# RESEARCH

# EURASIP Journal on Wireless Communications and Networking

# **Open Access**

# A positioning algorithm based on improved robust extended Kalman filter with NLOS identification and mitigation



Yan Wang<sup>1\*</sup>, Huikang Yang<sup>1</sup> and Yuxin Gong<sup>1</sup>

\*Correspondence: wangyan\_jgxy@neuq.edu.cn

<sup>1</sup> Department of Computer and Communication Engineering, Northeastern University, Qinhuangdao 066004, Hebei Province, China

# Abstract

With the development of the information age and the maturity of Internet of Things technology, wireless sensor network has been widely applied in indoor localization. However, the non-line-of-sight (NLOS) propagation in complicated environment and the inherent noise of the sensor will introduce errors in the measurements, which will seriously lead to inaccurate positioning. In this paper, a novel localization scheme based on the mean reconstruction method is proposed, which reconstructs the distance measurements from all beacon nodes by taking the average twice to weaken the adverse effects of NLOS. At the same time, the noise average is re-estimated when the distance difference is not too large. Next, the robust extended Kalman filter (REKF) is used to process the reconstructed distance measurements to obtain positioning results. To make the positioning results more accurate, hypothesis test is used as NLOS identification to classify the position estimates generated from all distance combinations by least-squares. Then, the residual weighting (RWGH) method is utilized to combine the position estimates that fall into the validation region. At last, we merge the results from RWGH and REKF. The simulation and experimental results show that the proposed algorithm has high positioning accuracy and strong positioning robustness.

**Keywords:** Wireless sensor network, Mean reconstruction, Robust extended Kalman filter, NLOS identification, Residual weighting

# **1** Introduction

A certain number of low-energy sensors are deployed in a specific range indoors or outdoors, thus forming a wireless sensor network (WSN). Using communication technology to receive relevant information about the target, such as time of arrival (TOA), time difference of arrival (TDOA), angle of arrival (AOA) and received signal strength (RSS). The researchers have extended wireless positioning technology, which has good effects on science, military and social life [1]. In WSN, the pre-fixed sensors are called beacon nodes (BNs), while the state of the object to be located is indeterminate, usually in motion, so we call it mobile node (MN). According to whether there are obstacles between the BN and the MN, the channel can be divided into two types: line-of-sight



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http:// creativeCommons.org/licenses/by/4.0/.

(LOS) and non-line-of-sight (NLOS) propagation. In the real world, obstacles such as humans, tables, cars, and plants will cause diffraction and refraction of the electric signal during the NLOS transmission process and produce the signal time delay, which make the measurements become larger than actual values [2]. Thus eventually lead to the deviation between the positioning trajectory and the actual mobile trajectory [3]. The errors of contaminated measurements in NLOS transmission are mainly NLOS errors, which make the measurements have positive errors that can't be ignored. The NLOS errors are adverse to the positioning accuracy, so it is especially crucial to alleviate and eliminate the NLOS errors.

Nowadays, several NLOS processing algorithms have been proposed. Some algorithms directly process all original measurement data [4–9]; while others identify the measurement data and divide them into several categories [10–16], even discarding the measurement values with larger errors and only utilizing the measurement values with smaller errors into subsequent steps. Therefore, we can divide these algorithms into NLOS mitigation and NLOS identification.

In this paper, we propose a positioning algorithm based on mean reconstruction REKF by using the TOA measurements, which combines NLOS mitigation and NLOS identification. The proposed algorithm has made some improvements in positioning performance and algorithm applicability. The main contributions in the proposed algorithm as follows:

- The proposed algorithm is based on REKF, which integrates robust regression into EKF and can be applied to nonlinear systems. Unlike the IMM framework, the REKF has a significant advantage that does not depend on the prior knowledge of the NLOS errors, so REKF is more in line with the actual circumstances.
- 2. In the worst case, our algorithm time complexity is  $(r + K) \cdot O(n)$ , where *r* is the number of Newton–Raphson iterations and *K* is the total number of distance combinations.
- 3. The proposed algorithm has a significant advantage in localization performance with high accuracy and strong robustness. In addition, the proposed algorithm is suitable for different NLOS error distributions and independent of BNs distribution. Particularly, the proposed algorithm works better for the environments with lots of NLOS errors.

The subsequent structure of the paper is arranged as follows: Sect. 2 introduces related works. Sections 3 and 4 describe the signal model and the proposed algorithm in detail respectively. We conducted simulations and experiments, and analyzed the results in Sect. 5. The conclusion is drawn in Sect. 6.

# 2 Related works

#### 2.1 NLOS mitigation

There are two strategies for NLOS mitigation: the first one is only to remedy range errors. The residual weighting (RWGH) method is to take the reciprocal of the distance difference as the residual, which removes the effect of outliers by giving less weight to points farther away from the measurements [4]. However, the complexity of

the method needs to be reduced, some researchers were committed to improving the residual weighting algorithm. For instance, by reserving the combinations with the smallest residuals and performing summation, the RWGH method becomes computationally simple [5]. The iterative minimum residual algorithm is proposed in [6], which sets a minimum residual threshold, and iteratively retains position estimates with residuals less than this threshold as final results. In TDOA positioning, solving optimization problems by convex relaxation is also a hot topic for NLOS mitigation. Through l(1)-norm robustification and neurodynamic optimization in NLOS, the convex optimization problems with inequality constraints are solved by redefining Augmented Lagrangian effectively [7].

The second strategy is to directly suppress the NLOS impact by using dedicated localization techniques. The authors in [8] proposes a NLOS nodes localization scheme based on Harris Hawke optimization algorithm to promote reliable data propagation between vehicle nodes in emergency situations. It can locate NLOS nodes based on their dynamic characteristics and adaptive positioning styles obtained through reference nodes. A novel self-learning LS positioning algorithm has been proposed in [9], which can not only solve the distortion problem of LS in NLOS environment, but also not be limited to the condition that LS requires at least three nodes in localization problems.

# 2.2 NLOS identification

The NLOS identification part is generally conducted before NLOS mitigation, which can recognize whether the propagation state between nodes is LOS or NLOS, after identification different error elimination technologies are used for different conditions. The NLOS identification methods include range, channel and location aspects. The identification method based on channel is challenging, and the method based on distance can be measured by UWB, which is relatively easy to obtain, and the position can also be calculated using distance. The method based on distance mainly use the probability density function (PDF) or variance of range estimates to distinguish between LOS and NLOS [10]. In location identification, probability theory [11] or machine learning [12] can be used to determine whether the position estimate is generated by NLOS measurements.

There exist some methods to identify whether the state is in LOS or NLOS. The Wylie method mainly calculates the measurement variance and judges whether the distance measurements are contaminated by NLOS errors [13]. The decision theory and z-test are used for identifying NLOS propagation [14], and then the NLOS measurement data are discarded. Even if the prior information of NLOS error is unknown, the algorithm is still effective. These identification methods are based on range. The modified probabilistic data association (MPDA) algorithm obtains a position estimate through different combinations of ranging by LS [15]. The NLOS detection is to identify the optimized position estimate which falls into the validation gate. Then the optimized position estimates are reserved and used to calculate the position of the MN by the corresponding association (PNI) technique based on support vector machine (SVM) is designed in [16], which can reduce the effects of transmission channel-dependent characteristics and optimize the classifier, so it's based on channel. Our algorithm improves the hypothesis test in the MPDA to determine whether the position estimate is from LOS or NLOS.

