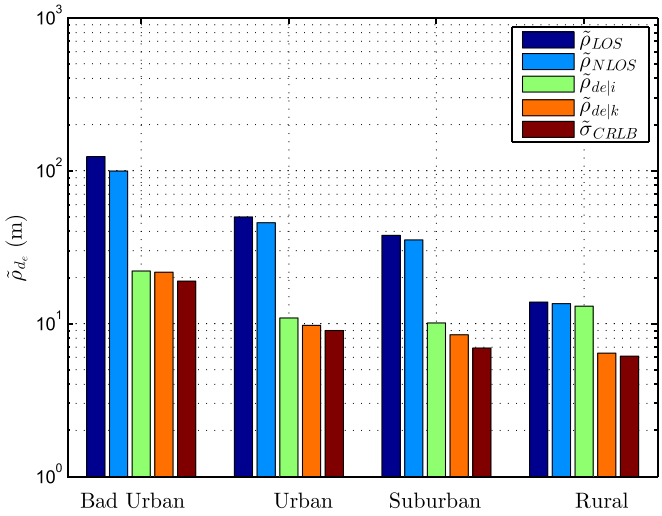


Fig. 8. Average RMSE of range error in different setups compared with average CRLB, obtained by averaging over 1000 user locations with $N = 100$.

TABLE II
SAMPLE VARIANCE COMPARED TO NOISE VARIANCE FOR DIFFERENT ENVIRONMENTS

Environment	$\sigma_{\tilde{r}}^2 = D^2 d_m(2\mu_z^2 - 1) + \sigma_n^2$	σ_n^2
Bad Urban	48577	3600
Urban	10796	3600
Suburban	7648	3600
Rural	4050	3600

TABLE III
PERCENTAGE OF THE NUMBER OF OVERLAPS C WITH $N = 100$

Environment	$C = 3$	$C = 2$	$C = 1$
Bad Urban	75.00	24.40	0.60
Urban	80.17	19.53	0.30
Suburban	81.29	18.46	0.24
Rural	80.30	19.36	0.34

compute $\sigma_{\tilde{r}}^2$ for the four environments and compare it with σ_n^2, as shown in Table II. It is evident that $\sigma_{\tilde{r}}^2$ in the rural case is relatively close to the value of σ_n^2, in contrast to the other three environments, where $\sigma_{\tilde{r}}^2$ is more pronounced and comparable to the noise variance. This means that the decision rule from [25] can easily identify link status for the bad urban, urban, and suburban environments, while making larger identification errors in the case of rural environment. With this said, it should be stressed that we did not address the identification problem in this paper, but used this method from [25] as is.

Although higher N can reduce the distance error (see Figs. 6 and 7), this would require more calculations and storage capability at the processing center, i.e., higher complexity. For this reason, we select $N = 100$ as a suitable trade-off for the subsequent results in this paper.

For a deeper look, we now compare the range estimation using different setups in Fig. 8. In the legend, $\tilde{\rho}_{\mathrm{LOS}}$ indicates that the results were obtained by the averaging estimator that assumes that all the links are LOS [31], i.e., ignoring the existence of NLOS paths. On the contrary, $\tilde{\rho}_{\mathrm{NLOS}}$ is obtained by the estimator in (15) that assumes that all the paths are NLOS, i.e., assumes measurement bias exists in all paths. $\tilde{\rho}_{de|i}$, $\tilde{\rho}_{de|k}$, and $\tilde{\sigma}_{\mathrm{CRLB}}$ are the square roots of the quantities defined in (43), (42), and (41), respectively.

It can be concluded from Fig. 8 that when link identification is implemented, the proposed estimator performs close to $\tilde{\sigma}_{\mathrm{CRLB}}$ with differences of 3.1, 1.9, 3.2, and 6.5 m in the bad urban, urban, suburban, and rural environments, respectively. Smaller gaps occur when the link status is perfectly known at the processing center, with the performance of the proposed technique approaching that of the CRLB for the suburban and rural areas. Moreover, notice that the error upper bound is well above the proposed estimator error. On an absolute measure, the proposed estimator provides range estimates with $\tilde{\rho}_{de|i}$ of 21.4, 10.9, 10.1, and 13.0 m in the four environments, respectively.

