BIG DATA MINING AND ANALYTICS
ISSN 2096-0654 01/06 pp155–170
Volume 3, Number 3, September 2020
DOI: 10.26599/BDMA.2020.9020001

Error Data Analytics on RSS Range-Based Localization

Shuhui Yang*, Zimu Yuan, and Wei Li

Abstract: The quality of measurement data is critical to the accuracy of both outdoor and indoor localization methods. Due to the inevitable measurement error, the analytics on the error data is critical to evaluate localization methods and to find the effective ones. For indoor localization, Received Signal Strength (RSS) is a convenient and low-cost measurement that has been adopted in many localization approaches. However, using RSS data for localization needs to solve a fundamental problem, that is, how accurate are these methods? The reason of the low accuracy of the current RSS-based localization methods is the oversimplified analysis on RSS measurement data. In this proposed work, we adopt a generalized measurement model to find optimal estimators whose estimated error is equal to the Cramér-Rao Lower Bound (CRLB). Through mathematical techniques, the key factors that affect the accuracy of RSS-based localization methods are revealed, and the analytics expression that discloses the proportional relationship between the localization accuracy and these factors is derived. The significance of our discovery has two folds: First, we present a general expression for localization error data analytics, which can explain and predict the accuracy of range-based localization algorithms; second, the further study on the general analytics expression and its minimum can be used to optimize current localization algorithms.

Key words: Cramér-Rao Lower Bound (CRLB); error data analytics; generalized least squares; Received Signal Strength (RSS)

1 Introduction

Location information is critical in many mobile device-based applications. With the increasing demands on location assisted services in both indoor and outdoor environments, location positioning approaches have progressed significantly recently. Localization approaches based on various wireless measurements, such as Received Signal Strength (RSS)[1], Angle Of Arrival (AOA)[2], Time Difference Of Arrival (TDOA)[3], Time Of Arrival (TOA)[4], and many others[5], have been proposed. Different from the Global Positioning System (GPS) for outdoor positioning, which uses satellites to provide universal services[6, 7], the performance of indoor positioning systems is sensitive to device deployment and signal collection and measurement. Additionally, hardware and software cost is an issue to consider when developing indoor positioning system. Among the above mentioned localization schemes, RSS-based positioning is a promising approach since the RSS signal measurement is easy to obtain without the extra hardware and software cost. We will be considering RSS-based positioning schemes in our work.

However, the accuracy of those proposed localization methods remains a changeling issue. In order to solve this issue, we need to accomplish two key tasks: The first one is to identify the optimal

• Shuhui Yang is with Department of Mathematics, Statistics, and Computer Science, Purdue University Northwest, Hammond, IN 46323, USA. E-mail: yang246@purdue.edu.
• Zimu Yuan is with the Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100864, China. E-mail: yuanzimu@iie.ac.cn.
• Wei Li is with Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100864, China. E-mail: liwei@ict.ac.cn.
* To whom correspondence should be addressed.
Manuscript received: 2020-01-31; accepted: 2020-02-07

localization algorithm, where its estimation error equals the Cramér-Rao Lower Bound (CRLB); then the second is to perform localization data analytics, through mathematical techniques, finding the analytics expression that describes the exact relationship between the localization error and the measurement error, and also how the measurement error propagates during the calculation. Most of the existing studies[8–11] used some simplified assumptions to analyze the theoretical accuracy of the range-based localization methods. On one hand, in those work, there is a strong assumption that the path loss exponent as well as the transmission power are known. In experiments, they can be obtained simply by calibrations, or some empirical values are used. However, in practical conditions, such an assumption may not hold. On the other hand, most of those work also assumed that the data errors of RSS measurement are Gaussian or Identical and Independent Distribution (IID). However, as our following study shows, these assumptions are not true in realistic environments. Therefore, the optimality of these studies holds only under special and ideal conditions. The error data analytics on these non-optimal algorithms cannot reflect the true accuracy of RSS range-based localization approaches.

In order to study the accuracy of the range-based localization methods, we adopt two key strategies: One is to select a general log-distance path loss range model, and the other is to assume RSS measurement errors being non-IID. Based on these two strategies, we discover that the Generalized Least Squares (GLS) method can be used as an effective and efficient estimator, with its estimation error being equal to CRLB. We then can use mathematical techniques to explore the propagation of the measurement errors in the GLS method. We derive an analytic expression for the localization error, which shows that the accuracy of range-based localization methods majorly depends on the following key factors: the amount of anchors, the relative geometry of the anchors, the measurement ranges, the errors in the measurement data, and the errors in the estimated parameters. In order to know how these factors affect the localization accuracy, we further study the minimum of the localization error, and discover that the final localization error is proportional to the measurement ranges, the errors in the measurement and estimated parameters, and inversely proportional to the square root of the amount of anchors. The major contributions of this paper are as follows: (1) We develop a general error data analytics expression for the GLS method, based on which we can explain the accuracy of the range-based localization algorithms, i.e., what magnitude errors should be expected; (2) the proposed analytics expression as well as its minimum provide insights in the possible directions regarding the optimization of current localization algorithms. The observed gap between the practical and the optimal accuracy implies the remarkable potential for such optimization.

The remainder of this paper is organized as follows. Section 2 gives the problem. Section 3 describes how to use the optimal and practical estimators to solve the range-based localization problem. Section 4 presents the theoretical analytics on the accuracy of the GLS method. Section 5 gives the experimental verification and Section 6 concludes the paper.

2 Problem and Related Work

RSS is defined as the power level of a signal at the receiver in wireless communication and networking. Based on the radio propagation theory, RSS follows the inverse-square law, that is, the quantity of RSS is inversely proportional to the square of the distance between the transmitter and the receiver. Researchers proposed some range models based on the above inverse-square law[12]. Among these models, the most widely used one is called the log-distance path loss model[13,14], which can be described as

$$\bar{P}_i = \alpha - 10\beta \lg \sqrt{(x_i - x_u)^2 + (y_i - y_u)^2} + \epsilon_i \quad (1)$$

where $\bar{P}_i$ indicates the measured RSS in dBm, which is received at the unknown location (x_u, y_u) (the blind node) from the anchor (x_i, y_i). β is the path loss exponent. In a free space, the value of β is 2, which indicates that the power loss is minimal. The signal strength received at 1 m is denoted as α, which can be regarded as an indicator of the transmission power of anchors. ϵ_i is the measurement error in $\bar{P}_i$, which is unknown in practice. Then, the problem of RSS range-based localization can be described as (note that RSS-based localization methods mainly include range-based and range-free schemes[15]. In this paper, we purely focus on the range-based approach.): given m anchors with known positions as $(x_i, y_i), i = 1, 2, \ldots, m$ and RSS measurements $(\bar{P}_1, \bar{P}_2, \ldots, \bar{P}_m)$, how to calculate (x_u, y_u)?

Many range-based localization methods have been designed to solve the above problem[8,10,16]. Using the range model in Eq. (1), systems that are non-linear and

over-determined were developed in Refs. [8, 10, 16]. To solve the above problem, two tasks here are to be accomplished: (1) the linearization of the non-linear system, and (2) the development of an optimal solution. As a matter of fact, several linearization techniques have been proposed, such as the classical Newton method. To avoid the local minima introduced to the system by the Newton method, alternative linearization algorithms[17], multidimensional scaling[18], semidefinite programming[19], second-order cone programming[20], and linearized least squares methods[21–23] were also proposed.

In order to find the most effective solution, however, we need to combine the above linearization techniques with optimal estimators. Two important issues here should be addressed: The first one is how to treat the parameters α and β, i.e., the indicator of the transmission power and the path loss exponent; the second is the statistical assumption on the measurement error ϵ_i. For α and β, some studies[24–26] applied the simply assumption that they are known parameters, assigning them with empirical values, or achieved by calibrations in the experiments. However, as has pointed out by some other studies[27, 28], these two parameters cannot be treated as known constant values in practice. This is because they may be volatile in different indoor environments. Based on the above discussion, we could classify the assumptions on α and β into two types: the simplified range model and the general range model, according to whether α or β is treated as known or not. For ϵ_i, in much existing work, it is basically assumed as Gaussian or IID. In other words, the two different measurement errors ϵ_i and ϵ_j are independent but identical random variables. As our following study shows, however, when the general range model is considered, ϵ_i and ϵ_j cannot be simply considered as identical.

Based on the above discussion, we classify the existing range-based localization methods into four types, as shown in Table 1. For the methods in Types I – III, they either used the simplified range model or assumed the measurement error is IID. When a general range model is adopted, these methods cannot guarantee their optimality. In fact, if the assumptions could not reflect the realistic statistical property of RSS measurements, the claimed optimal algorithms could not guarantee the optimal accuracy. Accordingly, the error analytics under such conditions may not reflect the true accuracy. Even when the measurement error is not significant, it leads to a large bias due to propagation, resulting in large bias. Different from the previous researches, our studies can be classified as Type IV. Specifically, we use the general log-distance path loss range model and identify that the measurement error is non-IID. Based on the assumption in Type IV, we recognize the GLS method as the optimal estimator. Then, we use mathematical techniques to analyze the propagation of the measurement and obtain the theoretical optimal accuracy of RSS range-based localization approach. From the study, we see that the efficient estimators for Types I – III in Table 1 cannot provide optimal solutions when the assumption in Type IV is adopted. Our study further verifies that the methods in Types I – III are special cases in Type IV.

Table 1 Classification of the assumptions.

