A Two-Stage Range-Free Localization Method for Wireless Sensor Networks

Wen Yingyou, Li Zhi, Meng Yinghui, and Zhao Dazhe

1 Ministry of Education Key Laboratory of Medical Image Computing, Northeastern University, Shenyang 110004, China
2 Neusoft Research, Shenyang, Liaoning 110179, China

Correspondence should be addressed to Li Zhi; remcall@163.com

Received 21 October 2014; Accepted 12 January 2015

Academic Editor: Qiangfu Zhao

Copyright © 2015 Wen Yingyou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Range-free localization plays an important role in low-cost and large scale wireless sensor networks. Many existing range-free localization methods encounter high localization error, especially for the network with a coverage hole. One reason for high localization error is unreasonable distance estimation method. Another reason is that unknown nodes use the shortest distance which has large cumulative distance error to estimate their positions. In this paper, a two-stage centralized range-free localization algorithm (TCRL) is proposed. In the first stage, we design a novel rational distance estimation method to alleviate the distance estimation error between neighbor nodes based on the connectivity information and geometric features. In the second stage, a novel neighborhood function is derived from the estimated distances between neighbor nodes. Then a new localization strategy is proposed based on greedy idea. Finally, the proposed algorithm is compared with the same type algorithms in two network scenarios, namely, random deployment and random deployment with a coverage hole. The simulation results show that TCRL achieves more accurate and reliable results than most of existing range-free methods in the two network scenarios.

1. Introduction

A wireless sensor network (WSN) is composed of many battery-powered, low-cost sensor nodes deployed for generating specific event reports [1]. In many envisioned and existent applications, such as environmental monitoring, precision agriculture, and vehicle tracking, accurate location of sensor nodes is an indispensable requirement [2–4]. In addition, most routing algorithms of WSNs also require position information [5]. However, WSNs are often designed to operate in inaccessible regions and sensor nodes are deployed randomly, so the positions of sensor nodes are always uncontrollable. Although location awareness can be enabled in principle by the use of a global positioning system (GPS) or manual configuration, they are not suitable for the low-cost and large scale WSNs [6]. Sensor node localization has become an important area which attracts significant research interest.

A number of localization algorithms have been proposed recently for WSNs, which are generally classified into range-based and range-free localization schemes [7, 8]. For the range-based localization, all neighbor nodes must be able to measure the distances or angles between neighbor nodes using some measurement technique such as Time of Arrival (TOA) [9], Time Difference of Arrival (TDOA) [10], Received Signal Strength Arrival (RSSI) [11], and Angle of Arrival (AOA) [12]. The range-free localization only uses the connectivity information between neighbor nodes. In theory, the range-based methods can achieve more accurate estimated positions than the range-free methods. However, range-based methods always need some additional measuring devices. Although RSSI algorithm needs no extra hardware, the RSSI measurement is subject to negative effects of radio interference, obstacles, and individual differences of transmitters and receivers. So the range-based localization methods are not suitable for some low-cost and large scale WSNs.

In this paper, we focus on the range-free localization algorithms. Existing range-free algorithms encounter high localization error for WSNs, especially the network with a coverage hole. We explore most of existing range-free algorithms and find two reasons which lead to high localization error. One reason is that the distance estimation method is...
not good enough. This causes the distance estimation error which also contributes to the high localization error. Another reason is the unreasonable localization strategy. They always use the shortest distance between unknown nodes and anchor nodes to calculate the estimated positions by the trilateration or other methods [13]. Because the shortest distance path between an unknown node and an anchor node is often a zigzag one especially in the network with a coverage hole, compared with the direct line segment, it leads to high cumulative distance error and high localization error.

This paper studies how to alleviate the two problems mentioned above in existing range-free algorithms and then proposes a novel two-stage centralized range-free localization algorithm, called TCRL algorithm. Centralized localization algorithms mean nodes transmit data to a central site where the calculation is performed. Generally speaking, the centralized localization algorithms can get more positioning accuracy than the distributed algorithms because the central site can perform more complicated calculation. In the first stage, we analyze the geometric relationship between distance and intersection area of neighbor nodes and design a method to estimate more accurate distance between neighbor nodes only using the connectivity information and geometric features. In the second stage, we make a deep analysis on the estimated distance of neighbor nodes between two successive localization results and design a neighborhood function. Then a new localization strategy based on greedy idea is proposed. The localization strategy uses the neighborhood function and the estimated distance generated by the first stage between neighbor nodes to estimate the positions of unknown nodes. Because the localization strategy only uses the distance between neighbor nodes instead of the shortest distance, it can significantly reduce the cumulative distance error and localization errors. What is more, TCRL only uses connectivity information and introduces no extra hardware to measure distances between neighbor nodes, and it also does not need that sensor nodes are uniform deployed and the communication range is regular.

The rest of this paper is organized as follows: Section 2 summarizes related work. Section 3 describes TCRL algorithm including network model, distance estimation method, neighborhood function, and localization strategy. Section 4 tests the impact of various factors on TCRL and compares TCRL with the latest same type of algorithms to verify advantage and practicality of TCRL algorithm. Finally, we summarize this paper in Section 5.

2. Related Work

In range-free localization algorithms, nodes have no ability to measure distance or angle relative to their neighbor nodes. The range-free algorithm can be further divided into two categories: local techniques and hop-counting techniques.