C. Location Estimation

After discussing the range estimation accuracy in Section VI-B as a first step, we now discuss the results of the second step: location estimation using the trilateration detailed in Section IV.

First, we present the frequency of the number of overlaps used for trilateration[4] in Table III. These values were obtained during the simulations campaign by counting the number of overlaps in each simulation iteration and dividing the total number of occurrences by the number of iterations, i.e., 10^5. As can be noticed, in the great majority of experiments, three overlaps are used to obtain a triangle centroid. On the other hand, two overlaps occur with lower probability, whereas one overlap occurs with a negligible percentage. These results were obtained for 10^5 experiments. Note that this table is obtained when the link statuses are unknown but identified.

[4]Measured as a percentage of the total number of experiments.

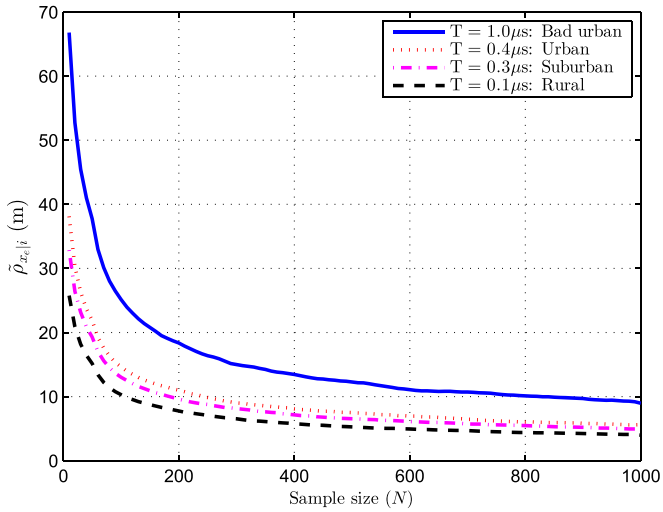


Fig. 9. Average RMSE of location estimation error as a function of N, obtained by averaging over 1000 user locations, with link status known.

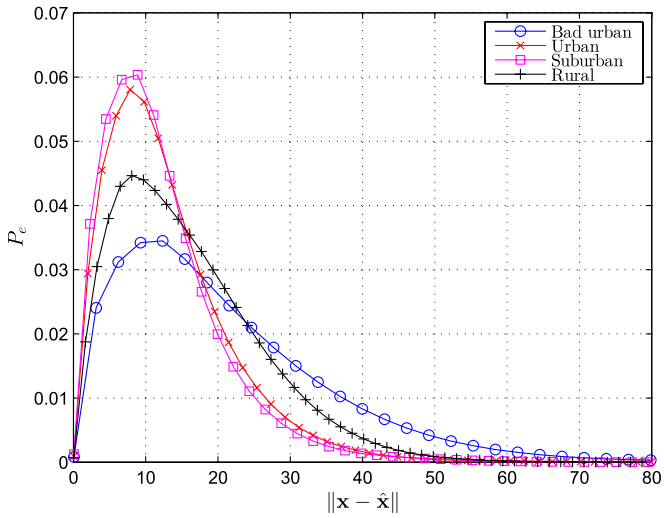


Fig. 11. PDF of the location error with $N = 100$, obtained by averaging over 1000 user locations, with link status identified.

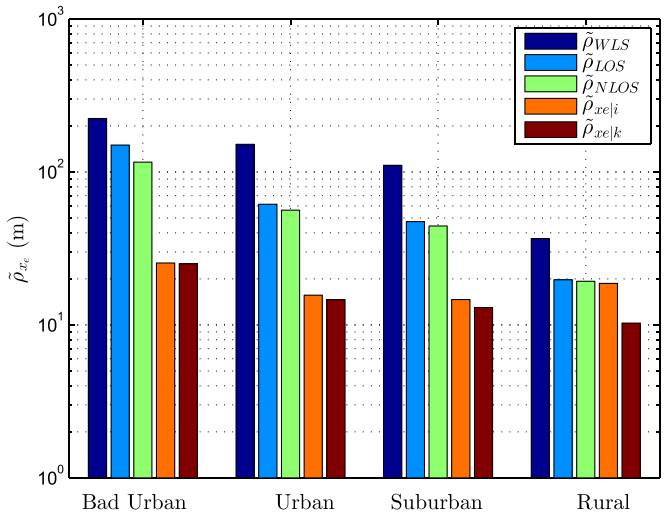


Fig. 10. Average RMSE of location error in different setups obtained by averaging over 1000 user locations with N = 100.

