SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

October 2019, Vol. 62 202301:1–202301:15
https://doi.org/10.1007/s11432-019-9909-5

Geometry-based non-line-of-sight error mitigation and localization in wireless communications

Jingyu HUA[1,2*], Yejia YIN[3], Anding WANG[1], Yu ZHANG[3] & Weidang LU[3]

[1]*School of Information and Electronic Engineering, Zhejiang Gongshang University, Hangzhou* 310018*, China;*
[2]*Zhejiang Provincial Key Laboratory of Information Processing, Communication and Networking, Zhejiang* 310027*, China;*
[3]*College of Information Engineering, Zhejiang University of Technology, Hangzhou* 310023*, China*

Received 9 January 2019/Revised 12 March 2019/Accepted 31 May 2019/Published online 26 August 2019

Abstract Recently, positioning services have received considerable attention. The primary source of the positioning error is non-line-of-sight (NLOS) propagation. To address this problem, we propose a novel NLOS mitigation scheme, in which the geometric relationship between a base station and a mobile station is used. This makes it possible to identify range measurements corrupted by NLOS errors, and the mobile station can then estimate its position through line-of-sight (LOS) measurements. Moreover, the threshold of the NLOS detector is derived via a hybrid method using both the analytical derivation and computer simulation, which significantly reduces the difficulty of identifying thresholds. After identifying the NLOS measurements, a two-step weighted-least-squares algorithm is used to obtain the localization, in which both range and angle measurements are considered. The simulation results reveal that the proposed algorithm yields a high identification probability of NLOS measurements, which results in improved localization performance.

Keywords wireless localization, non-line-of-sight error, geometry, cellular network, residual

Citation Hua J Y, Yin Y J, Wang A D, et al. Geometry-based non-line-of-sight error mitigation and localization in wireless communications. Sci China Inf Sci, 2019, 62(10): 202301, https://doi.org/10.1007/s11432-019-9909-5

1 Introduction

Recently, mobile communication has witnessed rapid developments, and 5G as well as multiple-input and multiple-output (MIMO) antenna technology have received considerable attention from researchers [1–3]. Moreover, accurate location-based services (LBS), such as emergency road assistance, location-based advertisements, location-based messaging and route navigation [4, 5], have become popular topics of research. Note that primary localization methods can be classified into satellite localization and terrestrial wireless localization. The former is useful for active applications in an outdoor environment, such as global position system (GPS)-based navigation, where non-military users can achieve a positioning error less than 10 m [6]; however, it is difficult to use GPS in indoor environments. Moreover, in certain passive applications, such as the safety and security applications like E-911, it is impossible to obtain the GPS locations of hostile targets. However, these positions can be obtained via a wireless signal, such as the cellular signal of the target. Thus, in this study, we focus on terrestrial wireless localization.

Terrestrial wireless localization methods can be classified into two groups: the range-free method and the range-based method. Generally, the former method maintains acceptable performance in a large-scale network that has a high base station density. However, the actual network scale and base station density may considerably vary. Moreover, the propagation environment may be variable such as indoor, urban

* Corresponding author (email: eehjy@163.com)

and suburban areas. Therefore, in this study, we focus on range-based localization because it is more robust to deal with the above mentioned variabilities.

Ranged-based localization depends on the measurement of localization parameters, such as the time of arrival (TOA) [7], received signal strength (RSS) [8], time difference of arrival (TDOA) [9,10] and angle of arrival (AOA) [11,12]. Furthermore, the range-based localization suffers from two types of non-ideal factors: Gaussian measurement noises and non-line-of-sight (NLOS) errors [11]. Generally, an NLOS error occurs when the direct path between the mobile station (MS) and base station (BS) is blocked by an obstacle. Gaussian noise has been extensively studied and greatly improved in previous studies; however, additional studies are required on the corruption of localization by NLOS errors. Note that NLOS propagation causes a long excess path delay, which leads to significant biases in TOA and AOA measurements [13,14]. In this study, the NLOS error is assumed to be larger than the measurement noise; this assumption is consistent with many previous studies [15,16].

The NLOS problem of cellular localization was studied in [11], in which Wang et al. indicated that the typical ranging error introduced by NLOS propagation was approximately half of the cell radius. To solve this problem, several NLOS mitigation techniques have been proposed. Yan et al. [17] attempted to solve the NLOS problem by first identifying the NLOS BSs, and then using the predetermined NLOS statistical information to calibrate the NLOS measurements. However, a reliable statistical model of NLOS information is difficult to obtain because of the complicated and variable wireless propagation environments; thus, in most situations, the above algorithm is not expected to be effective. Moreover, the absence of a reliable statistical model of NLOS error makes it difficult to theoretically analyze the positioning performance; thus, researchers usually analyze NLOS-corrupted localization via simulations.

Certain researchers have improved localization performance by properly weighting line-of-sight (LOS) and NLOS measurements, where the weighting factors are derived from the localization geometries or the range residuals [17,18]. These algorithms always improve the NLOS localization to some extents, however, the NLOS error cannot be completely eliminated, and the optimal weighting factors are not achievable. Thus, these estimates may not always be reliable. Shu et al. [19] proposed a weighted localization algorithm based on geographical constraints. In [20–22], to improve position estimation, the researchers pre-constructed several specified scattering models; however, the rapid-changing urban environment (i.e., appearance and disappearance of buildings) made it difficult to identify an accurate scattering model. In [23], the angle of departure (AOD) was used to identify the NLOS measurement as long as there were sufficient single reflected NLOS paths or the AOD could be measured at the MS. Theoretically, AOD methods can effectively mitigate the NLOS error; however, the single-reflection propagation model is not consistent with practical environments, and the AOD measurement in the MS does not have sufficient accuracy due to the array size limitation. Consequently, the AOD method does not perform well in real-world scenarios.

In addition to the above-mentioned methods, some researchers have attempted to perform identifying LOS measurements and estimating position using only LOS measurements. The identification criteria originated from the time-history-based hypothesis test [24], probabilistic model [23], maximum likelihood detection [25], and residual test [26]. The disadvantages of these methods are that the number of LOS BSs must be no less than three, and the probability of a false alarm must be low. Marano et al. [27] and Zhang et al. [28] proposed using multipath channel information to identify NLOS measurements, in which the statistics of a channel, such as kurtosis, were analyzed according to the channel measurements. However, these two position estimators may be sensitive to the scattering environment as was observed in [28]; moreover, they may require a large number of real-world measurements to cover a wide range of environments in the cellular network. Therefore, it is preferable to investigate NLOS mitigation using an analytical method rather than a measurement method. Yu and Dutkiewicz [29] collected multiple range estimates in the time domain to identify NLOS measurements in which the diversity gain in the time domain significantly increased identification probability. However, in addition to its complex processing and excellent performance, latency is incurred during the collection of range estimates to establish a history, and the NLOS status must be invariant over such durations, which limits its application in a changing environment. In addition to the above-mentioned recognition algorithm, some researchers have

proposed optimization algorithms to solve positioning problems at the cost of higher complexities. Shu et al. [30] proposed using the Lagrangian method to transform the constrained TDOA problem into the identification of a set of polynomials in the Lagrangian multiplier. Wang et al. [9] proposed transforming the positioning problem into a semi-definite programming problem.

To address above-mentioned insufficiencies, Hua et al. [31] proposed detecting the LOS-BS by algorithm-residuals and thereby estimating the MS position, where temporary position estimation was required to construct the residual, resulting in increased computation. Moreover, the method of [31] experiences performance degradation in the case of two LOS BSs. Unlike [31], in this study we used the ranging-residual to detect the LOS BS because this type of residual is more popular and useful for researchers. In particular, we first exploit the geometric relationship between the MS and BS to identify the LOS BS, where both TOA and AOA measurements are included. Second, we employ a two-step weighted-least-squares (TS-WLS) method to estimate the MS position using only LOS BSs. Unlike traditional methods, the proposed algorithm performs well for only two LOS BSs; thus, it is suitable for practical applications. In this study, we evaluate the proposed algorithm by Monte Carlo simulations that use randomly generated NLOS errors and MS positions. The results reveal that the proposed method has a higher detection probability and better localization performance.