# 2.3 NLOS identification and NLOS mitigation

Some positioning algorithms are divided into two steps: first identifying the NLOS error and then eliminating it. All of them play a certain role in improving the positioning accuracy. The Wylie method uses the time history of distance measurements from each BN, combined with the method of measuring noise standard deviation and residual analysis rank test, to determine whether there exists NLOS error in the distance measurements [13]. The second step is to eliminate the NLOS errors to achieve true TOA reconstruction by using the time history of TOA measurements. A distance measurement reconstruction mode was proposed in [17]. After identifying the NLOS error, Liu uses an orthogonal polynomial to optimize the distance measurements obtained by the biased Kalman filter (KF). In [18], an Bayesian sequential test is designed to identify whether the distance measurements are contaminated by NLOS errors. After identification, the original distance measurements, which are contaminated by NLOS errors, are optimized by a modified Kalman filter (MKF). In [19], according to the number of position estimates that fall into the validation gate, the environment at this moment is divided into mild NLOS and severe NLOS condition; then the average value of NLOS error is estimated by using the historical positioning results, and the distance measurements from each beacon node are reconstructed. Our algorithm reconstructs the distance measurements to reduce the NLOS errors and determines whether the position estimate is from LOS or NLOS, and finally combines the results.

#### 3 Signal model

 $X(n) = [x(n) y(n)\hat{x}(n) \hat{y}(n)]^T$  is the state vector of the MN, where *n* is the time step and *N* is the whole time steps. (x(n), y(n)) is the coordinate of the MN and  $(\hat{x}(n) \hat{y}(n))$  denotes the velocity of the MN. According to the force model, the state vector of the MN changes over time, which can be described as:

$$X(n) = FX(n-1) + Gu(n-1), \quad n = 1, 2, ..., N.$$
(1)

$$F = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad G = \begin{bmatrix} \Delta t^2/2 & 0 \\ 0 & \Delta t^2/2 \\ \Delta t & 0 \\ 0 & \Delta t \end{bmatrix},$$
(2)

where,  $\Delta t$  is the sampling period. u(n) is the driving noise, which is assumed to be a Gaussian white noise with a mean of zero and covariance matrix of Q(n). F is the state transition matrix of the MN, and G is the interference input matrix which is used to described the random acceleration of the MN causes by u(n). Within the specific area, L beacon nodes are randomly deployed to receive the signal from the MN. D(n) is the distance measurements based on TOA data between the BNs and MN at time step n, which can be defined as and calculated by follows:

$$\boldsymbol{D}(n) = [d_1(n), d_2(n), \cdots, d_L(n)]^T$$
(3)

$$D(n) = h(X(n)) + b(n), \quad n = 1, 2, ..., N$$
 (4)

where,  $\mathbf{h}(\mathbf{X}(n)) = [h_1(x(n)), h_2(x(n)), ..., h_L(x(n))]^T$  represents the actual distances between the BNs and the MN at time step *n* and  $\mathbf{b}(n) = [b_1(n), b_2(n), ..., b_l(n), ..., b_L(n)]^T$ is the noise vector. The actual distances  $\mathbf{h}(\mathbf{X}(n))$  are measured by Euclidean distance and  $h_l(\mathbf{X}(n))$  is the *l*-th component of  $\mathbf{h}(\mathbf{X}(n))$ .

$$h_l(X(n)) = \sqrt{(x(n) - x_B(l))^2 + (y(n) - y_B(l))^2}, \quad l = 1, \dots, L$$
(5)

where,  $(x_B(l), y_B(l))^T$  denotes the coordinate of the *l*-th BN.

The noise vector  $\boldsymbol{b}(n)$  consists of sensor noise and NLOS error, and the sensor noise is also called LOS error.  $\boldsymbol{b}(n)$  contains random variables with a variance describing Gaussian noise due to the NLOS propagation, and it can be defined as:

$$b_l(n) = \begin{cases} \sigma_{LOS} w(n) & \text{LOS} \\ \sqrt{\sigma_{LOS}^2 + \sigma_{NLOS}^2} w(n) + b_{NLOS} & \text{NLOS} \end{cases}$$
(6)

where, w(n) is Gaussian noise  $N(0, 1^2)$ .  $b_{NLOS}$  denotes the NLOS error, which changes with the environment and could follow different distribution.  $\sigma_{LOS}$  and  $\sigma_{NLOS}$  are the standard deviation of the sensor noise and the NLOS error respectively. The measurement covariance matrix is:

$$R(n) = diag\left[\sigma_1^2, \sigma_2^2, \dots, \sigma_L^2\right], \quad l = 1, 2, \dots, L$$

$$\tag{7}$$

where  $\sigma_L^2$  represents the variance of the noise.

In this paper, we assume that the standard deviation of the sensor noise  $\sigma_{LOS}$  is known beforehand, while  $\sigma_{NLOS}$  is unknown for us.

# 4 The proposed algorithm program

#### 4.1 General concept

The flowchart of the proposed algorithm is illustrated in Fig. 1. First, we obtain a predicted position through Kalman prediction with real initial position information. Then we calculate the distances from the predicted position to all beacon nodes and make difference with the original distance measurements. In the first averaging, the distance differences from all beacon nodes are averaged, and the second averaging is the historical time average. The average value of the two times is used as the estimation of noise average, and then the mean reconstruction is performed: we subtract the estimation of noise average from the original distance measurements, and the results are approximated as the distances from all beacon nodes to the mobile node, thereby reducing the error caused by the noisy environment. However, reconstruction may produce a negative error when the noise is small, so we re-estimate the noise. After that, the robust regression technology is introduced to improve the EKF, which does not need to assume the prior knowledge about the NLOS errors. To further enhance the positioning accuracy, we divide the distance measurements into multiple groups, and each group produces a position estimate by least-squares. Then by processing these position estimates with EKF, the filtered position estimates can be obtained. LOS identification is to recognize whether the position estimate is in LOS through hypothesis testing. The position estimates whose test statistic is not greater than the validation threshold can fall into



Fig. 1 The flowchart of the proposed algorithm

the validation region and are selected for residual weighting, and the other position estimates are discarded. Lastly, the result of REKF and the result of residual weighting are merged as the positioning result, which will be input into the next moment positioning.

# 4.2 Mean reconstruction method

In indoor localization, the sensors return observation value D(n), which consisted of the actual distances and the noise b(n) [20]. b(n) is mixed with the LOS error and the NLOS error. The LOS error is generally smaller than NLOS error. So we can use the noise estimation to estimate NLOS error. The NLOS error has a certain rule and follows a specific distribution, which has different values with different probabilities. We take the method of averaging multiple distance measurements differences, the random NLOS errors tend to offset each other in the linear dynamic system in the process of averaging [21].