Error distribution	Log-distance path loss range model	
	Simplified	General
IID	I. Refs. [29–33]	II. Refs. [27, 28, 34, 35]
Non-IID	III. Refs. [22, 23, 36–41]	IV. Our work

3 Location Estimation

3.1 Range-based localization problem

In this subsection, we will explain the range-based localization problem using the the assumption of Type IV in Table 1. Denote the four unknown parameters in Eq. (1) as a vector $\theta = (\alpha, \beta, x_u, y_u)^{\mathrm{T}}$, where T indicates the transpose of a vector or a matrix, then the RSS measurement can be expressed as

$$f_i(\theta) = \alpha - 10\beta \lg \sqrt{(x_i - x_u)^2 + (y_i - y_u)^2} \quad (2)$$

where $f_i(\theta)$ indicates the RSS measurement calculated by θ. Denote all $f_i(\theta)$ as the following vector:

$$F(\theta) = (f_1(\theta), f_2(\theta), \ldots, f_m(\theta))^{\mathrm{T}} \quad (3)$$

and all measured RSS values as the following vector:

$$\bar{P} = (\bar{P}_1, \bar{P}_2, \ldots, \bar{P}_m)^{\mathrm{T}} \quad (4)$$

When θ is the true solution, we actually have

$$\bar{P} - F(\theta) = e \quad (5)$$

where $e = (\epsilon_1, \epsilon_2, \ldots, \epsilon_m)^{\mathrm{T}}$. Since e is unknown, in practice, the following system is used to find a solution for θ:

$$\bar{P} - F(\theta) \approx 0 \quad (6)$$

That is, given a set of measured RSS values ($\bar{P}_1$, $\bar{P}_2, \ldots, \bar{P}_m$), we have the system described as a matrix form:

$$\begin{cases} \bar{P}_1 = \alpha - 10\beta \lg \sqrt{(x_1 - x_u)^2 + (y_1 - y_u)^2}, \\ \bar{P}_2 = \alpha - 10\beta \lg \sqrt{(x_2 - x_u)^2 + (y_2 - y_u)^2}, \\ \quad \vdots \\ \bar{P}_m = \alpha - 10\beta \lg \sqrt{(x_m - x_u)^2 + (y_m - y_u)^2} \end{cases} \quad (7)$$

Usually, the amount of anchors used in Eq. (7) is more than four, i.e., $m \geqslant 4$, there are four unknown parameters. In addition, more measurements could eliminate the effect of the random error too. Since Eq. (2) is non-linear, the system in Eq. (7) is non-linear and over-determined.

Basically, solving the system in Eq. (7) depends on two key operations: linearizing a non-linear system and finding a best estimation of θ for the over-determined system. For the first issue, as discussed in Section 2, several mathematical tools can be used. A typical approach is the Newton method that is selected as the linearization tool in our work (note that in this paper we do not consider the local minimum of the Newton method because our idea is to analyze the theoretical accuracy of range-based localization methods other than their practical performance). For the issue of the best estimation, it is critical to know the statistical property of e, i.e., the covariance matrix Ω of e:

$$\Omega = \mathrm{Cov}(e) = \begin{pmatrix} \sigma_{\epsilon_1}^2 & \sigma_{\epsilon_1\epsilon_2} & \cdots & \sigma_{\epsilon_1\epsilon_m} \\ \sigma_{\epsilon_2\epsilon_1} & \sigma_{\epsilon_2}^2 & \cdots & \sigma_{\epsilon_2\epsilon_m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{\epsilon_m\epsilon_1} & \sigma_{\epsilon_m\epsilon_2} & \cdots & \sigma_{\epsilon_m}^2 \end{pmatrix} \tag{8}$$

where Cov indicates the covariance matrix of a vector; σ_{ϵ_i} indicates the standard variance of the error ϵ_i on the i-th measurement $\bar{P}_i$; and $\sigma_{\epsilon_i\epsilon_j}$ indicates the covariance of the i-th and j-th measurement errors. When the errors in e are independent (i.e., $\sigma_{\epsilon_i\epsilon_j} = 0$) but not identical, the covariance matrix Ω becomes a weighted diagonal matrix:

$$\Omega = \begin{pmatrix} \sigma_{\epsilon_1}^2 & & & \\ & \sigma_{\epsilon_2}^2 & & \\ & & \ddots & \\ & & & \sigma_{\epsilon_m}^2 \end{pmatrix} \tag{9}$$

Moreover, when all errors in e are independent and identical, i.e., for any ϵ_i and ϵ_j, there is $\sigma_{\epsilon_i}^2 = \sigma_{\epsilon_j}^2 = \sigma_r^2$, we have

$$\Omega = I\sigma_r^2 \tag{10}$$

where I is the identical matrix. Note that for different cases of Ω, various optimal estimators can be used. For the case in Eqs. (8)–(10), the generalized least squares method, Weighted Least Squares (WLS) method, and Ordinary Least Squares (OLS) method will be the optimal estimators, respectively. As shown later, the property of Ω determines what optimal estimator should be selected, and the correctness of the error analytics also depends on what estimators are used.

3.2 Non-IID measurement error

As mentioned earlier, the measurement error ϵ_i cannot be simply regarded as IID. In this subsection, we will show that even when the simplified range model is used, the measurement error ϵ_i cannot be seen as IID. Consider an estimation $\bar{\theta} = (\bar{\alpha}, \bar{\beta}, \bar{x}_u, \bar{y}_u)^{\mathrm{T}}$ for the measurement $\bar{P}_i$. Putting $\bar{\theta}$ into Eq. (2), we have

$$\bar{P}_i = f_i(\bar{\theta}) = \bar{\alpha} - 10\bar{\beta}\lg\sqrt{(x_i-\bar{x}_u)^2+(y_i-\bar{y}_u)^2} = P_i + \epsilon_i \tag{11}$$

where P_i is the true RSS measurement. Then we get

$$\sigma_{\epsilon_i}^2 = \mathrm{Var}(\epsilon_i) = \mathrm{Var}(P_i + \epsilon_i) = \mathrm{Var}(f_i(\bar{\theta})) \tag{12}$$

where Var indicates the variance of a random variable or a function. Since the expression of f_i is known, we can use the delta method[42] to calculate the variance of f_i:

$$\mathrm{Var}(f_i(\bar{\theta})) \approx \nabla f_i(\theta)^{\mathrm{T}} \cdot \mathrm{Cov}(\bar{\theta}) \cdot \nabla f_i(\theta) \tag{13}$$

It is easy to know that $\bar{\alpha}$, $\bar{\beta}$, $\bar{x}_u$, and $\bar{y}_u$ are the mean of α, β, x_u, and y_u, respectively. Since we have

$$\begin{cases} \dfrac{\partial f_i(E(\bar{\alpha}))}{\partial\alpha} \approx 1, \\ \dfrac{\partial f_i(E(\bar{\beta}))}{\partial\beta} \approx 10\lg d_i, \\ \dfrac{\partial f_i(E(\bar{x}_u))}{\partial x_u} \approx \dfrac{10\beta(x_i - x_u)}{\ln(10)d_i^2}, \\ \dfrac{\partial f_i(E(\bar{y}_u))}{\partial y_u} \approx \dfrac{10\beta(y_i - y_u)}{\ln(10)d_i^2} \end{cases} \tag{14}$$

where E denotes the mean of a random variable and $d_i = \sqrt{(x_i - x_u)^2 + (y_i - y_u)^2}$. Then we get

$$\nabla f_i(\theta) = \left(1, 10\lg(d_i), \frac{10\beta(x_i - x_u)}{\ln(10)d_i^2}, \frac{10\beta(y_i - y_u)}{\ln(10)d_i^2}\right)^{\mathrm{T}} \tag{15}$$

Denoting the variance of $\overline{\alpha}$, $\overline{\beta}$, $\overline{x_u}$, and $\overline{y_u}$ as σ_α^2, σ_β^2, σ_x^2, and σ_y^2, then we have

$$\mathrm{Cov}(\bar{\theta}) = \begin{pmatrix} \sigma_\alpha^2 & & & \\ & \sigma_\beta^2 & & \\ & & \sigma_x^2 & \\ & & & \sigma_y^2 \end{pmatrix} \tag{16}$$

Putting Eqs. (15) and (16) into Eq. (13) and simplifying, we finally get

$$\sigma_{\epsilon_i}^2 = \sigma_\alpha^2 + (10\lg(d_i))^2\sigma_\beta^2 + \left(\frac{10\beta}{\ln(10)d_i}\right)^2(\sigma_x^2 + \sigma_y^2) \tag{17}$$

From Eq. (17), it is obvious that the measurement error cannot be assumed as the identical distribution due to the existence of d_i. In addition, when the simplified range model is considered, i.e., $\sigma_\alpha = \sigma_\beta = 0$, Eq. (17) becomes

$$\sigma_{\epsilon_i}^2 = \left(\frac{10\beta}{\ln(10)d_i}\right)^2(\sigma_x^2 + \sigma_y^2) \tag{18}$$

It is clear that when $d_i \neq d_j$, there is $\sigma^2_{\epsilon_i} \neq \sigma^2_{\epsilon_j}$. This indicates that even when the simplified range model is used, the measurement error cannot be considered as Gaussian or IID. Therefore, the optimality of the algorithms based on the IID assumption (Types I and III in Table 1) cannot hold.

3.3 Non-linear GLS estimator

As discussed earlier, the problem of range-based localization is to find a set of parameters best fitting Eq. (6). In this paper, we combine the Newton method with the GLS method to solve the problem (from the knowledge of estimation theory, we know that other methods such as maximum likelihood estimation may also be optimal estimators. The GLS method is selected since its statistical distribution can be expressed without knowing the exact value of the covariance matrix). First, we will use the Newton method to linearize the system in Eq. (6). Define the residual vector $r(\theta)$ between the measured vector $\bar{P}$ and the calculated vector $F(\theta)$ as

$$r(\theta) = \bar{P} - F(\theta) \tag{19}$$

Since the residual is non-linear, we need to linearize it and find a solution iteratively. By Taylor's theorem, the system in Eq. (19) can be linearized at each iteration:

$$r(\theta) \approx r(\theta^{(s)}) + J(\theta - \theta^{(s)}) \tag{20}$$

where $\theta^{(s)}$ indicates that it is the initial guess of θ or the solution calculated by the previous iteration, and s indicates the number of the iteration. J is the Jacobian matrix with

$$J_{(i,j)} = \frac{\partial r_i(\theta^{(s)})}{\partial \theta_j} = \frac{\partial F_i(\theta^{(s)})}{\partial \theta_j} \tag{21}$$

Denoting the final Jacobian matrix as J, we have

$$J = \begin{pmatrix} -10\beta \dfrac{x_1 - x_u}{d_1^2 \ln(10)} & -10\beta \dfrac{y_1 - y_u}{d_1^2 \ln(10)} & -10\lg(d_1) & 1 \\ -10\beta \dfrac{x_2 - x_u}{d_2^2 \ln(10)} & -10\beta \dfrac{y_2 - y_u}{d_2^2 \ln(10)} & -10\lg(d_2) & 1 \\ \vdots & \vdots & \ddots & \vdots \\ -10\beta \dfrac{x_m - x_u}{d_m^2 \ln(10)} & -10\beta \dfrac{y_m - y_u}{d_m^2 \ln(10)} & -10\lg(d_m) & 1 \end{pmatrix} \tag{22}$$