The remainder of this study is organized as follows. In Section 2, we introduce the localization model and the conventional localization algorithms. In Section 3, we propose a LOS identification algorithm and a semi-analytical threshold determining method. In Section 4, the numerical results are provided and analyzed. Finally, in Section 5, the conclusion of this study is provided.

2 Preliminaries

The actual range between the MS and the i-th BS can be expressed as follows [28]:

$$r_i = \sqrt{(x - x_i)^2 + (y - y_i)^2}, \tag{1}$$

where (x, y) and (x_i, y_i) denote the actual MS position and the ith BS coordinate, respectively. Because of the existence of measurement noise and NLOS error, the range measurement can be modeled as follows:

$$\hat{r}_i = r_i + e_i + n_i, \tag{2}$$

where $\hat{r}_i$, e_i, and n_i denote the range measurement, measurement noise, and NLOS error, respectively. If the propagation path is the LOS, then $n_i = 0$; otherwise, n_i is usually much larger than e_i. Moreover, the actual range is usually much larger than the measurement noise.

The tangent of the AOA at the ith BS without any disturbance can be written as follows:

$$\tan(\theta_i) = \frac{y - y_i}{x - x_i}. \tag{3}$$

Because the AOA measurement is corrupted by measurement noise and NLOS error, it should be rewritten as follows:

$$\hat{\theta}_i = \theta_i + \alpha_i + \beta_i, \tag{4}$$

where α_i and β_i denote the measurement noise and NLOS error in the AOA measurement, respectively. Generally, the direct path may be blocked a majority of the time, and the NLOS error of the AOA may be very large.

Based on the above-mentioned model, the following linear equation group (N BSs) can be easily derived according to the following geometric relation:

$$\boldsymbol{AX} = \boldsymbol{b}, \tag{5}$$

with

$$\boldsymbol{b}=\begin{bmatrix}\hat{r}_1^2-x_1^2-y_1^2\\ \vdots\\ \hat{r}_N^2-x_N^2-y_N^2\\ y_1-\tan\hat{\theta}_1x_1\\ \vdots\\ y_N-\tan\hat{\theta}_Nx_N\end{bmatrix},\quad \boldsymbol{X}=\begin{bmatrix}x\\ y\\ R\end{bmatrix},\quad \boldsymbol{A}=\begin{bmatrix}-2x_1,-2y_1,1\\ \vdots\\ -2x_N,-2y_N,1\\ -\tan\hat{\theta}_1,1,0\\ \vdots\\ -\tan\hat{\theta}_N,1,0\end{bmatrix},$$

where R denotes x^2+y^2. Note that the upper halves of $\boldsymbol{A}$ and $\boldsymbol{b}$ are derived from (1), while the lower halves represent the linear equations between (x_i,y_i) and (x,y) in (3).

Eq. (5) can be solved using the least square (LS) principle as follows:

$$\hat{\boldsymbol{X}}=(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}, \tag{6}$$

where $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{-1}$ denote the matrix transpose and inversion, respectively. Finally, the position estimation of the MS is the former two elements of $\hat{\boldsymbol{X}}$; this is referred as the TOA-AOA/LS algorithm in this study.

Because the LS algorithm produces a large bias in MS position estimation, an algorithm called the two-step WLS (TS-WLS) algorithm, such as that in [9], is introduced.

Assuming that the actual value of measurement noise is small, the position estimation error vector $\boldsymbol{b}$ resulting from measurement noise can be approximated as follows:

$$\psi\approx\begin{bmatrix}2r_1e_1\\ \vdots\\ 2r_Ne_N\\ r_1\alpha_1\\ \vdots\\ r_N\alpha_N\end{bmatrix}, \tag{7}$$

where r_i, e_i, and α_i are defined in (1), (2), and (4), respectively. In (7), we exploit the following relations:

$$(x-x_i)\left(\tan(\theta_i+\alpha_i)-\tan(\theta_i)\right)=(x-x_i)\frac{\sin(\alpha_i)}{\cos(\alpha_i)\cos(\theta_i)}\approx\frac{x-x_i}{\cos(\theta_i)}\alpha_i=r_i\alpha_i,$$

$$(r_i+e_i)^2-r_i^2\approx 2r_ie_i.$$

Then, we can rewrite (7) in the matrix form:

$$\psi=2\boldsymbol{B}\boldsymbol{z}, \tag{8}$$

where $\boldsymbol{B}=\frac{1}{2}\mathrm{diag}\{2r_1,2r_2,\ldots,2r_N,r_1,r_2,\ldots,r_N\}$ and $\boldsymbol{z}=[e_1,e_2,\ldots,e_N,\alpha_1,\alpha_2,\ldots,\alpha_N]^{\mathrm{T}}$. Here, $\mathrm{diag}\{\boldsymbol{v}\}$ represents a diagonal matrix with its diagonal elements equal to vector $\boldsymbol{v}$.

Because the elements of $\boldsymbol{B}$ are unknown in practice, they can only be approximated by range measurements. Accordingly, the covariance matrix of ψ can be expressed as follows:

$$\Psi=E\left[\psi\psi^{\mathrm{T}}\right]=4\boldsymbol{B}\boldsymbol{Q}\boldsymbol{B}, \tag{9}$$

where $\boldsymbol{Q}=\mathrm{diag}\{\sigma_{r,1}^2,\ldots,\sigma_{r,N}^2,\sigma_{\alpha,1}^2,\ldots,\sigma_{\alpha,N}^2\}$. Moreover, $\sigma_{r,i}^2$ and $\sigma_{\alpha,j}^2$ denote the variances of the ith range measurement and the jth AOA measurement, respectively. Eq. (5) can then be solved by the WLS method as follows [6]:

$$\hat{\boldsymbol{X}}_{\mathrm{WLS}}=(\boldsymbol{A}^{\mathrm{T}}\Psi^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\Psi^{-1}\boldsymbol{b}. \tag{10}$$

Eqs. (8)–(10) are iteratively performed, i.e., $\boldsymbol{B}$ is recalculated by an older estimation of (10), which leads to a new position estimation. This iteration is stopped when the previous MS position estimation

slightly differs from the most recent MS position estimation. Then, the covariance matrix of $\hat{\boldsymbol{X}}_{\mathrm{WLS}}$ can be obtained as follows [6]:

$$\mathrm{cov}(\hat{\boldsymbol{X}}_{\mathrm{WLS}}) = (\boldsymbol{A}^{\mathrm{T}}\Psi^{-1}\boldsymbol{A})^{-1}. \tag{11}$$

The estimation of (10) requires independent elements in $\hat{\boldsymbol{X}}_{\mathrm{WLS}}$; however, the elements are not truly independent. Therefore, another WLS process should be used to refine the position estimation, which is also performed in [6]. First, we assume that the errors of x, y, and R are s_1, s_2, and s_3, respectively. Thus, $\hat{\boldsymbol{X}}_{\mathrm{WLS}}$ can be expanded as follows:

$$\hat{\boldsymbol{X}}_{\mathrm{WLS}} = \begin{bmatrix} x + s_1 \\ y + s_2 \\ R + s_3 \end{bmatrix}. \tag{12}$$

Second, we define another error vector as follows:

$$\psi' = \boldsymbol{b}' - \boldsymbol{A}'\mathbf{X}', \tag{13}$$

where

$$\boldsymbol{A}' = \begin{bmatrix} 1, & 0 \\ 0, & 1 \\ 1, & 1 \end{bmatrix} \quad \text{and} \quad \boldsymbol{X}' = \begin{bmatrix} x^2 \\ y^2 \end{bmatrix}.$$

In (13),

$$\boldsymbol{b}' = \begin{bmatrix} (x + s_1)^2 \\ (y + s_2)^2 \\ R + s_3 \end{bmatrix}$$

can be computed according to $\hat{\boldsymbol{X}}_{\mathrm{WLS}}$. From (13), by assuming that residual estimation errors are small, we can derive

$$\psi' = \begin{bmatrix} 2xs_1 + s_1^2 \\ 2ys_2 + s_2^2 \\ s_3 \end{bmatrix} \approx \begin{bmatrix} 2xs_1 \\ 2ys_2 \\ s_3 \end{bmatrix}. \tag{14}$$

Thus, the covariance matrix of ψ' is formulated as follows:

$$\Psi' = E\left[\psi'\psi'^{\mathrm{T}}\right] = 4\boldsymbol{B}'\mathrm{cov}(\hat{\boldsymbol{X}}_{\mathrm{WLS}})\boldsymbol{B}' \tag{15}$$

with $\boldsymbol{B}' = \mathrm{diag}\{x, y, 0.5\}$. Finally, $\boldsymbol{X}'$ can be solved by another WLS process as follows [6]:

$$\hat{\boldsymbol{X}}'_{\mathrm{WLS}} = (\boldsymbol{A}'^{\mathrm{T}}\Psi'^{-1}\boldsymbol{A}')^{-1}\boldsymbol{A}'^{\mathrm{T}}\Psi'^{-1}\boldsymbol{b}'. \tag{16}$$

Using (16), MS position estimation is straightforward.