TABLE IV
LOCALIZATION ERROR CENTRAL TENDENCY MEASURES $N = 100$

Measure	Bad Urban	Urban	Suburban	Rural
Mean (m)	20.79	13.07	12.28	15.93
Median (m)	17.42	11.15	10.44	14.01
Mode (m)	10.97	8.14	7.90	8.16

Fig. 9 shows the location estimation accuracy in terms of $\tilde{\rho}_{x_e|k}$ for the four environments. Observations similar to Fig. 6 can be seen in Fig. 9. Particularly, notice that any increase in recorded data size beyond $N = 200$ offers a minor enhancement relative to the additional required complexity. For example, we would get 50% RMSE enhancement when the data size is five times larger.

We now benchmark the location estimation error, when the status is either known or identified at the processing center, with the cases when all the links are assumed to be either LOS or NLOS. This comparison is shown in Fig. 10, for $N = 100$. Remarkably, the proposed method—isolated from identification error—performs well below the NLOS and LOS estimation cases with RMSE values of 25.1, 14.6, 13.0, and 10.3 m, in the four environments, respectively. Furthermore, when the case of $\tilde{\rho}_{x|i}$ is considered, we get RMSE values of 25.4, 15.7, 14.6, and 18.7 m, respectively. Since the error of a stage propagates to the subsequent stage, it should be highlighted that the $\tilde{\rho}_{x|i}$ effect in Fig. 10 comprises three error components, namely, identification error, range estimation error, and location estimation error.

Finally, note that the proposed approach outperforms the WLS approach, which is denoted by $\tilde{\rho}_{\mathrm{WLS}}$ and outlined in (45) and (46), by a large margin. This is because our approach takes into account the bias distribution function, in contrast to that in [16]. Moreover, note that due to the bias being non-Gaussian, the numerator of (45) is non-Gaussian. Moreover, WLS requires a higher number of BSs for an accurate estimate. These two reasons cause WLS to have a greater estimation error than the other considered approaches in Fig. 10.

To gain better insight into the nature of the location error, Fig. 11 shows the approximate error distributions for the four environments obtained during the simulations with identification. Again, note that this is the resultant of the identification error, ranging error, and localization error. It can be seen that most of the error is concentrated at very low values. Table IV describes this quantitatively. Since the bad-urban environment has the highest median rms delay spread, T, among all the investigated environments, it exhibits the worst performance in the considered scenarios.

VII. CONCLUSION

In this paper, we have presented a two-stage closed-form NLOS mobile localization technique that is based on TOA estimation at the closest three BSs. We have used a

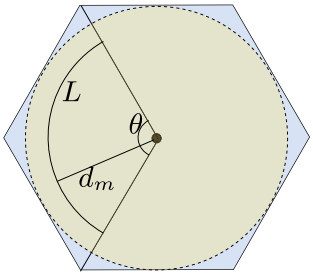


Fig. 12. $m = 1$: $0 \le d_m < (\sqrt{3}/2)R$.

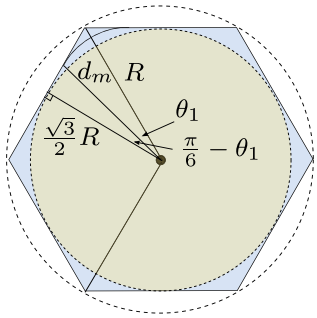


Fig. 13. $m = 1$: $(\sqrt{3}/2)R \le d_m < R$.

distance-dependent bias model to derive an asymptotically unbiased estimator to estimate the distance between each of the BS and the MS, at the first stage. Subsequently, the intersection points of the circles, which are defined by radii equal to the estimated distances, are used to define a triangle whose centroid is taken as the user location estimate. In addition, we have derived the MSE of the proposed range estimator. Moreover, we have obtained the CRLB numerically to benchmark the performance of the distance estimators. Furthermore, we have also derived the pdf of the distance between the three BSs and the MS. The simulations presented herein have demonstrated that with a cell radius of 500 m, our localization method is accurate with an average position error of 20.79 m, in the worst environment, i.e., the bad-urban environment. A key feature in our work is that it solves NLOS MS localization in closed form, in contrast to most works that are either search based or iterative. It is worth mentioning that in a rural environment, a more efficient identification rule is needed. The focus of this paper has been on localization using three BSs. We did not address the identification issue, neither did we address the case where more BSs are involved in the localization. We will consider these two issues in our future work.