In the local techniques, unknown nodes use the positions of anchor nodes to estimate their own positions. The paper [14] proposes the centroid algorithm. An unknown node estimates its position as the centroid of coordinates of its neighbor anchor nodes. A density-adaptive algorithm can reduce the localization error of the centroid algorithm if the anchor nodes are well deployed in paper [15]. The paper [16] proposes APIT algorithm. APIT divides the region into some small triangular regions between anchor nodes. Each unknown node estimates its position based on the center of gravity of the intersection of all the triangles that the unknown node resides in. The local range-free techniques always need a lot of anchor nodes and localization error is large.

The hop-counting techniques usually use some methods to estimate the distance between nodes. The paper [17] proposes DV-Hop algorithm. An unknown node can estimate its position by trilateration algorithm if it estimates its distances to three and more than three anchor nodes. DV-Hop algorithm suffers from the hop-distance ambiguity problem. Those nodes with the same hop count to an anchor node will have the same estimated distance, although they do have different distances to an anchor. The reason is that the value of hops between neighbor nodes is 1 independent of the distance between neighbor nodes. To resolve this problem, some improved DV-HOP algorithms are proposed. The paper [18] uses RSSI to differentiate nodes from the same hop-count. The paper [19] uses neighbor information of nodes. The paper [20] uses combination of RSSI and neighbor information. The paper [21] uses RSSI of a node to sort its 1-hop neighbor nodes by decreasing order to obtain a high-dimensional signature. They can achieve better result than DV-HOP, but they implicitly assume that all nodes are uniformly distributed in the deployed region or use RSSI. However, in practice, large-scale deployment of nodes with a random manner hardly guarantees such a global uniform distribution and the variability of RSSI in practical environment makes distance estimators hard to determine the models’ parameters. The paper [22] proposed a new proximity measure, named RND, to represent the distance between neighbor nodes and designed DV-RND. DV-RND uses RND to replace the hop between neighbor nodes and can get better localization accuracy than other peer classical range-free algorithms. However, DV-RND assumes that four anchor nodes are deployed at four locations close to the four corners of the field and the localization results are still not good enough, especially for the network with a coverage hole. The paper [23] proposes CMDS algorithm to improve the localization accuracy for the network with a coverage hole, but it requires that nodes can measure distance using RSSI which is always affected by appearance of sensor nodes and characteristics of the surrounding terrain.

3. TCRL Algorithm

TCRL consists of two stages. In the first stage, a novel distance estimation method is proposed to generate accurate estimated distance between neighbor nodes. In the second stage, a novel localization strategy, just using the estimated distance between neighbor nodes, is designed. In this section, we introduce network model and describe TCRL in detail.

3.1. Network Model. In this paper, we just consider the static WSNs. In order to make the algorithm more practical,
all nodes including unknown nodes and anchor nodes are randomly deployed in a 2-dimensional region. In addition, all nodes can only get connectivity information between neighbor nodes and do not need any additional hardware to measure RSSI or other pieces of information. Moreover, we make some assumptions which are not uncommon in many range-free localization schemes as follows: all nodes have the same communication radius, denoted by \( r \). All communication links between neighbor nodes are symmetric. Each node can get information of all its neighbor nodes. All sensor nodes are connected; that is to say at least one routing path exists between any pair of nodes.

Note that two nodes are neighbor nodes if and only if \( d_{ij} \leq r \), where \( d_{ij} \) is the Euclidean distance between two nodes. We use \( M_i = \{ j \mid j \neq i \text{ and } d_{ij} \leq r \} \) to denote the set of neighbor nodes of node \( i \). The network consists of \( n \) nodes, and there are \( m \) anchor nodes and \( n-m \) unknown nodes among them. Anchor nodes are aware of their coordinates. We use \( A \) to denote anchor nodes and \( U \) to denote unknown nodes.

### 3.2. The First Stage of TCRL

Distance estimation is of great importance for accurate localization. In the first stage, we design a novel distance estimation method only using connectivity information and geometric features between neighbor nodes. Figure 1 shows the distance model between two neighbor nodes \( i \) and \( j \).

In Figure 1, the black solid points are other neighbor nodes of nodes \( i \) and \( j \). We can observe that the distance \( d_{ij} \) between node \( i \) and its neighbor node \( j \) determines the size of the intersection area denoted by \( S_{ij} \), and \( S_{ij} \) is inversely proportional to \( d_{ij} \). \( S_{ij} \) can be calculated as follows:

\[
S_{ij} = 2r^2 \arccos\left(\frac{d_{ij}}{2r}\right) - d_{ij}\sqrt{r^2 - \frac{d_{ij}^2}{4}}. \tag{1}
\]

The ratio of \( S_{ij} \) and the communication area \( S \) of node \( i \) can be calculated by

\[
\frac{S_{ij}}{S} = \frac{S_{ij}}{\pi r^2} = \frac{2}{\pi} \arccos\left(\frac{d_{ij}}{2r}\right) - \frac{d_{ij}}{r}\sqrt{1 - \left(\frac{d_{ij}}{2r}\right)^2}. \tag{2}
\]

Equation (2) can also be written as

\[
y = \frac{S_{ij}}{\pi r^2}, \quad x = \frac{d_{ij}}{2r},
\]

\[
y = \frac{2}{\pi} \arccos(x) - \frac{2}{\pi} x \sqrt{1 - x^2}.
\tag{3}
\]

