*Research Article*

# A Distributed Anchor Node Selection Algorithm Based on Error Analysis for Trilateration Localization

**Yingsheng Fan ,[1,2] Xiaogang Qi ,[1,3] Baoguo Yu,[3] and Lifang Liu[4]**

[1]*School of Mathematics and Statistics, Xidian University, Xi'an, Shaanxi 71007, China*
[2]*Basic Courses Department, Zhejiang Police College, Hangzhou, Zhejiang 310053, China*
[3]*State Key Laboratory of Satellite Navigation System and Equipment Technology, Shijiazhuang 050000, China*
[4]*School of Computer Science and Technology, Xidian University, Xi'an, Shaanxi 710071, China*

Correspondence should be addressed to Xiaogang Qi; xgqi@xidian.edu.cn

This paper proposes a distributed anchor node selection algorithm based on error analysis for trilateration localization (EATL). The influence of distance measurement error on localization accuracy is discussed from two aspects: condition number of triangle formed by the three anchor nodes and the relative position between the unknown node and the three anchor nodes. Based on the error analysis, three principles for optimizing the selection of anchor nodes are given and then an algorithm for selecting anchor nodes on the ring is proposed.

## 1. Introduction

In a distributed sensor network, for most applications, such as target tracking, environmental monitoring [1], the geographical information of sensor nodes needs to be known. Estimation of node position is a fundamental requirement in distributed sensor networks. One possible solution is to install GPS receiver for each sensor node (or similar system, such as BeiDou Navigation Satellite System), but this scheme is limited by the characteristics of distributed sensor network itself. On the one hand, the cost of node with GPS system will be two orders of magnitude higher than that of ordinary node [2]. On the other hand, according to the different applications, sensor nodes are often deployed in the interior, city buildings, or even forest environment, and satellite signals are easily affected by many factors such as multipath interference and occlusion [3]. The localization accuracy is poor or even affecting its usability.

Cooperative localization [4, 5] is a new idea to realize high accuracy positioning in GPS denied environments. The basic idea is to use the following information to assist node localization, such as the distance information obtained from the communication between sensor nodes, the relative velocity information via Doppler shift measurement in the dynamic network. In large-scale distributed networks, due to the limited communication capability, nodes only can interact with their neighbors. All nodes will form a connected multihop network. According to the different application scenarios, coordinates here can be standard coordinates, such as latitude and longitude, or relative coordinate system. On the one hand, arranging a large number of anchor nodes is expensive. On the other hand, in some applications such as battlefield environments, anchor nodes only can be deployed around the network. This does not guarantee that all unknown nodes are adjacent to anchor nodes and obtain enough localization information. Consider an application in a two-dimensional network as shown in Figure 1; anchor nodes are arranged around the network, and unknown nodes are arranged in the network. Initially, anchor nodes broadcast the coordinate information. Due to the limitation of the communication radius, only a few unknown nodes can obtain enough localization information to estimate their own coordinates. The unknown node which has completed its own location will be an anchor node to assist other unknown nodes to be located. Through this kind of information
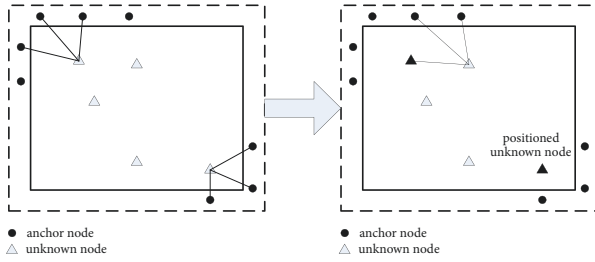
FIGURE 1: Cooperative localization.



FIGURE 2: The ideal model of Trilateration.

interaction between nodes, cooperative localization of the whole network will be completed.

Localization techniques of the distributed sensor network can be divided into two categories: range-based localization algorithm and range-free localization algorithm. This paper only considers range-based localization algorithm. This technique includes three main categories: Trilateration-based algorithm [6], Maximum Likelihood-based algorithm [7], and Multidimensional Scaling (MDS)-based algorithm. In Trilateration-based algorithm, unknown nodes measure the distances to three neighbor anchor nodes and then use this information to estimate their locations. The specific mathematical model will be given in the next section. When there are more than three neighbor anchor nodes, the unknown node should choose three to estimate its location. Maximum Likelihood-based algorithm is different from Trilateration-based algorithm. When the unknown node has more than three neighbor anchor nodes, it will utilize the information of all anchor nodes to construct an overdetermined equation and find the least square solution. The MDS algorithm based on distance measurement is divided into two main categories. The first category is classical MDS [8, 9]. This algorithm provides relative coordinate to a seemingly nonconvex localization problem, using only singular value decomposition. Within the classical MDS framework, the complete Euclidean distance matrix is needed, but this matrix is often very difficult to be obtained. The second category is to construct a pressure function [10–12] and use SMACOF [13] algorithm to minimize the pressure function and get the relative coordinate estimate. SMACOF algorithm is an iterative solution which will lead to high computational complexity. Although Trilateration- based algorithm and Maximum Likelihood-based algorithm need a certain number of anchor nodes, their computational complexity is low and they are suitable for cooperative localization of large-scale distributed networks.

The main distance measurement techniques of range-based localization algorithm are Received Signal Strength Indication (RSSI) [14], Time-of-Arrival (TOA) [15], and Ultrawideband (UWB) [16]. There is also a tradeoff between device cost and range accuracy. Using RSSI technique is cheap but the accuracy of measurement may be low, using TOA technique needs to guarantee time synchronization, and using UWB technique can achieve accuracies on the order of centimeters, but at the expense of high device and energy costs.
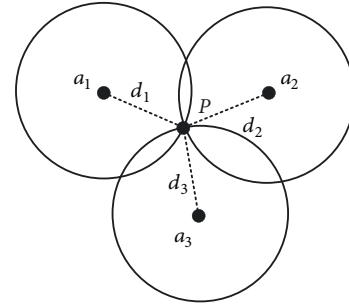
It should be noted that even in idealized setups with no obstacles or other external factors, relatively small error from noisy sensor measurements can induce much larger errors in node position estimate [17]. In Trilateration-based algorithms, effect of range measurement error on localization accuracy is mainly related to the selection of anchor node. In [6], an anchor node selection scheme based on the minimum condition number is proposed to improve the localization accuracy. However, this scheme only considers the influence of the distribution of three anchor nodes on localization accuracy. Actually, the relative position between the unknown node and the three anchor nodes will also have a great impact on localization accuracy. In Trilateration-based algorithms, when the unknown node completes the estimation of its location, it will become an anchor node to assist other unknown nodes to be located. In this process, the iterative error will be produced. Accumulation of measurement error will increase rapidly if the anchor nodes are not selected properly. Therefore, the selection of anchor nodes will have an important influence on localization accuracy.

The main contributions of this paper are as follows:

(1) The influence of distance measurement error on localization accuracy is discussed from two aspects: condition number of triangle formed by the three anchor nodes and the relative position between the unknown node and the three anchor nodes.

(2) Based on error analysis, an anchor nodes selection scheme (EATL) is proposed, which can effectively improve the localization accuracy.

The rest of the paper is organized as follows. The mathematical model of Trilateration localization is provided in Section 2 and error analysis is provide in Section 3. Based on the analysis in Section 3, an anchor nodes selection algorithm is given in Section 4. Simulation analysis is provided in Section 5, while Section 6 concludes the paper.