APPENDIX
DERIVATION OF THE PROBABILITY DISTRIBUTION FUNCTION OF USER RANGE

Assuming that the MS occupies an infinitesimal area that can be approximated by a point, the pdf in this case is given by the length of the arc, i.e., L, of radius d_m and angle θ, divided by the area of the sector, i.e., A, of radius R, where $L = d_m\theta$, and $A = (\sqrt{3}/2)R^2$.

A. *Hosting Cell* ($m = 1$)

1) $0 \le d_m < (\sqrt{3}/2)R$*:* In Fig. 12, $\theta = 2\pi/3$, and

$$f_D(d_m) = \frac{L}{A} = \frac{4\pi d_m}{3\sqrt{3}R^2}, \quad 0 \le d_m < \frac{\sqrt{3}}{2}R.$$

2) $(\sqrt{3}/2)R \le d_m < R$*:* In Fig. 13

$$\theta_1 = \frac{\pi}{6} - \cos^{-1}\left(\frac{\sqrt{3}R}{2d_m}\right).$$

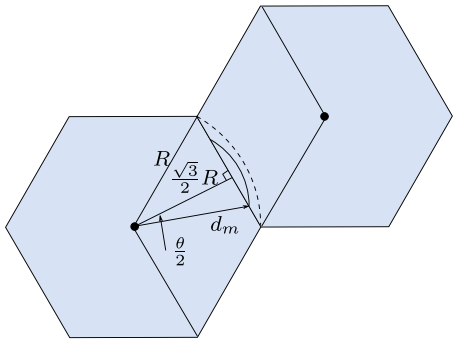


Fig. 14. $m = 2, 3$: $(\sqrt{3}/2)R \le d_m < R$.

From the symmetry

$$f_D(d_m) = \frac{4\theta_1 d_m}{A} = \frac{8d_m}{\sqrt{3}R^2}\left[\frac{\pi}{6} - \cos^{-1}\left(\frac{\sqrt{3}R}{2d_m}\right)\right], \quad \frac{\sqrt{3}}{2}R \le d_m < R.$$

B. *Neighboring Cells* ($m = 2, 3$)

1) $(\sqrt{3}/2)R \le d_m < R$*:* In Fig. 14, the arc angle can be computed from

$$\frac{\theta}{2} = \cos^{-1}\left(\frac{\sqrt{3}R}{2d_m}\right).$$

Consequently

$$f_D(d_m) = \frac{4d_m}{\sqrt{3}R^2}\cos^{-1}\left(\frac{\sqrt{3}R}{2d_m}\right), \quad \frac{\sqrt{3}}{2}R \le d_m < R.$$

2) $R \le d_m < \sqrt{3}R$*:* In Fig. 15, using the law of sine

$$\frac{\sin\theta}{R} = \frac{\sin\left(\frac{2\pi}{3}\right)}{d_m}$$

$$\theta = \sin^{-1}\left(\frac{\sqrt{3}R}{2d_m}\right)$$

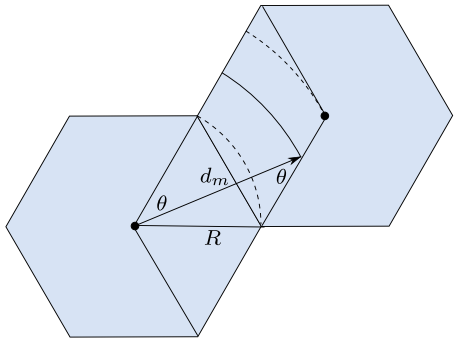


Fig. 15. $m = 2, 3$: $R \leq d_m < \sqrt{3}R$.