We can see that Eq. (20) is a linear system with the known $\theta^{(s)}$ and the unknown θ. Now the linear least square method can be used to find a solution for θ. Since $r(\theta)$ has the same statistical property with e, the GLS method can be used as an efficient estimator, which will find a solution to minimize the squared Mahalanobis length of the residual vector $r(\theta)$. Therefore, an efficient estimation of θ is given by

$$\theta - \theta^{(s)} = (J^{\mathrm{T}} \Omega^{-1} J)^{-1} J^{\mathrm{T}} \Omega^{-1} r(\theta^{(s)}) \tag{23}$$

With Eq. (23), we get a formula to refine the solution iteratively. Basically, by an initial guess $\theta^{(0)}$, the formula can be iteratively used to find a best solution until the residual is less than a threshold or other terminating conditions are met. When the iteration is finished, the estimated θ will be approximately treated as the true solution. Denoting the final estimation error as $\Delta\theta = \theta - \theta^{(s)}$, we have

$$\Delta\theta = (J^{\mathrm{T}} \Omega^{-1} J)^{-1} J^{\mathrm{T}} \Omega^{-1} e \tag{24}$$

According to the covariance law, we have

$$\mathrm{Cov}(\Delta\theta) = \mathrm{Cov}((J^{\mathrm{T}} \Omega^{-1} J)^{-1} J^{\mathrm{T}} \Omega^{-1} e) = (J^{\mathrm{T}} \Omega^{-1} J)^{-1} \tag{25}$$

In addition, we also have

$$\mathrm{Cov}(\Delta\theta) = \begin{pmatrix} \sigma_\alpha^2 & & & \\ & \sigma_\beta^2 & & \\ & & \sigma_x^2 & \\ & & & \sigma_y^2 \end{pmatrix} \tag{26}$$

Denoting the localization error as σ^2 and taking Eqs. (25) and (26) together, we have

$$\sigma^2 = \sigma_x^2 + \sigma_y^2 = (J^{\mathrm{T}} \Omega^{-1} J)^{-1}_{(3,3)} + (J^{\mathrm{T}} \Omega^{-1} J)^{-1}_{(4,4)} \tag{27}$$

In Eq. (27), the subscripts (3, 3) and (4, 4) indicate the third and the fourth elements of the diagonal line of $(J^{\mathrm{T}} \Omega^{-1} J)^{-1}$, respectively. Since $(J^{\mathrm{T}} \Omega^{-1} J)$ is a 4×4 matrix, it is hard to find the analytic expression of its inverse matrix. This means that if we want to get the analytic expression for σ^2, alternative methods should be considered.

Note that from the knowledge of the GLS estimator, we know it is unbiased, consistent, efficient, and asymptotically normal. The meaning of efficient here is that the estimation error of the GLS method equals CRLB. In other words, the GLS method can achieve the best accuracy.

3.4 Practical estimators

As discussed earlier, the GLS method cannot be used in practice since the covariance matrix Ω is unknown. In this subsection, we will introduce some practical localization algorithms, which will be compared with the GLS estimator in Section 5 in order to know how close the practical accuracy is to the best accuracy. In fact, when the IID assumption is adopted, these practical algorithms are regarded as efficient estimators. Here, we will briefly introduce several commonly used practical algorithms, such as Feasible Generalized Least Squares (FGLS), OLS, and the Levenberg-Marquardt (LM) algorithm[43].

For the FGLS method, it is used as an implementable version of the GLS estimator. In our case, the goal is

to find an available approximation for Ω. Actually, the approximated Ω could be initially estimated by applying inefficient estimators, and then the Newton method can be used to calculate the residuals and build the consistent FGLS estimator during each of its iterative steps. In this paper, we choose the estimated covariance matrix calculated by the LM method. Denoting the estimated covariance matrix as Ω_{LM}, we use the following equation at each iteration:

$$\theta - \theta^{(s)} = (J^{\mathrm{T}}\Omega_{\mathrm{LM}}^{-1}J)^{-1}J^{\mathrm{T}}\Omega_{\mathrm{LM}}^{-1}r(\theta^{(s)}) \tag{28}$$

For the OLS method, it uses the Gauss-Newton algorithm directly. However, as conducted in Section 4.2, we know that the RSS measurement error is not spherical. Thus, the OLS method may be statistically inefficient with misleading inferences. Therefore, the result calculated by the OLS method may not be optimal. Similar to the non-linear GLS method, the OLS method should also be combined with an iterative approach to refine the solution repeatedly. Actually, for the OLS method, the equation used for the iteration will be

$$\theta - \theta^{(s)} = (J^{\mathrm{T}}J)^{-1}J^{\mathrm{T}}r(\theta^{(s)}) \tag{29}$$

For the LM method, it can be considered as a robust version of the OLS method in dealing with the challenging issues of global convergence. In mathematics and computing, the LM method is also known as the Damped Least-Squares (DLS) method, which is used to solve non-linear least squares problems. The LM method interpolates between the Gauss-Newton algorithm and the method of gradient descent, which is more robust than the Gauss-Newton algorithm. In many cases, it finds a solution even if it starts very far off the final minimum. The interested readers can find more in Ref. [43].

4 Theoretical Error Data Analytics

4.1 Basic idea

From Eq. (27), it is easy to see that the localization accuracy depends on the Jacobian matrix J and the covariance matrix Ω. However, it is hard to use Eq. (27) to analyze the relationship between σ^2 and $\sigma_{\epsilon_i}^2$ since it involves calculating the inverse of a high-order matrix. Therefore, we conquer the problem by dividing it into a three-step procedure. The first step is to convert RSS measurements to measurement ranges, then we will analyze the relationship between RSS measurement errors and range estimation errors caused by the conversion. In the second step, based on the converted measurement ranges, we use the multilateration method to calculate the unknown location, that is, using m converted ranges to estimate the unknown location. Through analyzing the propagation of the range estimation errors in the multilateration method, we finally get an analytic expression of the estimation error. From the expression, we can see that the localization error depends on several key factors. However, the error expression does not provide enough information on how these factors affect the accuracy. In the third step, we further derive the minimal localization error, by which we can clearly see how different factors affect the localization accuracy.

4.2 Errors in RSS measurement conversion

In the first step, we will investigate the relationship between measurement errors and range estimation errors when RSS measurements are converted to measurement ranges. Denote the distance between (x_u, y_u) and (x_i, y_i) as d_i, i.e.,

$$d_i = \sqrt{(x_i - x_u)^2 + (y_i - y_u)^2} \tag{30}$$

Let $\omega = (\alpha, \beta, P_i)^{\mathrm{T}}$, we get a function h_i for d_i:

$$d_i = h_i(\omega) = 10^{\frac{\alpha - P_i}{10\beta}} \tag{31}$$

Consider an estimation $\bar{\theta} = (\bar{x}_u, \bar{y}_u, \bar{\alpha}, \bar{\beta})^{\mathrm{T}}$ for the measurement $\bar{P}_i$. Putting $\bar{\alpha}$, $\bar{\beta}$, and $\bar{P}_i$ into Eq. (31), we get an estimated distance $\bar{d}_i$ for d_i, that is

$$\bar{d}_i = h_i(\bar{\omega}) = 10^{\frac{\bar{\alpha} - \bar{P}_i}{10\bar{\beta}}} = d_i + \xi_i \tag{32}$$

where ξ_i is the error in $\bar{d}_i$ and $\bar{\omega} = (\bar{\alpha}, \bar{\beta}, \bar{P}_i)^{\mathrm{T}}$. Denoting the variance of $\bar{d}_i$ as $\sigma_{\xi_i}^2$, we get

$$\sigma_{\xi_i}^2 = \mathrm{Var}(\xi_i) = \mathrm{Var}(d_i + \xi_i) = \mathrm{Var}(h_i(\bar{\omega})) \tag{33}$$

Then we can calculate the variance of $h_i(\bar{\omega})$ by

$$\mathrm{Var}(h_i(\bar{\omega})) \approx \nabla h_i(\omega)^{\mathrm{T}} \cdot \mathrm{Cov}(\bar{\omega}) \cdot \nabla h_i(\omega) \tag{34}$$

Since we have

$$\begin{cases} \dfrac{\partial h_i(E(\bar{\alpha}))}{\partial \alpha} \approx d_i \dfrac{\ln(10)}{10\beta}, \\ \dfrac{\partial h_i(E(\bar{\beta}))}{\partial \beta} \approx d_i \dfrac{\ln(10)}{10\beta}(10\lg(d_i)), \\ \dfrac{\partial h_i(E(\bar{P}_i))}{\partial P_i} \approx -d_i \dfrac{\ln(10)}{10\beta} \end{cases} \tag{35}$$

Then we get

$$\nabla h_i(\omega) = \left(d_i \frac{\ln(10)}{10\beta}, d_i \frac{\ln(10)}{10\beta}(10\lg(d_i)), d_i \frac{\ln(10)}{10\beta} \right)^{\mathrm{T}} \tag{36}$$

Since

$$\mathrm{Cov}(\Delta\theta) = \begin{pmatrix} \sigma_\alpha^2 & & \\ & \sigma_\beta^2 & \\ & & \sigma_{\epsilon_i}^2 \end{pmatrix} \tag{37}$$

Putting Eqs. (36) and (37) into Eq. (34), we finally get

$$\sigma_{\xi_i}^2 \approx \left(\frac{\ln(10)}{10\beta}\right)^2 d_i^2 \tilde{\sigma}_i^2 \tag{38}$$

where

$$\tilde{\sigma}_i^2 = \sigma_{\epsilon_i}^2 + \sigma_\alpha^2 + (10\lg(d_i))^2 \sigma_\beta^2 \tag{39}$$

From Eq. (38), we can see that the error in the converted ranges depends on several factors: the true distance d_i, the measurement error in $\bar{P}_i$, and the estimation error of both α and β. Since the measurement error ϵ_i is assumed as independent to others, it is easy to know that ξ_i is also independent to other range estimation errors. From Eq. (38), we also know that ξ_i does not have the identical distribution. Therefore, for the multilateration method, we also need to use the GLS algorithm to achieve an optimal solution. Note that if α and β are perfectly estimated (i.e., $\sigma_\alpha^2 = \sigma_\beta^2 = 0$), and if the measurement error is IID (i.e., $\sigma_{\epsilon_i}^2 = \sigma_r^2$), then Eq. (38) becomes

$$\sigma_{\xi_i}^2 \approx d_i^2 \left(\frac{\ln(10)}{10\beta}\right)^2 \sigma_r^2 \tag{40}$$

which is the case described in Ref. [38]. This also verifies that previous works in Ref. [38] are special cases of Type IV in Table 1.