Using Taylor series expansion, \( \arccos(x) \) and \( x \sqrt{1 - x^2} \) can be written as

\[
\arccos(x) = \frac{\pi}{2} - x - \frac{1}{6} x^3 - \frac{3}{40} x^5 - \cdots,
\]

\[
x \sqrt{1 - x^2} = x - \frac{3}{8} x^5 - \cdots.
\tag{4}
\]

According to (3) and (4), (2) can be written as

\[
\frac{S_{ij}}{\pi r^2} = 1 - \frac{2}{\pi} \left(\frac{d_{ij}}{r}\right) + \frac{1}{12\pi} \left(\frac{d_{ij}}{r}\right)^3 + \frac{1}{320\pi} \left(\frac{d_{ij}}{r}\right)^5 + \cdots.
\tag{5}
\]

Because \( d_{ij} \) is the distance between neighbor nodes, the value range of \( d_{ij}/r \) must be \( 0 \leq d_{ij}/r \leq 1 \). From (5), we can observe that \( S_{ij}/\pi r^2 \) is mainly determined by \( 2(d_{ij}/r) \), so it is approximate linear function relationship between \( S_{ij}/\pi r^2 \) and \( d_{ij}/r \). When \( d_{ij}/r \) is denoted by \( p \), Figure 2 gives the corresponding value of \( S_{ij}/\pi r^2 \) denoted by \( q \) in (2).

Figure 2 also verifies that it is an approximate linear relationship between \( S_{ij}/\pi r^2 \) and \( d_{ij}/r \). In (2), when \( d_{ij}/r \) is 0, \( S_{ij}/\pi r^2 \) is 1. When \( d_{ij}/r \) is 1, \( S_{ij}/\pi r^2 \) is 0.391. We can get a linear function as follows:

\[
\frac{d_{ij}}{r} = \frac{1}{0.609} \left(1 - \frac{S_{ij}}{\pi r^2}\right). \tag{6}
\]

The value of \( S_{ij} \) is unknown; however, the density of sensor nodes is high for range-free wireless sensor networks and Figure 1 shows the area is direct ratio to the number of nodes, so \( S_{ij} \) can be estimated by

\[
S_{ij} = \sigma \cdot \frac{N_j}{N_i} \cdot \pi r^2, \tag{7}
\]
where $N_{ij} = |M_i \cap M_j| + 2$ is the number of nodes within the intersection area $S_{ij}$ and $N_i = |M_i| + 1$ is the number of nodes within the communication range of node $i$. $\sigma$ is a correction parameter to make (7) more accurate and the concrete value will be given in the following test. In practice, $N_{ij}$ and $N_i$ can be easily obtained by exchanging the neighbor information between nodes $i$ and $j$. Finally, we can get an important equation from (6) and (7):

$$NDR_{ij} = \frac{d_{ij}}{r} = \frac{1}{0.609} \left( 1 - \sigma \cdot \frac{N_{ij}}{N_i} \right), \quad (8)$$

where $NDR_{ij}$ is used to denote the neighbor distance relationship from node $i$ to its neighbor node $j$. However, because all nodes are randomly deployed, it is not uncommon that $N_i \neq N_j$. That is to say it is not uncommon that $NDR_{ij} \neq NDR_{ji}$. Since the bigger the $N_i$ is, the more accurate the estimated $S_{ij}$ in (7) is, we use the equation as follows to estimate $NDR$ between neighbor nodes $i$ and $j$ in this paper:

$$NDR_{ij} = NDR_{ji} = \frac{d_{ij}}{r} = \frac{1}{0.609} \left( 1 - \sigma \cdot \frac{N_{ij}}{\max(N_i, N_j)} \right), \quad (9)$$

where the function $\max(N_i, N_j)$ is used to take the maximum of $N_i$ and $N_j$. From the (9) we can estimate the distance between neighbor nodes $d_{ij}$ as follows:

$$\tilde{d}_{ij} = r \cdot NDR_{ij}. \quad (10)$$

However, because $NDR_{ij}$ is an estimated value and not accurate enough, the estimated distance is also not accurate. Hence, after we get the $NDR_{ij}$ between neighbor nodes, we can use the Floyd-Warshall shortest path algorithm to calculate the shortest NDR-path which is the one with the minimum value of all NDR-path between two anchor nodes. Then we can compute a NDR correction factor $\lambda_{NDR}$ as follows:

$$\lambda_{NDR} = \frac{\sum_{k=1}^{n} \sum_{s=1}^{m} d_{ks}}{\sum_{k=1}^{m} \sum_{s=1}^{m} \min NDR_{ks}}, \quad (11)$$

where $d_{ks}$ is the Euclidean distance between anchor nodes $k$ and $s$. $\min NDR_{ks}$ is the value of the shortest NDR-path between anchor nodes $k$ and $s$. $m$ is the number of anchor nodes. Finally, we can estimate the distance between two neighbor nodes as follows:

$$\tilde{d}_{ij} = \bar{d}_{ij} = \lambda_{NDR} \cdot NDR_{ij}. \quad (12)$$

3.3. The Second Stage of TCRL. From the first stage, we can get the estimated distances between neighbor nodes. The goal of localization algorithm is to estimate the positions of unknown nodes as accurate as possible. In existing range-free algorithms, the distance between neighbor nodes is obtained; an unknown node estimates the shortest-distances to more than three anchor nodes and then uses trilateration or other algorithms to estimate its position. However, the shortest path between two nodes is often zigzag, especially in a network with coverage hole as shown in Figure 3.