3 Identification of NLOS and LOS range measurement

Similar to many conventional localization algorithms, both the TOA-AOA/LS and TOA-AOA/TS-WLS algorithms suffer greatly from NLOS propagation. Thus, in this section we investigate how to identify NLOS and LOS range measurements. Then, the detected LOS measurements are used in the localization algorithm, which explicitly mitigates the influence of NLOS. Moreover, it is difficult to analyze the proposed NLOS detector in theory because a reliable model for statistical NLOS information is absent in our study and the existing literature. Thus, we present detailed simulations in Section 4 to analyze the proposed algorithm.

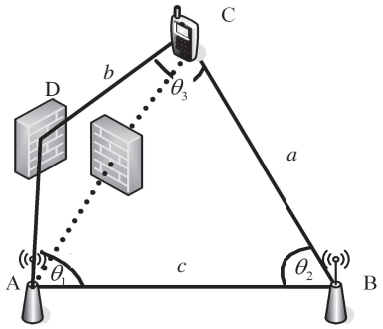

Figure 1 Example of NLOS and LOS propagation.

3.1 Localization geometry between MS and BSs

Figure 1 presents an example of NLOS and LOS propagation, where A, B, and C denote BS-A, BS-B, and the MS, respectively. Figure 1 illustrates that the propagation channel between C and A is blocked by an obstacle D, leading to an NLOS-corrupted range measurement b and angle measurement θ_1. Because there is no obstacle in the channel (C $\rightarrow$ B), the range measurement a and angle measurement θ_2 are LOS measurements, i.e., they are only affected by measurement noise and are close to their actual values. Because the spatial locations of BSs are invariant, the distance between A and B can be predetermined and denoted c. Now, the range (distance) between A and C can be obtained using the law of cosine:

$$b' = \sqrt{a^2 + c^2 - 2ac\cos\theta_2}. \tag{17}$$

Moreover, if the channel between A and C is LOS, the range from C to B can be computed as follows:

$$a' = \sqrt{b^2 + c^2 - 2bc\cos\theta_1}. \tag{18}$$

Next, we can define residuals as follows:

$$\delta_1 = |a - a'|, \quad \delta_2 = |b - b'|. \tag{19}$$

From Figure 1 and (17)–(19), it can be seen that if there is no NLOS channel, the measured a and b must be close to the calculated a' and b', i.e., the residual must be small. Otherwise, at least a or b must be corrupted by the NLOS error; consequently, the residual must be much larger than that of the LOS case with a high probability. Next, this geometric characteristic is used to design a high-performance detector to distinguish NLOS and LOS measurements. This requires only two LOS measurements, although conventional methods requires at least three. To clearly explain NLOS/LOS identification, we briefly introduce the parameters used in this subsection as follows.

- a, b: BS-MS distance directly computed by TOA measurements;
- a', b': BS-MS distance computed by the cosine law, in which both TOA and AOA measurements are employed;
- a_0, b_0: actual BS-MS distance;
- $\Delta a, \Delta b$: measurement noise of $\{a, b\}$;
- $\Delta a', \Delta b'$: calculation error of $\{a', b'\}$ in (17) and (18);
- θ_1, θ_2: AOA measurements;
- $\theta_{1,0}, \theta_{2,0}$: actual AOAs;
- $\Delta\theta_1, \Delta\theta_2$: measurement noise of $\{\theta_1, \theta_2\}$.

3.2 Identification strategy

Based on the above discussion, the identification of LOS/NLOS measurements can be achieved by a process of threshold detection; specifically, if the residuals in (19) are smaller than the threshold, both

BS-A and BS-B are considered as LOS BSs, and a, b, θ_1, and θ_2 are then LOS measurements. Otherwise, there is at least one NLOS BS. Finally, the proposed identification scheme is structured as follows:

(1) Assume that there are N BSs, and the measured range and AOA are $\hat{r}_i$ and $\hat{\theta}_i$, respectively.

(2) Divide all BSs into several groups with two BSs per group, which produces C_N^2 groups for N BSs.

(3) Calculate the residuals of a group using (17)–(19) and compare the residuals with a certain threshold. If the residual is smaller than the threshold, the range and AOA measurements of this group are estimated from LOS environments, which can later be used to localize the MS. Otherwise, the measurements are questionable.

(4) Repeat step (3) C_N^2 times until each group has been addressed.

(5) Compare and analyze all threshold detection results and identify all LOS measurements.

Based on the identification result of each group, all LOS measurements can be determined. Then, all of the identified LOS measurements, which may have been identified by different groups, are used to determine the location of the MS using the TS-WLS method.

Naturally, the group number depends on the value of N, i.e., a larger N leads to larger group number and therefore higher computation. Generally, conventional studies use $N = 5$ or 7, which produces an acceptable group number.

3.3 Threshold derivation

Because it is difficult to develop a reliable and universal statistical model for NLOS error, it is challenging to theoretically determine the threshold mentioned in Subsection 4.2. Moreover, it is impossible to determine this threshold by an exhaustive search. Therefore, in this subsection, we propose a semi-analytical method that combines analytical and simulated approaches to address this challenge. First, a perturbation analysis is performed to determine the relation between the residual and localization parameter. Second, a targeted simulation is performed to identify the optimal thresholds in various scenarios. Finally, a comprehensive analysis of these thresholds is performed to determine an appropriate threshold that can account for the varying propagation environment. Note that the selected threshold is not optimal for all scenarios; however, it always leads to an acceptable performance.

By exploiting the above hybrid simulation/derivation strategy, the threshold can be determined in an averaged sense. Therefore, although the simulation stage only covers limited scenario parameters, such as network topology, MS position, TOA precision, AOA precision, and maximum NLOS error, the determined threshold yields acceptable performance for a wide range of scenarios. Next, we demonstrate that the threshold can be connected with the propagation environment by introducing the standard deviation of measurement noise.

If the propagation channels (C→A) and (C→B) are both LOS, the residual must be small, and the measured a, b, θ_1, and θ_2 in (17) and (18) contain measurement noise. Thus, we have

$$\begin{aligned} &a = a_0 + \Delta a, \quad b = b_0 + \Delta b, \\ &\theta_1 = \theta_{1,0} + \Delta\theta_1, \quad \theta_2 = \theta_{2,0} + \Delta\theta_2, \end{aligned} \tag{20}$$

where $a_0, b_0, \theta_{1,0}$ as well as $\theta_{2,0}$ represent the corresponding real values, and Δa, Δb, $\Delta\theta_1$ as well as $\Delta\theta_2$ represent the measurement noise.