4.3 Errors in location estimation

After converting RSS measurements to measurement ranges, we use the multilateration approach to calculate (x_u, y_u). Note that only when an optimal estimator is used, the localization error will be the same with that in Eq. (27). Next, we will analyze how range errors propagate in the multilateration approach. For the distance $\overline{d}_i$ estimated by $\overline{P}_i$, $\overline{\alpha}$, and $\overline{\beta}$, we have

$$\overline{d}_i = \sqrt{(x_i - x_u)^2 + (y_i - y_u)^2} + \xi_i \tag{41}$$

Note that the statistical property of ξ_i is shown in Eq. (38). Denoting the unknown location as a vector $\vartheta = (x_u, y_u)^{\mathrm{T}}$ and writing Eq. (41) as a function k_i, we have

$$d_i = k_i(\vartheta) = \sqrt{(x_i - x_u)^2 + (y_i - y_u)^2} \tag{42}$$

Denote all ranges $\bar{d}_i$ calculated by $\bar{P}_i$, $\bar{\alpha}$, and $\bar{\beta}$ with Eq. (31) as

$$\bar{d} = (\overline{d}_1, \overline{d}_2, \ldots, \overline{d}_m)^{\mathrm{T}} \tag{43}$$

and all ranges calculated by (ϑ) with Eq. (42) as

$$k(\vartheta) = (k_1(\vartheta), k_2(\vartheta), \ldots, k_m(\vartheta))^{\mathrm{T}} \tag{44}$$

When ϑ indicates the true location, we have

$$\bar{d} - k(\vartheta) = g \tag{45}$$

where $g = (\xi_1, \xi_2, \ldots, \xi_m)^{\mathrm{T}}$. Since the error vector g is unknown, we use the following system to find a solution:

$$\bar{d} - \bar{k}(\vartheta) \approx 0 \tag{46}$$

Writing Eq. (46) in a matrix form, we have

$$\begin{cases} \bar{d}_1 = \sqrt{(x_1 - x_u)^2 + (y_1 - y_u)^2}, \\ \bar{d}_2 = \sqrt{(x_2 - x_u)^2 + (y_2 - y_u)^2}, \\ \vdots \\ \bar{d}_m = \sqrt{(x_m - x_u)^2 + (y_m - y_u)^2} \end{cases} \tag{47}$$

Let Ω_d be the covariance matrix of g. Since the errors in any two ranges are uncorrelated, we have $\sigma_{\xi_i \xi_j} = 0$ when $i \neq j$. Then the covariance matrix of g becomes

$$\Omega_d = \begin{pmatrix} \sigma_{\xi_1}^2 & & & \\ & \sigma_{\xi_2}^2 & & \\ & & \ddots & \\ & & & \sigma_{\xi_m}^2 \end{pmatrix} \tag{48}$$

Putting Eq. (38) into Eq. (48), we have

$$\Omega_d = \left(\frac{\ln(10)}{10\beta}\right)^2 \begin{pmatrix} d_1^2\tilde{\sigma}_1^2 & & & \\ & d_2^2\tilde{\sigma}_2^2 & & \\ & & \ddots & \\ & & & d_m^2\tilde{\sigma}_m^2 \end{pmatrix} \tag{49}$$

Knowing the covariance matrix of g, we can use the non-linear GLS method to find a best estimation for Eq. (46). Define the residual vector between $\bar{d}$ and $k(\vartheta)$ as

$$r_d(\vartheta) = \bar{d} - k(\vartheta) \tag{50}$$

After linearizing the above system, we have

$$r_d(\vartheta) \approx r_d(\vartheta^{(s)}) + J_d(\vartheta - \vartheta^s) \tag{51}$$

where the Jacobian matrix J_d has

$$(J_d)_{(i,j)} = \frac{\partial (r_d)_i(\vartheta^{(s)})}{\partial \vartheta_j} = \frac{\partial k_i(\vartheta^{(s)})}{\partial \vartheta_j} \tag{52}$$

that is

$$J_d = \begin{pmatrix} \dfrac{x_1 - x_u}{d_1} & \dfrac{y_1 - y_u}{d_1} \\ \dfrac{x_2 - x_u}{d_2} & \dfrac{y_2 - y_u}{d_2} \\ \vdots & \vdots \\ \dfrac{x_m - x_u}{d_m} & \dfrac{y_m - y_u}{d_m} \end{pmatrix} \tag{53}$$

Then an efficient estimation of ϑ can be calculated by

$$\vartheta - \vartheta^{(s)} = (J_d^{\mathrm{T}} \Omega_d^{-1} J_d)^{-1} J_d^{\mathrm{T}} \Omega_d^{-1} r_d(\vartheta^{(s)}) \tag{54}$$

When the iteration is finished, the estimation error $\Delta\vartheta = \vartheta - \vartheta^{(s)}$ can be expressed as

$$\Delta\vartheta = (J_d^{\mathrm{T}} \Omega_d^{-1} J_d)^{-1} J_d^{\mathrm{T}} \Omega_d^{-1} g \tag{55}$$

By the knowledge of the covariance law, we have

$$\Omega_{\Delta\vartheta} = \mathrm{Cov}((J_d^{\mathrm{T}} \Omega_d^{-1} J_d)^{-1} J_d^{\mathrm{T}} \Omega_d^{-1} g) = (J_d^{\mathrm{T}} \Omega_d^{-1} J_d)^{-1} \tag{56}$$

Putting Eq. (49) into Eq. (56), we have

$$\Omega_{\Delta\vartheta}=\left(\frac{\ln(10)}{10\beta}\right)^2\left(J_d^{\mathrm{T}}\begin{pmatrix}\frac{1}{d_1^2\tilde{\sigma}_1^2} & 0 & \cdots & 0\\ 0 & \frac{1}{d_2^2\tilde{\sigma}_2^2} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \frac{1}{d_m^2\tilde{\sigma}_m^2}\end{pmatrix}J_d\right)^{-1} \tag{57}$$

Since a multiplication of two 2×2 matrices can be solved by the analytic method, after simplifying, we have

$$\Omega_{\Delta\vartheta}=\left(\frac{\ln(10)}{10\beta}\right)^2\begin{pmatrix}\psi_1 & \psi_2\\ \psi_2 & \psi_3\end{pmatrix}^{-1} \tag{58}$$

where

$$\psi_1=\sum_{i=1}^{m}\frac{(x_i-x_u)^2}{d_i^4\tilde{\sigma}_i^2}=\sum_{i=1}^{m}\frac{\cos^2\phi_i}{d_i^2\tilde{\sigma}_i^2} \tag{59}$$

$$\psi_2=\sum_{i=1}^{m}\frac{(x_i-x_u)(y_i-y_u)}{d_i^4\tilde{\sigma}_i^2}=\sum_{i=1}^{m}\frac{\cos\phi_i\sin\phi_i}{d_i^2\tilde{\sigma}_i^2} \tag{60}$$

$$\psi_3=\sum_{i=1}^{m}\frac{(y_i-y_u)^2}{d_i^4\tilde{\sigma}_i^2}=\sum_{i=1}^{m}\frac{\sin^2\phi_i}{{d_i}^2\tilde{\sigma}_i^2} \tag{61}$$

where ϕ_i is the angle between the line by the anchor (x_i,y_i), the blind node (x_u,y_u), and X-axis. Since $(J_d^{\mathrm{T}}\Omega_d^{-1}J_d)$ in Eq. (56) is a 2×2 matrix, we can use the analytical method to calculate its inverse matrix. That is, we have

$$\begin{pmatrix}\psi_1 & \psi_2\\ \psi_2 & \psi_3\end{pmatrix}^{-1}=\frac{1}{\psi_1\psi_3-\psi_2^2}\begin{pmatrix}\psi_3 & -\psi_2\\ -\psi_2 & \psi_1\end{pmatrix} \tag{62}$$

Also because we have

$$\Omega_{\Delta\vartheta}=\begin{pmatrix}\sigma_x^2 & \\ & \sigma_y^2\end{pmatrix} \tag{63}$$

we finally get the following analytic expression of σ^2:

$$\sigma^2=\sigma_x^2+\sigma_y^2=\left(\frac{\ln(10)}{10\beta}\right)^2\frac{\psi_1+\psi_3}{\psi_1\psi_3-\psi_2^2} \tag{64}$$

From the above study, we know that the accuracy of the multilateration method is the same with that of the calculation in Section 3.3. That is, the error expression in Eq. (64) is the same with the localization error in Eq. (27). From Eq. (64), we can see that the accuracy depends on the following factors: the measurement range d_i, the relative geometry of anchors ϕ_i, the amount of anchors m, the measurement error $\sigma_{\epsilon_i}^2$, and the errors in estimated α and β, i.e., σ_α^2 and σ_β^2. In addition, when β and α are perfectly estimated, i.e., $\sigma_\beta^2=\sigma_\alpha^2=0$, and if the measurement error is IID, Eq. (64) turns into the case described in Ref. [38]. However, Eq. (64) does not provide enough information on how these factors affect the accuracy. Therefore, we need to find other ways to clarify how these factors affect the accuracy of the GLS method.