In Figure 3, $i$ is an unknown node and $j$ is an anchor node. The dotted line between $i$ and $j$ is the real distance, and the solid curve between $i$ and $j$ is the shortest distance path. We can see the error between the shortest distance and real distance is very large. The localization error must be large by using the shortest-distance. Considering this, we propose a novel localization strategy only using the distances between neighbor nodes instead of the shortest-distance between nodes which are not neighbor nodes.

We consider localization problem as a combinatorial optimization problem. The purpose of the proposed algorithm is to optimize the objective function $CF$ shown in

$$CF = \sum_{i=m+1}^{n} \sum_{j \in M_i} \left( \bar{d}_{ij} - \tilde{d}_{ij} \right)^2, \quad (13)$$

where $CF$ is the squared error between the estimated distances and corresponding correct distances of neighbor nodes. $M_i$ is the set of neighbor nodes of unknown node $i$. $\bar{d}_{ij}$ is the estimated distance obtained from the first stage. In this paper, we consider the estimated distance $\bar{d}_{ij}$ as the correct distance. $\tilde{d}_{ij}$ is estimated distance calculated by the estimated positions. $(\bar{x}_i, \bar{y}_i)$ and $(\bar{x}_j, \bar{y}_j)$ are the estimated positions of nodes $i$ and $j$. Note that if $j$ is an anchor node, $(\bar{x}_i, \bar{y}_i) = (x_j, y_j)$. Next, we will give the design of neighborhood function.

The purpose of neighborhood function is to generate a new solution from an old one. In this paper, a solution is a set of estimated positions of unknown nodes. We know the correct distance $\tilde{d}_{ij}$ obtained from the first stage. We assume that the estimated positions of two neighbor unknown nodes $i$ and $j$ in the new solution are correct, which are denoted by
International Journal of Distributed Sensor Networks

They must meet the following relationship:

\[
\begin{align*}
(x_i^{\text{new}} - x_i^{\text{new}})^2 + (y_i^{\text{new}} - y_j^{\text{new}})^2 &= d_{ij}^2 = \left[ \frac{d_{ij}}{d_{ij}} \right]^{2} \left( x_i^{\text{old}} - x_j^{\text{old}} \right)^2 + \left[ \frac{d_{ij}}{d_{ij}} \right]^{2} \left( y_i^{\text{old}} - y_j^{\text{old}} \right)^2, \\
\overline{d_{ij}} &= \sqrt{\left( x_i^{\text{old}} - x_j^{\text{old}} \right)^2 + \left( y_i^{\text{old}} - y_j^{\text{old}} \right)^2},
\end{align*}
\]

(14)

where \( (x_i^{\text{old}}, y_j^{\text{old}}) \) and \( (x_j^{\text{old}}, y_j^{\text{old}}) \) denote the estimated positions of nodes \( i \) and \( j \) in the old solution. \( \overline{d_{ij}} \) is the distance calculated by the old solution. From (14), we can obtain the relation of abscissa and ordinate as follows:

\[
\begin{align*}
x_i^{\text{new}} - x_i^{\text{old}} &= \frac{d_{ij}}{d_{ij}} \cdot (x_i^{\text{old}} - x_j^{\text{old}}), \\
y_i^{\text{new}} - y_i^{\text{old}} &= \frac{d_{ij}}{d_{ij}} \cdot (y_i^{\text{old}} - y_j^{\text{old}}).
\end{align*}
\]

(15)

We can see that (15) are the sufficient condition of (14). If an unknown node \( i \) meets (15) with all its neighbor nodes including unknown nodes and anchor nodes, we consider that the estimated position of unknown node \( i \) must be the correct position because the value of \( CF \) is 0 and meets the relationship:

\[
\text{correct position because the value of } CF \text{ is 0 and meets the relationship:}
\]

\[
(N1 + N2) \cdot U^{\text{new}} - NA \cdot A = (D1 + D2) \cdot U^{\text{old}} - D3 \cdot A,
\]

(17)

where the matrix \( A \) is the coordinate of anchor nodes and \( NA \) denotes the correct neighbor relationship between unknown nodes and anchor nodes. \( na_{ik} \) is 1 if an unknown node \( i \) and anchor node \( k \) are neighbor nodes, otherwise 0. \( U^{\text{old}} \) and \( U^{\text{new}} \), respectively, denote the estimated coordinates matrixes of all unknown nodes of old and new solutions. The parameter matrixes \( N1, N2, D1, D2, \) and \( D3 \) can be calculated as follows:

\[
\begin{align*}
N1_{ij} &= \begin{cases} 
nu_{ij} & i = j \\
-nu_{ij} & i \neq j,
\end{cases} \\
N2_{ij} &= \begin{cases} 
n申请 & i = j \\
0 & i \neq j,
\end{cases} \\
D1_{ij} &= \begin{cases} 
nu_{ij} \frac{d_{ij}}{d_{ij}} & i = j, \overline{d_{ij}} \neq 0 \\
0 & otherwise,
\end{cases} \\
D2_{ij} &= \begin{cases} 
nm \frac{d_{ik}}{d_{ik}} & i = j, \overline{d_{ik}} \neq 0 \\
0 & otherwise,
\end{cases} \\
D3_{ik} &= \begin{cases} 
nm \frac{d_{ik}}{d_{ik}} & \overline{d_{ik}} \neq 0 \\
0 & otherwise,
\end{cases}
\end{align*}
\]

