WSNs node localization algorithm based on multi-hop distance vector and error correction

Ke Zhang · Guang Zhang · Xiuwu Yu · Shaohua Hu · Youcui Yuan

Accepted: 24 August 2022 / Published online: 21 September 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
Wireless sensor networks (WSNs) have broad application prospects in various industries, and node localization technology is the foundation of WSN applications. Recently, many range-free node localization algorithms have been proposed, but most of them suffer from low accuracy. In order to improve the localization accuracy, in this paper we proposed the node localization algorithm based on multi-hop distance vector and error correction (MDV-EC). In terms of distance estimation, firstly the MDV-EC algorithm calculates the neighbor distance according to node neighbor relationship, then estimates the distance between unknown node and anchor node in multi-hop manner, and finally calibrates the distance refer to distance correction coefficient. In view of similarity of localization errors of nodes in similar regions, an error correction scheme is also investigated, which corrects the node initial estimated locations of nodes refer to the localization error vector of nearby anchor node. Simulation results show that our proposed MDV-EC has better performance than the other two algorithms in terms of node localization accuracy, and the error correction scheme can effectively reduce the localization errors.

Keywords Wireless sensor networks (WSNs) · Range-free · Node localization · Error correction · Local similarity

1 Introduction
Wireless sensor networks (WSNs) are composed of many sensor nodes with wireless communication capability, which is a key technology for the Internet of Things (IoT) and has broad application prospects in many fields [1–3]. Node localization is one of the critical issues of WSNs system, and the location information of nodes is crucial for WSNs, as the sensor data without location information will become meaningless [4, 5]. While traditional localization technologies such as GPS can provide good location information, equipping all nodes with GPS chips will undoubtedly increase the cost of system deployment. In addition, sensor nodes have limited energy and usually cannot be charged in a timely manner. Equipping the nodes with GPS will consume a lot of energy and reduce the lifetime of the system. Therefore, in practice, only a few nodes called anchor nodes are equipped with GPS to obtain their locations, while others obtain their locations with the assistance of localization algorithms and anchor nodes [6].

In recent years, researchers have proposed many node localization algorithms, which can be roughly divided into two categories: range-based and range-free localization algorithms. The ranging-based localization algorithm uses ranging techniques to obtain the distance between nodes and estimate the position based on the coordinates of the anchor nodes and distance estimation information. Popular ranging techniques, such as time of arrival (TOA), time difference of arrival (TDOA), angle of arrival (AOA) and received signal strength (RSS) [7]. Among them, TOA, TDOA and AOA have good ranging accuracy, but these methods require additional hardware, which raises the system deployment cost and is not suitable for large-scale deployment. RSS-based ranging is used to estimate the distance based on the propagation attenuation characteristics of wireless signals in the environment. This approach has attracted the attention of many researchers as it does not require additional hardware and has the advantage of low cost. However, the robustness of RSS-based ranging technique is poor, and its ranging
accuracy is easily affected by environmental factors. The range-free algorithm investigates the network connectivity information to estimate the node locations directly or obtains the distance between nodes based on connectivity information (e.g., hop count) to further obtain the estimated node locations. Compared to range-based localization algorithms, range-free algorithms are less affected by the environment, more robust, and do not require additional hardware, making it more adaptable to large-scale network deployment environments. Among the range-free localization algorithms, distance vector-based localization is one of the most popular approaches, but the method still has the disadvantage of low localization accuracy. For the distance vector node localization algorithm, the estimated distance between nodes is related to the network connectivity of the path region, and the neighboring nodes have similar network connectivity, which leads to the similarity of the neighboring node localization errors. From this aspect, in order to improve the node localization accuracy, we proposed the MDV-EC in this paper. The MDV-EC algorithm consists of two parts: initial location estimation and error correction. In the initial position estimation phase, the distance between neighboring nodes is first estimated based on the neighbor relationship, the estimated distance between the unknown node and the anchor node is obtained by exploring the shortest path between nodes, and finally, the initial estimated position of the node is obtained by the least squares method. In the error correction phase, all anchor nodes calculate their estimated positions according to the node position estimation method, and the error vector of each anchor node is obtained with their estimated and real positions. Given the similarity of the error vectors, the error vectors of the anchor nodes are used to correct the initial estimated locations of the unknown nodes around the anchor nodes, further reducing the localization errors. Simulation results show that the localization accuracy of our proposed MDV-EC algorithm outperforms two representative distance vector localization algorithms, DV-Hop and DV-RND, in different scenarios.