## 2. Mathematical Model of Trilateration Localization

We use $P(x, y)$ to represent the unknown node as shown in Figure 2; three anchor nodes within the communication radius of $P$ are $a_1(x_1, y_1)$, $a_2(x_2, y_2)$, and $a_3(x_3, y_3)$, respectively.
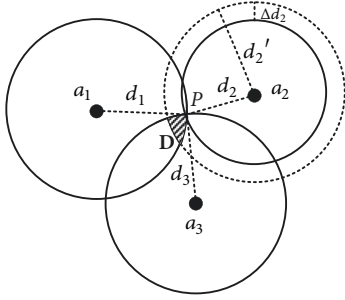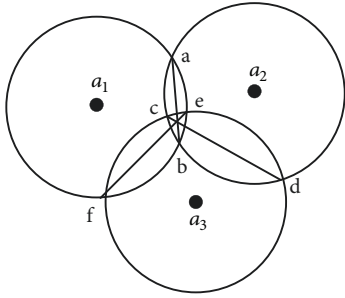
FIGURE 3: Distance measurement error of Trilateration.



FIGURE 4: The actual model of Trilateration.

The exact Euclidean distances without any noise between $P$ and $a_1$, $P$ and $a_2$, $P$ and $a_3$ are $d_1$, $d_2$, and $d_3$, respectively. The coordinates of $P$ can be obtained by solving (1).

$$(x - x_1)^2 + (y - y_1)^2 = d_1{}^2$$
$$(x - x_2)^2 + (y - y_2)^2 = d_2{}^2 \qquad (1)$$
$$(x - x_3)^2 + (y - y_3)^2 = d_3{}^2$$

We get (2) by using the third formula minus the first formula and the third formula minus the second formula in (1).

$$2(x_3 - x_1)x + 2(y_3 - y_1)y = \ldots$$
$$d_1{}^2 - d_3{}^2 + x_3{}^2 - x_1{}^2 + y_3{}^2 - y_1{}^2$$
$$\qquad (2)$$
$$2(x_3 - x_2)x + 2(y_3 - y_2)y = \ldots$$
$$d_1{}^2 - d_3{}^2 + x_3{}^2 - x_1{}^2 + y_3{}^2 - y_1{}^2$$

In practice, the measured distances are modeled as a noisy version of the actual node distances. For example, the actual distance between anchor node $a_2$ and $P$ is $d_2$, the measured distance is $d_2'$, and $\Delta d_2$ denotes the distance measurement error. In this case, the three circles intersect to form an area $D$ as the shaded area shown in Figure 3. In Figure 4, the three circles intersect with one another. From elementary geometry, we know that the lines $ab$, $cd$, and $ef$ will intersect at one point. When we solve (2), we get the coordinates of the intersection point actually.

TABLE 1: An example of ill-conditioned equation.

| | (x, y) | Actual Distance | Measurement Distance |
|---|---|---|---|
| Anchor 1 | (815.83,10.60) | 149.220 | 149.215 |
| Anchor 2 | (836.01,13.08) | 151.560 | 151.500 |
| Anchor 3 | (869.17,17.28) | 160.724 | 160.761 |

## 3. Error Analysis

When the distance measurement error $\Delta d$ exists, the coordinates of the unknown node $P$ obtained by solving (2) will also be inaccurate. Even if $\Delta d$ is small, the coordinates of $P$ obtained by inappropriate anchor node combination may also have great errors. See the example in Table 1; $(x, y)$ refers to the coordinates of the three anchor nodes, the third column is the actual distance between the unknown node and the anchor node, and the fourth column is the measurement distance with errors.

Based on actual distances, we can get the exact coordinates of $P$ as $(790, 157)$ by solving (2). But from measurement distances, the solution is $(760, 402)$. It can be seen from Table 1 that the differences between actual distances and measurement distances are negligible, but the differences of solutions, especially ordinates, are several hundred meters. Therefore, it is very important to study which factors will affect the solution of (2). To simplify the analysis, we rewrite (2) as follows:

$$AX = b \qquad (3)$$

where $A$ is a coefficient matrix, $X$ denotes the coordinates of $P$, and $b$ is a column vector. We have

$$A = \begin{bmatrix} 2(x_3 - x_1) & 2(y_3 - y_1) \\ 2(x_3 - x_2) & 2(y_3 - y_2) \end{bmatrix},$$

$$X = \begin{bmatrix} x \\ y \end{bmatrix} \qquad (4)$$

$$\text{and } b = \begin{bmatrix} d_1{}^2 - d_3{}^2 + x_3{}^2 - x_1{}^2 + y_3{}^2 - y_1{}^2 \\ d_1{}^2 - d_3{}^2 + x_3{}^2 - x_1{}^2 + y_3{}^2 - y_1{}^2 \end{bmatrix}.$$

$A$ is completely described by the coordinates of the three anchor nodes and can also be understood as the distribution of the three anchor nodes in the two-dimensional space. The elements in $b$ are completely determined by the coordinates of the three anchor nodes and the distances from the unknown node to the anchor nodes. That is, the relative position of the unknown node and the three anchor nodes will have a decisive influence on $b$. In this paper, the error analysis will be made from the distribution of three anchor nodes and the relative positions of the unknown nodes and the three anchor nodes. In Section 3.1.1 we discuss the situation when there is only distance measurement error; that is, we assume that the coordinates of the three anchor nodes $a_1$, $a_2$, and $a_3$ are accurate. In Section 3.1.2 we will discuss the situation when both the anchor nodes position error and

the distance measurement error exist. We assume that the distance measurement error $\Delta d$ is a random variable which obeys uniform distribution, $\Delta d \sim U(-\varepsilon, \varepsilon)$.

### 3.1. The Influence of the Distribution of Anchor Nodes on Localization Accuracy

#### 3.1.1. The Influence of Distance Measurement Error.
In this paper, we always assume that a matrix norm $\|\cdot\|$ is 2-norm; that is to say, $\|\cdot\|_2 = \sqrt{\lambda_{\max}(A^T A)}$. According to the property of matrix norm, we have $\|AB\| \leq \|A\| \cdot \|B\|$.

When the distance measurement error exists, the error is mainly reflected on the column vector $b$ in (3); $b$ can be rewritten as follows:

$$b' = \begin{bmatrix} (d_1 + \Delta d_1)^2 - (d_3 + \Delta d_3)^2 + x_3^2 - x_1^2 + y_3^2 - y_1^2 \\ (d_1 + \Delta d_2)^2 - (d_3 + \Delta d_3)^2 + x_3^2 - x_1^2 + y_3^2 - y_1^2 \end{bmatrix}, \quad (5)$$

and, thus, using $b'$, it follows that

$$A(X + \Delta X) = b + \Delta b \quad (6)$$

where $\Delta X$ is the error of solution and

$$\Delta b = b' - b$$
$$= \begin{bmatrix} (2\Delta d_1 \cdot d_1 - 2\Delta d_3 \cdot d_3) + (\Delta d_1^2 - \Delta d_3^2) \\ (2\Delta d_2 \cdot d_2 - 2\Delta d_3 \cdot d_3) + (\Delta d_2^2 - \Delta d_3^2) \end{bmatrix}. \quad (7)$$

Here we assume that the matrix $A$ is nonsingular and we can obtain the following.

$$\Delta X = A^{-1} (\Delta b) \quad (8)$$

From (3) and (8), we can get $1/\|X\| \leq \|A\|/\|b\|$ and $\|\Delta X\| \leq \|A^{-1}\| * \|\Delta b\|$. Then

$$\frac{\|\Delta X\|}{\|X\|} \leq \|A^{-1}\| * \|A\| \frac{\|\Delta b\|}{\|b\|} \quad (9)$$

where $\|\Delta X\|/X$ denotes relative error of solution, $\|\Delta b\|/\|b\|$ denotes relative error of column vector $b$, and $\|A^{-1}\| * \|A\|$ denotes the condition number of coefficient matrix $A$, where

$$cond(A) = \|A^{-1}\| * \|A\| \quad (10)$$

$\|A^{-1}\| * \|A\|$ and $\|\Delta b\|/\|b\|$ will determine the upper bound of $\|\Delta X\|/\|X\|$.