(18)

where \( nu_{ij} \) is 1 if unknown nodes \( i \) and \( j \) are neighbor nodes, otherwise 0. From (17), we can get the neighborhood function as follows:

\[
U^{\text{new}} = \text{inv}(N1 + N2) \cdot \left[ (D1 + D2) \cdot U^{\text{old}} + (NA - D3) \cdot A \right].
\]

(19)

Equation (19) can generate a new solution from an old one. Through the preceding descriptive analysis, the new solution is not always better than the old solution, so we use the objective function \( CF \) to distinguish the quality of the two solutions. Next, we will use this neighborhood function to propose a localization strategy based on the greedy ideas.

We have introduced the objective function and the design of neighborhood function. Then, the localization strategy...
is proposed. Algorithm 1 is the structure of localization strategy. At the beginning, some known data is loaded. Such as the correct neighbor relationship between nodes and the estimated distances between neighbor nodes obtained from the first stage. Then, an initial solution $U^{\text{old}}$ (line 5) is generated. In the loop (lines 9–13), a local optimal solution $U^{\text{new}}$ is generated. Here, $\text{comCF}(U^{\text{old}})$ is used to calculate the value of CF of an old solution. $\text{generateSolution}(U^{\text{old}})$ is used to generate a new solution by an old one based on the neighborhood function. $\text{minDiff}(\text{minDiff} \geq 0)$ denotes the minimum difference between two CF values of old and new solutions. We use CF to distinguish the quality of the two successive solutions. Because the localization strategy is based on the greedy idea, we just accept the better solution (a solution is better than another one if and only if its value of CF is smaller). The value of $\text{minDiff}$ cannot be small enough. $\text{maxNum}$ is the maximum iteration number.

In order to get more accurate estimated positions, we add a correction operation $\text{correction}(U^{\text{old}})$ in line 15. We can use the estimated positions of unknown nodes and real positions of anchor nodes to calculate the estimated neighbor relationship between all nodes. If the estimated neighbor relationship of an unknown node is the same with the correct neighbor relationship, this unknown node is positioning accuracy in most cases and elevated to an anchor node. Otherwise it will be identified as a wrong localization node and need to be localized again. $\text{correction}(U^{\text{old}})$ is an iterative process. The steps of $\text{correction}(U^{\text{old}})$ are as follows.

Step 1. If the estimated neighbor relationship of an unknown node is correct, the unknown node is elevated to an anchor node. Find all unknown nodes that can be elevated to anchor nodes. The estimated positions of these nodes are the finally location coordinates. If there is no unknown node that can be elevated to an anchor node, go to Step 3. Otherwise, go to Step 2.

Step 2. A new set of anchor nodes and unknown nodes is generated after Step 1. Using the loop (line 5–13) and the new set of anchor nodes, we can generate another set of estimated positions of the new set of unknown nodes. Go to Step 1.

Step 3. Output the estimated positions of all unknown nodes.

From the description above, we can see that the localization strategy is a local search algorithm. The result is a local optimal solution. To make the result more accurate, we can perform the localization strategy for $p$ times (line 4) with different initial solutions and choose the final solution by the lowest value of CF.

4. Simulation and Results

In order to evaluate the performance of TCRL algorithm, many simulations are performed using MATLAB. We just discuss the comparison with the DV-HOP [17] and DV-RND [22] algorithms, because DV-HOP algorithm is the forerunner of range-free algorithm using hop-counting technique and DV-RND algorithm is the latest range-free algorithm.

We consider a $100 \times 100$ m deployment field with 200 sensor nodes. In order to make the test more close to the real environment, all sensor nodes including the unknown nodes and anchor nodes are randomly distributed and all nodes can only get connectivity information between neighbor nodes and do not need any additional hardware to measure RSSI or other pieces of information.

We compare their performances for distance estimation and node localization in two network scenarios: random network deployment shown in Figure 4(a) and random network deployment with a coverage hole shown in Figure 4(b). The coverage hole is $20 \times 60$ m. We also test various factors of WSNs and evaluate their influence on algorithms, such as the number of anchor nodes and different communication
range $r$. We simulate 100 different network deployments to obtain relatively fair results.

The test is divided into two sections. The first section is used to verify the effectiveness of distance estimation method of TCRL proposed by this paper in Section 3.2. The second is to verify the effectiveness of localization strategy of TCRL described in Section 3.3.

4.1. Comparison of Estimated Distance Error. In this section, we compare the estimated distance error of three algorithms, DV-HOP, DV-RND, and TCRL. To evaluate the performance of these algorithms, we define estimated distance error (EDE). EDE is the average absolute difference between the estimated distances and corresponding real inter-node distances:

$$EDE = \frac{1}{r} \sum_{i=0}^{n} \sum_{j \in M_i} \left| d_{ij} - \overline{d}_{ij} \right| \times 100\%,$$

where $|M_i|$ is the number of neighbor nodes of unknown node $i$; $d_{ij}$ is the estimated distance generated by the distance estimation method; $(x_i, y_i)$ and $(x_j, y_j)$, respectively, denote the true positions of nodes $i$ and $j$. In order to make the test more comprehensive, we test the impact of different communication range and number of anchor nodes on the EDEs. All results are averaged over 100 different network deployments.