The rest of this paper is organized as follows. Section 2 provides an overview of related works, Sect. 3 illustrates the network model and problem statement, Sect. 4 describes our proposed MDV-EC localization algorithm, Sect. 5 evaluates the performance of the proposed algorithm, and the last section concludes the paper and discusses future research directions.

2 Related works

For range-free algorithms, the most popular range-free algorithms are the centroid algorithm [8], approximate triangular point (APIT) [9], and distance vector hopping (DV-Hop) [10]. Among them, DV-Hop algorithm has received wide attention due to its high robustness and simplicity. The DV-Hop algorithm is a typical distance vector node localization algorithm. The basic idea is that firstly, the average hop distance in the network is calculated based on the anchor nodes and network connectivity, secondly, the unknown node gets its distance from each anchor node based on the number of hops and the average hop distance, and finally, the estimated position of the unknown node is obtained by the least squares method. Although the DV-Hop algorithm has the advantages of computational simplicity and environmental adaptability, its localization accuracy is far from satisfactory. Therefore, many researchers have proposed a number of improved algorithms for DV-Hop. Some researchers tried to enhance the localization performance by improving the accuracy of the average hop distance estimation and hop count estimation. Shahzad et al. in [11] proposed the DV-MaxHop algorithm. The DV-MaxHop defines the maximum number of hops to constrain the information propagation in the network, discards the information of anchor nodes with larger hops, obtains more accurate average hop distance, and improves the accuracy of localization. Compared with DV-Hop, the accuracy of DV-MaxHop algorithm has been improved, but its positioning error is still unsatisfactory. In order to improve the accuracy of hop estimation, paper [12] and [13] proposed an RSS-based hop estimation method, which uses RSS information to serialize the hop values and avoid the large errors of integer hop numbers. However, the accuracy of this hop estimation method is susceptible to environmental factors. Some researchers argue that the reliability of different anchor nodes is different during localization, so by referring to some reliable anchor nodes, better localization accuracy can be obtained. From this point of view, Cao et al. proposed the dynamic anchor nodes set algorithm (DANS) in [14]. The DANS algorithm transforms the anchor selection problem into a combinatorial optimization problem and solves it using a binary particle swarm algorithm to obtain the set of reliable anchor points. After that, the unknown node calculates the best estimated coordinates based on the information of the reliable anchor nodes. Chen et al. proposed an enhanced DV-Hop algorithm in [15]. First, the algorithm abstracts the relationship between nodes in the network as a spring model and corrects the distance between the unknown node and the anchor node by considering the combined influence of each node. After that, reliable anchor nodes are selected according to the random sampling consensus, and the coordinates of the unknown nodes are calculated using the reliable anchor nodes. To accommodate anisotropic node deployment, in [16] Zhang et al. proposed a boundary-based anchor selection scheme, which determines the reliability of anchor nodes based on the number of boundary nodes contained in the path, and selects reliable anchor nodes to provide information for localization. Although the above methods can identify reliable anchor
nodes in the network to provide information for localization and avoid the negative effect of some unreliable anchor nodes on localization accuracy, they may fail when the number of anchor nodes is small. With the emergence of heuristic algorithms, some researchers have tried to use heuristic optimization algorithms such as differential evolution (DE) [17], grey wolf optimization (GWO) [18], particle swarm optimization (PSO) [19, 20] and so on to solve for the estimated coordinates of nodes. In addition to the heuristic optimization approach, Jin et al. proposed an iterative localization algorithm CVX-DV-hop in [21]. The CVX-DV-hop transforms the location estimation problem into a problem with convex functions and nonconvex constraints. Then, designed a continuous convex approximation method to solve the optimization problem iteratively. Though the solution accuracy of the above methods is greatly improved compared with the least squares method, they are too computationally intensive and their localization accuracy is limited by the distance estimation accuracy. To further improve the node localization accuracy, researchers proposed some multi-hop distance estimation methods based on neighbor connectivity. In [22], Wu et al. proposed a new distance vector localization algorithm called DV-RND, which estimates the distance of each hop based on the connectivity of neighbors, obtains the estimated distance between the unknown node and the anchor node based on the multi-hop path, and finally calculates the coordinates of the node. Wang et al. in [23] proposed an expected hop estimation method based on node density to improve the accuracy of per-hop distance estimation. Wu et al. [24] applied the expected hop progress node localization algorithm in anisotropic sensor deployment scenario. The above-mentioned works improve the accuracy of distance estimation and reduce the localization error, but most of them ignore the error similarity introduced by distance vector localization methods. The algorithm first explores the neighbor relationship to improve the multi-hop distance estimation accuracy, and then makes full use of the error similarity to correct the estimated coordinates and further reduce the localization error. Next section we will introduce the network model and problem statements.