#### 3.1.2. The Influence of Distance Measurement Error and Anchor Nodes Position Error.
When both the anchor nodes position error and the distance measurement error exist, the error is mainly reflected in the coefficient matrix $A$ and the column vector $b$. Equation (6) should be rewritten as follows.

$$(A + \Delta A)(X + \Delta X) = b + \Delta \overline{b} \quad (11)$$

Here $\Delta \overline{b}$ is expressed as follows.

$$\Delta \overline{b} = \begin{bmatrix} (2\Delta d_1 \cdot d_1 - 2\Delta d_3 \cdot d_3) + (\Delta d_1^2 - \Delta d_3^2) + (2\Delta x_3 \cdot x_3 - 2\Delta x_1 \cdot x_1) + (\Delta x_3^2 - \Delta x_1^2) + (2\Delta y_3 \cdot y_3 - 2\Delta y_1 \cdot y_1) + (\Delta y_3^2 - \Delta y_1^2) \\ (2\Delta d_2 \cdot d_2 - 2\Delta d_3 \cdot d_3) + (\Delta d_2^2 - \Delta d_3^2) + (2\Delta x_3 \cdot x_3 - 2\Delta x_2 \cdot x_2) + (\Delta x_3^2 - \Delta x_2^2) + (2\Delta y_3 \cdot y_3 - 2\Delta y_2 \cdot y_2) + (\Delta y_3^2 - \Delta y_2^2) \end{bmatrix} \quad (12)$$

We assume that $\|\Delta A\|$ is small and can satisfy $\|A^{-1}\| \cdot \|\Delta A\| < 1$. If $\|\Delta A\|$ is very large, the coordinates error of the anchor node itself is very large. Substituting (3) in (11) yields

$$\Delta X = A^{-1} \left[ \Delta \overline{b} - \Delta A \cdot X - \Delta A \cdot \Delta X \right]. \quad (13)$$

Computing the norm we have

$$\|\Delta X\| \leq \|A^{-1}\| \left( \|\Delta \overline{b}\| + \|\Delta A\| \cdot \|X\| + \|\Delta A\| \cdot \|\Delta X\| \right). \quad (14)$$

And then

$$\left( 1 - \|A^{-1}\| \cdot \|\Delta A\| \right) \|\Delta X\|$$
$$\leq \|A^{-1}\| \left( \|\Delta \overline{b}\| + \|\Delta A\| \cdot \|X\| \right). \quad (15)$$

So we have the following.

$$\left( 1 - \|A^{-1}\| \cdot \|\Delta A\| \right) \frac{\|\Delta X\|}{\|X\|}$$
$$\leq \|A^{-1}\| \left( \frac{\|A\| \cdot \|\Delta \overline{b}\|}{\|b\|} + \|\Delta A\| \right) \quad (16)$$

From inequality (16), we obtain the following.

$$\frac{\|\Delta X\|}{\|X\|}$$
$$\leq \frac{\|A\| \cdot \|A^{-1}\|}{1 - \|A\| \cdot \|A^{-1}\| (\|\Delta A\| / \|A\|)} \left( \frac{\|\Delta \overline{b}\|}{\|b\|} + \frac{\|\Delta A\|}{\|A\|} \right) \quad (17)$$

In (17), $\|\Delta \overline{b}\|/\|b\|$ denotes relative error of column vector $b$ and $\|\Delta A\|/\|A\|$ denotes relative error of coefficient matrix
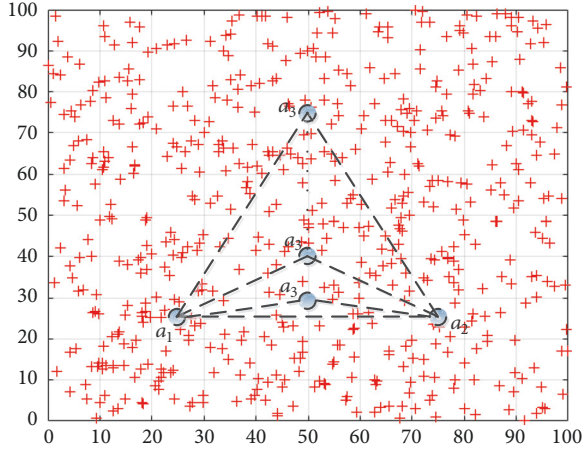
FIGURE 5: The distribution of three anchor nodes.



FIGURE 6: $\overline{\Delta'}$ for different condition number.

*A.* When $\|\Delta A\|$ is small, inequality (17) is similar to (9), and the upper bound of $\|\Delta X\|/\|X\|$ is determined by the condition number of coefficient matrix $A$ and $\|\Delta \overline{b}\|/\|b\|$.

*3.1.3. The Influence of cond(A) on Localization Accuracy.* From (9) and (17), the upper bound of $\|\Delta X\|/\|X\|$ is dependent on $\|A^{-1}\| * \|A\|$. In (1), we use the third formula minus the first and the second formula, respectively, and then get the coefficient matrix

$$A_{a_1} = \begin{bmatrix} 2(x_1 - x_2) & 2(y_1 - y_2) \\ 2(x_1 - x_3) & 2(y_1 - y_3) \end{bmatrix}. \tag{18}$$
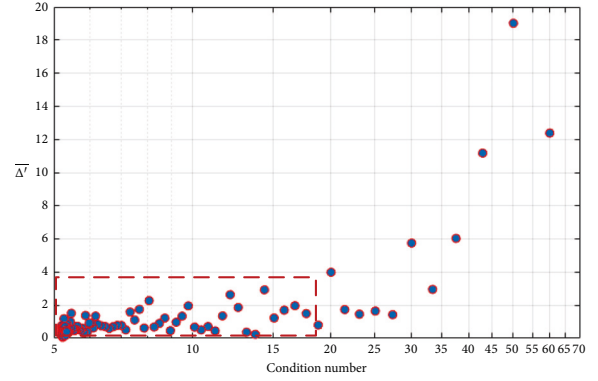
Similarly, we can also use the first formula minus the second and the third formula, respectively. Then we will get

$$A_{a_2} = \begin{bmatrix} 2(x_2 - x_1) & 2(y_2 - y_1) \\ 2(x_2 - x_3) & 2(y_2 - y_3) \end{bmatrix}$$

$$\text{and } A_{a_3} = \begin{bmatrix} 2(x_3 - x_1) & 2(y_3 - y_1) \\ 2(x_3 - x_2) & 2(y_3 - y_2) \end{bmatrix}. \tag{19}$$

$cond(A_{a_1}) + cond(A_{a_2}) + cond(A_{a_3})$ indicates the collinear degree of three anchor nodes. From the theory of ill-conditioned matrix, the greater the condition number of $A$ is, the more sensitive the linear equation with coefficient matrix $A$ is to $\Delta A$ and $\Delta b$. When three anchor nodes constitute an equilateral triangle, the sum of the condition numbers will obtain a minimal value 5.1963.