The impact of communication range on EDE is shown in Figure 5.

In Figure 5, we set the number of anchor nodes to 20. The variation of communication range is from 13 m to 25 m. If the communication range is less than 13 m, sometimes the network is not connected. Figure 5(a) is the result of network without a coverage hole and Figure 5(b) is the result of network with a coverage hole.

In fact, all range-free localization algorithms for sensor network are sensitive to connectivity. Connectivity of sensor network is related to three main factors, namely, number of nodes, node distribution, and communication range of node. As range-free localization algorithm is usually used in low-cost and random deployed large scale wireless sensor networks, number of nodes and node distribution are not key factors of connectivity. Under these circumstances, communication range has a more important effect on localization algorithm because it determines the coverage density of the network which is essential for network connectivity.

Figure 5 shows that the EDEs of DV-RND and TCRL algorithms decrease as the communication range increases in the two network deployments. Connectivity improvement has a positive impact on these two algorithms. The EDEs of DV-HOP have little change as the communication range increases because increased communication range confused node when hop count is carried out. Figure 5 also shows that the distance estimation method of TCRL is always better than DV-HOP and DV-RND in both two network deployments independent of the communication range. It is observed that TCRL always outperforms the DV-HOP and DV-RND in EDE.

We also test the impact of the correction parameter $\sigma$ in (9) on the distance estimation method of TCRL algorithm. Figure 6 shows the results.

As mentioned above, although range-free localization algorithm is usually applied to large scale wireless sensor networks, number of nodes may not large enough and well-distributed to ensure an accurate estimation of intersection.

---

**Algorithm 1:** The structure of the localization strategy of TCRL.

\[
\begin{align*}
(1) & \text{localization} (\text{minDif}, \text{maxNum}, p) \\
(2) & \text{load some known data; finalSolution is null} \\
(3) & \text{minCF = inf} \\
(4) & \text{for i = 1; p} \\
(5) & \text{Unold = randomly generate an initial solution;} \\
(6) & \text{CFnew = comCF(Unold);} \\
(7) & \text{k = 0; Unew = Unold;} \\
(8) & \text{While k <= maxNum \&\& k <= maxNum) \\
(9) & \text{k = k + 1; CFnew = comCF(Unold);} \\
(10) & \text{Unold = generateSolution(Unold);} \\
(11) & \text{CFnew = comCF(Unnew);} \\
(12) & \text{end} \\
(13) & \text{end} \\
(14) & \text{Unold = comCF(Unold);} \\
(15) & \text{Unnew = correction(Unold);} \\
(16) & \text{CFnew = comCF(Unnew);} \\
(17) & \text{if minCF \geq CFnew} \\
(18) & \text{finalSolution = Unnew;} \\
(19) & \text{minCF = CFnew;} \\
(20) & \text{end} \\
(21) & \text{end}
\end{align*}
\]
area $S_{ij}$ using (7). Considering this, we adopt $\sigma$ as a correction parameter. From Figure 6, we can see that the EDEs of TCRL increase when $\sigma > 1$ and the EDEs of TCRL with $\sigma = 0.9$ are always smaller. This verifies the effectiveness of the correction parameter $\sigma$ and we set $\sigma = 0.9$ in subsequent test. Next, we test the impact of the number of anchor nodes on EDE, the results is shown in Figure 7.

In Figure 7, the communication range is 25 m, and $\sigma$ of TCRL is 0.9. The variation range of number is from 3 to 20. Figure 7(a) is the test results of network without a coverage hole and Figure 7(b) is the test results of network with a coverage hole. We can see that the EDEs of the three algorithms are almost invariant as the number of anchor nodes increases. It is also clearly observed that the distance...
estimation method of TCRL always has smaller EDEs than that of DV-HOP and DV-RND in both network deployments independent of the number of anchor nodes.

In general, no matter how much the communication range is and how many anchor nodes there are, TCRL algorithm always has smaller EDEs than DV-HOP and DV-RND in both network deployments. This test section verifies that the distance estimation method of TCRL described in Section 3.2 is always more accurate than DV-HOP and DV-RND, especially for the network with a coverage hole.

4.2. Comparison of Localization Error. In this section, we compare the localization error of four algorithms, DV-HOP, DV-RND, DV-NDR, and TCRL. Note that DV-NDR uses the estimated distances generated by the first stage of TCRL to calculate the estimated positions of unknown nodes by traditional trilateration. By comparison of DV-NDR and TCRL algorithms, we can verify the effectiveness of localization strategy of TCRL described in Section 3.3. To evaluate the performance of localization algorithms, we define localization error (LE) as follows:

\[
LE = \frac{1}{r(n-m)} \sum_{i=m+1}^{n} \sqrt{(x_i - \bar{x})^2 + (y_i - \bar{y})^2} \times 100\%.
\]  

(21)

Because the localization strategy of TCRL described in Section 3.3 is a local optimization algorithm, the positioning result has direct relation to the initial solution when the neighborhood function is fixed. Before comparing the performance of the four algorithms, we first test the impact of different types of initial solution on TCRL algorithm. In this paper, we design two types of initial solution. One type, denoted by same-position, is that the initial positions of all unknown nodes are in the same random position. The other type, which is denoted by different-position, is that all the initial unknown nodes are in the different random positions. Figures 8(a) and 8(b) show the results.