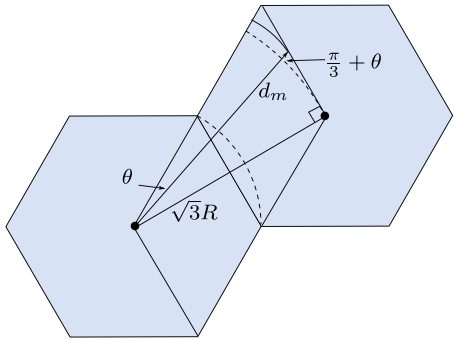


Fig. 16. $m = 2, 3$: $\sqrt{3}R \leq d_m < 2R$.

leading to

$$f_D(d_m) = \frac{2d_m}{\sqrt{3}R^2} \sin^{-1}\left(\frac{\sqrt{3}R}{2d_m}\right), \quad R \leq d_m < \sqrt{3}R.$$

3) $\sqrt{3}R \leq d_m < 2R$*:* In Fig. 16

$$\sin\left(\theta + \frac{\pi}{3}\right) = \frac{\sqrt{3}R}{d_m}$$

$$\theta = \sin^{-1}\left(\frac{\sqrt{3}R}{d_m}\right) - \frac{\pi}{3}.$$

This implies that

$$f_D(d_m) = \frac{2d_m}{\sqrt{3}R^2}\left[\sin^{-1}\left(\frac{\sqrt{3}R}{d_m}\right) - \frac{\pi}{3}\right], \ \sqrt{3}R \leq d_m < 2R.$$

To sum up, for $m = 1$, this distribution is given by

$$f_D(d_m) = \begin{cases} \frac{4\pi d_m}{3\sqrt{3}R^2}, & 0 \leq d_m < \frac{\sqrt{3}R}{2} \\ \frac{8d_m}{\sqrt{3}R^2}\left[\frac{\pi}{6} - \cos^{-1}\left(\frac{\sqrt{3}R}{2d_m}\right)\right], & \frac{\sqrt{3}}{2}R \leq d_m < R. \end{cases} \tag{47}$$

For $m = 2$ and 3, this distribution is given by

$$f_D(d_m) = \begin{cases} \frac{4d_m}{\sqrt{3}R^2} \cos^{-1}\left(\frac{\sqrt{3}R}{2d_m}\right), & \frac{\sqrt{3}R}{2} \leq d_m < R \\ \frac{2d_m}{\sqrt{3}R^2} \sin^{-1}\left(\frac{\sqrt{3}R}{2d_m}\right), & R \leq d_m < \sqrt{3}R \\ \frac{2d_m}{\sqrt{3}R^2}\left[\sin^{-1}\left(\frac{\sqrt{3}R}{d_m}\right) - \frac{\pi}{3}\right], & \sqrt{3}R \leq d_m < 2R. \end{cases} \tag{48}$$

REFERENCES

[1] FCC:9-1-1 Service, Fed. Commun. Commission, Washington, DC, USA, Jan. 2015. [Online]. Available: http://transition.fcc.gov/pshs/services/911-services/

[2] J. Caffery and G. Stüber, "Overview of radiolocation in CDMA cellular systems," *IEEE Commun. Mag.*, vol. 36, no. 4, pp. 38–45, Apr. 1998.

[3] W. Kim, J. G. Lee, and G. I. Jee, "The interior-point method for an optimal treatment of bias in trilateration location," *IEEE Trans. Veh. Technol.*, vol. 55, no. 4, pp. 1291–1301, Jul. 2006.

[4] Y. T. Chan, H. Y. C. Hang, and P. C. Ching, "Exact and approximate maximum likelihood localization algorithms," *IEEE Trans. Veh. Technol.*, vol. 55, no. 1, pp. 10–16, Jan. 2006.

[5] C. Chang and A. Sahai, "Estimation bounds for localization," in *Proc. 1st Annu. IEEE Commun. Soc. Conf. Sensor Ad Hoc Commun. Netw.*, Oct. 2004, pp. 415–424.

[6] 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, 3rd Quart. 2009.

[7] K. Yu and Y. Guo, "Statistical NLOS identification based on AOA, TOA, and signal strength," *IEEE Trans. Veh. Technol.*, vol. 58, no. 1, pp. 274–286, Jan. 2009.