But in most cases, it is difficult to find three anchor nodes that just form an equilateral triangle. In order to illustrate the influence of the condition number on localization accuracy, we do the experiment as follows. As shown in Figure 5, 600 unknown nodes are randomly deployed in a two-dimensional area of size 100*100. The coordinates of anchor nodes $a_1$ and $a_2$ are (25, 25) and (75, 25), respectively. The x-coordinate of $a_3$ is 50, and the y-coordinate starts from 27, sampling every 0.5 meters, sampling 100 times, and observing the influence of the condition number of triangles formed by the three anchor nodes on localization accuracy. We assume that each

unknown node can communicate directly with any anchor node and distance measurement error $\Delta d$ obeys the uniform distribution $U(-2, 2)$.

The MAE of the location estimates is given by

$$\Delta_i = \sqrt{(x_i - \widehat{x}_i)^2 + (y_i - \widehat{y}_i)^2},$$

$$\overline{\Delta} = \frac{\sum_{i=1}^{N} \Delta_i}{N}, \tag{20}$$

where $N$ is the number of unknown nodes, $(x_i, y_i)$ are the exact coordinates of unknown node $i$, and $(\widehat{x}_i, \widehat{y}_i)$ are the estimation coordinates. $\Delta_i$ denotes the distance between the estimated position and the real position of unknown node $i$. $\overline{\Delta}$ denotes MAE.

To give a dimensionless form, we define $\Delta_i'$ as follows:

$$\Delta_i' = \frac{\Delta_i}{(1/3) \sum_1^3 |\Delta d_{ij}|},$$

$$\overline{\Delta'} = \frac{\sum_{i=1}^{N} \Delta_i'}{N}, \tag{21}$$

where $\Delta d_{ij}$ denotes the distance measurement error between unknown node $i$ and anchor node $j$. The relationship between $\overline{\Delta'}$ and the condition number is shown in Figure 6.

As shown in Figure 6, $\overline{\Delta'}$ has an increased trend with the increase of condition number, but it is not a completely monotonically increasing relationship. This shows that $\overline{\Delta'}$ is not completely determined by condition number of the triangle formed by the anchor nodes. When the condition number is less than 18, $\overline{\Delta'}$ is less than 3.6. In the process of computer simulation, we find that the condition number of the triangle formed by the three anchor nodes has the following function relationship with the degree of inner angles $(\alpha_1, \alpha_2, \alpha_3)$ of the triangle.

$$cond\left(A_{a_1}\right) + cond\left(A_{a_2}\right) + cond\left(A_{a_3}\right)$$

$$= \cot\left(\frac{\alpha_1}{2}\right) + \cot\left(\frac{\alpha_2}{2}\right) + \cot\left(\frac{\alpha_3}{2}\right) \tag{22}$$

FIGURE 7: The condition numbers of two isosceles right-angle triangles.
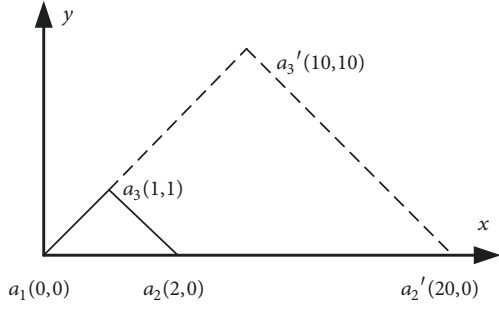


FIGURE 8: Two isosceles right-angle triangles.

From this equation, when the condition number is 18, the minimum interior angle is $13°$.

When the shape of a triangle is determined, its condition number is also determined. As shown in Figure 7, the condition number of isosceles right-angle triangle formed by $a_1$, $a_2$, and $a_3$ is 6.2361 and the condition number of isosceles right-angle triangle formed by $a_1'$, $a_2'$, and $a_3'$ is also 6.2361.

Although condition numbers of the two triangles are the same, their influences on error $\Delta X$ are different.

From

$$A_{a_1} = \begin{bmatrix} 2(x_1 - x_2) & 2(y_1 - y_2) \\ 2(x_1 - x_3) & 2(y_1 - y_3) \end{bmatrix}, \tag{23}$$

we obtain the following.

$$A_{a_1}^{-1}$$
$$= \frac{1}{2(x_3 - x_1)(y_3 - y_1) - 2(x_3 - x_2)(y_3 - y_2)} \tag{24}$$
$$\cdot \begin{bmatrix} (y_3 - y_2) & -(y_3 - y_1) \\ -(x_3 - x_2) & (x_3 - x_1) \end{bmatrix}$$

Substituting $a_1(0,0)$, $a_2(2,0)$, and $a_3(1,1)$ in (24) yields $A_{a_1}^{-1} = \begin{bmatrix} 0.2500 & -0.2500 \\ 0.2500 & 0.2500 \end{bmatrix}$, but when we use $a_1'(0,0)$, $a_2'(20,0)$, and $a_3'(10,10)$, it follows that $(A_{a_1}')^{-1} = \begin{bmatrix} 0.0250 & -0.0250 \\ 0.0250 & 0.0250 \end{bmatrix} = (1/10)A_{a_1}^{-1}$. In the triangle formed by three anchor nodes, the larger the shortest edge, the smaller the elements in $A^{-1}$.

Using the data in Table 1, we will get $A = \begin{bmatrix} 106.686 & 13.3518 \\ 66.3182 & 8.3877 \end{bmatrix}$ and $cond(A) = 1709.6$.

The large condition number indicates that the matrix is ill-conditioned and very sensitive to disturbances. Further we can obtain $A^{-1} = \begin{bmatrix} 0.8946 & -1.4241 \\ -7.0733 & 11.3787 \end{bmatrix}$. The large elements in $A^{-1}$ will lead to a large error of location estimate.

### 3.2. Influence of the Relative Positions of the Unknown Nodes and Anchor Nodes on Localization Accuracy.

In this section, we only discuss the situation in Section 3.1.1; that is, the coordinates of the anchor nodes are precise, and only distance measurement error exists. According to the expression of $\Delta \overline{b}$ in Section 3.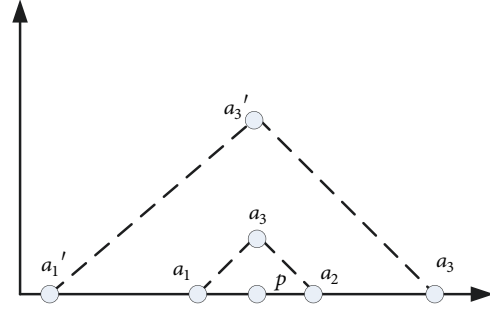1.2, when $\|\Delta A\|$ is small, the elements in $\Delta \overline{b}$ are mainly affected by the relative position of unknown node and the three anchor nodes. This is similar to the case of $\Delta b$.

When estimating the coordinates of $p$, there are two groups of anchor nodes: $a_1, a_2, a_3$ and $a_1', a_2', a_3'$ as shown in Figure 8. The triangle $a_1'a_2'a_3'$ is the expansion of the triangle $a_1a_2a_3$. We assume that line $a_1'a_2'$ is $N$ times as long as line $a_1a_2$.