### 3 Network model and problem statements

In the real deployment scenario, the communication coverage area of nodes is not a standard circle considering environmental factors. In order to simulate the signal propagation in real scenarios, the radio irregularity model is used in this paper to simulate the communication coverage of sensor nodes in the network [13]. According to the radio propagation model, nodes have different signal strengths in different directions, and the signals in adjacent directions have a certain correlation. The correlation of adjacent direction signals is determined by the degree of irregularity (DOI). Dividing the space into 360 directions, the signal strength of each direction can be calculated by Eq. (1).

\[
P_r = P_t - P(d_0) - 10 \times \eta \times \log_{10}\left(\frac{d_i}{d_0}\right) \times K_i
\]

where \(P_r\) represents the signal strength received by the node, \(P_t\) represents the signal strength of the message sent, \(\eta\) represents the signal attenuation constant (path loss exponent), \(d_0\) represents the reference distance, \(P(d_0)\) is the received signal strength at the reference distance \(d_0\) and \(K_i\) is the direction parameter of DOI, which is calculated by Eq. (2).

\[
K_i = \begin{cases} 
1, & \text{if } i = 0 \\
K_{i-1} + \text{rand}_i & \text{else, } i \in [1, 2, \ldots, 360]
\end{cases}
\]

where \(\text{rand}_i\) is the direction random parameter, \(\text{rand}_i \in [-\text{DOI}, \text{DOI}]\). The DOI value controls the stability of the signal; a larger value means a more unstable signal and a smaller value means a more stable signal. When DOI is equal to 0, the coverage area of the signal is a standard circle. To describe the model more specifically, we set the communication radius \(R = 30\), and plot the signal coverage areas with DOI = 0.01 and DOI = 0.02, respectively, as shown in Fig. 1.

In this paper, we assume that WSNs system satisfy the following conditions.

1. All nodes are randomly and uniformly arranged in deployment area, forming a connectivity graph in which any two nodes can communicate in a single-hop or multi-hop manner.
2. Some of the nodes in the network have determined and known locations, and they are called anchor nodes.
3. All nodes in the network have the same communication radius \(R\). The communication model of all nodes is the DOI model.
4. The connectivity of any two nodes in the network is symmetric, i.e., when node \(i\) can communicate directly with node \(j\), then node \(j\) must be able to communicate directly with node \(i\).

The node localization problem in this paper is to obtain the geographic location of all sensor nodes with the assistance of anchor nodes and the network connectivity information.

### 4 Proposed node localization algorithm

#### 4.1 Neighbor distance estimation

As shown in Fig. 2, node \(u\) and \(i\) are two neighboring nodes, and their communication radius is \(R\). The common communication coverage area is the light blue part. In a network with uniform random deployment of sensor nodes, the closer two
adjacent nodes are, the larger their communication common coverage area is and the more nodes fall into the common coverage area. Similarly, the farther two nodes are from each other, the smaller their communication common coverage area is and the fewer nodes fall into the common coverage area. As a result, we can estimate the distance of two adjacent nodes by exploring their neighbor information.

In random uniform deployment network, according to the geometric relationship, the distance $d_{ui}$ between two neighboring nodes $u, i$ is a function of the common coverage area $A(d_{ui})$, as shown in Eq. (3) [22, 25].

$$A(d_{ui}) = 2R^2 \arccos\left(\frac{d_{ui}}{2R}\right) - \frac{d_{ij}}{2} \sqrt{4R^2 - d_{ui}^2} \tag{3}$$

Let $x = d_{ui}/2R$ and bring it into Eq. (3) to obtain the following equation.

$$\frac{A(d_{ui})}{R^2} = 2 \arccos(x) + 2x \sqrt{1 - x^2} \tag{4}$$

The Taylor expansion of the right part of the Eq. (4) at $x = 0$ gives formula (5).