Assuming that the initial state information is known and the movement is entered from the door of the room, the position of the door can be used as the initial position. Based on this accurate data, the Kalman prediction position at the first moment is close to the real position. We calculate the distances from the predicted position to all beacon nodes, and the differences can be considered as the noise. The average of the differences from all beacon node can be approximated as the mean value of noise. On the basis of the above operations, the Kalman filter values of the subsequent time steps are also relatively accurate. Therefore, we can estimate the noise and average the differences in subsequent time steps.

#### 4.2.1 Kalman prediction

The state vector X(0|0) and covariance matrix P(0|0) are initialized. The state vector and covariance matrix of current time step n can be predicted by the state and the covariance matrix at previous time step n - 1:

$$X(n|n-1) = FX(n-1|n-1)$$
(8)

$$\boldsymbol{P}(n|n-1) = \boldsymbol{F}\boldsymbol{P}(n-1|n-1)\boldsymbol{F}^{T} + \boldsymbol{G}\boldsymbol{Q}\boldsymbol{G}^{T}$$
(9)

 $D_{e,l}(n)$  is the distance between the predicted position and the *l*-th BN. We subtract  $D_{e,l}(n)$  from the distance measurements  $d_l(n)$  to get  $d_{e,l}(n)$ ,  $d_{e,l}(n)$  is the noise estimate of the *l*-th BN.

$$D_{e,l}(n) = \left\| BX(n|n-1) - [x_B(l), y_B(l)] \right\|_2$$
(10)

$$d_{e,l}(n) = \left\| d_l(n) - D_{e,l}(n) \right\|_1$$
(11)

where, **B** is  $\begin{bmatrix} 1000\\ 0100 \end{bmatrix}$ , and  $\|\bullet\|_n$  denotes *n*-norm.

#### 4.2.2 First average

To overcome the randomness of the noise,  $d_{e,l}(n)$  from the *L* BNs are averaged to obtain M(n), which can be approximately considered of the noise at time step *n*.

$$A(n) = \sum_{l=1}^{L} d_{e,l}(n)$$
(12)

$$M(n) = A(n)/L \tag{13}$$

#### 4.2.3 Second average

It is assumed that the NLOS errors follow the same distribution in a separate indoor environment with a small range and similar obstacles. So we perform the second average, historical data from initial time step to current time step n are averaged as the estimation of noise average:

$$m(n) = \sum_{i=1}^{n} M(i)/n$$
(14)

With the process of sampling, the average value can represent the digital characteristics of the noise in this environment.

#### 4.2.4 Mean reconstruction

We mainly let the distance measurements directly subtract the noise average estimation. The noise is mixed with the sensor noise and the NLOS error, and the NLOS error is generally larger than sensor noise. Noise average estimation is mainly determined by the large NLOS error, therefore the distance measurement in NLOS minus the estimated mean value m(n) will be closer to the actual distance. To mitigate the influence of NLOS error, we reconstruct the distance measurements by subtracting the recorded mean value:

$$d_l(n) = d_l(n) - m(n) \tag{15}$$

However, the distance measurement in LOS with a little noise error, subtracting the estimated mean value m(n), may bring about a negative error. If the following conditions are met, it can be roughly considered as a little noise error.

$$d_{e,l}(n) < \sigma_{LOS}, d_{e,l}(n) < M(n), \quad l = 1, \dots, L$$
 (16)

For this case, similar to formula (12) and (13), re-estimate the noise to reduce the influence of negative error and make the reconstructed distances close to the actual distances:

$$\tilde{A}(n) = \sum_{l=1}^{T} d_{e,l}(n)$$
(17)

$$\tilde{M}(n) = \tilde{A}(n)/T \tag{18}$$

where, *T* is the number that  $d_{e,l}(n)$  is judged to be a little error. For little noise error, we reconstruct the distance measurements by subtracting the re-estimated value of the noise:

$$d_l(n) = d_l(n) - \tilde{M}(n) \tag{19}$$

#### 4.3 REKF algorithm

# 4.3.1 Deformation of the EKF formula

In the EKF, the Jacobian matrix H(n), the innovation v(n), the innovation covariance matrix S(n) and the extended Kalman gain K(n) at time step n can be calculated as follows:

$$H(n) = \frac{\partial h(X(n|n-1))}{\partial X(n|n-1)}$$
(20)

$$\boldsymbol{\nu}(n) = \boldsymbol{D}(n) - \boldsymbol{h}(\boldsymbol{X}(n|n-1)) \tag{21}$$

$$S(n) = H(n)P(n|n-1)HT(n) + R(n)$$
<sup>(22)</sup>

$$\boldsymbol{K}(n) = \boldsymbol{P}(n|n-1)\boldsymbol{H}^{T}(\boldsymbol{S}(n))^{-1}$$
(23)

where D(n) is the reconstructed distance measurements. The update of the state and covariance matrix are:

$$X(n|n) = X(n|n-1) + K(n)v(n)$$
(24)

$$\boldsymbol{P}(n|n) = E\left[ (\boldsymbol{X}(n) - \boldsymbol{X}(n|n)) (\boldsymbol{X}(n) - \boldsymbol{X}(n|n))^T \right]$$
(25)

In the next steps, we replace X(n) with X. The EKF equation is rewritten and transformed into linear regression model by applying the robust technology. The state Eq. (1) and the measurement (4) are rewritten as follows:

$$\begin{bmatrix} I_4 \\ H(n) \end{bmatrix} X(n) = \begin{bmatrix} FX(n-1|n-1) \\ D(n) - h(X(n|n-1)) + H(n)X(n|n-1) \end{bmatrix} + e(n)$$
(26)

$$\boldsymbol{e}(n) = \begin{bmatrix} \boldsymbol{F}(\boldsymbol{X}(n-1) - \boldsymbol{X}(n-1|n-1)) + \boldsymbol{G}\boldsymbol{w}(n-1) \\ -\boldsymbol{v}(n) \end{bmatrix}$$
(27)

$$E(\boldsymbol{e}(n)\boldsymbol{e}^{T}(n)) = \begin{bmatrix} \boldsymbol{P}(n|n-1) & 0\\ 0 & \boldsymbol{R}(n) \end{bmatrix} = \boldsymbol{C}(n)\boldsymbol{C}^{T}(n)$$
(28)

where,  $E(\cdot)$  represents the expectation and C(n) is obtained by using Cholesky decomposition for  $E(e(n)e^{T}(n))$ . The linear regression model is constructed by multiplying (26) with  $C^{-1}(n)$ :

$$\tilde{y} = AX + bf(\tilde{\nu}) + \tilde{\nu} \tag{29}$$

$$\tilde{\boldsymbol{y}} = \boldsymbol{C}^{-1}(n) \begin{bmatrix} \boldsymbol{X}(n|n-1) \\ \boldsymbol{D}(n) - \boldsymbol{h}(\boldsymbol{X}(n|n-1)) + \boldsymbol{H}(n)\boldsymbol{X}(n|n-1) \end{bmatrix}$$
(30)