In the case of fixed distance measurement error $\Delta d$, for triangle $a_1a_2a_3$, we have

$$\Delta b = \begin{bmatrix} (2\Delta d_1 \cdot d_1 - 2\Delta d_3 \cdot d_3) + (\Delta d_1^2 - \Delta d_3^2) \\ (2\Delta d_2 \cdot d_2 - 2\Delta d_3 \cdot d_3) + (\Delta d_2^2 - \Delta d_3^2) \end{bmatrix}$$

and $\Delta X$

$$= A^{-1} \tag{25}$$
$$\cdot \begin{bmatrix} (2\Delta d_1 \cdot d_1 - 2\Delta d_3 \cdot d_3) + (\Delta d_1^2 - \Delta d_3^2) \\ (2\Delta d_2 \cdot d_2 - 2\Delta d_3 \cdot d_3) + (\Delta d_2^2 - \Delta d_3^2) \end{bmatrix}.$$

For triangle $a_1'a_2'a_3'$, we have

$$\Delta b'$$
$$= \begin{bmatrix} N \cdot (2\Delta d_1 \cdot d_1 - 2\Delta d_3 \cdot d_3) + (\Delta d_1^2 - \Delta d_3^2) \\ N \cdot (2\Delta d_2 \cdot d_2 - 2\Delta d_3 \cdot d_3) + (\Delta d_2^2 - \Delta d_3^2) \end{bmatrix}$$

and $\Delta X'$

$$= A^{-1} \tag{26}$$
$$\cdot \begin{bmatrix} (2\Delta d_1 \cdot d_1 - 2\Delta d_3 \cdot d_3) + \dfrac{1}{N}(\Delta d_1^2 - \Delta d_3^2) \\ (2\Delta d_2 \cdot d_2 - 2\Delta d_3 \cdot d_3) + \dfrac{1}{N}(\Delta d_2^2 - \Delta d_3^2) \end{bmatrix}.$$

As $\Delta d \sim U(-\varepsilon, \varepsilon)$, the expectation and variance of $\Delta d$ are $E(\Delta d) = 0$ and $D(\Delta d) = \varepsilon^2/3$. The expectation of the elements in $\Delta X$ and $\Delta X'$ is 0. To compare the variance of the elements in $\Delta X$ and $\Delta X'$, we just need to compare $D(\Delta d_1^2 - \Delta d_3^2)$ and $D((1/N)(\Delta d_1^2 - \Delta d_3^2))$. As $D(\Delta d_1^2 - \Delta d_3^2) = (8/45)\varepsilon^4$ and $D((1/N)(\Delta d_1^2 - \Delta d_3^2)) = (8/(45 \cdot N^2))\varepsilon^4$, the variance of the elements in $\Delta X'$ is smaller than $\Delta X$; we select triangle $a_1'a_2'a_3'$ to estimate the coordinates of $p$.

In fact, when the positions of the three anchor nodes are determined (that is, matrix $A^{-1}$ is determined), the elements in $\Delta b$ are mainly determined by the relative positions of the unknown nodes and the three anchor nodes. In geometry, the Fermat point of a triangle, also called the Torricelli point or Fermat–Torricelli point, is a point such that the total distance from the three vertices of the triangle to the point is the minimum possible. $d_1$, $d_2$, and $d_3$ must be satisfied: $d_1 + d_2 + d_3 \geq d_{\min}$. $d_{\min}$ is the minimum total distance.

Since the value of $\Delta d_1{}^2 - \Delta d_3{}^2$ is small, we assume

$$\Delta b_{a3} = \begin{bmatrix} 2\Delta d_1 \cdot d_1 - 2\Delta d_3 \cdot d_3 \\ 2\Delta d_2 \cdot d_2 - 2\Delta d_3 \cdot d_3 \end{bmatrix}, \tag{27}$$

and in the same way we have

$$\Delta b_{a1} = \begin{bmatrix} 2\Delta d_2 \cdot d_2 - 2\Delta d_1 \cdot d_1 \\ 2\Delta d_3 \cdot d_3 - 2\Delta d_1 \cdot d_1 \end{bmatrix}$$

$$\text{and } \Delta b_{a2} = \begin{bmatrix} 2\Delta d_1 \cdot d_1 - 2\Delta d_2 \cdot d_2 \\ 2\Delta d_3 \cdot d_3 - 2\Delta d_2 \cdot d_2 \end{bmatrix}. \tag{28}$$

From $\Delta b_{a3}$ we have $\|\Delta b_{a3}/2\|_2{}^2 = (\Delta d_1 \cdot d_1 - \Delta d_3 \cdot d_3)^2 + (\Delta d_2 \cdot d_2 - \Delta d_3 \cdot d_3)^2$. In the same way we have $\|\Delta b_{a1}/2\|_2{}^2$ and $\|\Delta b_{a2}/2\|_2{}^2$.

Then we obtain the following.

$$E\left(\left\|\frac{\Delta b_{a1}}{2}\right\|_2^2\right) + E\left(\left\|\frac{\Delta b_{a2}}{2}\right\|_2^2\right) + E\left(\left\|\frac{\Delta b_{a3}}{2}\right\|_2^2\right)$$
$$= \frac{4}{3}\varepsilon^2 \left(d_1{}^2 + d_2{}^2 + d_3{}^2\right) \tag{29}$$

To get the minimum value of (29), the problem can be transformed into the following optimization problem.

$$\begin{aligned} \min \quad & d_1{}^2 + d_2{}^2 + d_3{}^2 \\ s.t. \quad & d_1 + d_2 + d_3 - d_{\min} \geq 0, \\ & d_1 \geq 0, \\ & d_2 \geq 0, \\ & d_3 \geq 0 \end{aligned} \tag{30}$$

The objective function of (30) is a convex function, and the inequality constraints are all linear functions, so the K-T point of (30) must be the global optimal solution. The K-T point can be obtained as $d_1 = d_2 = d_3 = d_{\min}/3$.

According to the above analysis, the distances between the three anchor nodes and the unknown node should be as similar as possible.

## 4. Anchor Node Selection Algorithm Based on Error Analysis

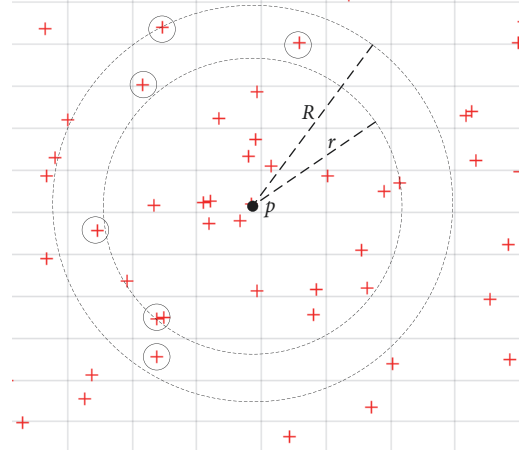*4.1. Design Principle.* The algorithm presented in this paper (EATL) abides by the following three principles:



FIGURE 9: Selection of anchor nodes.

(1) The minimum internal angle of the triangle formed by the three anchor nodes should be larger than 13°.

(2) The shortest edge of the triangle formed by the three anchor nodes should be as long as possible.

(3) The distances between the three anchor nodes and the unknown node should be as similar as possible.

Based on the above principles, we can select the anchor nodes on the ring centered on the unknown node. As shown in Figure 9, $p$ is the unknown node, nodes marked in red are anchor nodes, $R$ is communication radius, and $r$ is the inner radius of the ring.