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

$$x \sqrt{1 - x^2} = x - \frac{1}{2}x^3 - \frac{1}{8}x^5 \cdots \tag{5}$$

Bring Eq. (5) into Eq. (4), we get Eq. (6).

$$\frac{A(d_{ui})}{R^2} = \pi - 4x + \frac{2}{3}x^3 + \frac{1}{10}x^5 \cdots \tag{6}$$

Since nodes $u$ and $i$ are each other’s neighbor nodes, their distance $d_{ui}$ is smaller than the communication radius $R$. Because $x \leq d_{ui}/2R$, we have $x \leq 0.5$. Since $x$ is less than 0.5, the value of the higher order term of $x$ is very small. Therefore, ignoring the $x^3$ and higher terms in Eq. (6), taking $x = d_{ui}/2R$ into Eq. (6) and simplifying it, we obtain Eq. (7) as follows.

$$d_{ui} = \frac{\pi R^2 - A(d_{ui})}{2R} \tag{7}$$

According to Eq. (7), we can estimate the neighbor node distance $d_{ui}$ based on the communication radius $R$ and the common coverage area $A(d_{ui})$. The node communication radius $R$ is predefined, and the common coverage area $A(d_{ui})$ can be estimated from the neighbor node relationship. Next, we estimate the common coverage area $A(d_{ui})$ of the communication range of the two neighboring nodes $u, i$ based on their neighbor relationship. The node density of $A(d_{ui})$ can be obtained by dividing the number of common neighbor nodes of $u, i$ by the area of $A(d_{ui})$. Similarly, the node
density of the communication coverage area can be obtained by dividing the number of neighbor nodes by the area of the node communication range. Since the nodes are randomly and uniformly arranged, the node density is the same in all regions. According to this feature, the node density of the common communication coverage area is the same as the node density within the communication range of each node, which leads to Eq. (8).

\[
\frac{A(d_{ui})}{CN_{ui}} = \frac{\pi R^2}{M_u}
\]  

(8)

where \( CN_{ui} \) represents the number of common neighbor nodes, \( M_u \) represents the number of neighbors of node \( u \). It is worth noting that the number of both \( CN_{ui} \) and \( M_u \) contain node \( u, i \) itself. From Eq. (8), we know that the estimated coverage area \( \hat{A}(d_{ui}) \) of two nodes \( u, i \) can be calculated according to the number of neighboring nodes of \( u, i \) and the communication radius \( R \), as shown in Eq. (9).

\[
\hat{A}(d_{ui}) = \frac{CN_{ui}}{M_u} \pi R^2
\]

(9)

Equation (9) uses the number of neighboring nodes of \( u \) and the number of common neighbor nodes of \( u, i \) to estimate the common coverage area \( \hat{A}(d_{ui}) \). Similarly, we can estimate the \( \hat{A}(d_{ui}) \) according to the number of neighboring nodes of \( i \) and the number of common neighboring nodes. To avoid estimation bias, we take the average of the two estimation methods as the final estimate of \( \hat{A}(d_{ui}) \), as shown in Eq. (10).

\[
A(d_{ui}) = \frac{\hat{A}(d_{ui}) + \hat{A}(d_{ui})}{2} = \frac{(NB_u + NB_i)CN_{ui}}{2NB_uNB_i} \pi R^2
\]

(10)

To simplify the presentation, we define the intersection degree \( I_{ui} \) associated with nodes \( u \) and \( i \) as shown in Eq. (11).

\[
I_{ui} = \frac{(NB_u + NB_i)CN_{ui}}{2NB_uNB_i}
\]

(11)

Taking Eq. (11) into Eq. (10) yields Eq. (12).

\[
A(d_{ui}) = \pi R^2 I_{ui}
\]

(12)

Simultaneous Eqs. (7) and (12), after eliminates \( A(d_{ui}) \) we can obtain the functions of \( d_{ui} \) and \( I_{ui} \) as shown in Eq. (13).

\[
d_{ui} = \frac{(1 - I_{ui}) \pi R}{2}
\]

(13)

According to Eq. (13), we can use the intersection degree \( I_{ui} \) of two nodes to calculate the estimated distance \( d_{ui} \) of two nodes, the \( I_{ui} \) can be obtained from the adjacency information. Referring to Eq. (13), all nodes can estimate the distance to their neighbor nodes.