4.4 Minimum of localization error

From Eq. (64), we obtain the analytic expression of the localization error, i.e., the optimal solution in an analytic form. Next, we will derive its minimum to explore how different factors affect the accuracy. $\dfrac{\psi_1+\psi_3}{\psi_1\psi_3-\psi_2^2}$ in Eq. (64) can be written as

$$\frac{\sum_{i=1}^{m}\frac{\cos^2\phi_i}{d_i^2\tilde{\sigma}_i^2}+\sum_{i=1}^{m}\frac{\sin^2\phi_i}{d_i^2\tilde{\sigma}_i^2}}{\left(\sum_{i=1}^{m}\frac{\cos^2\phi_i}{d_i^2\tilde{\sigma}_i^2}\right)\left(\sum_{i=1}^{m}\frac{\sin^2\phi_i}{d_i^2\tilde{\sigma}_i^2}\right)-\left(\sum_{i=1}^{m}\frac{\cos\phi_i\sin\phi_i}{d_i^2\tilde{\sigma}_i^2}\right)^2} \tag{65}$$

After simplifying, we get

$$\frac{\sum_{i=1}^{m}\frac{1}{d_i^2\tilde{\sigma}_i^2}}{\sum_{i=1,j=1}^{m}\frac{(\sin\phi_j\cos\phi_i-\cos\phi_j\sin\phi_i)^2}{d_i^2d_j^2\tilde{\sigma}_i^2\tilde{\sigma}_j^2}} \tag{66}$$

and it can be rewritten as

$$\frac{\sum_{i=1}^{m}\frac{1}{d_i^2\tilde{\sigma}_i^2}}{\sum_{i=1,j=1}^{m}\frac{\sin^2(\phi_j-\phi_i)}{d_i^2d_j^2\tilde{\sigma}_i^2\tilde{\sigma}_j^2}} \tag{67}$$

Since $\sin^2(\phi_j-\phi_i)\leqslant 1$, we have

$$\frac{\sum_{i=1}^{m}\frac{1}{d_i^2\tilde{\sigma}_i^2}}{\sum_{i=1,j=1}^{m}\frac{\sin^2(\phi_j-\phi_i)}{d_i^2d_j^2\tilde{\sigma}_i^2\tilde{\sigma}_j^2}}\geqslant\frac{\sum_{i=1}^{m}\frac{1}{d_i^2\tilde{\sigma}_i^2}}{\sum_{i=1,j=1}^{m}\frac{1}{d_i^2d_j^2\tilde{\sigma}_i^2\tilde{\sigma}_j^2}} \tag{68}$$

Let the minimum of $(d_1,d_2,\ldots,d_m)$ be $d_{\min}$, the minimum of $(\sigma_{\epsilon_1}^2,\sigma_{\epsilon_2}^2,\ldots,\sigma_{\epsilon_m}^2)$ be $\sigma_{\epsilon_{\min}}^2$, and let $\tilde{\sigma}_{\min}^2=\sigma_{\epsilon_{\min}}^2+\sigma_\alpha^2+(10\lg(d_{\min}))^2\sigma_\beta^2$. Since $d_j^2\tilde{\sigma}_j^2\geqslant d_{\min}^2\tilde{\sigma}_{\min}^2$, we have

$$\frac{\sum_{i=1}^{m}\frac{1}{d_i^2\tilde{\sigma}_i^2}}{\sum_{i=1,j=1}^{m}\frac{1}{d_i^2d_j^2\tilde{\sigma}_i^2\tilde{\sigma}_j^2}}\geqslant\frac{\sum_{i=1}^{m}\frac{1}{d_i^2\tilde{\sigma}_i^2}}{\sum_{i=1,j=1}^{m}\frac{1}{d_i^2\tilde{\sigma}_i^2d_{\min}^2\tilde{\sigma}_{\min}^2}} \tag{69}$$

Since

$$\frac{\sum_{i=1}^{m}\frac{1}{d_i^2\tilde{\sigma}_i^2}}{\sum_{i=1,j=1}^{m}\frac{1}{d_i^2\tilde{\sigma}_i^2d_{\min}^2\tilde{\sigma}_{\min}^2}}=\frac{d_{\min}^2\tilde{\sigma}_{\min}^2\sum_{i=1}^{m}\frac{1}{d_i^2\tilde{\sigma}_i^2}}{m\sum_{i=1}^{m}\frac{1}{d_i^2\tilde{\sigma}_i^2}} \tag{70}$$

and

$$\frac{d_{\min}^2\tilde{\sigma}_{\min}^2\sum_{i=1}^{m}\frac{1}{{d_i}^2}}{m\sum_{i=1}^{m}\frac{1}{d_i^2}}=\frac{1}{m}d_{\min}^2\tilde{\sigma}_{\min}^2 \tag{71}$$

we finally get

$$\frac{\psi_1 + \psi_3}{\psi_1\psi_3 - \psi_2^2} \geqslant \frac{1}{m} d_{\min}^2 \tilde{\sigma}_{\min}^2 \tag{72}$$

Then, putting Eqs. (72) and (64) together, we get

$$\sigma^2 \geqslant \left(\frac{\ln(10)}{10\beta}\right)^2 \frac{1}{m} d_{\min}^2 \tilde{\sigma}_{\min}^2 \tag{73}$$

We can see that the right side of Eq. (73) is the minimum of the localization error of the GLS method. The inequality means that the real localization error of range-based localization methods cannot be smaller than that minimum.

From Eq. (73), we can see that the minimum of the localization error is proportional to $d_{\min}^2$, $\tilde{\sigma}_{\min}^2$, and the inverse of m. We see that when the measurement range d_i and the measurement error $\tilde{\sigma}_i^2$ increase, $d_{\min}^2$ and $\tilde{\sigma}_{\min}^2$ will probably increase. This implies that the measurement ranges and errors have a proportional relationship with the localization error, i.e., the minimal range and measurement error will probably increase when all ranges or measurement errors increase. Similarly, increasing the amount of anchors will decrease the localization error. Therefore, from the minimum, we know how factors in the previous section affect the localization accuracy (note that the minimum does not contain the information of ϕ_i, which means the effect of the relative geometry should be studied in future work).

5 Experiment and Verification

5.1 Verification method

When verifying the theoretical analytics in Section 4, we prefer using the simulation approach other than real experiments. There are three reasons for this strategy: Firstly, when comparing different localization algorithms, we need to know the optimal accuracy that can be achieved. However, since the practical algorithms such as the OLS method cannot guarantee the best estimation, we have to use the GLS method to get the best accuracy by simulation. In other words, only with the simulation data can we calculate the covariance matrix Ω, which is unknown in practice. Secondly, in order to verify the correctness of our analytics in Section 4, we need to know the accurate value of σ_α, σ_β, σ_{ϵ_i}, and β. However, in practical environments, we cannot use non-optimal algorithms to get their accurate values. Therefore, only by simulation data can we know their correct values. Thirdly, in our experiments, we need to change the values of different factors to verify their relationship with the localization error. However, in real environments, several factors are not adjustable, such as σ_α, σ_β, and σ_{ϵ_i}. Therefore, only by simulation can we freely adjust these factors to see how their changes affect the localization accuracy. Due to the above reasons, we choose the simulation approach to verify the theoretical error data analytics in Section 4.

5.2 Simulation setting

Before presenting the experimental study, we first introduce the settings of the simulation environment. We can define the blind node at the location (0, 0). The anchors are randomly deployed in a circular area with the radius of 20 m. In order to meet the requirement of the model as described in Eq. (1), the minimal measurement range is set to 1 m. We set $\alpha = -25$ dBm, i.e., we assume all anchors use homogenous devices whose transmission powers are equal. For β, we set it as 2, which means an ideal propagation condition.

We generate two different RSS measurement data for the verification of the correctness of the proposed error data analytics in Section 4. The first one is the data with the non-IID measurement error, which satisfies the relationship in Eq. (17). In order to generate $\bar{P}_i$ with the non-IID error, we need to add errors to the true parameters $\theta = (\alpha, \beta, x_u, y_u)^{\mathrm{T}}$. Since there are m measurements, for i-th measurement $\bar{P}_i$, we denote its errors added to every parameter as $\Delta\alpha_i$, $\Delta\beta_i$, Δx_i, and Δy_i, respectively. Let $\Delta\theta_i = (\Delta\alpha_i, \Delta\beta_i, \Delta x_i, \Delta y_i)^{\mathrm{T}}$. Adding $\Delta\theta_i$ to the true vector θ, we get $\theta + \Delta\theta_i = (\alpha + \Delta\alpha_i, \beta + \Delta\beta_i, x_u + \Delta x_i, y_u + \Delta y_i)^{\mathrm{T}}$. Putting $\theta + \Delta\theta_i$ into Eq. (2) as the input, i.e., $f_i(\theta + \Delta\theta_i)$, we can obtain an RSS measurement $\bar{P}_i$. Note that we use the variances σ_α^2, σ_β^2, σ_x^2, and σ_y^2 to generate $\Delta\alpha_i$, $\Delta\beta_i$, Δx_i, and Δy_i, respectively, here $4 \leqslant i \leqslant m$. If there are no special remarks, we have $\sigma_\alpha = 2.5$ dBm, $\sigma_\beta = 0.2$, $\sigma_x = 1$ m, and $\sigma_y = 1$ m as default. By m measurements, we obtain the vector of the RSS measurement $\bar{P}$. Furthermore, by $\bar{P} - F(\theta)$, we also get the error vector e, which is needed for the calculation of the covariance matrix Ω for the GLS method.

The second type of simulation data will be generated with the IID error. For this type, we generate a vector of the measurement error $e = (\epsilon_1, \epsilon_2, \ldots, \epsilon_m)$ with the Gaussian distribution, that is, the variance of e is σ_r^2. Here, the default setting for σ_r is 2.5 dBm. Then we use Eq. (5) to generate the vector of the RSS measurements $\bar{P}$, i.e., $\bar{P} = F(\theta) + e$. With this method, we get m RSS measurements with the IID error. Since e is also known, we can construct the covariance matrix Ω for

the GLS method. However, this type of data cannot be used to calculate the analytic error expression and its minimum because the correct value of σ_α and σ_β cannot be calculated.

5.3 Localization algorithms for comparison

In this subsection, we will introduce the practical algorithms and the optimal algorithm to be compared. Through comparison, we can investigate their accuracy gap and find the opportunity to optimize current algorithms or design new algorithms. In order to verify the correctness of the theoretical error data analytics in Section 4, i.e., the correctness of the analytic expression of the localization error and its minimum, we will also explain the method of calculating these two formulas (we will not investigate how the relative geometry affects the localization accuracy since we do not know the exact relationship from the expression). We introduce the configuration of the algorithms first as follows:

• **Levenberg-Marquardt method.** Among the proposed techniques, we apply the Levenberg-Marquardt method[43], which is widely regarded as the most robust one in dealing with the challenging issues of global convergence. This algorithm requires setting an initial scale for each parameter. In our experiments, we set the initial scale for x_u and y_u as $(-20, 20)$, which means the blind node will not locate outside the testing area. For α, the initial scale is set as $(-20, -30)$, and for β, we set its initial scale as $(1, 3)$.

• **Ordinary least squares.** The OLS method assumes the measurement error being IID, for both types of the simulation data. We simply use the Gauss-Newton method to find the solution. It uses the iterative procedure to refine the solution. Similar to the LM method, the OLS method also need to set initial values for each parameter. In our experiments, in order to eliminate the local minimal, we set the initial value for $(x_u, y_u, \alpha, \beta)$ as $(0, 0, -25, 2)$.