$$\boldsymbol{A} = \boldsymbol{C}^{-1}(n) \begin{bmatrix} \boldsymbol{I}_4 \\ \boldsymbol{H}(n) \end{bmatrix}$$
(31)

$$\tilde{\boldsymbol{\nu}} = -\boldsymbol{C}^{-1}(n)\boldsymbol{e}(n) \tag{32}$$

# 4.3.2 Robust regression

Referring to the linear regression model (29), the maximum likelihood estimation of X at time step n can be obtained by solving the following coupled equation:

$$\sum_{i=1}^{M+\dim(X)} [\mathbf{A}]_{ij} \times \varphi\left(\tilde{\mathbf{y}}_i - \sum_{j'}^{\dim(X)} [\mathbf{A}]_{ij'} X_{j'}\right) = 0$$
(33)

where, dim(·) is the operation to count the dimension of *X* and j' = 1, 2, ..., dim(X). However, the pdf  $f(\tilde{v})$  of the  $\tilde{v}$  is asymmetric because the mean of the NLOS error is positive. Therefore, we use the descending score function  $\varphi(\cdot)$ , which is referred to [22]:

$$\varphi(\theta) = \begin{cases} \theta & |\theta| < \alpha_1 \\ \beta_1 \tanh[0.5b_1(\alpha_2 - |\theta|)] \operatorname{sgn}(\theta) & \alpha_1 \le |\theta| \le \alpha_2 \\ 0 & |\theta| > \alpha_2 \end{cases}$$
(34)

where,  $\beta_1$  is chosen to be 1.739; according to [22], we set  $\alpha_1$  and  $\alpha_2$  as 1.5 and 3 respectively. The state estimate *X* obtained above still contains a large error, so we use Newton–Raphson method to optimize *X*:

At first, we need to obtain the initial state estimate  $X_0$  by least squares (LS)s:

$$X_r = \left(A^T A\right)^{-1} A^T \tilde{y} \quad r = 0 \tag{35}$$

where,  $X_r$  is the state estimate in *r*-th iteration.

Next, we determine the error residual  $\hat{v}$  through deformation by shifting (29) and estimate the error scale  $\sigma_v$  by (37):

$$\tilde{\nu} = \tilde{y} - AX_r \tag{36}$$

$$\sigma_{\nu} = 1.483 mean \left\{ \left| \hat{\tilde{\nu}} - mean(\hat{\tilde{\nu}}) \right| \right\}$$
(37)

Then, the update step of the robust Kalman filter is performed, in which the state estimate  $X_r$  is updated iteratively by Newton–Raphson method. The iteration will continue until the preset convergence is met, otherwise the iteration will be broken, and the final result will be reserved.

$$X_{r+1} = X_r + \mu \left( A^T A \right)^{-1} A^T \varphi(\hat{\tilde{\nu}} / \sigma_{\nu})$$
(38)

$$\mu = 1 \left/ \left( 1.25 \max\left( \left| \varphi'(\hat{\tilde{\nu}}/\sigma_{\nu}) \right| \right) \right)$$
(39)

$$\|X_{r+1} - X_r\|_2 < \varepsilon \tag{40}$$

where, the preset threshold  $\varepsilon$  is set to 0.5.

At last,  $X_{r+1}$  is determined as the state vector X at this time step n and the two-dimension coordinates of the position estimates Z(n) can be obtained by X multiplying the matrix B:

$$Z(n) = BX(n) \tag{41}$$

#### 4.4 LOS identification strategy

If the noise is small, the reconstruction has little change to the original distance measurements and the reconstruction distance measurements have limited improvement to the positioning accuracy. The enhancement of the accuracy should be improved from the positioning method. We use multiple position estimates generated by distance measurements to eliminate outliers, and retain the position estimates under the condition of LOS with small errors.

We firstly divide the reconstructed distance measurements into *K* combinations and each combination generate a corresponding position estimates via least-squares. Then by processing these position estimates with extended Kalman filter (EKF), filtered positions  $\hat{Z}_i(n)$  can be obtained which is expressed as follow:

$$\widehat{\boldsymbol{Z}}_{i}(n) = \left[\widehat{\boldsymbol{x}}_{i}(n), \widehat{\boldsymbol{y}}_{i}(n)\right]^{T}, i = 1, \dots, K$$
(42)

where

$$K = \sum_{i=3}^{L} C_L^i \tag{43}$$

Next, we obtain the coordinate Z(n|n-1) of the predicted position and calculate the innovation  $\hat{v}_i(n)$  between the position estimate of the *i*-th combination and the predicted position:

$$\boldsymbol{Z}(n|n-1) = \boldsymbol{B}\boldsymbol{X}(n|n-1) \tag{44}$$

$$\boldsymbol{a}_{i}(n) = \boldsymbol{Z}_{i}(n) - \boldsymbol{Z}(n|n-1)$$
(45)

LOS identification: If the NLOS error in position estimate of each combination has been mostly eliminated after the distance measurements reconstruction, the condition of the position estimate can be considered in LOS. Furthermore, there are:

$$\boldsymbol{a}_i(n) \sim N(0, \widehat{\boldsymbol{S}}_i(n)), \quad i = 1, \dots, K$$
(46)

where,  $\hat{S}_i(n)$  is the innovation covariance matrix of  $\hat{Z}_i(n)$ .

$$\widehat{\boldsymbol{S}}_{i}(n) = \boldsymbol{B}\boldsymbol{P}(n|n-1)\boldsymbol{B}^{T} + \sigma_{G}^{2}(\widehat{\boldsymbol{H}}_{i}^{T}(n)\widehat{\boldsymbol{H}}_{i}(n))^{-1}$$
(47)

$$\widehat{H}_{i}(n) = \frac{\partial h(Z_{i}(n))}{\partial \widehat{Z}_{i}(n)}$$
(48)

*K* hypotheses and alternatives are defined to validate (46). The hypothesis  $\zeta_{0,i}$  holds true when the *i*-th position estimate is determined in LOS, otherwise the hypothesis  $\zeta_{1,i}$  holds true [15].

$$\zeta_{0,i}:\boldsymbol{a}_i(n) \sim N\left(0, \widehat{\boldsymbol{S}}_i(n)\right), \quad i = 1, \dots, K$$
(49)

$$\zeta_{1,i}: \operatorname{not} \zeta_{0,i}, \quad i = 1, \dots, K$$
 (50)

$$\boldsymbol{T}_{i}(n) = \boldsymbol{a}_{i}^{T}(n)\widehat{\boldsymbol{S}}_{i}^{-1}(n)\boldsymbol{a}_{i}(n)$$
(51)

where,  $T_i(n)$  is the test statistic, which is used to compare with the threshold value  $\gamma$  of the validation gate to judge whether the position estimate  $\hat{Z}_i(n)$  is in the validation region. If  $T_i(n)$  isn't larger than  $\gamma$ , the hypothesis  $\zeta_{0,i}$  holds; otherwise  $\zeta_{0,i}$  is rejected. The threshold value  $\gamma$  is related to tracking threshold probability  $P_g$ , which represents the probability that the position estimate that falls into the verification area can be correctly detected, and it can be obtained from the following formula:

$$\int_{0}^{\gamma} f_{\chi^{2}(2)}(x) dx = P_{g} = 1 - P_{Fa}$$
(52)

where,  $f_{\chi^2(2)}(\cdot)$  is the probability density function of chi-square distribution, which has two degrees of freedom.  $P_{Fa}$  is the false alarm rate, while the sum of  $P_g$  and  $P_{Fa}$  is 1.