### 4.2 Preliminary location estimate

After initialization, all nodes obtain their neighboring information by data exchange and calculate the estimated distance to neighbor nodes according to the method illustrated in Sect. 3.1. After the neighbor distance estimation, all anchor nodes flood their information to the whole network, and all nodes get the estimated distance to each anchor node during flooding. To improve the accuracy of distance estimation, each anchor node calculates its distance correction coefficient (DCC) based on the anchor position information and the estimated distance. The unknown node corrects the estimated distance from the anchor node based on DCC. The preliminary unknown estimates of nodes are as follows.

Suppose there are \( n \) anchor nodes in the deployment area and the set of anchor nodes is \([A_1, A_2, ..., An]\). After deployment, all nodes broadcast a “hello” message containing their ID number and receive “hello” messages from neighboring nodes. In this way, node neighbor node information is collected. Then, exchange neighbor node information with all neighboring nodes and estimate the distance between the node itself and its neighboring nodes. After neighboring node distance estimation, all anchor nodes broadcast “flood message” message and flood it to the whole network. The format of the “flood message” is \( \text{flooding_msg} \{ Ai, d_{ac}, NID \} \), where \( Ai \) represents the ID of the anchor node to which the flooding message belongs, \( d_{ac} \) represents the cumulative distance of the message propagation, i.e., the cumulative estimated distance of the message from the anchor node, and \( NID \) represents the set of relay nodes. Before the flooding starts, the node initializes the estimated multi-hop distance \([d_{A1}, d_{A2}, ..., d_{An}]\) with each anchor node, where all elements are initialized to null. When unknown node \( k \) receives the “flooding_msg” message belonging to anchor node \( Ai \), it first refers the estimated distance \( d_{i} \) between the node itself and the relay node and calculates the cumulative estimated distance \( FD_{Ai} \) of anchor node \( Ai \) in the “flooding_msg” propagation path according to Eq. (14).

\[
FD_{Ai} = d_{ac} + d_{i}
\]

(14)

After updating the cumulative distance, the node checks whether there is multi-hop estimated distance information of anchor node \( Ai \) in memory, if there is no such anchor node information, it saves the updated \( FD_{Ai} \) as the shortest multi-hop estimated distance \( d_{Ai} \) of anchor node \( Ai \), updates \( d_{ac} \) in “flooding_msg” as \( d_{ai} \), and \( NID \) is updated to the set of neighbor nodes of the receiving node itself, and forwards it after the update. If a node has multi-hop estimated distance
All nodes broadcast messages with communication radius $R$, and receive broadcast messages from neighboring nodes to obtain information about neighboring nodes.

**Initialize the set of estimated distances to each anchor node** $[d_{A1}, d_{A2}, \ldots, d_{Ai}]$

- **Is it an anchor node?**
  - **N**
  - **Y**
    - **Folding the message** $\text{flooding\_msg}(A_i, d_{A}, N_{id})$ to the network
  - **Maintain listening state**

- **Received a flooding message??**
  - **N**
  - **Y**
    - **Calculate the cumulative estimated distance** $F_D_{A_i}$ in the propagation path according to equation (14)

- **$d_{ui}$ value is empty?**
  - **Y**
    - **$d_{ui} < F_D_{A_i}$?**
      - **N**
      - **Y**
        - **Update** $d_{ui}$ to the cumulative estimated distance $F_D_{A_i}$
        - **Update the** $d_a$ and $N_{id}$ in flooding message, and forward the message.
    - **Discard this flooding message**
  - **End**

Information of anchor node $A_i$ in its memory, then it compares $F_D_{A_i}$ with $d_{Ai}$ and keeps the minimum value. In this way, all nodes can get their shortest estimated distances from all anchor nodes. The flowchart of this process is shown in Fig. 3.

After the network flooding, all nodes get the estimated distance from all anchor nodes, and all anchor nodes calculate their respective $DCC$ based on the estimated distance from other anchor nodes according to Eq. (15).

$$DCC(A_i) = \frac{\sum_{i \in AN, i \neq A_i} \sqrt{(x_{Ai} - x_i)^2 + (y_{Ai} - y_i)^2}}{\sum_{i \in AN, i \neq A_i} d_{ki}}$$  \hspace{1cm} (15)
obtains the final estimated distance according to Eq. (16).

\[ \hat{d}_{ki} = DCC(Ai) \times d_{ki} \]  