We only select anchor nodes on the ring shown in Figure 9. On the one hand, it reduces the complexity of the algorithm. In Trilateration algorithms, if unknown node selects the optimal combination among all neighbor anchor nodes, $C_N^3$ calculations will be performed, where $N$ is the average number of neighbor anchor nodes. If the unknown node only selects anchor nodes on the ring, then $N$ will be reduced to $N' = N \cdot (R^2 - r^2)/R^2$. On the other hand, it ensures that the distances between the unknown node and the three anchor nodes are as similar as possible, which also satisfies principle (3).

The six anchor nodes marked with black circle are available for $p$ in Figure 9. Thus, 6 anchor nodes will have $C_6^3 = 20$ combinations. After narrowing down the selection range of anchor nodes, we will select 3 among these 6 anchor nodes according to principle (1) and principle (2).

For principle (1), the minimum internal angle of the triangle formed by the three anchor nodes should be larger than 13°. This principle is to reduce the collinearity of the three anchor nodes. In order to guarantee principle (2), we set the shortest side length threshold *threshold_a*, as shown in Table 2. Among all combinations satisfying the threshold, we select the maximum $d_{\min}(l)$, as shown in Table 2.

It should be noted that the inner radiuses $r$ and *threshold_a* are given in Section 5 through simulation experiments, where $r = 0.6R$, *threshold_a* $= 0.8R$.

*4.2. Symbol Description.* The main parameters and variables used in the algorithm are shown in Table 2.

TABLE 2: Explanation of parameters and variables.

| Notation | Description |
|---|---|
| $node1(i)$ | Unknown node $i$ |
| $R$ | Communication radius |
| $r$ ring | The ring shown in Figure 9 |
| $node1(i).neighbor$ | Neighbor anchor nodes set of unknown node $i$ on $r$ ring, recording distance and location information of anchor nodes |
| $N(i)$ | The number of neighbor anchor nodes of unknown node $i$ |
| $node1(i).flag1$ | Localization flag of unknown node $i$. The initial value is 0, update to 1 after localization |
| $node1(i).flag2$ | Record the time when the unknown node $i$ is located |
| $\theta_{\min}(l)$ | Minimum internal angle of the $l$th combination of anchor node, $l = 1 \cdots C^3_{N(i)}$ |
| $d_{\min}(l)$ | The shortest edge of the $l$th combination of anchor node, $l = 1 \cdots C^3_{N(i)}$ |
| $threshold\_a$ | The shortest edge threshold |
| $t$ | Waiting time threshold |
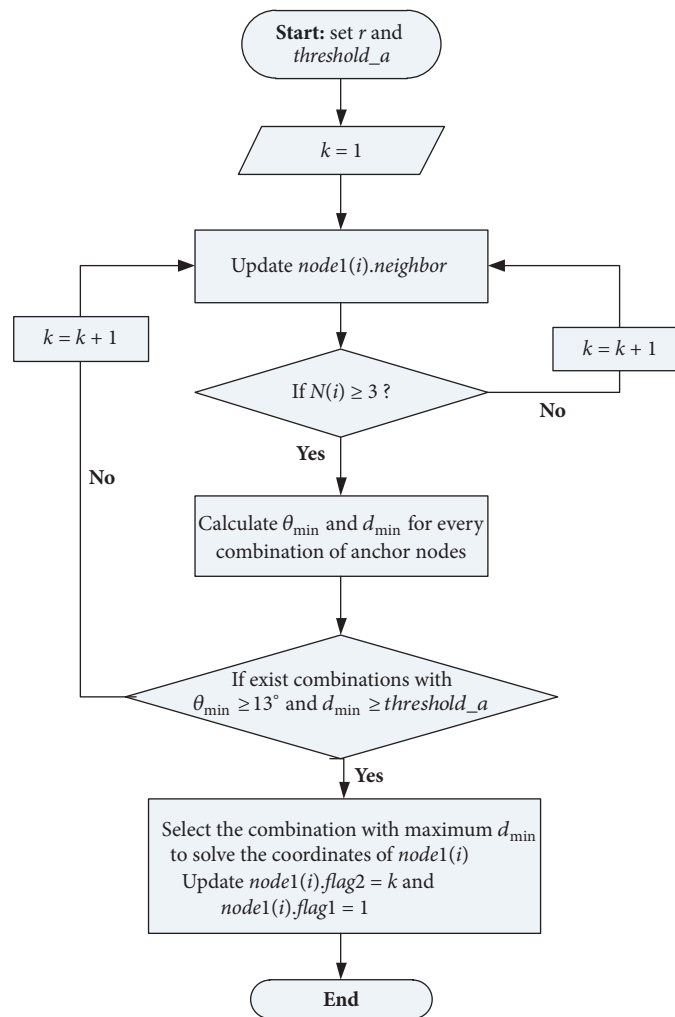


FIGURE 10: Algorithm flowchart.

*4.3. Algorithm Procedure.* In the beginning, initial anchor nodes broadcast their location information, and unknown nodes collect the information of neighbor anchor nodes and measure the distance.

The unknown node starts executing the anchor nodes selection algorithm. The algorithm flowchart is shown in Figure 10. After completing localization procedure, the unknown node updates $node1(i).flag1 = 1$ and becomes an anchor node. Then, it will broadcast its own location information. If the unknown node fails to complete the localization procedure, for example, no anchor node information is collected, or the anchor node information does not satisfy
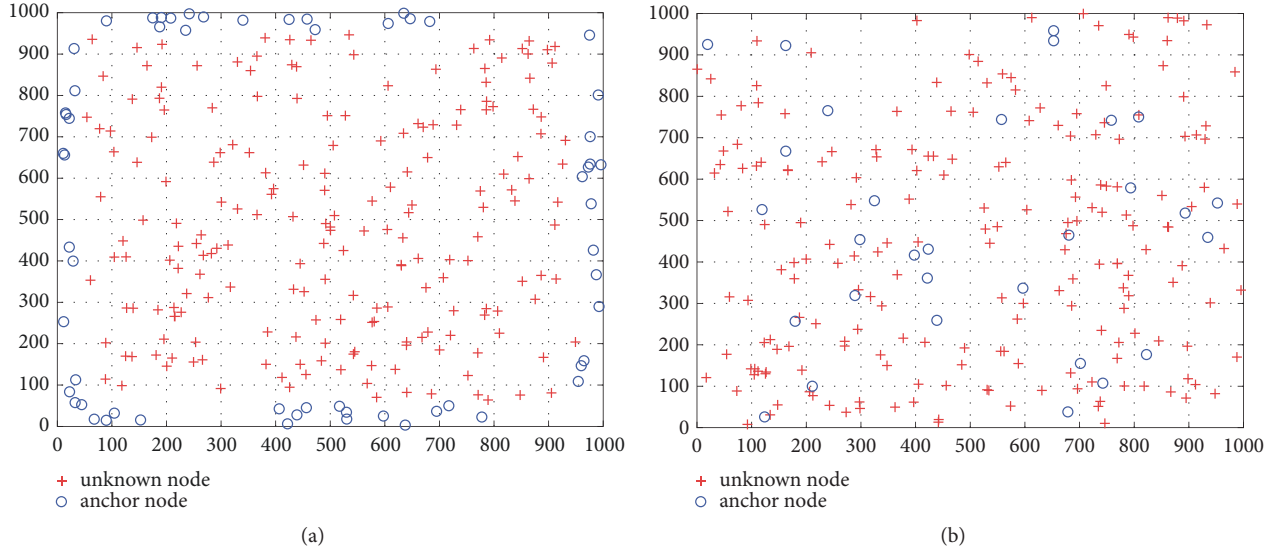
Figure 11: Two kinds of anchor nodes arrangement. (a) Around the square area. (b) In the square area.

requirement of the flowchart, then it will enter the waiting state and wait for receiving enough information to complete the localization procedure.