We record the position estimates that passed the hypothesis test and the number of them, which are denoted as  $\overline{Z}_j(n)$  and k(n) respectively.  $\overline{Z}_j(n)$  is expressed as follows:

$$\overline{\mathbf{Z}}_{j}(n) = \left[\overline{x}_{j}(n), \overline{y}_{j}(n)\right]^{T}, \quad j = 1, \dots, k(n)$$
(53)

where, k(n) is a non-negative integer no greater than *K*.

#### 4.5 Residual weighting method

When k(n) is more than zero, the position estimates fall into the validation region in the previous procedure are selected and used for residual weighting in this procedure. While the other position estimates are discarded. It may also occur that k(n) is equal to zero, which indicates that all position estimates do not meet the hypothesis and are not identified in LOS, then we use the original position estimates  $\hat{Z}_i(n)$ .

Residual weighting method is performed to prevent reconstruction failure. Assuming that the previous steps did not introduce large errors, the predicted point is relatively accurate and is near the real position of mobile node. Therefore, we take the predicted position X(n|n-1) as the center, and the definition of residual [4] is modified as follows:

When k(n) = 0:

$$res_i(n) = \left\| BX(n|n-1) - \left[ \widehat{x}_i(n), \widehat{y}_i(n) \right] \right\|_2, \quad i = 1, \dots, K$$
(54)

When k(n) > 0:

$$\operatorname{res}_{j}(n) = \left\| BX(n|n-1) - \left[ \overline{x}_{j}(n), \overline{y}_{j}(n) \right] \right\|_{2}, \quad j = 1, \dots, k(n)$$
(55)

The position estimate which is close to the predicted position is given a larger weight by taking the reciprocal of the corresponding residual. After calculating the residual of each position estimate, the residual weighted results of them are combined as the positioning result of this time step *n*:

When k(n) = 0:

$$W(n) = \frac{\sum_{i=1}^{K} \widehat{Z}_{i}(n)(res_{i}(n))^{-1}}{\sum_{i=1}^{K} (res_{i}(n))^{-1}}$$
(56)

When k(n) > 0:

$$W(n) = \frac{\sum_{j=1}^{k(n)} \overline{Z}_j(n) (res_j(n))^{-1}}{\sum_{j=1}^{k(n)} (res_j(n))^{-1}}$$
(57)

#### 4.6 Results merging

In this subsection, the positioning results of the above two methods are merged by weighting. Due to the distance measurements contaminated by noise, the filtered distance measurements are closer to the actual distances. We compute the distances from the BNs to the positioning results, which are expressed as h(Z(n)) and h(W(n)). The greater the differences between them and the distance measurements D(n) are, the more effective of this method can be considered. Therefore, we utilize the ideas in [4], and revise the definitions:

$$g_1(n) = \boldsymbol{D}(n) - h(\boldsymbol{Z}(n)) \tag{58}$$

$$g_2(n) = \boldsymbol{D}(n) - h(\boldsymbol{W}(n)) \tag{59}$$

Then the weighted sum of the two results at time step *n* can be obtained after computing the weighting coefficient  $\beta_i(n)$ , which is defined as O(n):

$$\beta_i(n) = \frac{g_i(n)}{\sum_{i=1}^2 g_i(n)}$$
(60)

$$\boldsymbol{O}(n) = \beta_1(n)\boldsymbol{Z}(n) + \beta_2(n)\boldsymbol{W}(n)$$
(61)

At the same time, we substitute O(n) into the first and second components of the state vector X(n|n):

$$[X(n|n)(1), X(n|n)(2)] = O(n)$$
(62)

# 5 Simulation and experimental results

# 5.1 Simulation scene and parameter settings

In this subsection, we mainly introduce the simulation scene and various parameters. The simulation was carried out on the MATLAB platform. And the MN had a fixed trajectory in a 80 × 80 m simulation area, where six beacon nodes were randomly distributed in Fig. 2. We sampled 80 times at a sampling interval of 0.5 s. The initial state of the MN was  $[1m78m \text{ 1m/s } 2.5\text{m/s}]^T$  and the initial covariance matrix of MN was  $P(0) = I_4$ . u(n) was the Gaussian white noise  $N(0, 1^2)$ , and  $R(0) = \sigma_l^2 I_6$  was the covariance matrix of measurement noise. In this paper, the threshold probability



Fig. 2 A Monte Carlo simulation scene

was  $P_g = 0.99$ , so the false alarm rate was  $P_{Fa} = 0.01$ . 1000 Monte Carlo simulations were conducted in each environment.

We use root mean square error (RMSE) and cumulative distribution function (CDF) of the localization error to characterize the localization performance of these algorithms, and RMSE is calculated as follows:

$$RMSE = \sqrt{\frac{1}{MC} \frac{1}{N-1} \sum_{k=1}^{MC} \sum_{n=2}^{N} \|O_k(n) - A(n)\|^2}$$
(63)

where, A(n) is the actual location, and  $O_k(n)$  is the estimated position at time step n in the k-th Monte Carlo simulation, which is obtained from (61). MC = 1000 is the total Monte Carlo runs times, and N = 80 denotes the number of whole samples in a Monte Carlo run.

To illustrate the general applicability of the proposed algorithm, we complicate the motion trajectory and set it as a cubic curve  $y = -0.00063x^3 + 40.63$ ,  $x \in [1, 80][1, 80]$ , where the inflection point is A(n)(1) = 40. Figure 2 shows the present route of the MN and the random distribution of six BNs in a Monte Carlo simulation.

The NLOS error may follow different distributions according to the environment, and the distribution of NLOS error may be related to the type and distribution of indoor obstacles. We did simulations when the NLOS error follow Gaussian, uniform and exponential distribution respectively. We mainly compared our algorithm with the conventional positioning algorithm: EKF [23], REKF [24], IMM-EKF [25] and probabilistic data association IMM (PIMM) [26]. The specific positioning performances of these algorithms in simulation are shown in the following section.

Parameter	Symbol	Values
Number of BNs	L	6
NLOS probability	P <sub>NLOS</sub>	0.5
Sensor noise	$N(0, \sigma_{LOS}^2)$	N(0, 1 <sup>2</sup> )
NLOS error	$N(\mu_{ extsf{NLOS}},\sigma_{ extsf{NLOS}}^2)$	$N(5, 4^2)$

 Table 1
 The parameters in Gaussian distribution



Fig. 3 The RMSE comparison of probability P<sub>NLOS</sub> under Gaussian distribution

#### 5.2 Simulation results and discussion

# 5.2.1 Gaussian distribution

On the condition that NLOS error and sensor noise followed Gaussian distribution  $N(\mu_{NLOS}, \sigma_{NLOS}^2)$  and  $N(0, \sigma_{LOS}^2)$  respectively, we set the default parameters as shown in the Table 1. The NLOS error probability is defined as the proportion of NLOS error to the noise composed of LOS error and NLOS error.