• **Feasible generalized least squares.** The FGLS method is similar to the OLS method, and the only difference is the equation used to calculate iteratively. For FGLS, we also need to set a proper initial value for all parameters. Here we use the same configuration with that of the OLS method. For the FGLS method, we use the LM method to construct an approximated covariance matrix for Ω, which means that the solution calculated by the LM method is adopted to generate and calculate its covariance matrix to replace Ω.

• **Generalized least squares.** For the GLS method, in practice, we cannot get the covariance matrix of the measurement error. Therefore, in our experiments, we need to obtain the measurement error added to the true value in order to calculate the covariance matrix. Since the GLS method also uses an iterative procedure to find the solution, we use the same configuration for unknown parameters as the OLS method. Therefore, for both simulation data, we use the generated RSS measurement error e to calculate $\Omega = \mathrm{Cov}(e)$. Note that, since Ω is unknown in practice, the GLS method is valid theoretically for our cases.

For all the above algorithms, we need to properly set the initial guess for x_u, y_u, α, and β to avoid the local minimal, that is, we set the initial guess close enough to their true values. Next, we will introduce how to calculate the analytic expression of the localization error and its minimum. For the analytic error expression, in addition to knowing (x_i, y_i), (x_u, y_u), and d_i, we also need to know $\tilde{\sigma}_i^2$, i.e., $\sigma_{\epsilon_i}^2$, σ_α^2, and σ_β^2. Since σ_α^2 and σ_β^2 have been assigned with the default value, we only need to determine $\sigma_{\epsilon_i}^2$. However, since $\sigma_{\epsilon_i}^2$ is different under varying RSS measurement $\overline{P_i}$ from the blind node to the anchor (x_i, y_i), it is hard to get $\sigma_{\epsilon_i}^2$ for all possible positions of anchors. Therefore, in our experiments, we use ϵ_i^2 as $\sigma_{\epsilon_i}^2$ in each location calculation. To eliminate the possible random noise caused by this approach, we repeat the calculation many times, i.e., when the number of repeating is large enough, the variance of ϵ_i will be close to $\sigma_{\epsilon_i}^2$. For the minimum of CRLB, we need to know $d_{\min}$ and $\tilde{\sigma}_{\min}^2$. It is easy to get $d_{\min}$ by finding out the minimal one in $(d_1, d_2, \ldots, d_m)$. For $\tilde{\sigma}_{\min}^2$, we select the minimum in $(\epsilon_1^2, \epsilon_2^2, \ldots, \epsilon_m^2)$.

5.4 Effect of measurement ranges

In this subsection, we will explore how measurement ranges affect the localization accuracy. In order to know the effect on the changing of d_i, we need to keep other factors fixed. To keep the relative geometry of anchors fixed, i.e., all ϕ_i fixed, we will extend the measurement ranges proportionally. In addition, we also keep $\tilde{\sigma_i}$ fixed, and therefore, the localization error depends only on the measurement ranges d_i. Here, we define an extending coefficient λ to indicate how measurement ranges change. That is, given a specific value of λ , the coordinates of all anchors are multiplied by λ. Then we use the extended coordinates of the anchors to calculate the localization of the blind node. In this experiment,

the extending coefficient changes from 1 to 3 with the step of 0.1. Other settings are assigned with the default values as introduced in Section 5.2.

Figure 1 shows the localization error of four localization algorithms (LM, OLS, FGLS, and GLS for short), the Analytic Error Expression (AEE), and the Minimum of the Error Expression (MoEE). From Fig. 1, we have the following observations: (1) When the extending coefficient λ increases, the localization error of four algorithms, the error expression and its minimum all increase proportionally. This verifies our error data analytics in Section 4.4, i.e., when other factors are fixed, the localization error is proportional to the measurement ranges. (2) When we compare different algorithms in terms of accuracy, we can see that the GLS method has the best performance, and the performances of other methods are similar. Additionally, we also observe a significant gap on accuracy between the practical and the optimal results. This indicates a great opportunity to improve the current practical algorithms. (3) From Fig. 1, we can see that the analytic error expression is close to the accuracy of the GLS method, which verifies the correctness of the expression in Eq. (64). (4) For the minimum of the localization error, we can see it is far below the optimal accuracy. This also means that it is possible to find novel algorithms if we can find the conditions that make the actual localization error close to the minimum.

Figure 2 shows the effect of the measurement ranges for the IID data. We can see that, when the measurement ranges increase, the localization error of the four algorithms increases proportionally. This also verifies the results in the experiments on the non-IID data. For the accuracy comparisons on different algorithms, the GLS method also has the best accuracy. Other algorithms have similar accuracy and their accuracy has a large gap to the optimal accuracy. An important conclusion from

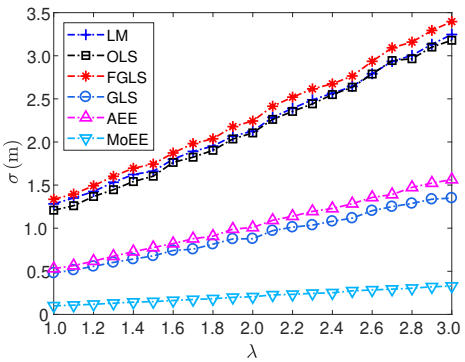


Fig. 1 Effects of measurement ranges with non-IID error.

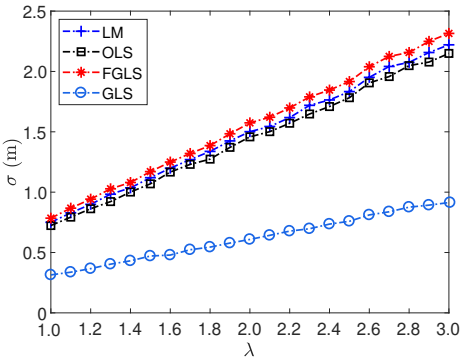


Fig. 2 Effects of measurement ranges with IID error.

this gap is that, even when the measurement error is IID, the claimed optimal algorithms, such as the OLS method, cannot obtain the best solution. This means that these algorithms are inefficient no matter what type of the simulation data is used. Therefore, the simplified IID assumption on the measurement error in Table 1 cannot hold.

5.5 Effect of measurement errors

Next, we will investigate how measurement errors affect the localization performance. Here, we change the measurement error with the other factors fixed. From Eq. (1) in Section 2, we know that the measurement error depends on the error in estimated α, β, x_u, and y_u. For σ_α and σ_β, we use their default values and keep them fixed (the effect of the change of σ_α and σ_β will be studied in later subsections). Then, the measurement error only depends on the error added to the true x_u and y_u. We increase Δx_i and Δy_i proportionally to generate the non-IID RSS measurement data. Here, Δx_i and Δy_i change in the same ratio to let the added location error change proportionally. Specifically, we increase the standard variance of Δx_i and Δy_i from 1 to 5 m with the step of 0.25 m.

However, for the non-IID data, we use m RSS measurements, whose error is $e = (\epsilon_1, \epsilon_1, \ldots, \epsilon_m)$ generated by the change of Δx_i and Δy_i. Since measurement errors may not be equal to each others, we cannot determine the relationship between e and the localization error σ^2, i.e., they do not have one to one mapping relationship. Therefore, in this experiment, we bring a new metric to measure the overall change of the measurement errors. We define the average of all measurement errors in e as

$$\hat{\sigma}_\epsilon = \frac{1}{m}\sum_{i=1}^{m}\sigma_{\epsilon_i} \tag{74}$$

Therefore, for the non-IID data, we will analyze the

relationship between σ and $\hat{\sigma}_\epsilon$.

Figure 3 shows the comparison on the non-IID data. Note that, $\hat{\sigma}_\epsilon$ changes from 4.26 to 7.06 dBm, which is caused by the change of Δx_i and Δy_i. We can see that when the average measurement error, i.e., $\hat{\sigma}_\epsilon$ increases, the localization error of the four algorithms, the error expression and its minimum increase proportionally. This verifies our error data analytics in Section 4, i.e., when other factors are fixed, the change of the localization error is proportional to the change of the measurement error. The accuracy comparison on different algorithms has the similar result with the previous experiments. The GLS method is the best and its performance is close to the analytic error expression. Additionally, the minimum of the localization error is also significantly smaller than the error of the optimal algorithm.

For the IID data, we change the value of σ_r from 2.5 to 10 dBm. Figure 4 shows the result. From Fig. 4, we observe the similar phenomenon with the experiments on the non-IID data. That is, the localization error of all algorithms increases when the measurement error σ_r increases. We also observe that the accuracy of the practical algorithms is significantly lower than that of the optimal algorithm. This also verifies that the practical algorithms cannot reach optimal accuracy when the measurement error is assumed as IID. In practical environments, it is inappropriate to use the assumption in Types I and II, i.e., measurement errors are IID.

5.6 Effect of errors in estimated α and β

As we discussed earlier, the variances of the estimated α and β are assumed as constants. However, in different environments and time periods, they may have different means and variances. Therefore, knowing how these two parameters affect the localization accuracy may help us find better policies to handle the environment-related issues. We will investigate how the changing of σ_α and σ_β affects the localization accuracy. Note that since σ_α and σ_β should be adjustable in our experiments, only the non-IID data is generated for comparison. In the tests, we change the value of σ_α from 0 to 5 dBm with the step of 0.25 dBm and the value of σ_β from 0.1 to 1 with the step of 0.05.

Figures 5 and 6 illustrate how the changes of σ_α and σ_β affect the localization accuracy. We see that in all algorithms, the analytic error expression and its minimum increase when σ_α or σ_β increases. In addition, the result of the accuracy comparison is similar as the result in previous experiments, i.e., the GLS method has

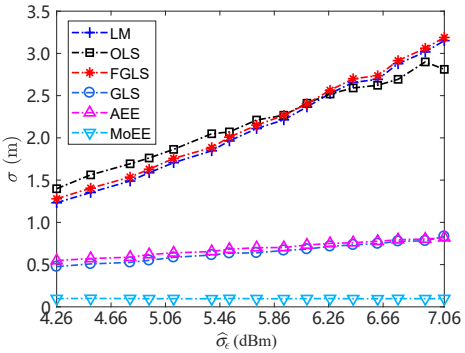


Fig. 3 Effects of measurement error with non-IID error.