[8] S. Mazuelas *et al.*, "Prior NLOS measurement correction for positioning in cellular wireless networks," *IEEE Trans. Veh. Technol.*, vol. 58, no. 5, pp. 2585–2591, Jun. 2009.

[9] C. Ma, R. Klukas, and G. Lachapelle, "A nonline-of-sight error-mitigation method for TOA measurements," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 641–651, Mar. 2007.

[10] Z. Wang and S. A. Zekavat, "Omnidirectional mobile NLOS identification and localization via multiple cooperative nodes," *IEEE Trans. Mobile Comput.*, vol. 11, no. 12, pp. 2047–2059, Dec. 2012.

[11] M. Wylie-Green and S. Wang, "Robust range estimation in the presence of the non-line-of-sight error," in *Proc. IEEE 54th Veh. Technol. Conf. Fall*, 2001, vol. 1, pp. 101–105.

[12] S. Venkatesh and R. Buehrer, "A linear programming approach to NLOS error mitigation in sensor networks," in *Proc. 5th Int. Conf. Inf. Process. Sensor Netw.*, 2006, pp. 301–308.

[13] K. Yu and Y. Guo, "Improved positioning algorithms for nonline-of-sight environments," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2342–2353, Jul. 2008.

[14] S. Venkatraman, J. Caffery, and H. R. You, "Location using LOS range estimation in NLOS environments," in *Proc. IEEE 55th Veh. Technol. Conf. Spring*, 2002, vol. 2, pp. 856–860.

[15] H. Miao, K. Yu, and M. Juntti, "Positioning for NLOS propagation: Algorithm derivations and Cramer–Rao bounds," *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 2568–2580, Sep. 2007.

[16] S. Gezici and Z. Sahinoglu, "UWB Geolocation Techniques for IEEE 802.15.4a personal area networks," Mitsubishi Electric Res. Labs., Cambridge, MA, USA, MERL Tech. Rep. 2, Aug. 2004.

[17] P. C. Chen, "A non-line-of-sight error mitigation algorithm in location estimation," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Sep. 1999, pp. 316–320.

[18] G. L. Sun and W. Guo, "Bootstrapping M-estimators for reducing errors due to Non-Line-of-Sight (NLOS) propagation," *IEEE Commun. Lett.*, vol. 8, no. 8, pp. 509–510, Aug. 2004.

[19] S. Yousefi, X. W. Chang, and B. Champagne, "Distributed cooperative localization in wireless sensor networks without NLOS identification," in *Proc. 11th WPNC*, Mar. 2014, pp. 1–6.

[20] T. Qiao and H. Liu, "Improved least median of squares localization for non-line-of-sight mitigation," *IEEE Commun. Lett.*, vol. 18, no. 8, pp. 1451–1454, Aug. 2014.

[21] F. Yin, C. Fritsche, F. Gustafsson, and A. Zoubir, "EM- and JMAP-ML based joint estimation algorithms for robust wireless geolocation in mixed LOS/NLOS environments," *IEEE Trans. Signal Process.*, vol. 62, no. 1, pp. 168–182, Jan. 2014.

[22] F. Yin, C. Fritsche, F. Gustafsson, and A. Zoubir, "TOA-based robust wireless geolocation and Cramer–Rao Lower Bound analysis in harsh LOS/NLOS environments," *IEEE Trans. Signal Process.*, vol. 61, no. 9, pp. 2243–2255, May 2013.

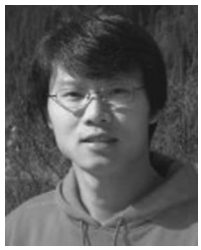
[23] Y. Xie, Y. Wang, P. Zhu, and X. You, "Grid-search-based hybrid TOA/AOA location techniques for NLOS environments," *IEEE Commun. Lett.*, vol. 13, no. 4, pp. 254–256, Apr. 2009.

[24] B. Sieskul, F. Zheng, and T. Kaiser, "A hybrid SS-TOA wireless NLOS geolocation based on path attenuation: TOA estimation and CRB for mobile position estimation," *IEEE Trans. Veh. Technol.*, vol. 58, no. 9, pp. 4930–4942, Nov. 2009.