Figure 3 demonstrates the change of RMSE under different probabilities of the NLOS error  $P_{NLOS}$  which varied from 0.1 to 1. The RMSE curve of the EKF, REKF and IMM-EKF had the same trend, and the RMSE value of them from largest to smallest were the EKF, REKF and IMM-EKF. The PIMM had the largest RMSE value and poorest positioning performance when  $P_{NLOS}$  was above 0.6. The proposed algorithm had the lowest average value of RMSE, which was 2.1047 m and at least 39.90% lower than that of the other four algorithms. Figure 4 displays the change in positioning performance when  $\mu_{NLOS}$  varied from 2 to 11 m. The RMSE of the EKF, REKF and IMM-EKF had the same uptrend and they were nearly close, but their RMSE value were not small. No matter how many  $\mu_{NLOS}$  was, the proposed algorithm always had the smallest RMSE. The average values of the RMSE were the EKF, the REKF, the IMM-EKF, the PIMM and the proposed algorithm in descending order, and the specific values of them were 4.3656 m, 4.2116 m, 4.1703 m, 4.1438 m and 2.5052 m respectively.



Fig. 4 The RMSE comparison of mean value  $\mu_{\it NLOS}$  under Gaussian distribution



Fig. 5 The RMSE comparison of standard deviation  $\sigma_{NLOS}$  under Gaussian distribution

Figure 5 displays the change in positioning performance when  $\sigma_{NLOS}$  varied from 2 to 11 m. The RMSE of the EKF, REKF and IMM-EKF had the same uptrend, but their RMSE value were not smaller than the other two algorithms. The proposed algorithm had the smallest RMSE among these algorithms. The average values of the RMSE were the EKF, the REKF, the IMM-EKF, the PIMM and the proposed algorithm in descending order, and the specific values of them were 4.9315 m, 4.7162 m, 4.7013 m, 4.5468 m and 2.9892 m respectively. As shown in Fig. 6, the RMSE curve of the proposed algorithm is steeper than other algorithms, which is far from the other curves.



Fig. 6 The RMSE comparison of standard deviation  $\sigma_{NLOS}$  under Gaussian distribution

The 50% localization errors of the proposed algorithm was about 1.771 m, which was at least 11.5% lower than that of the other four algorithms. And the 90% localization errors of the proposed algorithm were about 3.383 m, which was at least 31.81% lower than that of the other algorithms.

With the increase of the probability, mean or standard deviation of the NLOS error, the true value of NLOS error at a certain time will increase, and the interference to the measured value will intensify. The localization errors of the position estimates processed by the location algorithm will be greater. Therefore, in Figs. 3, 4 and 5, the RMSE of each algorithm has an upward trend. Since the proposed algorithm adopted mean reconstruction method, the greater the probability of NLOS error was, the more times NLOS situation occurred, the more the mean value can reflect the actual situation. The proposed algorithm in Fig. 3 had small RMSE and good positioning performance in the case of high probability of NLOS error. When the probability of the NLOS error  $P_{NLOS}$ was small, the NLOS interference was little, so the RMSE was also small. The PIMM algorithm associated multiple position estimates with Kalman predicted position. These position estimates directly affected the results, and they were generated by LS. The effectiveness of the PIMM algorithm largely depends on the number of excellent measurements, so in the case of high NLOS probability, the PIMM performance was not outstanding. The CDF reflects the distribution of overall localization errors, which can be seen from Fig. 6. Most of the localization errors of the proposed algorithm were less than 3 m, and only few outliers appeared in the positioning.

#### 5.2.2 Uniform distribution

On the condition that NLOS error followed uniform distribution  $U(0, \max)$ , we set the default parameters as shown in the Table 2.

We studied the positioning effect of different localization algorithms when the NLOS error probability was increasing. Figure 7 is similar to in Fig. 4, in which the

Parameter	Symbol	Values
Number of BNs	L	6
NLOS probability	P <sub>NLOS</sub>	0.5
Sensor noise	$N(0, \sigma_{LOS}^2)$	$N(0, 1^2)$
NLOS error	U(min, max)	U(0, 10)

Table 2	The parameters	in uniform	distribution
---------	----------------	------------	--------------



Fig. 7 The RMSE comparison of the probability P<sub>NLOS</sub> under uniform distribution

RMSE of the proposed algorithm changed slightly. Especially when  $P_{NLOS}$  was above 0.3, its value was stable at around 2 m. The RMSE of the EKF, REKF and IMM-EKF were rising linearly with  $P_{NLOS}$ . The proposed algorithm had the lowest average value of RMSE, which was 1.9576 m and at least 40.44% lower than that of the other four algorithms. Then, we let the parameter max change from 6 to 15 m. Figure 8 shows that the RMSE were positively correlated to the parameter max. The EKF, REKF and IMM-EKF had the largest RMSE and the largest growth rate. The average values of the RMSE were the EKF, the REKF, the IMM-EKF, the PIMM and the proposed algorithm in descending order, and the specific values of them were 3.5689 m, 3.4480 m, 3.4335 m, 3.3649 m and 2.9892 m respectively. The influence of the parameter L on the positioning effect of different algorithms was shown in Fig. 9. No matter how many BNs there were, our algorithms. The proposed algorithm had the lowest average value of RMSE, which was 2.0197 m and at least 37.36% lower than that of the four other algorithms.

In the case of uniform distribution, the larger the upper limit of NLOS error was, the greater the probability of NLOS error appeared with a larger value, so the interference to positioning became stronger. As shown in the Fig. 8 the RMSE of each algorithm increased with parameter max. The proposed algorithm determined the



Fig. 8 The RMSE comparison of parameter max under uniform distribution



Fig. 9 The RMSE comparison of the number of BNs under uniform distribution

position estimates and discarded the position estimates that did not fall into the verification gate. Therefore, the proposed algorithm was not susceptible to NLOS errors and had good robustness with smallest RMSE curve growth rate. The more the number of beacon nodes were, the more information about the mobile node can be measured, and the more equations in the maximum likelihood location method can be established, so the the position estimates obtained by LS will be more accurate. As shown in the Fig. 9, the more BNs were, the better these algorithms performed.

Parameter	Symbol	Values
Number of BNs	L	6
NLOS probability	P <sub>NLOS</sub>	0.5
Sensor noise	$N(0, \sigma_{LOS}^2)$	$N(0, 1^2)$
NLOS error	$E(1/\lambda)$	E(1/4)

Table 3 The parameters in exponential distribution



Fig. 10 The RMSE under the probability P<sub>NLOS</sub> under exponential distribution

#### 5.2.3 Exponential distribution

On the condition that NLOS error followed exponential distribution  $E(1/\lambda)$ , we set the default parameters as shown in the Table 3.