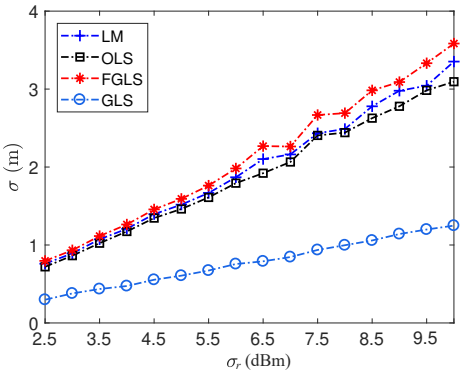


Fig. 4 Effects of measurement error with IID error.

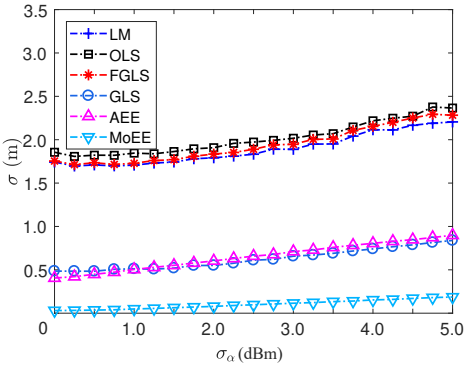


Fig. 5 Effects of error with estimated α.

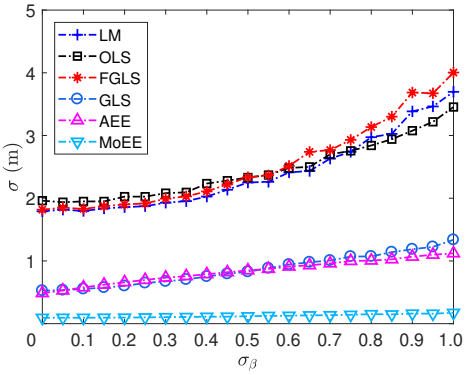


Fig. 6 Effects of error with estimated β.

the best accuracy which is close to the analytic error expression. The accuracy of the practical algorithms also has a significant gap with the optimal accuracy. This verifies that the simplified range model in Types I and III cannot hold.

5.7 Effect of the amount of anchors

As analyzed in Section 4.4, we know that the localization error is inversely proportional to the square root of the amount of anchors. In this test, we use both the non-IID data and the IID data to verify this analytics. Here, we keep other factors fixed and change the amount of anchors from 8 to 24.

Figure 7 shows the result of the experiments on the non-IID data. From Fig. 7, we can see that when the amount of anchors increases, the localization error of the four algorithms, the analytic error expression and its minimum all decrease. Note that the relationships are not linear, which verifies that the localization error is a power function of the amount of anchors. We also see that, when the amount of anchors is small, the performance of different algorithms has some fluctuations. However, when the amount of anchors is large enough, the localization accuracy will be stable and the result of the accuracy comparison is the same with that in previous experiments. In addition, we can see that the error expression is very close to the error of the GLS method.

Figure 8 shows the comparison when the IID data is used. Similar to the experiments on the non-IID data, the localization error will decrease when the amount of anchors increases. The result of the accuracy comparison is also similar to the results on the non-IID data. From Fig. 8, we also observe the significant gap on the accuracy between the practical algorithms and the optimal algorithm. The result also verifies that the assumption of the IID measurement error cannot hold.

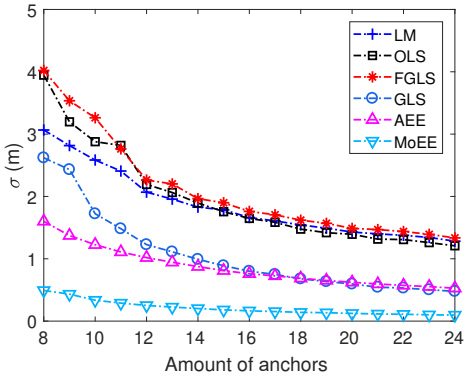


Fig. 7 Effects of the amount of anchors with non-IID error.

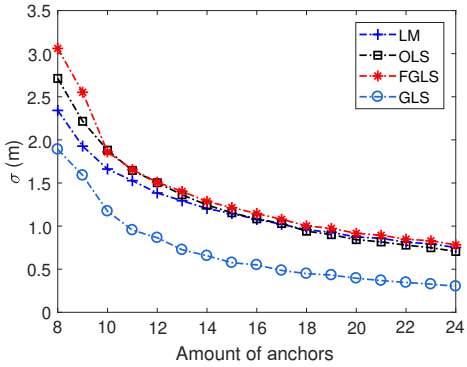


Fig. 8 Effects of the amount of anchors with IID error.

5.8 Summary and discussion

In summary, our experiments verify the correctness of the theoretical error data analytics in Section 4. In one hand, our experiments support the proposed analytics on the error expression and the lower bound, that is, what are the key factors affecting the localization accuracy and how these factors affect the accuracy. In addition, our experiments show that there is a significant gap between the practical accuracy and the optimal accuracy, which means there may have a good opportunity to improve the accuracy of the current algorithms.

Based on the study, we further discuss the possibility of improving the current algorithms. For example, from the test on the measurement ranges, we possibly can use anchors closer to the blind node to increase the localization accuracy. Similarly, from the test on the amount of anchors, we may also obtain better accuracy by using more anchors for measurements. Since the errors of estimated α and β also affect the localization accuracy, we may use less volatile devices and indoor environments to decrease σ_α and σ_β, but increase the accuracy.

Note that in this paper, we did not consider the effect of the relative geometry of the anchors. However, from the experiments, we observe that in all tests, the lower bound is significantly smaller than the localization error of the optimal algorithm. Considering the lower bound is obtained by a condition $\sin^2(\phi_i - \phi_j) \leqslant 1$, we can expect that the relative geometry plays an important role. That is, a specific geometry may bring better accuracy compared with the random geometry used in this paper. Therefore, an important work in future is to study how the relative geometry will affect the localization accuracy.

6 Conclusion and Future Work

In this work, we focus on a long-term unresolved problem: the accuracy of the RSS range-based localization methods. We adopt a generalized measurement model to find optimal estimators whose estimation error equals the Cramér-Rao lower bound. Through mathematical techniques, the key factors for the accuracy of RSS-based localization methods are revealed, and the analytics expression that discloses the proportional relationship between the localization accuracy and these factors is derived. Our experiment through simulation also verifies the correctness of the proposed theoretical error data analytics.

In the future, our work will focus on two further issues to complete the study of the localization accuracy problem: (1) analyzing how the relative geometry of anchors affects the localization accuracy and finding out the optimal geometries for the localization anchor placement. (2) Finding out practical algorithms whose accuracy may be equal to or close to the optimal algorithm, i.e., the GLS method.

Acknowledgment

This work was partially supported by the National Key Research and Development Program of China (No. 2016YFE0121800).

References

[1] M. Hata, Empirical formula for propagation loss in land mobile radio services, *IEEE Transactions on Vehicular Technology*, vol. 29, no. 3, pp. 317–325, 1980.

[2] Q. Zhou and Z. S. Duan, Weighted intersections of bearing lines for AOA based localization, in *2014 17th Int. Conf. Information Fusion*, Salamanca, Spain, 2014, pp. 1–8.

[3] K. H. Yang, G. Wang, and Z. Q. Luo, Efficient convex relaxation methods for robust target localization by a sensor network using time differences of arrivals, *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2775–2784, 2009.

[4] R. M. Vaghefi and R. M. Buehrer, Asynchronous time-of-arrival-based source localization, in *2013 IEEE Int. Conf. Acoustics*, *Speech and Signal Processing*, Vancouver, Canada, 2013, pp. 4086–4090.

[5] Z. M. Yuan, S. H. Yang, and W. Li, A two-tier positioning algorithm for wireless networks with diverse measurement types, in *2013 IEEE Global Communications Conf.*, Atlanta, GA, USA, 2013, pp. 134–139.

[6] W. Li, D. Li, S. H. Yang, Z. W. Xu, and W. Zhao, Design and analysis of a new GPS algorithm, in *2010 IEEE 30th Int. Conf. Distributed Computing Systems*, Genova, Italy, 2010, pp. 40–51.

[7] W. Li, Z. M. Yuan, B. Chen, and W. Zhao, Performance comparison of positioning algorithms for complex GPS systems, in *2012 32nd Int. Conf. Distributed Computing Systems Workshops*, Macau, China, 2012, pp. 273–278.

[8] G. Chandrasekaran, M. A. Ergin, J. Yang, S. Liu, Y. Y. Chen, M. Gruteser, and R. P. Martin, Empirical evaluation of the limits on localization using signal strength, in *2009 6th Annual IEEE Communications Society Conf. Sensor*, *Mesh and Ad Hoc Communications and Networks*, Rome, Italy, 2009, pp. 1–9.

[9] A. Zanella, Best practice in RSS measurements and ranging, *IEEE Communications Surveys & Tutorials*, vol. 18, no. 4, pp. 2662–2686, 2016.

[10] E. Elnahrawy, X. Y. Li, and R. P. Martin, The limits of localization using signal strength: A comparative study, in *2004 1st Annu. IEEE Communications Society Conf. Sensor and Ad Hoc Communications Networks*, Santa Clara, CA, USA, 2004, pp. 406–414.

[11] Z. M. Yuan, W. Li, J. D. Zhu, and W. Zhao, A cost-efficiency method on beacon nodes placement for wireless localization, in *2015 Int. Conf. Computing*, *Networking and Communications*, Garden Grove, CA, USA, 2015, pp. 546–550.

[12] T. K. Sarkar, Z. Ji, K. Kim, A. Medouri, and M. Salazar-Palma, A survey of various propagation models for mobile communication, *IEEE Antennas and Propagation Magazine*, vol. 45, no. 3, pp. 51–82, 2003.

[13] S. Phaiboon, An empirically based path loss model for indoor wireless channels in laboratory building, in *2002 IEEE Region 10 Conf. Computers*, *Communications*, *Control and Power Engineering*, Beijing, China, 2002, pp. 1020–1023.

[14] S. S. Ghassemzadeh, R. Jana, C. W. Rice, W. Turin, and V. Tarokh, A statistical path loss model for in-home UWB channels, in *2002 IEEE Conf. Ultra Wideband Systems and Technologies*, Baltimore, MD, USA, 2002, pp. 59–64.

[15] S. P. Singh and S. C. Sharma, Range free localization techniques in wireless sensor networks: A review, *Procedia Computer Science*, vol. 57, pp. 7–16, 2015.