Similarly, the calculated ranges can be formulated as

$$a' = a_0 + \Delta a', \quad b' = b_0 + \Delta b', \tag{21}$$

where $\Delta a'$ and $\Delta b'$ denote calculation errors. We then substitute (20) and (21) into (17) and square both sides. Eq. (17) can then be reformulated as

$$(b_0 + \Delta b')^2 = a^2 + c^2 - 2ac\cos\theta_2 = (a_0 + \Delta a)^2 + c^2 - 2(a_0 + \Delta a)c\cos(\theta_{2,0} + \Delta\theta_2)\,. \tag{22}$$

By expanding both sides of (22), we obtain

$$b_0^2 + 2b_0\Delta b' + \Delta b'^2 = a_0^2 + 2a_0\Delta a + \Delta a^2 + c^2 - 2(a_0 + \Delta a)c(\cos\theta_{2,0}\cos\Delta\theta_2 - \sin\theta_{2,0}\sin\Delta\theta_2)\,. \tag{23}$$

Because the value of $\Delta\theta_2$ is very small in LOS environments, we have

$$\cos\Delta\theta_2 \approx 1, \quad \sin\Delta\theta_2 \approx \Delta\theta_2.$$

By ignoring the second-order term of the measurement noise and calculation error in (23), we obtain

$$\begin{aligned} &b_0^2 + 2b_0\Delta b' \approx a_0^2 + 2a_0\Delta a + c^2 - 2(a_0+\Delta a)c(\cos\theta_{2,0}\cos\Delta\theta_2 - \sin\theta_{2,0}\sin\Delta\theta_2) \\ &= a_0^2 + c^2 - 2a_0c\cos\theta_{2,0} + 2a_0\Delta a - 2(-a_0c\sin\theta_{2,0}\Delta\theta_2 + \Delta ac\cos\theta_{2,0} - \Delta ac\Delta\theta_2\sin\theta_{2,0}). \end{aligned} \tag{24}$$

Note that the value of b_0^2 is equal to

$$b_0^2 = a_0^2 + c^2 - 2a_0c\cos\theta_{2,0}. \tag{25}$$

Eq. (24) can then be further simplified as follows:

$$2b_0\Delta b' \approx 2a_0\Delta a + 2a_0c\sin\theta_{2,0}\Delta\theta_2 - 2\Delta ac\cos\theta_{2,0} + 2\Delta ac\Delta\theta_2\sin\theta_{2,0}. \tag{26}$$

Thus, $\Delta b'$ is given by

$$\Delta b' \approx (a_0\Delta a + a_0c\sin\theta_{2,0}\Delta\theta_2 - \Delta ac\cos\theta_{2,0} + \Delta ac\Delta\theta_2\sin\theta_{2,0})/b_0. \tag{27}$$

Considering

$$|b-b'| = |b_0+\Delta b - b_0 - \Delta b'| = |\Delta b - \Delta b'| \leqslant |\Delta b'| + |\Delta b|\,, \tag{28}$$

we can then define the threshold as follows:

$$\mathrm{TH}_2 = |\Delta b'| + |\Delta b|\,. \tag{29}$$

By denoting σ_r and σ_a as the standard deviation of the range measurement and AOA measurement, respectively, we obtain the following approximations:

$$|\Delta b'| \leqslant \frac{\lambda}{b_0}\left|\begin{array}{l} a_0\sigma_r + a_0c\sigma_a\sin\theta_{2,0} - \\ \sigma_r c\cos\theta_{2,0} + \sigma_r c\lambda\sigma_a\sin\theta_{2,0} \end{array}\right|, \tag{30a}$$

$$|\Delta b| \leqslant \lambda\,|\sigma_r|\,, \tag{30b}$$

where λ determines the confidence interval of a Gaussian variable. Note that the above inequality is reliable with a high probability if a suitable λ is selected, and an appropriate λ is determined in subsequent simulations. The above expressions are more complex than the threshold derivation in [31]. However, the proposed residual is more robust than the residual of [31] because the proposed residual introduces more NLOS-dependent variables in the threshold calculation. Therefore, it is more sensitive to NLOS errors. In particular, when there are only two LOS BSs, the advantage is clear, as demonstrated in Subsection 4.2.

For more accurate NLOS identification, the value of λ is determined by targeted simulations, and the detecting threshold is then computed according to (29) and (30). In practice, the actual a_0 and b_0 are unknown; therefore, they can only be replaced by the measured a and b. This indicates that it is necessary to search for an appropriate λ with the help of simulations. Similarly, the actual value of $\theta_{2,0}$ is also unavailable and may deviate from its measurement to a large degree. Moreover, the range measurement is more stable than the angle measurement in the NLOS environment; thus, we can make an alternate approximation of $\theta_{2,0}$ using the cosine law as follows:

$$\begin{aligned} &\cos(\theta_{2,0}) \approx (a^2 + c^2 - b^2)/(2ac), \\ &\sin(\theta_{2,0}) \approx \sqrt{1-\cos^2(\theta_{2,0})}. \end{aligned} \tag{31}$$

Now, using (29)–(31), the threshold TH_2 can be easily computed. Similarly, the threshold of $|a-a'|$ can be formulated as follows:

$$\mathrm{TH}_1 = |\Delta a'| + |\Delta a|\,, \tag{32}$$

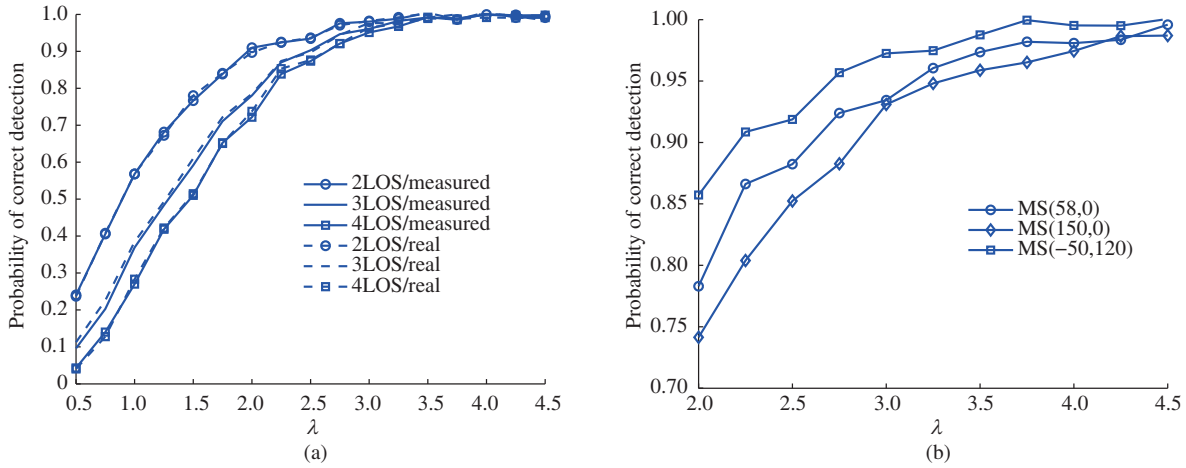

Figure 2 (Color online) Influence of λ on the probability of correct identification. Comparison of different (a) LOS BS numbers and (b) MS locations.

where

$$|\Delta a'| \leqslant \frac{\lambda}{a_0} \left| \begin{array}{l} b_0\sigma_r + b_0 c\sigma_a \sin(\theta_{1,0}) - \\ \sigma_r c \cos\theta_{1,0} + \sigma_r c\lambda\sigma_a \sin\theta_{1,0} \end{array} \right|$$

and $|\Delta a| \leqslant \lambda\,|\sigma_r|$. Given the above-mentioned thresholds, the NLOS detector can be defined as follows:

$$\begin{cases} \delta_1 < \mathrm{TH}_1 \quad \text{and} \quad \delta_2 < \mathrm{TH}_2 \rightarrow \text{Both LOS BSs}, \\ \delta_1 > \mathrm{TH}_1 \quad \text{or} \quad \delta_2 < \mathrm{TH}_2 \rightarrow \text{at least one NLOS BS existed}. \end{cases} \tag{33}$$

It can be seen in (29) and (32) that the threshold depends on the standard deviation of the range measurement (SDR) and the standard deviation of the angle measurement (SDA). Estimating these two parameters from the range and angle measurements is beyond the scope of this study and is left for future work. In this study, we assume that these two parameters are already known.