In a relatively sparse network, some unknown nodes may wait for a long time to find the anchor node information which can satisfy the requirement of the flowchart. Therefore, according to different applications, a waiting time threshold $t$ may be set. When reaching the waiting time threshold and the localization procedure is still not completed, the conditions in the flowchart can be properly relaxed. For example, the value of $r$ can be reduced. Of course this may reduce localization accuracy.

## 5. Performance Evaluation

*5.1. Simulation Setup.* Our experiments are run on various topologies of networks in Matlab R2017a. The 200 unknown nodes are placed randomly with a uniform distribution within a $1000*1000$ square area. The $m$ anchor nodes are placed (a) randomly with a uniform distribution around the square area as shown in Figure 11(a), or (b) randomly with a uniform distribution within the square area as shown in Figure 11(b). In order to observe the influence of network average connectivity on localization accuracy, we change communication radius $R$ between 200 and 300. We assume that $\Delta d \sim U(-1\%R, 1\%R)$. The performance of different algorithms is compared using mean absolute error (MAE) of the location estimates. We also calculate error bar defined by standard deviation to compare the stability of the algorithms. The standard deviation is given as follows.

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\Delta_i - \overline{\Delta}\right)^2} \tag{31}$$

The performance of the proposed algorithm (EATL) is compared with that of the Maximum Likelihood-based (ML)
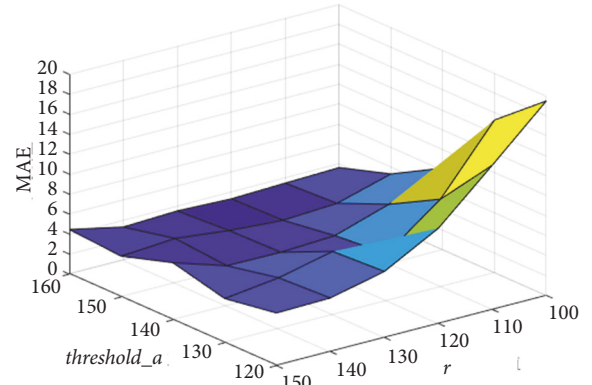


Figure 12: MAE varies with $r$ and *threshold_a*.

localization algorithm and the minimum condition number-based (FMMC) localization algorithm [6]. In ML approach, the unknown node to be located requires a minimum of seven neighbor anchor nodes. When comparing the performance of the three algorithms, we mainly do simulation on the network topology shown in Figure 11(a). For this type of network topology, the anchor nodes are placed around the network, the inside unknown nodes require more iterations to complete location which may result in greater iteration error. In many application scenarios, the anchor nodes can only be randomly placed around the network, such as the battlefield environment. It is more practical to simulate the network topology in Figure 11(a).

*5.2. r and threshold_a.* We set $m = 75$, $R = 200$, and $\Delta d \sim U(-2, 2)$. The network topology is shown in Figure 11(a). We change $r$ from $0.5R$ to $0.75R$, sampling every $0.05R$, sampling 6 times. Similarly, *threshold_a* is changed from $0.6R$ to $0.8R$, sampling every $0.05R$, sampling 5 times. Thus, There are a total of 30 combinations of $r$ and *threshold_a*. We do experiment to solve the MAE of every combination. The value of MAE varies with $r$ and *threshold_a* is shown in Figure 12.
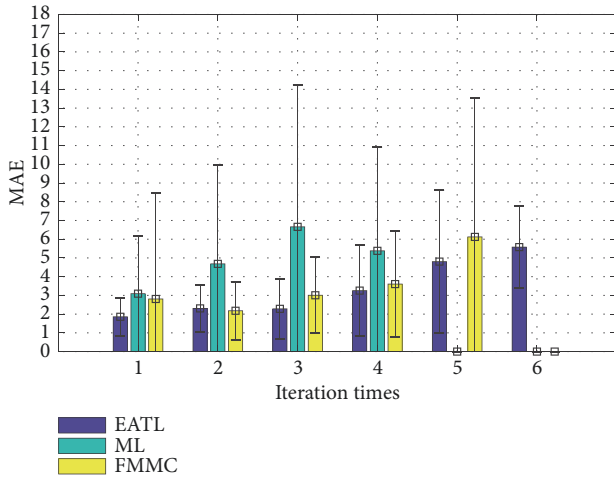
Figure 13: MAE for different iteration times.

From Figure 12, with the increase of $r$ and $thereshold\_a$, that is, the constraints becoming more and more stringent, the MAE is generally on a downward trend. For example, when $r = 0.5R$ and $threshold\_a = 0.6R$, $MAE = 19.47$. When $r = 0.6R$ and $threshold\_a = 0.8R$, $MAE = 3.62$. However, as $r$ increases, the area where unknown nodes select anchor nodes continuously decreases, which lead to a long waiting time of the unknown nodes to be located. The entire network requires more iteration times. When $r = 0.7R$, it takes more than 10 iteration times to complete the entire network localization. The increase in the number of iterations means the accumulation of iteration errors. From Figure 12, it can also be seen that when $r$ increases to $0.65R$, continued increase of $r$ does not significantly reduce the MAE, but the entire network localization time increases significantly. After balancing the localization time and localization accuracy, in the following simulations, if there are no special instructions, we take $r = 0.6R$ and $threshold\_a = 0.8R$.

*5.3. MAE Performance.* We set $m = 75$ and $R = 200$ to observe that MAE varies with iteration times of the three algorithms. The network topology is shown in Figure 11(a).

As shown in Figure 13, The MAE of all three algorithms tends to increase with the number of iterations, which is due to the accumulation of errors. According to the error bar of each iteration, the standard deviation of the proposed algorithm is smaller than that of ML and FMMC, and the proposed algorithm is more stable.

Overall, ML algorithm requires 4 iteration times to complete the localization and the MAE is 5.07. FMMC algorithm requires 5 iteration times and the MAE is 4.12. The proposed algorithm requires 6 iteration times and the MAE is 3.62. Compared with ML and FMMC, the MAE of proposed algorithm decreased by 28.6% and 12.1%, respectively.

*5.4. Impact of Network Connectivity.* We set $m = 75$ and the values of $R$ are 200, 225, 250, 275, and 300, respectively. The network topology is shown in Figure 11(a). Figure 14(a) shows

that, with the increase of the communication radius of nodes, the network connectivity increases. From Figure 14(b), we can see that, under different communication radius, the MAE and standard deviation of the proposed algorithm are lower than that of ML algorithm and FMMC algorithm. This shows that the proposed algorithm has good stability and scalability.

*5.5. Impact of Different Number of Anchor Nodes.* We set $R = 200$ and the values of $m$ are 20, 25, 30, 35, 40, 45, and 50, respectively. The network topology is shown in Figure 11(b). From Figure 15, with the increase in the number of anchor nodes, the MAE and standard deviation are gradually reduced. When the number of anchor nodes exceeds 30, the increase in the number of anchor nodes has no obvious effect on improving the localization accuracy of the network. When the number of anchor nodes in the network is less than 20, the process takes too long and the localization accuracy is low.