In Figure 8, we set that the number of anchor nodes is 20, \( r = 25 \) m, and \( \sigma = 0.9 \). We test 10 different initial solutions for each type of initial solution. Figure 8(a) shows the test results in random network deployment and Figure 8(b) shows the results in the network with a coverage hole. It is clearly seen that the positioning results generated by the same-position initial solution are always more accurate and stable than the different-position initial solution, especially for the network with a coverage hole. So we use the initial solution generated by the same-position way in later tests. Figure 8 also shows that the different initial solutions result in different Le's. In order to get more accurate positioning results, TCRL is performed \( p \) trials with \( p \) different initial solutions and generate \( p \) sets of different positioning results. Then we choose the final positioning results by the lowest value of CF. In this paper, the value of \( minDif \) is 0.001; \( maxNum \) is 2000 and \( p \) is 5 in Algorithm 1. Next, we will test the impact of different communication ranges on LE. Figures 9(a) and 9(b) show the results.

In Figure 9, we set the number of anchor nodes to 20 and \( \sigma = 0.9 \). The variation range of communication range is from 13 m to 25 m. Figures 9(a) and 9(b) show the Le's versus different communication range. Note that DV-NDR uses the estimated distances generated by the first stage of TCRL to calculate the estimated positions of unknown nodes by traditional trilateration. We can see that the Le's of four algorithms decrease as the communication range increases. Meanwhile, the Le's of TCRL are always lower than DV-HOP, DV-RND and DV-NDR, especially in network with
a coverage hole. Through the comparison of TCRL and DV-DNR, we can approve the effectiveness of localization strategy described in Section 3.3. Next, we will test the impact of the different number of anchor nodes on LE. Figures 10(a) and 10(b) show the results.

In Figure 10, the communication range is 25 m and $\sigma = 0.9$. Note that the variation range of the number of anchor nodes is from 4 to 20 in Figure 10. Because the LEs are very large when the number of anchor nodes is 3, we use Table 1 to show the results. We can see that the LEs are very large except TCRL.

Figure 10(a) shows that the LEs of the four algorithms decrease as the number of anchor nodes increases and the changing trend is very smooth in network deployment without a coverage hole. Figure 10(b) shows that the LEs of TCRL decrease as the number of anchor nodes increases and...
Table 1: The average computation time and average LE.

| Algorithm | Without a coverage hole | With a coverage hole |
|-----------|-------------------------|----------------------|
| DV-HOP    | 418.707                 | 399.594              |
| DV-RND    | 191.893                 | 144.012              |
| DV-NDR    | 164.932                 | 130.915              |
| TCRL      | 39.137                  | 37.421               |

In summary, TCRL algorithm always achieves more accurate positioning results than DV-HOP and DV-RND in the two random network scenarios independent of communication range and number of anchor nodes. Through the comparison of TCRL and DV-NDR, we can also verify the effectiveness of the localization strategy of TCRL described in Section 3.3.

5. Conclusions

In this paper, we explore existing range-free algorithms and find two reasons which lead to high localization error. In order to alleviate the two problems in existing range-free algorithms, we propose a novel two-stage centralized range-free localization algorithm, called TCRL. In the first stage, we analyze the relationship between distance and intersection area of neighbor nodes and approve that it is an approximate linear function between them. We use this linear function to get a new neighbor distance relationship NDR between neighbor nodes. We also calculate a NDR correction factor $\lambda_{NDR}$. Finally, we can obtain the estimated distance between neighbor nodes by $\lambda_{NDR}$ and NDR. Because the shortest distance path between an unknown node and an anchor node is often a zigzag one, especially in the network with a coverage hole, compared with the direct line segment, it leads to distance overestimation and high localization errors by trilateration method. Hence, in the second stage of TCRL, we make an in-depth analysis on the distance between neighbor nodes and design a novel neighborhood function. Then we propose the localization strategy of TCRL based on greedy idea. Using the estimated distance generated by the first stage of TCRL and the localization strategy described by the second stage of TCRL, we can get a set of accurate estimated positions of unknown nodes. Finally, we do a lot of experiments. To make the test scenario closer to the reality, all the sensor nodes are randomly deployed in a region and nodes can only get the connectivity information and do not need any additional hardware to measure RSSI or other pieces of information. We test the impact of different communication range and number of anchor nodes on the estimated distance errors and localization errors to separately verify the effectiveness of distance estimation method and localization strategy of TCRL. The test results show that TCRL algorithm always outperforms DV-HOP and DV-RND in the two network deployments, no matter how much the communication range is and how many anchor nodes there are, especially for the network deployment with a coverage hole.

Conflict of Interests

The authors declare that they do not have any commercial or associative interest that represents a conflict of interests in connection with the work submitted.
Acknowledgments

This work was supported by the Special Fund from the Central Collegiate Basic Scientific Research Bursary of China (Grant no. 110818001 and Grant no. 100218001) and in part by grants from the National Natural Science Foundation of China (Grant nos. 60903159 and 61173153).