4 Simulations and analysis

Without loss of generality, we construct a classic base station topology with five BSs located at (0, 0), (1000, 1000), (−1000, −1000), (1000, −1000), and (−1000, 1000), where the measurement unit is 'meter'. Moreover, both range measurement noise and angle measurement noise are modeled as zero-mean Gaussian variables. However, the NLOS error of the range measurement is modeled as a uniform distributed random variable. It takes a value of 50–500 m, which is used throughout this paper; moreover, because NLOS effects on AOA measurements can be very large, in this study we assume that the NLOS error of AOA is uniformly distributed from $-\pi$ to π, which represents a highly-poor measurement result.

4.1 Determination of λ

As demonstrated in the above Subsection 3.3, the selection of λ plays an important role in NLOS identification. Therefore, a tradeoff should be made between the false alarm probability and missing probability. Here, under the guidance of the analytical study presented in Subsection 3.3, targeted simulations are performed to determine an appropriate λ and the detecting thresholds. Note that each simulation runs independently 2000 times. Moreover, the standard deviation of the zero-mean Gaussian measurement noise is 10 m for the range measurement and 1° for the angle measurement.

Figure 2(a) shows the probability of correct identification with different λ, where the dotted curve nLOS/real signifies that the real values of a_0, b_0, $\theta_{2,0}$, and $\theta_{1,0}$ are used in (29) and (32). The measured variables are used in the solid curve nLOS/measured, and the coordinate position of the MS is (58.5,

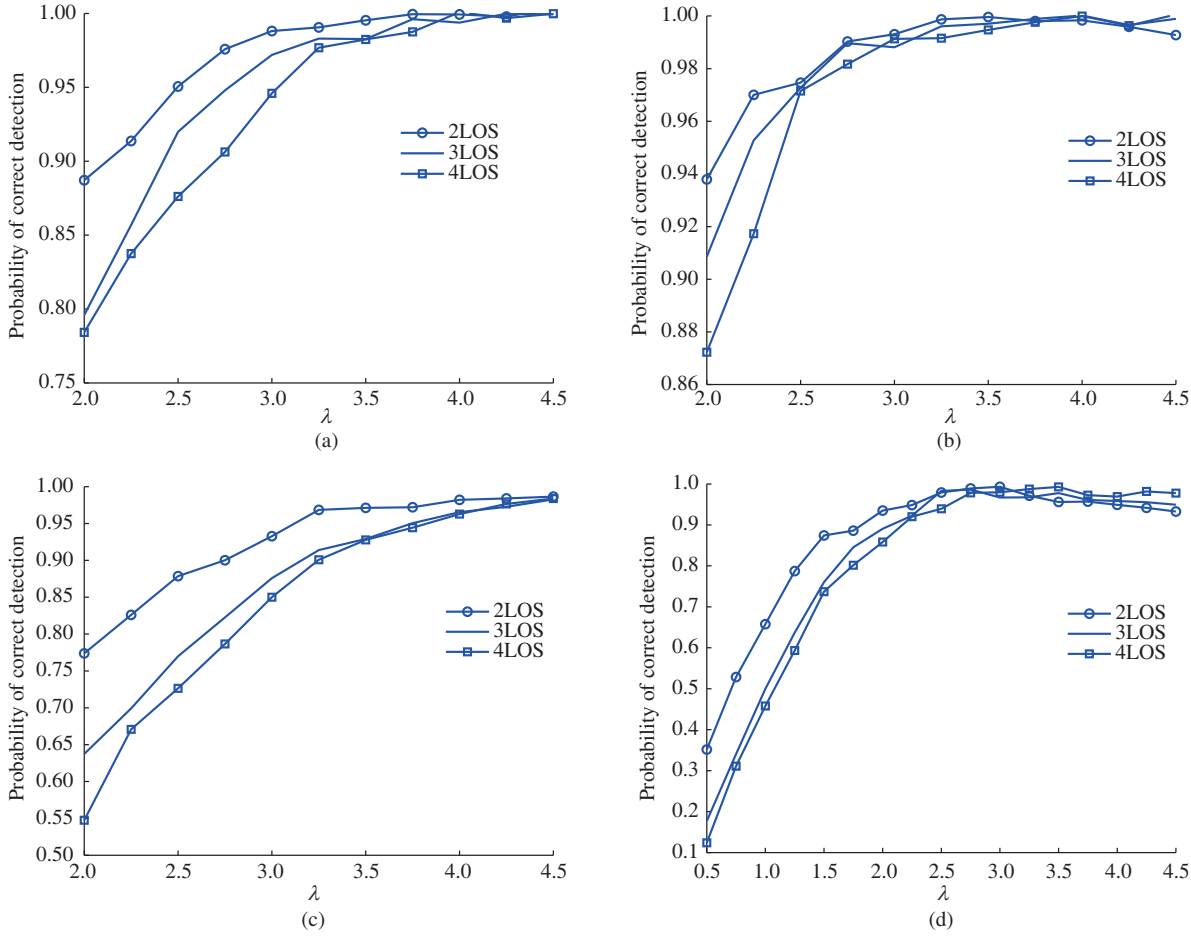

Figure 3 (Color online) Probability of correct detection with different λ, SDR, and SDA. (a) SDR = 5 m, SDA = 0.5°; (b) SDR = 4 m, SDA = 1°; (c) SDR = 16 m, SDA = 1°; (d) SDR = 10 m, SDA = 2°.

0). Note that nLOS means that there are n LOS measurements. Figure 2(a) demonstrates that both the measured variable value and real variable value produce almost identical curves, which confirms the validity of the analytical derivation in Subsection 4.3. Moreover, because a large number of LOS measurements may result in a larger probability of false alarms when the value of λ is small, it is much easier to identify NLOS measurements when there are only two LOS measurements, as can be seen in Figure 2(a). This may be because when λ is small, there are more opportunities to identify the LOS group as an NLOS group, and a higher number of LOS groups leads to a higher probability of error. Figure 2(b) shows the relationship between λ and the detection probability when the MS is in different positions. The resulting simulation reveals that, although different MS positions lead to fluctuations in detection performance, they are still acceptable within a certain range of λ values. Generally, the probability of correct detection (identification) is high when λ decreases into the range of 2.5–4.5. Yet, this area must be thoroughly searched to achieve an appropriate selection of λ.

Next, with different SDA and SDR setups, the probability of correct identification is presented for several λ, and the result is helpful in determining λ for a wide and common range of SDA and SDR values.

Figure 3 shows the probability of correct identification for different values of λ, SDR, and SDA. As shown in Figure 3(a), to achieve a high detection probability of no less than 95%, we selected the λ region 3.25–4.5 as the candidate are, A_1. Similarly, as shown in Figure 3(b), we defined the candidate area A_2 as $2.5 \leqslant \lambda \leqslant 4.5$. Furthermore, the candidate area A_3, as seen in Figure 3(c), is $3.6 \leqslant \lambda \leqslant 4.5$, and the candidate area A_4, as seen in Figure 3(d), is $2.75 \leqslant \lambda \leqslant 3.75$. Finally, a feasible λ region is as follows:

$$f = A_1 \cap A_2 \cap A_3 \cap A_4 = [3.6, 3.75]. \tag{34}$$

Table 1 Detection probability of RT algorithm with SDA= 1° and SDR = 9 and 18 m

LOS BS number	7	6	5	4	3	2
SDR = 9 m	0.940	0.947	0.944	0.943	0.847	–
SDR = 18 m	0.932	0.941	0.949	0.927	0.721	–

Table 2 Performance of proposed algorithm with different SDA: SDR = 9 m

LOS BS number	7	6	5	4	3	2
SDA = 0.1	1	1	1	0.999	0.998	0.994
SDA = 0.5	0.995	0.991	0.987	0.985	0.988	0.976
SDA = 1	0.979	0.968	0.951	0.945	0.937	0.935

Table 3 Performance of proposed algorithm for different SDA: SDR = 18 m

LOS BS number	7	6	5	4	3	2
SDA = 0.1	1	1	0.999	0.997	0.996	0.994
SDA = 0.5	0.994	0.990	0.989	0.988	0.982	0.983
SDA = 1	0.985	0.967	0.956	0.938	0.937	0.941

Based on (34), further simulations are run to determine a more appropriate value of λ, and the final choice is 3.68. Note that the detailed simulation is not provided because of space limitations. Moreover, simulations are run for different MS positions, and the results are very similar to those of Figures 2 and 3, which signifies that the determination of λ is insensitive to MS position. The following subsections use random MS positions, and the resulting superior performance demonstrates that the determination of λ is insensitive to MS position. Moreover, although the simulation setup is specified in this subsection, note that the determined λ is not sensitive to the specified simulation setup, as the derivation in Section 3 does not require strict constraints on the environmental setup.