## 6. Conclusion

This paper proposes an anchor selection algorithm based on error analysis, starting from an example of ill-conditioned linear equation to show that selecting the right anchor nodes combination will make a big difference in localization accuracy. The influence of distance measurement error on localization accuracy is discussed from two aspects: condition number of triangle formed by the three anchor nodes and the relative position between the unknown node and the three anchor nodes. Then an algorithm of selecting anchor nodes on a ring is proposed. The values of $r$ and $thereshold\_a$ are given through simulation experiments. Simulation also shows that the performance of the proposed algorithm in MAE and standard deviation are better than those of ML algorithm and FMMC algorithm.

## Data Availability

The data used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments
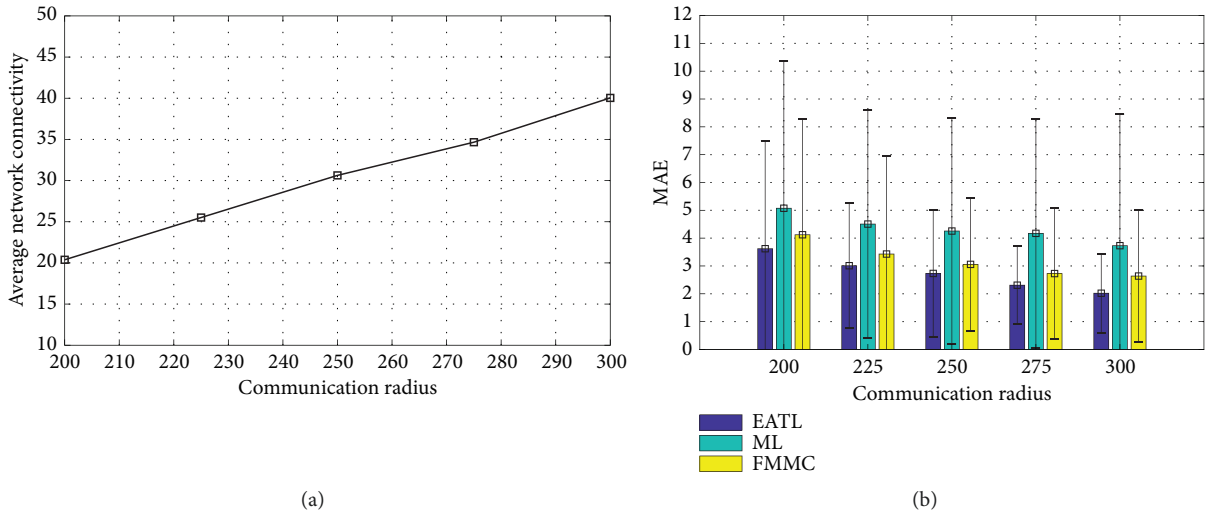
(a)



EATL
ML
FMMC

(b)

FIGURE 14: Impact of network connectivity. (a) Network connectivity for different communication radius. (b) MAE for different network connectivity.
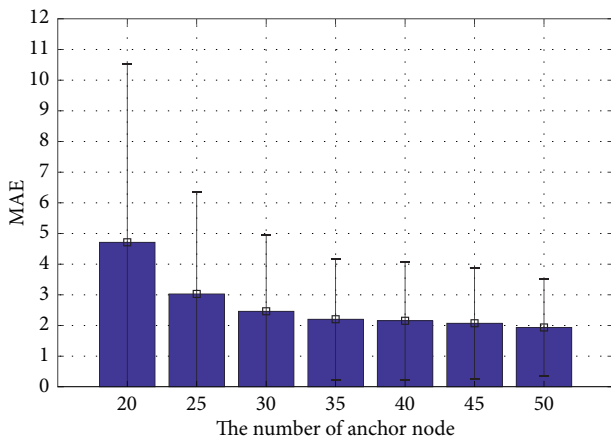


FIGURE 15: Impact of different number of anchor nodes.

College) for his valuable comments during the revision of the paper.

## References

[1] Z. Kashino, J. Y. Kim, G. Nejat, and B. Benhabib, "Spatiotemporal adaptive optimization of a static-sensor network via a non-parametric estimation of target location likelihood," *IEEE Sensors Journal*, vol. 17, no. 5, pp. 1479–1492, 2017.

[2] F.-B. Wang, L. Shi, and F.-Y. Ren, "Self-localization systems and algorithms for wireless sensor networks," *Journal of Software*, vol. 16, no. 5, pp. 857–868, 2005.

[3] P.-H. Tseng, Z. Ding, and K.-T. Feng, "Cooperative self-navigation in a mixed LOS and NLOS environment," *IEEE Transactions on Mobile Computing*, vol. 13, no. 2, pp. 350–363, 2014.

[4] L. Dong, "Cooperative localization and tracking of mobile Ad Hoc networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3907–3913, 2012.

[5] F. Yin, C. Fritsche, D. Jin, F. Gustafsson, and A. M. Zoubir, "Cooperative localization in WSNs using gaussian mixture modeling: Distributed ECM algorithms," *IEEE Transactions on Signal Processing*, vol. 63, no. 6, pp. 1448–1463, 2015.

[6] X. Du, L. Sun, F. Xiao, and R. Wang, "Localization algorithm based on minimum condition number for wireless sensor networks," *Journal of Electronics (China)*, vol. 30, no. 1, pp. 25–32, 2013.

[7] I. Guvenc, C.-C. Chong, and F. Watanabe, "Analysis of a linear least-squares localization technique in LOS and NLOS environments," in *Proceedings of the IEEE 65th Vehicular Technology Conference (VTC2007-Spring)*, pp. 1886–1890, April 2007.

[8] Y. Shang, W. Ruml, Y. Zhang, and M. Fromherz, "Localization from connectivity in sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 15, no. 11, pp. 961–974, 2004.

[9] H. Jamali-Rad and G. Leus, "Dynamic multidimensional scaling for low-complexity mobile network tracking," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 4485–4491, 2012.

[10] N. Saeed and H. Nam, "Cluster based multidimensional scaling for irregular cognitive radio networks localization," *IEEE Transactions on Signal Processing*, vol. 64, no. 10, pp. 2649–2659, 2016.

[11] J. A. Costa, N. Patwari, and A. O. Hero III, "Distributed weighted-multidimensional scaling for node localization in sensor networks," *ACM Transactions on Sensor Networks*, vol. 2, no. 1, pp. 39–64, 2006.

[12] S. Kumar, R. Kumar, and K. Rajawat, "Cooperative localization of mobile networks via velocity-assisted multidimensional scaling," *IEEE Transactions on Signal Processing*, vol. 64, no. 7, pp. 1744–1758, 2016.

[13] I. Borg and P. J. F. Groenen, *Modern Multidimensional Scaling, Theory and Applications*, Springer, New York, NY, USA, 2005.

[14] S. Tomic, M. Beko, R. Dinis, and P. Montezuma, "Distributed algorithm for target localization in wireless sensor networks using RSS and AoA measurements," *Pervasive and Mobile Computing*, vol. 37, pp. 63–77, 2017.

[15] N. Patwari, A. O. Hero III, 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.

[16] K. Guo, Z. Qiu, W. Meng, L. Xie, and R. Teo, "Ultra-wideband based cooperative relative localization algorithm and experiments for multiple unmanned aerial vehicles in GPS denied environments," *International Journal of Micro Air Vehicles*, vol. 9, no. 3, pp. 169–186, 2017.

[17] A. Savvides, W. L. Garber, R. L. Moses, and M. B. Srivastava, "An analysis of error inducing parameters in multihop sensor node localization," *IEEE Transactions on Mobile Computing*, vol. 4, no. 6, pp. 567–577, 2005.