Figure 10 demonstrates the RMSE values of each algorithm under different probabilities  $P_{NLOS}$ . When  $P_{NLOS}$  was 0.3, the performance of the proposed algorithm was better than the EKF, the REKF and the IMM-EKF, but it was not as good as the PIMM. When the probability  $P_{NLOS}$  was above 0.3, the proposed algorithm performed well in localization, and the PIMM was sub-optimal. The proposed algorithm had the lowest average value of RMSE, which was 2.2132 m and at least 28.46% lower than that of the other four algorithms. Figure 11 demonstrates the RMSE values of each algorithm under different parameters  $\lambda$ . The average values of the RMSE were the EKF, the IMM-EKF, the REKF, the PIMM and the proposed algorithm in descending order, and the specific values of them were 3.1050 m. 2.9599 m, 2.9554 m, 2.7823 m and 2.1185 m respectively. The CDF of localization error of different algorithms are shown in Fig. 12. The 90% localization errors of these algorithms were the EKF, the IMM-EKF, the PIMM and the proposed algorithm in descending order, and the specific values of  $P_{NLOS} = P_{NLOS} =$ 

In the case of uniform distribution, with the increase of parameter  $\lambda$ , the maximum value of NLOS error increases exponentially. Compared with Gaussian and uniform



Fig. 11 The RMSE comparison of parameter  $\lambda$  under exponential distribution



Fig. 12 The CDF of localization error under exponential distribution

distribution of Figs. 4, 5 and 8, as shown in the Fig. 11, the growth rate of RMSE curve is larger and the curve is steeper. Although there are three different NLOS error distributions, the proposed algorithm showed good positioning performance and robustness.

#### 5.3 Experimental site and equipment

The real environment is often complex and changeable, so on the basis of simulation, we conducted two experiments in real indoor environments. Ultra-wide band (UWB) technique is popular for researchers to use in indoor positioning, due to its high accuracy and low energy consumption [27]. In the experiment, the UWB nodes were used



Fig. 13 The physical image of UWB node



Fig. 14 The actual scene of the university classroom

as radio frequency (RF) devices and sensor nodes in WSN, and Fig. 13 displays a UWB node. The MN was connected with the power supply and it was strapped to the walker's shoes; to maintain the same height with MN, the BNs were placed on the ground. At the same time, the BNs transmitted the distance measurements to the laptop through the universal serial bus (USB) cable and the laptop displayed the distance measurements on the screen.

As shown in Figs. 14 and 15, the experimental site was a university classroom with a large number of tables, seats and students, which can be considered as obstacles. Therefore, the measured data obtained from the UWB nodes were easily contaminated by NLOS errors, containing large errors. We randomly placed six UWB nodes and moved at a constant speed of 1.2 m/s, and sampled 30 times in total. We took the corner of the lower left corner as the origin; took the long side of the classroom as the x-axis, and the short side as the y-axis, to establish a two-dimensional coordinate system, and the coordinate of the initial position was  $[0.6m \ 1.2m]^T$ . The motion trajectory was roughly a rectangle. We first moved in the vertical direction, so the horizontal speed was 0m/s. The entire positioning process was sampled 30 times in total, during the sampling time ( $\Delta t = 1s$ ), each beacon node collected more than 20



Fig. 15 The schematic diagram of the university classroom



Fig. 16 The actual scene of the research group office

distance measurements. After sampling, the average value of them was obtained as the input of each localization algorithm.

Figures 16 and 17 show the actual scenario and schematic diagram of the research group office. The difference was that the scope of this scene was smaller, so there were only four BNs, and the track distance was relatively short, so only 15 times were sampled. As shown in Fig. 17, we set an origin at the bottom right of the room, extending the y-axis in the left direction of the plan view and the x-axis in the upper direction of the plan view to establish a coordinate system, and the coordinate of the initial position was [7.2 m 1.8 m]<sup>*T*</sup> (Figs. 18, 19, 20, and 21).



Fig. 17 The schematic diagram of the research group office



Fig. 18 The data before and after the algorithms processing in the university classroom

# 5.4 Experimental results and discussion

To illustrate the filtering effect of the algorithms, the data before and after the algorithm processing were shown in Figs. 18 and 21. Because the proposed algorithm sub-tracted a value during reconstruction, the processed results may be smaller than the



Fig. 19 The trajectories in the university classroom



Fig. 20 The CDF versus the localization error in the university classroom

actual distances, but they were close to the actual distances, so the position estimates obtained were still more accurate than that in other algorithms. The positioned trajectories in the university classroom and the research group office are shown in Figs. 19 and 22, and the track of the proposed algorithm was close to the real track. In addition, the CDF of the localization errors of the proposed algorithm were smaller than other algorithms in Figs. 20 and 23. Because the interference of NLOS errors caused by lots of obstacles in the experimental site were large, the localization performance of the EKF, REKF, IMM-EKF and PIMM were relatively poor. Since the classroom was not as large as the simulation scene, and the noise here was not as complicated as expected, so the final positioning results were slightly different from the simulation.



Fig. 21 The data before and after the algorithms processing in the research group office



Fig. 22 The trajectories in the research group office

The algorithm and data processing were conducted in MATLAB2018b platform of the laptop whose computer system is a 64-bit operating system and the processor is Intel (R) Core (TM) i5-7300HQ CPU @ 2.50 GHz. Tables 4 and 5 show the running time that each algorithm estimated the position at a sampling time in the university classroom and the research group office. In order to ensure the high precision of positioning, it is inevitable to increase the amount of calculation. The sampling frequency of this experiment was set to 1 Hz. Although the proposed algorithm consumed more time, the processing time for each sampling point was still less than one second, which equalled to the interval between single sample. Therefore, the algorithm can achieve real-time tracking and positioning.



Fig. 23 The CDF versus the localization error in the research group office

	Table 4	Running time	of each s	sampling	point in the	university	/ classroom
--	---------	--------------	-----------	----------	--------------	------------	-------------

Algorithm	Running time/s
EKF	0.0132
IMM-EKF	0.0361
REKF	0.0368
PIMM	0.2357
PROPOSED	0.3243

Table 5	Running	time of	each san	npling	point in <sup>.</sup>	the	research	group	office

Algorithm	Running time/s
EKF	0.0168
IMM-EKF	0.0422
REKF	0.0451
PIMM	0.1411
PROPOSED	0.1894

# 6 Conclusion

In this paper, we proposed a mean reconstruction method to improve the REKF, which can detect and mitigate the NLOS errors to enhance the positioning accuracy. The specific operation of the algorithm is to use the Kalman prediction position to calculate the differences from the distance measurements, which are nearly equivalent to the noise. After averaging the data to estimate the mean value of the noise, the distance measurements are reconstructed by subtracting the estimation of the noise average. At the same time, the results of the residual weighting can ensure the effectiveness of the reconstruction method. The algorithm can significantly reduce the NLOS errors, especially in the environment with severe NLOS errors. The simulation

and experimental results both demonstrated that the proposed algorithm had higher localization accuracy with strong robustness, which had a prominent advantage over traditional localization algorithms. While, there are few shortcomings in the proposed algorithm.