4.2 NLOS identification performance

In Subsection 4.1, the selection of λ is determined. Here, we compare the proposed NLOS detection algorithm with the residual test (RT) algorithm [32]. Moreover, to demonstrate the influence of BS topologies, the proposed algorithm is tested in another classic BS topology, in which BSs are located at $(0,0),(\sqrt{3}r,0),(\frac{\sqrt{3}}{2}r,\frac{3}{2}r),(-\frac{\sqrt{3}}{2}r,\frac{3}{2}r),(-\sqrt{3}r,0),\ (-\frac{\sqrt{3}}{2}r,-\frac{3}{2}r)$, and $(\frac{\sqrt{3}}{2}\mathrm{r},-\frac{3}{2}\mathrm{r})$, where $r = 1000$ m. Moreover, the MS position is generated randomly in the topology.

Table 1 shows that the RT algorithm can only deal with the case in which the LOS BS number is no less than three. However, when there are three LOS BSs, the probability of successful detection cannot meet the requirements of real-world applications. Tables 2 and 3 show that the probability of successful detection is higher than 0.93 even if there are two LOS BSs. Moreover, when SDA = 0.1°, the AOA estimate is accurate, and the probability of successful detection is at least 99%.

Moreover, we determined via simulation that the final localization performance rapidly degrades with a detection probability of less than 0.9, which signifies that the proposed method performs well for a common range of SDA. Previous studies have demonstrated that the AOA estimator of the BS can limit its SDA within one degree in practice [33]; therefore, the proposed method is effective in real-world scenarios. Note that a smaller BS number leads to a higher identification probability, i.e., the identification performance of a five-BS setup is superior to that of a seven-BS setup by a small amount. The reason for this is as follows:

• For a given LOS measurement number, a larger BS number results in more NLOS measurement groups, and therefore a higher false alarm probability.

Accordingly, we use a five-BS setup in this study, except in the comparison with the RT algorithm. Moreover, we recommend that engineers use a five-BS setup for use with our proposed algorithm. This will be a simple task in a cellular network because an MS is currently capable of constructing links with multiple remote units.

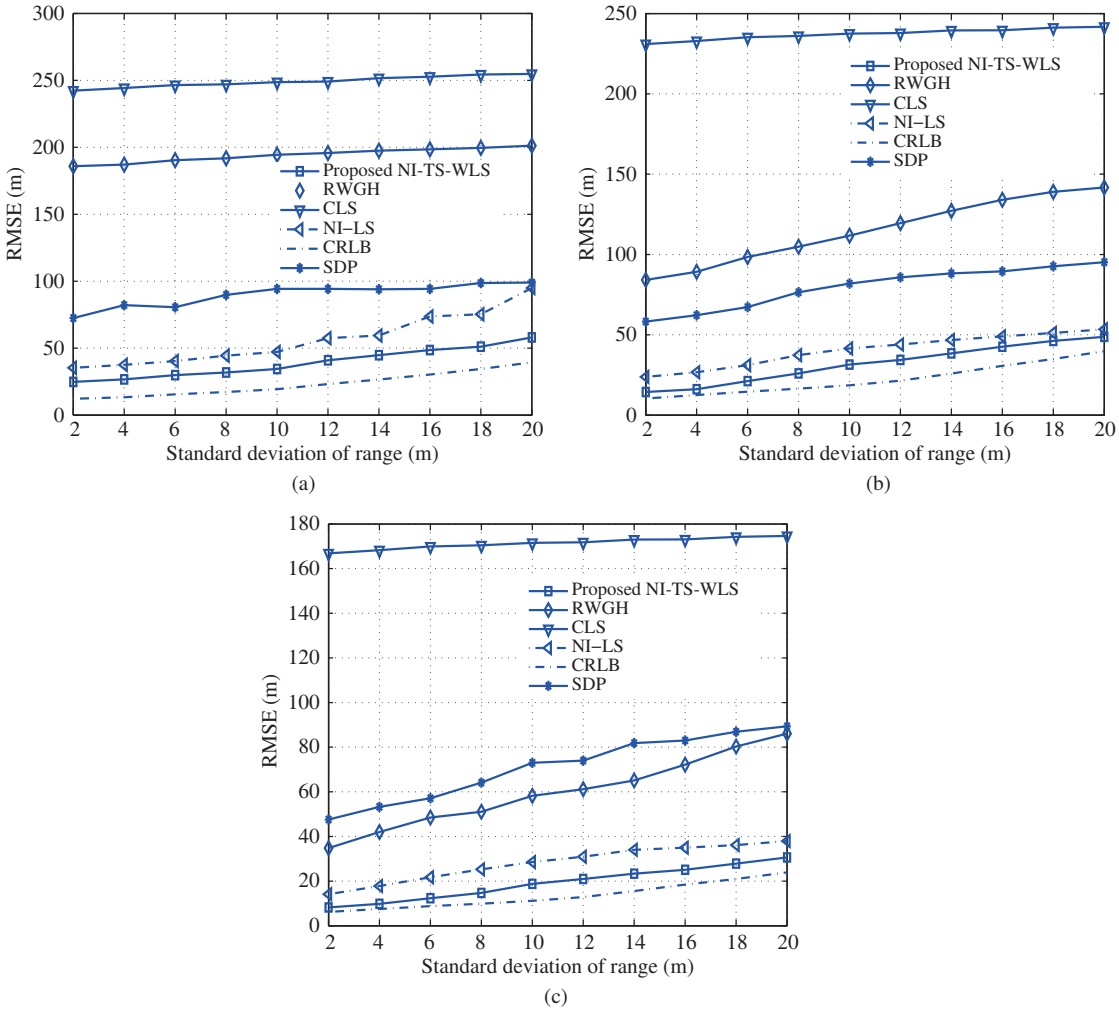

Figure 4 (Color online) RMSE performance comparison for different SDR: SDA = 1°. (a) Case 5/2; (b) case 5/3; (c) case 5/4.

4.3 Performance of NLOS identification and TS-WLS localization

In this subsection, six algorithms are compared under the same simulation conditions as in Subsection 4.1. The first algorithm is NLOS identification plus the least squares algorithm (NI-LS) [31]; the second algorithm is the proposed NLOS identification plus the TS-WLS localization algorithm, denoted NI-TS-WLS; the third algorithm is the residual weighting algorithm (RWGH) [34]; and the final algorithms are optimization algorithms such as the constrained least square (CLS) algorithm [35], opt-LLOP algorithm [36], and SDP algorithm [9]. The Cramer-Rao lower bound (CRLB) [37] is presented as a performance benchmark representing the best achievable performance of the TOA/AOA-based position estimator without diversity in time or other domains.

As shown in Figure 1, the BS number is five; thus, there are three cases in this subsection, as follows:

- Case 5/2: two measurements are LOS.
- Case 5/3: three measurements are LOS.
- Case 5/4: four measurements are LOS.

Each case runs independently 1000 times, and the root-mean-square-error (RMSE) is then calculated.