[25] J. Borras, P. Hatrack, and N. Mandayam, "Decision theoretic framework for NLOS identification," in *Proc. 48th IEEE Veh. Technol. Conf.*, May 1998, vol. 2, pp. 1583–1587.
[26] L. Greenstein, V. Erceg, Y. Yeh, and M. Clark, "A new path-gain/delay-spread propagation model for digital cellular channels," *IEEE Trans. Veh. Technol.*, vol. 46, no. 2, pp. 477–485, May 1997.
[27] A Channel Model for Positioning, Coperation Europenne dans le Domaine de la Recherche Scientifique et Technique (COST), Dec. 2014. [Online]. Available: http://www.lx.it.pt/cost259/g98___20.htm
[28] Z. Deng, Y. Yu, W. Guan, and L. He, "NLOS error mitigation based on modified Kalman filter for mobile location in cellular networks," in *Proc. Int. Conf. WCSP*, Oct. 2010, pp. 1–5.
[29] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[30] A. L. Garcia, *Probability and Random Processes for Electrical Engineering*. New York, NY, USA: Addison-Wesley, 1994.
[31] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, 1st ed., ser. Fundamentals of Statistical Signal Processing. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.
[32] A. W. Bowman and A. Azzalini, *Applied Smoothing Techniques for Data Analysis*. New York, NY, USA: Oxford Univ. Press, 1997.
[33] M. Hlynka and D. Loach, "Generating uniform random points in a regular n-sided polygon," Windsor Math and Statistics Report, Dec. 2013. [Online]. Available: http://web2.uwindsor.ca/math/hlynka/Ngon.pdf

Zohair Abu-Shaban (S'14) received the Bachelor's degree in electrical engineering in Palestine at the top of his class in 2007. Subsequently, he worked as a teaching assistant in the same department for two years before winning a scholarship to pursue the Master's degree with Imperial College London, London, U.K. In 2010, he received the M.Sc. degree in communication and signal processing (with distinction) from the Department of Electrical and Electronic Engineering, Imperial College London, where he received the Department Prize for Outstanding Achievement. He is currently working toward the Ph.D. degree with the Australian National University (ANU), Canberra, Australia, where he is doing research on signal processing techniques for mobile localization.

From January 2012 to May 2014, he was an early-stage researcher with the University of Luxembourg, Luxembourg City, Luxembourg, working in the field of interference cancelation and synchronization for broadcasting satellites sponsored by the National Research Fund, Luxembourg (FNR).



Xiangyun Zhou (M'11) received the Ph.D. degree in telecommunications engineering from the Australian National University, Canberra, Australia, in 2010.

From 2010 to 2011, he was a Postdoctoral Fellow with the University Graduate Center, University of Oslo, Oslo, Norway. In 2011, he joined the Australian National University, where he is currently a Senior Lecturer. His research interests include communication theory and wireless networks.

Dr. Zhou currently serves on the Editorial Board of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and the IEEE COMMUNICATIONS LETTERS. In 2015, he served as a Guest Editor for the IEEE COMMUNICATIONS MAGAZINE feature topic on wireless physical-layer security. From 2013 to 2014, he was the Chair of the Australian Capital Territory Chapter of the IEEE Communications Society and Signal Processing Society. In 2011, he received the Best Paper Award at the IEEE International Conference on Communications.

Thushara D. Abhayapala (M'00–SM'08) received the B.E. degree (with Honors) in engineering and the Ph.D. degree in telecommunications engineering from the Australian National University (ANU), Canberra, Australia, in 1994 and 1999, respectively.

He is a Professor with ANU. He was the Director of the Research School of Engineering at ANU from January 2010 to October 2014 and the Leader of the Wireless Signal Processing Program with the National ICT Australia from November 2005 to June 2007. He has supervised over 30 research students and coauthored over 200 peer-reviewed papers. His research interests include spatial audio and acoustic signal processing, space–time signal processing for wireless communication systems, and array signal processing.

Dr. Abhayapala is an Associate Editor of the IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING. From 2011 to 2015, he was also a member of the Audio and Acoustic Signal Processing Technical Committee of the IEEE Signal Processing Society. He is a Fellow of Engineers Australia.