- 1. The mean reconstruction method is required to be used in a small range of area where obstacles are similar, and the NLOS errors follow the same distribution.
- 2. The reconstruction may bring negative errors, which make the reconstructed distances smaller than the actual distances, and the position estimates appear on both sides of the actual trajectory.
- 3. The proposed algorithm is more computationally intensive. First, in the RWGH step, the number of combinations selected is larger than other algorithms; second, the Newton–Raphson iteration will also consume a certain amount of computation.

#### Abbreviations

WSN	Wireless sensor network
LOS	Line-of-sight
NLOS	Non-line-of-sight
EKF	Extended Kalman filter
REKF	Robust extended Kalman filter
RWGH	Residual weighting
TOA	Time of arrival
TDOA	Time difference of arrival
AOA	Angle of arrival
RSS	Received signal strength
BNs	Beacon nodes
MN	Mobile node
IMM	Interactive multiple model
LS	Least squares
MPDA	Modified probabilistic data association
PDF	Probability density function
PNI	Practical NLOS identification
PIMM	Probabilistic data association IMM
SVM	Support vector machine
MKF	Modified Kalman filter
RMSE	Root mean square error
CDF	Cumulative distribution function
USB	Universal serial bus
RF	Radio frequency

#### Author contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by YW, HY and YG. The first draft of the manuscript was written by YW and HY, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

#### Funding

This work was supported by the National Natural Science Foundation of China under Grant No.62273083 and No. 61803077; Natural Science Foundation of Hebei Province under Grant No. F2020501012.

#### Availability of data and materials

The data will be available on request.

#### Declarations

#### **Competing interests**

The authors declare that there is no competing interests regarding the publication of this paper.

#### References

- 1. Q. Tang, J. Li, L. Wang, Y. Jia, G. Cui, Multipath imaging for nlos targets behind an L-shaped corner with single-channel UWB radar. IEEE Sens. J. 22(2), 1531–1540 (2022)
- S. Angarano, V. Mazzia, F. Salvetti, G. Fantin, M. Chiaberge, Robust ultra-wide band range error mitigation with deep learning at the edge. Eng. Appl. Artif. Intell. 102, 104278 (2021)
- X. Yang, J. Wang, D. Song, B. Feng, H. Ye, A novel NLOS error compensation method based IMU for UWB indoor positioning system. IEEE Sens. J. 21(9), 11203–11212 (2021)
- 4. P.C. Chen, A non-line-of-sight error mitigation algorithm in location estimation, in Proceedings of the IEEE Wireless Communications Networking Conference (WCNC), New Orleans, LA, USA, 21–24, vol. 1, pp. 316–320 (1999).
- L. Jiao, J. Xing, X. Zhang, LCC-RWGH: a NLOS error mitigation algorithm for localization in wireless sensor network, in Proceedings of the 2007 IEEE International Conference on Control and Automation, Guangzhou, China, pp. 1354–1359 (2007).
- 6. X. Li, An Iterative NLOS Mitigation Algorithm for Location Estimation in Sensor Networks. Mob. Wirel. Commun. (2006).
- W. Xiong, C. Schindelhauer, H.C. So, J. Bordoy, A. Gabbrielli, J. Liang, TDOA-based localization with NLOS mitigation via robust model transformation and neurodynamic optimization. Signal Process. 178, 107774 (2021)
- C.J. Malar, D.M. Priya, S. Janakiraman, Harris Hawk optimization algorithm-based effective localization of non-line-ofsight nodes for reliable data dissemination in vehicular ad hoc networks. Int. J. Commun. Syst. 32(2), 65–78 (2020)
- B. You, X. Li, X. Zhao and Y. Gao, A novel robust algorithm attenuating non-line-of-sight errors in indoor localization, in IEEE International Conference on Communication Software and Networks (2015), pp. 6–11.
- J. Borrás, P. Hatrack, and N.B. Mandayam, Decision theoretic framework for NLOS identification, in Proc. VTC'98, (Ottawa, Canada, 1998), pp. 1583–1587
- 11. S. Wu, S. Zhang, K. Xu, Probability weighting localization algorithm based on NLOS identification in wireless network. Wirel. Commun. Mob. Comput. **2019**, 1 (2019)
- 12. T. Suzuki, Y. Amano, NLOS multipath classification of GNSS signal correlation output using machine learning. Sensors 21, 2503 (2021)
- M.P. Wylie, J. Holtzman, The non-line of sight problem in mobile location estimation, in Proceedings of the IEEE ICUPC-5th International Conference on Universal Personal Communications, Cambridge, MA, USA, pp. 827–831 (1996)
- W. Zhao, R. He, B. Ai, Z. Zhong, H. Zhang, Vehicle localization based on hypothesis test in NLOS scenarios. IEEE Trans. Veh. Technol. 71(2), 2198–2203 (2022)
- U. Hammes, A.M. Zoubir, Robust mobile terminal tracking in NLOS environments based on data association. IEEE Trans. Signal Process. 58(11), 5872–5882 (2010)
- M. Liu, X. Lou, X. Jin, NLOS Identification for Localization Based on the Application of UWB. Wirel. Pers. Commun. 119, 3651–3670 (2021)
- S. Chen, J. Liu, and L. Xue, A robust non-line-of-sight error mitigation method in mobile position location. Adv. Neural Netw. ISNN (2009).
- L. Yan, Y. Lu, Y. Zhang, An improved NLOS identification and mitigation approach for target tracking in wireless sensor networks. IEEE Access 5, 2798–2807 (2017)
- L. Cheng, S. Huang, M. Xue, Y. Bi, A robust localization algorithm based on NLOS identification and classification filtering for wireless sensor network. Sensors 20(22), 6634 (2020)
- J. Zhu, S.S. Kia, Decentralized cooperative localization with LoS and NLoS UWB inter-agent ranging. IEEE Sens. J. 22(6), 5447–5456 (2022)
- C.M. Bishop, in *Pattern Recognition and Machine Learning*, Information Science and Statistics (Springer Science + Business Media, LLC, New York, USA, 2006). pp. 635–636.
- 22. J.R. Collins, Robust estimation of a location parameter in the presence of asymmetry. Ann. Stat. 4, 68–85 (1976)
- 23. W. Cui, B. Li, L. Zhang, Robust mobile location estimation in NLOS environment using GMM, IMM, and EKF. IEEE Syst. J. **13**(3), 3490–3499 (2019)
- U. Hammes, A.M. Zoubir, Robust MT tracking based on M-estimation and interacting multiple model algorithm. IEEE Trans. Signal Process. 59, 3398–3409 (2011)
- B.-S. Chen, C.-Y. Yang, F.-K. Liao, J.-F. Liao, Mobile location estimator in a rough wireless environment using extended Kalman-based IMM and data fusion. IEEE Trans. Veh. Technol. 58(3), 1157–1169 (2009)
- L. Cheng, M. Xue, Z. Liu, Y. Wang, A robust tracking algorithm based on a probability data association for a wireless sensor network. Appl. Sci. 10, 6 (2020)
- M. Ridolfi, J. Fontaine, B.V. Herbruggen, UWB anchor nodes self-calibration in NLOS conditions: a machine learning and adaptive PHY error correction approach. Wirel. Netw. 27, 3007–3023 (2021)

#### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.