(16)

where, \( d_{ki} \) and \( \hat{d}_{ki} \) represent estimated distance and final estimated distance respectively. After obtaining the final estimated distances to all anchor nodes, the preliminary estimated positions of the nodes are obtained by using the least squares method. The initial location estimation procedure is as follows.

\[
\begin{align*}
(x_{A1} - x_k)^2 + (y_{A1} - y_k)^2 &= \hat{d}_{k1}^2, \\
(x_{A2} - x_k)^2 + (y_{A2} - y_k)^2 &= \hat{d}_{k2}^2, \\
\ldots, \\
(x_{An} - x_k)^2 + (y_{An} - y_k)^2 &= \hat{d}_{kAn}^2
\end{align*}
\]  

(17)

where \((x_{Ai}, y_{Ai})\) represents the coordinate of the anchor node \(Ai\), \(\hat{d}_{ki}\) is the final estimated distance between the unknown node \(k\) and the anchor node \(Ai\), and \((x_k, y_k)\) is the coordinate of node \(k\). The first \(n-1\) terms of Eq. (17) are subtracted from the last term to obtain Eq. (18). The Eq. (18) can be written in the form of \(AX = B\), where \(A\), \(B\), and \(X\) are shown in Eqs. (19)–(21).

\[
\begin{align*}
-2(x_1 - x_{An})x_k - 2(y_{A1} - y_{An})y_k &= \hat{d}_{k1}^2 - \hat{d}_{kAn}^2 \\
-x_1^2 + x_{An}^2 - y_1^2 + y_{An}^2 \\
-2(x_2 - x_{An})x_k - 2(y_{A2} - y_{An})y_k &= \hat{d}_{k2}^2 - \hat{d}_{kAn}^2 \\
-x_2^2 + x_{An}^2 - y_2^2 + y_{An}^2 \\
\ldots \\
-2(x_{(A(n-1))} - x_{An})x_k - 2(y_{A(n-1)} - y_{An})y_k &= \hat{d}_{k(A(n-1))}^2 - \hat{d}_{kAn}^2 \\
-\hat{d}_{kAn}^2 - x_{An}^2 + x_{An}^2 - y_{An}^2 + y_{An}^2
\end{align*}
\]  

(18)

\[
A = -2 \begin{bmatrix}
x_1 - x_{An}, y_1 - y_{An} \\
x_2 - x_{An}, y_2 - y_{An} \\
\vdots \\
x_{A(n-1)} - x_{An}, y_{A(n-1)} - y_{An}
\end{bmatrix}
\]  

(19)

\[
B = \begin{bmatrix}
\hat{d}_{k1}^2 - \hat{d}_{kAn}^2 - x_1^2 + x_{An}^2 - y_1^2 + y_{An}^2 \\
\hat{d}_{k2}^2 - \hat{d}_{kAn}^2 - x_2^2 + x_{An}^2 - y_2^2 + y_{An}^2 \\
\ldots \\
\hat{d}_{k(A(n-1))}^2 - \hat{d}_{kAn}^2 - x_{An}^2 + x_{An}^2 - y_{An}^2 + y_{An}^2
\end{bmatrix}
\]  

(20)

\[
X = \begin{bmatrix}
x_k \\
y_k
\end{bmatrix}
\]  

(21)

The linear equation \(AX = B\) is solved by the least-squares method to obtain the solution of \(X\) as in shown in formula (22).

\[
X = (AA^T)^{-1}A^TB
\]  

(22)

### 4.3 Location error vector correction

The distance vector range-free localization algorithms estimate the distance between the unknown node and the anchor node by investigating the network connectivity, then calculate the unknown node location according to the estimated distance. For these localization algorithms, there is similarity in the network connection between two nodes with similar locations. Therefore, there is also similarity in the distance estimation errors between neighboring nodes and anchor nodes, which leads to similarity in their localization errors. To illustrate this phenomenon, we set 200 nodes to be randomly deployed in a 100 m × 100 m square area, of which 20 are anchor nodes. Apply two representative distance vector range-free localization algorithms DV-Hop and DV-RND to locate unknown nodes in the network. The localization error vectors of all sensor nodes are shown in Fig. 4.

In Fig. 4, the blue circle represents the true unknown node location, the red asterisk represents the anchor node location, and the blue line represents the error vector of the node. One end of the blue line connects the real position of the node and the other end is the estimated position of the node. It can be seen from Fig. 4 that nodes that are geographically very close to each other have a large similarity in their error vectors in terms of error magnitude and direction. Based on the above observations, we propose the following error correction method.