References

[1] J. Yick, B. Mukherjee, and D. Ghosal, “Wireless sensor network survey,” Computer Networks, vol. 52, no. 12, pp. 2292–2330, 2008.
[2] J. Lee, W. Chung, and E. Kim, “A new range-free localization method using quadratic programming,” Computer Communications, vol. 34, no. 8, pp. 998–1010, 2011.
[3] X. Q. Yu, P. T. Wu, W. T. Han, and Z. L. Zhang, “A survey on wireless sensor network infrastructure for agriculture,” Computer Standards and Interfaces, vol. 35, no. 1, pp. 59–64, 2013.
[4] G. Han, J. Jiang, L. Shu, Y. Xu, and F. Wang, “Localization algorithms of underwater wireless sensor networks: a survey,” Sensors, vol. 12, no. 2, pp. 2026–2061, 2012.
[5] Q. Shi, C. Comaniciu, D. D. Wang, and U. Tureli, “Cross-layer MAC design for location-aware wireless sensor networks,” International Journal of Communication Systems, vol. 24, no. 7, pp. 872–888, 2011.
[6] A. S. Velimirovic, G. L. Djordjevic, M. M. Velimirovic, and M. D. Jovanovic, “Fuzzy ring-overlapping range-free (FRORF) localization method for wireless sensor networks,” Computer Communications, vol. 35, no. 13, pp. 1590–1600, 2012.
[7] M. Simek, P. Moravek, D. Komosny, and M. Dusik, “Distributed recognition of reference nodes for wireless sensor network localization,” Radioengineering, vol. 21, no. 1, pp. 89–98, 2012.
[8] L. Xue, B. Yang, X. Guan, and Z. Liu, “An interactive and energy-efficient node localization scheme for wireless sensor networks,” Wireless Personal Communications, vol. 69, no. 4, pp. 1481–1502, 2013.
[9] A. Boukerche, H. A. B. F. Oliveira, E. F. Nakamura, and A. A. F. Loureiro, “Localization systems for wireless sensor networks,” IEEE Wireless Communications, vol. 14, no. 6, pp. 6–12, 2007.
[10] M. S. Alam, S. Alsharif, and N. Haq, “Efficient CDMA wireless position location system using TDOA method,” International Journal of Communication Systems, vol. 24, no. 9, pp. 1230–1242, 2011.
[11] P. Moravek, D. Komosny, M. Simek, D. Girbau, and A. Lazaro, “Energy analysis of received signal strength localization in wireless sensor networks,” Radioengineering, vol. 20, no. 4, pp. 937–945, 2011.
[12] Y. S. Lee, J. W. Park, and L. Barolli, “A localization algorithm based on AOA for Ad-hoc sensor networks,” Mobile Information Systems, vol. 8, no. 1, pp. 61–72, 2012.
[13] A. Savvides, C.-C. Han, and M. B. Strivastava, “Dynamic fine-grained localization in Ad-hoc networks of sensors,” in Proceedings of the 7th Annual International Conference on Mobile Computing and Networking, pp. 166–179, Rome, Italy, July 2001.
[14] N. Bulusu, J. Heidemann, and D. Estrin, “GPS-less low-cost outdoor localization for very small devices,” IEEE Personal Communications, vol. 7, no. 5, pp. 28–34, 2000.
[15] N. Bulusu, J. Heidemann, and D. Estrin, “Adaptive beacon placement,” in Proceedings of the 21st IEEE International Conference on Distributed Computing Systems, pp. 489–498, Mesa, Ariz, USA, April 2001.
[16] T. He, C. D. Huang, B. M. Blum, J. A. Stankovic, and T. Abdelzaher, “Range-free localization schemes for large scale sensor networks,” in Proceedings of the 9th Annual International Conference on Mobile Computing and Networking (MobiCom ’03), pp. 81–95, San Diego, Calif, USA, September 2003.
[17] D. Niculescu and B. Nath, “Ad hoc positioning system (APS) using AOA,” in Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies, pp. 1734–1743, April 2003.
[18] S. W. Yang, J. Y. Yi, and H. J. Cha, “HCRL: a hop-count-ratio based localization in wireless sensor networks,” in Proceedings of the 4th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks, pp. 31–40, San Diego, Calif, USA, 2007.
[19] D. Ma, M. J. Er, and B. Wang, “Analysis of hop-count-based source-to-destination distance estimation in wireless sensor networks with applications in localization,” IEEE Transactions on Vehicular Technology, vol. 59, no. 6, pp. 2998–3011, 2010.
[20] B. Q. Huang, C. B. Yu, B. D. O. Anderson, and G. Q. Mao, “Connectivity-based distance estimation in wireless sensor networks,” in Proceedings of the IEEE Global Telecommunications Conference, pp. 1–5, IEEE, Miami, Fla, USA, December 2010.
[21] Z. Zhong and T. He, “RSD: a metric for achieving range-free localization beyond connectivity,” IEEE Transactions on Parallel and Distributed Systems, vol. 22, no. 11, pp. 1943–1951, 2011.
[22] G. Wu, S. Wang, B. Wang, Y. Dong, and S. Yan, “A novel range-free localization based on regulated neighborhood distance for wireless Ad hoc and sensor networks,” Computer Networks, vol. 56, no. 16, pp. 3581–3593, 2012.
[23] M. Shon, M. Jo, and H. Choo, “An interactive cluster-based MDS localization scheme for multimedia information in wireless sensor networks,” Computer Communications, vol. 35, no. 15, pp. 1921–1929, 2012.
Submit your manuscripts at
http://www.hindawi.com