Figure 4 shows the influence of SDR on the localization algorithm. This figure demonstrates that the CLS and RWGH methods produce large estimation errors, while the SDP, NI-LS, and proposed algorithm NI-TS-WLS display better performance. However, although the SDP algorithm significantly improves its performance, its performance is poorer than that of the proposed algorithm, e.g., when SDR = 10 m

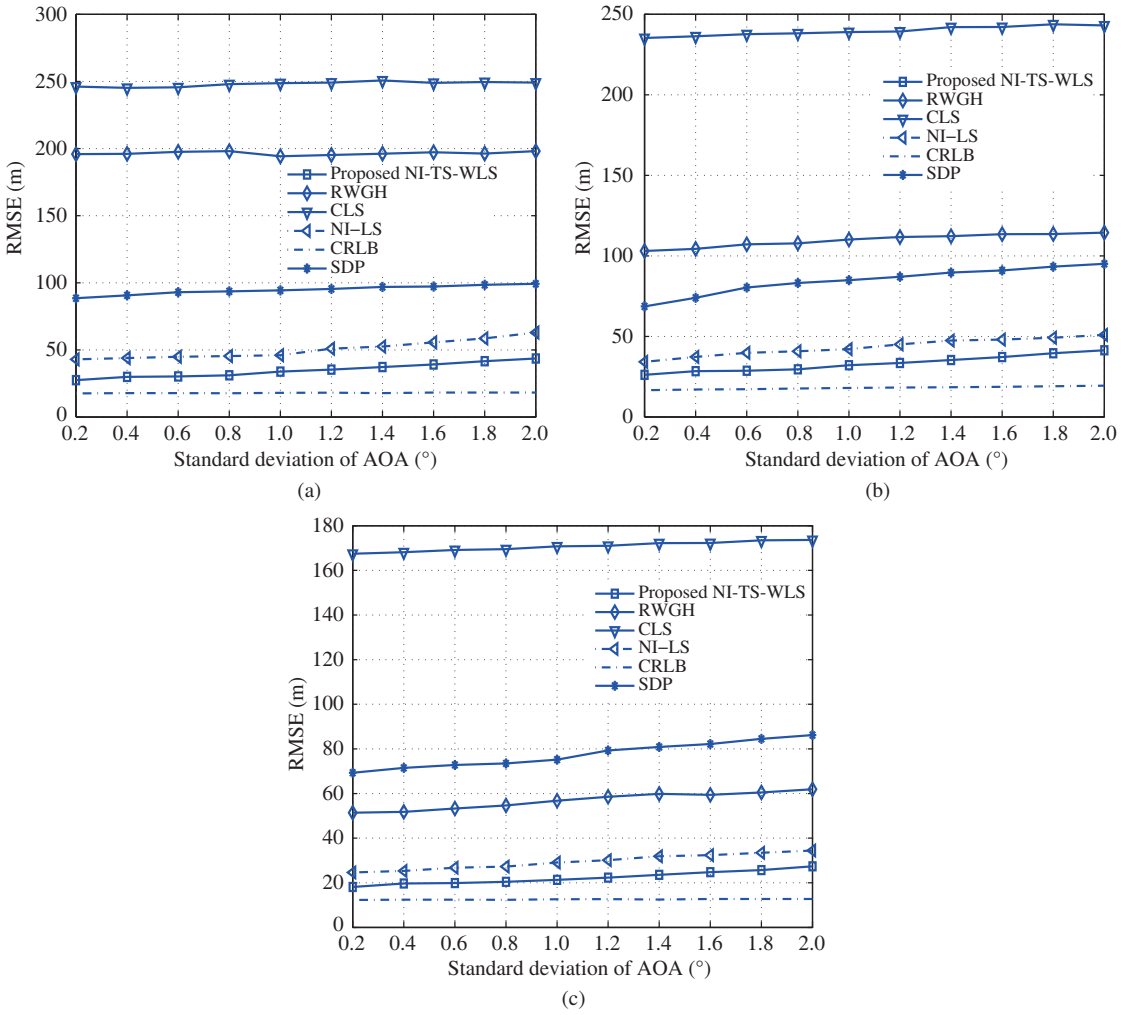

Figure 5 (Color online) RMSE performance comparison for different SDA: SDR = 10 m. (a) Case 5/2; (b) case 5/3; (c) case 5/4.

in case 5/2, the positioning error of the SDP and NI-LS algorithms is 98 and 49 m, respectively, while it is 33 m for the proposed algorithm. Moreover, the NI-LS method experiences a clear RMSE increase in case 5/2 because of a low detection probability. Furthermore, with an increasing number of LOS measurements, the curve of NI-TS-WLS tends to approach the CRLB; this signifies that the proposed NLOS identification algorithm performs well. Few mistakes occur in the proposed NLOS identification algorithm even with only two LOS BS measurements. Nevertheless, a large SDR leads to lower localization accuracy for all algorithms, which may have two causes. First, the increase in SDR leads to more difficult NLOS identification. Second, even if identification is successful, the larger SDR degrades the accuracy of the algorithm.

In addition to the SDR study, we investigated of the SDA, as shown in Figure 5. From this figure, it can be seen that the SDA has a small effect on all algorithms. The CLS method has the poorest performance, while the RWGH algorithm is significantly improved for a larger number of LOS BSs. Of the tested methods, the SDP algorithm is relatively stable; however, its performance is poorer than that of the NI-LS algorithm and proposed NI-TS-WLS algorithm. The proposed NI-TS-WLS algorithm has the best performance of all tested algorithms, e.g., when there are two LOS BSs and SDR = 10 m and SDA = 1, the RMSE of the proposed algorithm is 50 m, which is about 5% of the cell radius. Thus, it can be concluded that the proposed NLOS identification performs well. Note that few mistakes occur in the proposed NLOS identification even for only two LOS BS measurements (SDA $\leqslant 1^\circ$).

We present a more challenging environment to test our proposed algorithm, where we must determine

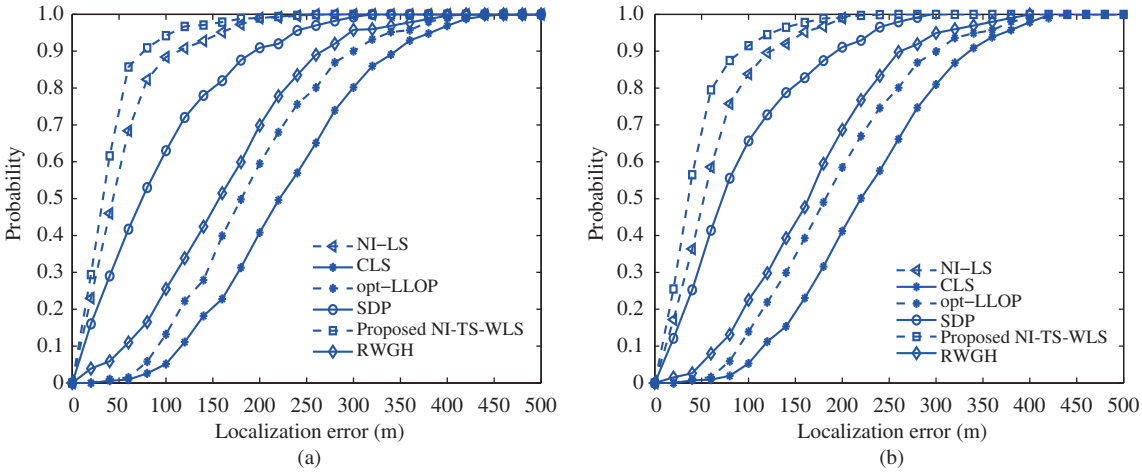

Figure 6 (Color online) Cumulated distribution function (CDF) of tested algorithms under harsh environments for two LOS BSs. (a) SDR = 40 m, SDA = 1°; (b) SDR = 40 m, SDA = 2°

that the NLOS error is uniformly distributed within 200–700 m. From this, it can be seen that the proposed NI-TS-WLS algorithm results in the best performance, e.g., as illustrated in Figure 6(b), the CDF is about 0.9 when the positioning error is 90 m, which is about 9% of the cell radius. Because real-world measurements may not be as harsh as those used in the previous figures, this error ratio will be further reduced.