Define the location error vector (LEV) of the anchor node as the vector of the estimated position of the anchor node pointing to the true position. The direction of \(LEV\) is from the estimated location to the true location. Before calculating the anchor node \(LEV\), the anchor node itself is considered as an unknown node, and the position of itself is estimated refer to the coordinate information and estimated distance of all anchor nodes (including itself), where the estimated distance from itself is set to 0. The \(LEV\) of the anchor node \(i\) is \(V(i)\), which is determined by Eq. (23).

\[
V(i) = (\hat{x}_{Ai} - x_{Ai}, y_{Ai} - \hat{y}_{Ai})
\]  

(23)

where \((\hat{x}_{Ai}, \hat{y}_{Ai})\) represents the estimated position of anchor node \(i\) and \((x_{Ai}, y_{Ai})\) represents the true position of anchor node \(i\). Considering that the magnitude and direction of the localization error between sensor nodes in close proximity are highly correlated, based on this property, we can correct the initial estimated position of the unknown node using the \(LEV\) of the anchor node. Assuming the initial estimated position of unknown node \(k\) is \(L_p(k) = (\hat{x}_k, \hat{y}_k)\) whose nearest anchor node is anchor node \(i\), and the \(LEV\) of this anchor node is \(V(i)\), then the final estimated position of unknown node \(k\) \(L_F(k)\) is calculated by Eq. (24).

\[
L_F(k) = L_p(k) + c\hat{d}_{ik} \times V(i)
\]  

(24)
where $c$ is the error correction factor, which is a constant between $[0, 1]$, and $\hat{d}_{ik}$ is the estimated distance between the unknown node $k$ and its reference anchor node $i$. Since the error correction vector of a node is negatively correlated with the distance between two nodes, i.e., the closer the nodes are, the higher the similarity of the error vectors, while the longer distant they are, the lower the similarity in error vectors. Therefore, the magnitude of the correction value is determined by the distance between the unknown node and its reference node; the closer the distance, the larger the correction value, and the farther the distance, the smaller the correction value, in order to avoid over-correction. When the distance between node $i$ and anchor node $k$ is 0, the correction value is equal to $V(i)$. The pseudo-code of our proposed localization error vector correction method is shown below.

![Fig. 4 Localization error description](image)

5 Experimental results and discussion

5.1 Parameter setting and evaluation metrics

We deploy 200 sensor nodes in a $100 \times 100$ m square area to test the performance of the localization algorithm and compare the proposed MDV-EC with two distance vector localization algorithms, DV-Hop [10] and DV-RND [22]. We also test the effectiveness of the proposed error correction method by comparing the performance of MDV (the proposed algorithm without distance vector correction) with our proposed MDV-EC algorithm. We complete all simulation experiments in MATLAB2018 platform. The default parameter settings in the simulation are shown in Table 1.

| Parameters | Value |
|------------|-------|
| Deployment area | $100 \times 100$ m |
| Node number ($N$) | 200 |
| Anchor number ($Na$) | 20 |
| Communication radius ($R$) | 30 |
| DOI | 0.01 |
| Error correction factor ($c$) | 0.95 |

In this paper, we use the average localization error ($ALE$) and average localization accuracy ($ALA$) to evaluate the performance of the localization algorithm. The smaller $ALE$ and $ALA$ represent the higher accuracy of the localization algorithm, which can be calculated by Eqs. (25), (26) respectively.

\[
ALE = \sum_{i \in N, i \notin Na} \sqrt{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2} \tag{25}
\]

\[
ALA = \frac{ALE}{R} \tag{26}
\]

where $(\hat{x}_i, \hat{y}_i)$ represents the final estimated position of unknown node $i$, $(x_i, y_i)$ represents the real position of unknown node $i$, and $R$ is the communication radius of the node.
5.2 Performance evaluation

Figures 5 and 6 test the effect of error correction factor $c$ on the average localization error of the proposed MDV-EC for different node densities and different number of anchor nodes respectively. The correction factor is a constant between [0, 1]. When the correction factor is small, the correction value of the node position is smaller, and the smaller correction value has less effect on the node position estimation and the error variation is smaller. In order to distinguish the impact of the correction coefficient $c$ on the average positioning error effectively, the range of the error correction coefficient $c$ is set to [0.75, 1] in Fig. 5 and 6.