[16] B. Dil, S. Dulman, and P. Havinga, Range-based localization in mobile sensor networks, in *European Workshop on Wireless Sensor Networks*, K. Römer, H. Karl, and F. Mattern, eds. Berlin, Germany: Springer, 2006, pp. 164–179.

[17] D. Denkovski, M. Angjelicinoski, V. Atanasovski, and L. Gavrilovska, Practical assessment of RSS-based localization in indoor environments, in *MILCOM 2012–2012 IEEE Military Communications Conf.*, Orlando, FL, USA, 2012, pp. 1–6.

[18] Y. Shang, W. Ruml, Y. Zhang, and M. P. Fromherz, Localization from mere connectivity, in *Proc. 4th ACM Int. Symp. Mobile ad Hoc Networking & Computing*, Annapolis, MD, USA, 2003, pp. 201–212.

[19] L. Doherty, K. S. J. Pister, and L. El Ghaoui, Convex position estimation in wireless sensor networks, in *20th Annu. Joint Conf. IEEE Computer and Communications Societies*, Anchorage, AK, USA, 2001, pp. 1655–1663.

[20] S. Tomic, M. Beko, and R. Dinis, RSS-based localization in wireless sensor networks using convex relaxation: Noncooperative and cooperative schemes, *IEEE Transactions on Vehicular Technology*, vol. 64, no. 5, pp. 2037–2050, 2015.

[21] L. X. Lin, H. C. So, and Y. T. Chan, Accurate and simple source localization using differential received signal strength, *Digital Signal Processing*, vol. 23, no. 3, pp. 736–743, 2013.

[22] H. C. So and L. X. Lin, Linear least squares approach for accurate received signal strength based source localization, *IEEE Transactions on Signal Processing*, vol. 59, no. 8, pp. 4035–4040, 2011.

[23] N. Salman, M. Ghogho, and A. H. Kemp, Optimized low complexity sensor node positioning in wireless sensor networks, *IEEE Sensors Journal*, vol. 14, no. 1, pp. 39–46, 2014.

[24] M. R. Gholami, R. M. Vaghefi, and E. G. Ström, RSS-based sensor localization in the presence of unknown channel parameters, *IEEE Transactions on Signal Processing*, vol. 61, no. 15, pp. 3752–3759, 2013.

[25] R. M. Vaghefi, M. R. Gholami, R. M. Buehrer, and E. G. Strom, Cooperative received signal strength-based sensor localization with unknown transmit powers, *IEEE Transactions on Signal Processing*, vol. 61, no. 6, pp. 1389–1403, 2013.

[26] Y. M. Xu, J. G. Zhou, and P. Zhang, RSS-based source localization when path-loss model parameters are unknown, *IEEE Communications Letters*, vol. 18, no. 6, pp. 1055–1058, 2014.

[27] S. Y. Cho, Localization of the arbitrary deployed APs for indoor wireless location-based applications, *IEEE Transactions on Consumer Electronics*, vol. 56, no. 2, pp. 532–539, 2010.

[28] G. Wang, H. Chen, Y. M. Li, and M. Jin, On received-signal-strength based localization with unknown transmit power and path loss exponent, *IEEE Wireless Communications Letters*, vol. 1, no. 5, pp. 536–539, 2012.

[29] M. Veletić and M. Šunjevarić, On the Cramer-Rao lower bound for RSS-based positioning in wireless cellular networks, *AEU-International Journal of Electronics and Communications*, vol. 68, no. 8, pp. 730–736, 2014.

[30] J. Gribben and A. Boukerche, Location error estimation in wireless ad hoc networks, *Ad Hoc Networks*, vol. 13, pp. 504–515, 2014.

[31] N. Salman, M. Ghogho, and A. H. Kemp, On the joint estimation of the RSS-based location and path-loss exponent, *IEEE Wireless Communications Letters*, vol. 1, no. 1, pp. 34–37, 2012.

[32] X. R. Li, RSS-based location estimation with unknown pathloss model, *IEEE Transactions on Wireless Communications*, vol. 5, no. 12, pp. 3626–3633, 2006.

[33] S. Tomic, M. Beko, R. Dinis, G. Dimic, and M. Tuba, Distributed RSS-based localization in wireless sensor networks with node selection mechanism, in *Technological Innovation for Cloud-Based Engineering Systems*, L. Camarinha-Matos, T. Baldissera, G. Di Orio, and F. Marques, eds. Costa de Caparica, Portugal: Springer International Publishing, 2015, pp. 204–214.

[34] Z. M. Yuan, W. Li, A. C. Champion, and W. Zhao, An efficient hybrid localization scheme for heterogeneous wireless networks, in *2012 IEEE Global Communications Conf.*, Anaheim, CA, USA, 2012, pp. 372–378.

[35] S. Tomic, M. Beko, and D. Rui, Distributed RSS-AoA based localization with unknown transmit powers, *IEEE Wireless Communications Letters*, vol. 5, no. 4, pp. 392–395, 2016.

[36] N. Salman, Y. J. Guo, A. H. Kemp, and M. Ghogho, Analysis of linear least square solution for RSS based localization, in *2012 Int. Symp. Communications and Information Technologies*, Gold Coast, Australia, 2012, pp. 1051–1054.

[37] A. J. Weiss and J. S. Picard, Network localization with biased range measurements, *IEEE Transactions on Wireless Communications*, vol. 7, no. 1, pp. 298–304, 2008.

[38] N. Patwari, A. O. Hero, M. Perkins, N. S. Correal, and R. J. O'Dea, Relative location estimation in wireless sensor networks, *IEEE Transactions on Signal Processing*, vol. 51, no. 8, pp. 2137–2148, 2003.

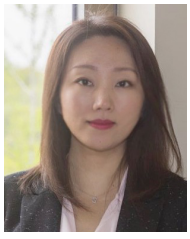
[39] L. Mailaender, On the CRLB scaling law for received signal strength (RSS) geolocation, in *2011 45th Annu. Conf. Information Sciences and Systems*, Baltimore, MD, USA, 2011, pp. 1–6.

[40] P. Tarrío, A. M. Bernardos, and J. R. Casar, Weighted least squares techniques for improved received signal strength based localization, *Sensors*, vol. 11, no. 9, pp. 8569–8592, 2011.

[41] J. Werner, J. Wang, A. Hakkarainen, D. Cabric, and M. Valkama, Performance and Cramer–Rao bounds for DoA/RSS estimation and transmitter localization using sectorized antennas, *IEEE Transactions on Vehicular Technology*, vol. 65, no. 5, pp. 3255–3270, 2016.

[42] G. Casella and R. L. Berger, *Statistical Inference*, 2nd ed. Pacific Grove, CA, USA: Duxbury, 2002.

[43] C. Kanzow, N. Yamashita, and M. Fukushima, WITHDRAWN: Levenberg–Marquardt methods with strong local convergence properties for solving nonlinear equations with convex constraints, *Journal of Computational and Applied Mathematics*, vol. 173, no. 2, pp. 321–343, 2005.



Shuhui Yang received the PhD degree from Florida Atlantic University in 2007. She is an associate professor of computer science in the Department of Mathematics, Statistics, and Computer Science at PNW. She worked as a postdoc research associate at Rensselear Polytechnic Institute before she joined Purdue University Northwest. Her research interests include wireless communication, mobile computing, parallel and distributed systems, wireless security and privacy, cloud and green computing, and social network applications. She had more than 40 publications in peer-reviewed conference proceedings and journals, including *IEEE Transactions*. Her research work is partially funded by the National Science Foundation. She is the guest editor for the *EURASIP Journal on Wireless Communications and Networking*

and special issue on wireless network security. She was the general chair of the 3rd and 4th National Workshop for REU research in Networking and Systems (REUNS); registration and finance chair of the 14th, 15th, and 16th IEEE International Conference on Mobile Ad-hoc Sensor Systems (IEEE MASS); general chair of the 1st and 2nd IEEE International Workshop on Big Data and Cloud Applications (BDCloudApp 2016); and program chair of the 1st and 2nd International Workshop on Smart City Communication and Networking (SmartCityCom). She serves as the program committee member for many conferences, including IEEE INFOCOM and IEEE ICCCN. She also serves as the program vice chair for the 15th IEEE International Conference on Green Computing and Communications (GreenCom) and 2020 International Conference on Edge Computing. She is a member of the IEEE and IEEE Computer Society.

Wei Li completed the undergraduate program in computer science and engineering at Beijing Institute of Technology, Beijing, China in 1996 and received the master degree there in 1999. In the same year, he joined the Institute of Computing Technology, Chinese Academy of Sciences (ICT, CAS) as research assistant. In 2007, he got the PhD degree from ICT, CAS. He is an associate professor at ICT, CAS. He is interested in distributed systems and operating systems. He proposed a resource locating method in distributed environments based on routing-transferring mechanism, and the related results were published in Grid 2002 International Workshop. He also proposed a secured resource sharing mechanism in a distributed system with multiple administrative domains, and the related results were published in Cluster 2003 and HPC Asia 2005 International Conference. He also works on service oriented architecture design for distributed system. His research team built a prototype of a service-oriented virtual machine as a developing and runtime environment for service oriented applications. Two related papers have been published in SCC 2005 and SCC 2006 International Conference. He is also interested in address space model for distributed systems. He has participated in several important national research projects since 2000, such as the project Grid System Software (sponsored by the National High-Tech Research and Development (863) Program of China (No. 2002AA104310)), Network Computing Environment for Scientific Research sponsored by NSFC, and Virtual Computing Environment sponsored by the National Key Basic Research and Development (973) Program of China. In 2005, his application for a project service-oriented distributed virtual machine was approved by NSFC. From 2002, he began to manage the affairs of grid research team at ICT. During 2005 – 2007, he served as the deputy director at the Grid and Service Computing Research Center of ICT. He is a member of the IEEE and IEEE Computer Society.

Zimu Yuan received the PhD degree from University of Chinese Academy of Sciences in 2015. He is an associate professor at Institute of Information Engineering, Chinese Academy of Sciences (IIE, CAS). His research interests include wireless communication and localization, software reverse engineering and security, and building knowledge system for vulnerability detection. In wireless communication and localization, he proposed localization mechanisms for fusing results of multiple measurement techniques and placement method of anchor nodes for minimizing localization errors. He had more than 30 publications in international peer-reviewed proceedings and journals, including IEEE/ACM International Conference on Automated Software Engineering (ASE), IEEE International Conference on Software Analysis, Evolution, and Reengineering (SANER), IEEE Global Communications Conference (GLOBECOM), and *IEEE Communications Letters*.