5 Conclusion

In this study, we present two investigations. First, we propose a novel NLOS identification algorithm that uses a novel ranging residual, and we then exploit a semi-analytical method to determine the detection threshold. Second, we use a two-step WLS algorithm to estimate the MS position using only the identified LOS measurements. Simulation results indicate that the proposed algorithm performs well in identifying NLOS measurements; in particular, the probability of successful detection is greater than 93%. Moreover, the proposed NLOS identification algorithm relaxes the requirements on the number of LOS measurements, which is more suitable for real-world scenarios. Moreover, the simulation results demonstrate that the proposed NI-TS-WLS algorithm performs better than conventional methods and has superior localization accuracy.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant No. 61471322) and Open Project of Zhejiang Provincial Key Laboratory of Information Processing, Communication and Networking, Zhejiang, China.

References

1 You X H, Zhang C, Tan X S, et al. AI for 5G: research directions and paradigms. Sci China Inf Sci, 2019, 62: 021301
2 Ji X S, Huang K Z, Jin L, et al. Overview of 5G security technology. Sci China Inf Sci, 2018, 61: 081301
3 Li L M, Wang D M, Niu X K, et al. mmWave communications for 5G: implementation challenges and advances. Sci China Inf Sci, 2018, 61: 021301
4 Liang K Q, Huang Z, He J Z. A passive localization method of single satellite using TOA sequence. In: Proceedings of IEEE International Conference on Computer and Communications (ICCC), Chengdu, 2016. 1795–1798
5 Takeno R, Seki Y, Sano M, et al. A route navigation system for reducing risk of traffic accidents. In: Proceedings of Global Conference on Consumer Electronics, Kyoto, 2016. 1–5
6 Rezaifard E, Abbasi P. Inertial navigation system calibration using GPS based on extended Kalman filter. In: Proceedings of 2017 Iranian Conference on Electrical Engineering (ICEE), Tehran, 2017. 778–782
7 Kim E, Kim K. Distance estimation with weighted least squares for mobile beacon-based localization in wireless sensor networks. IEEE Signal Process Lett, 2010, 17: 559–562
8 Etemadyrad N, Nelson J K. A sequential detection approach to indoor positioning using RSS-based fingerprinting. In: Proceedings of IEEE Global Conference on Signal and Information Processing (GlobalSIP), Washington, 2016. 1127–1131

9 Wang W, Wang G, Zhang F, et al. Second-order cone relaxation for TDOA-based localization under mixed LOS/NLOS conditions. IEEE Signal Process Lett, 2016, 23: 1872–1876
10 Chen J Y, Zhao Y S, Zhao C, et al. Improved two-step weighted least squares algorithm for TDOA-based source localization. In: Proceedings of 2018 19th International Radar Symposium (IRS), Bonn, 2018. 1–6
11 Wang D, Zhang L, Wu Y. The structured total least squares algorithm research for passive location based on angle information. Sci China Ser F-Inf Sci, 2009, 52: 1043–1054
12 Shu F, Qin Y, Liu T, et al. Low-complexity and high-resolution DOA estimation for hybrid analog and digital massive MIMO receive array. IEEE Trans Commun, 2018, 66: 2487–2501
13 Yang T C, Jin L. Single station location method in NLOS environment: the circle fitting algorithm. Sci China Inf Sci, 2011, 54: 381–385
14 Xiao Z, Hei Y Q, Yu Q, et al. A survey on impulse-radio UWB localization. Sci China Inf Sci, 2010, 53: 1322–1335
15 Xie Y, Wang Y, Wu B, et al. Localization by hybrid TOA, AOA and DSF estimation in NLOS environments. In: Proceedings of IEEE Vehicular Technology Conference - Fall, Ottawa, 2010. 1–5
16 Li Y, Williams S, Moran B, et al. High-dimensional probabilistic fingerprinting in wireless sensor networks based on a multivariate Gaussian mixture model. Sensors, 2018, 18: 2602
17 Yan J, Wu L, Zhu W. A base station identification algorithm for SFN positioning systems in NLOS environment. In: Proceedings of 2013 IEEE 78th Vehicular Technology Conference (VTC Fall), Las Vegas, 2013. 1–5
18 Jian G, Zhang L. A time-difference-of-arrival location method based on residual sorting and weighted matrix amended. In: Proceedings of International Conference on Wireless Communications and Signal Processing (WCSP), Nanjing, 2011. 1–5
19 Shu F, Yang S P, Lu J H, et al. On impact of earth constraint on TDOA-based localization performance in passive multisatellite localization systems. IEEE Syst J, 2018, 12: 3861–3864
20 Chen X, Tian Y G, Dong C Z, et al. An improved state space approach based method for extracting the target scattering center. In: Proceedings of 2017 International Applied Computational Electromagnetics Society Symposium (ACES), Suzhou, 2017. 1–3
21 Zhaounia M, Landolsi M A, Bouallegue R. A novel scattering distance-based mobile positioning algorithm. In: Proceedings of 2009 Global Information Infrastructure Symposium, Hammemet, 2009. 1–4
22 Qu Q Y, Guo K Y, Sheng X Q. Scattering centers induced by creeping waves on streamlined cone-shaped targets in bistatic mode. Antenn Wirel Propag Lett, 2015, 14: 462–465
23 Papakonstantinou K, Slock D. Hybrid TOA/AOD/Doppler-shift localization algorithm for NLOS environments. In: Proceedings of 2009 IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications, Tokyo, 2009. 1948–1952
24 Hasan S, Ukkusuri S V. Reconstructing activity location sequences from incomplete check-in data: a semi-Markov continuous-time Bayesian network model. IEEE Trans Intell Transp Syst, 2018, 19: 687–698
25 Mohammad F, Seyed A G, Reza S. A novel maximum likelihood based estimator for bearing-only target localization. In: Proceedings of 2014 22nd Iranian Conference on Electrical Engineering (ICEE), Tehran, 2014. 1522–1527
26 Zhang T, Yan X. Locating steganographic payloads by combining weighed-stego residuals and maximum a posteriori probability cover estimator. In: Proceedings of 2014 11th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), Xiamen, 2014. 798–802
27 Marano S, Gifford W M, Wymeersch H, et al. NLOS identification and mitigation for localization based on UWB experimental data. IEEE J Sel Areas Commun, 2010, 28: 1026–1035
28 Zhang J, Salmi J, Lohan E S. Analysis of kurtosis-based LOS/NLOS identification using indoor MIMO channel measurement. IEEE Trans Veh Technol, 2013, 62: 2871–2874
29 Yu K, Dutkiewicz E. NLOS identification and mitigation for mobile tracking. IEEE Trans Aerosp Electron Syst, 2013, 49: 1438–1452
30 Shu F, Yang S P, Qin Y L, et al. Approximate analytic quadratic-optimization solution for TDOA-based passive multi-satellite localization with earth constraint. IEEE Access, 2016, 4: 9283–9292
31 Hua J Y, Yin Y J, Lu W D, et al. NLOS identification and positioning algorithm based on localization residual in wireless sensor networks. Sensors, 2018, 18: 2991
32 Chan Y T, Tsui W Y, So H C, et al. Time-of-arrival based localization under NLOS conditions. IEEE Trans Veh Technol, 2006, 55: 17–24
33 Gerok W, El-Hadidy M, El Din S A, et al. Influence of the real UWB antennas on the AoA estimation based on the TDoA localization technique. In: Proceedings of IEEE Middle East Conference on Antennas and Propagation (MECAP 2010), Cairo, 2010. 1–6
34 Wang W, Wang G, Zhang J, et al. Robust weighted least squares method for TOA-based localization under mixed LOS/NLOS conditions. IEEE Commun Lett, 2017, 21: 2226–2229
35 Hua Z, Hang L, Yue L, et al. Geometrical constrained least squares estimation in wireless location systems. In: Proceedings of the 2014 4th IEEE Network Infrastructure and Digital Content (IC-NIDC), Beijing, 2014. 159–163
36 Zheng X, Hua J, Zheng Z, et al. LLOP localization algorithm with optimal scaling in NLOS wireless propagations. In: Proceedings of the 2013 IEEE 4th International Conference on Electronics Information and Emergency Communication (ICEIEC), Beijing, 2013. 45–48
37 Cheung K W, So H C, Ma W K, et al. A constrained least squares approach to mobile positioning: algorithms and optimality. EURASIP J Adv Signal Process, 2006, 2006: 020858