As we can see in Fig. 5, the ALE of the proposed MDV-EC algorithm firstly decreases and then increases as $c$ increases in different node densities. The main reason for this phenomenon is that when $c$ is small, the correction vector value of each unknown node is small, and the correction has less influence on the estimated position. With the continuous increase of $c$, the value of unknown nodes correction vectors gradually increases and the localization error of the estimated location gradually decreases, which indicates the validity of our proposed error correction. The algorithm achieves optimal localization accuracy when $c$ increases to about 0.95, and the ALE of the algorithm increases rapidly as $c$ continues to increase to 1. This is because when $c$ equals to 1, the error correction vector of each unknown node is equal to the LEV of nearby anchor node, and using this correction vector to correct the initial location will lead to over-correction, increase the localization error. In different node density scenarios, the higher the node density, the higher the localization accuracy, so according to Fig. 5 the algorithm has the highest node localization accuracy at $N = 300$. When $N$ is 100 and 200, the best localization accuracy is achieved at $c = 0.95$, and the ALE of the MDV-EC is 5.32 m and 5.21 m, respectively; when $N = 300$, the best localization accuracy is achieved at $c = 0.9$, and the ALE of the MDV-EC is 4.76 m. According to the figure, the localization accuracy is more sensitive to the error correction factor in the low node density scenarios. In the case of low node density deployment, the error of the node is larger, so the error vector of the anchor node is also larger, leading to a larger base of the reference correction vector, and therefore the ALE is more sensitive to $c$.

Figure 6 compares the impact of the error correction factor $c$ on ALE for different number of anchor node scenarios. According to the figure, in all cases, the ALE values gradually decrease as the correction factor $c$ increases from 0.75 to 0.95, while the ALE values increase as $c$ increases from 0.95 to 1. Under the condition of the same correction factor, the more the number of anchor nodes in the network, the higher the localization accuracy. There are two main aspects leading to this phenomenon. On one hand, the more anchor nodes, the more provided localization information, which is beneficial to localization accuracy; on the other hand, the more anchor nodes, the closer the average distance between the unknown node and the reference node, and the higher the reliability of the reference node. In addition, the figure shows that the ALE is more sensitive to $c$ in high anchor node density scenario. This is mainly because in high anchor node density scenario, the distance between the unknown node and its reference node is shorter, the correction vector of the unknown node is larger, and therefore the value of ALE is more sensitive to the change of $c$. The closer the average distance between the unknown node and its reference node, the more reliable the error vector will be.

Figure 7 compares the average localization accuracy of the proposed MDV-EC algorithm with the DV-Hop, DV-RND, and MDV algorithms in different node densities. It can be
seen from the figure that the error of the DV-Hop algorithm is insensitive to the node density and its ALA increases slowly with increasing node density, while the ALA of the proposed MDV-EC, DV-RND and MDV algorithms decreases gradually with increasing node density. Among all the algorithms, the proposed MDV-EC has the highest localization accuracy, which is better than MDV and DV-RND, where MDV has better localization accuracy than DV-RND. Therefore, the proposed MDV-EC outperformed the distance vector localization algorithms DV-Hop and DV-RND, meanwhile the proposed error correction method can make full use of the similarity of local errors to further reduce the node localization errors and improve the localization accuracy.

Figure 8 compares the localization accuracy of the proposed MDV-EC algorithm with the DV-Hop, DV-RND, and MDV algorithms for scenarios with different numbers of anchor nodes. As can be seen from the figure, the ALA values of all four algorithms gradually decrease with the increase of the number of anchor nodes, where the proposed MDV-EC algorithm has the highest localization accuracy and the classical DV-Hop algorithm has the worst localization accuracy. When the number of anchor nodes is less than or equal to 20, the MDV algorithm outperformed DV-RND, while when the number of anchor nodes is greater than 20, the accuracy of DV-RND is superior to the proposed MDV algorithm. Therefore, in practical application scenarios, we can improve the localization accuracy by increasing the number of anchor nodes.

Figure 9 compares the ALA of the four algorithms at different communication radius. It can be seen from the figure that the localization accuracy of the proposed MDV-EC algorithm is optimal in all scenarios. The ALA of DV-Hop slowly decreases as the communication radius $R$ increases, while the ALA of DV-RND, MDV and MDV-EC decrease firstly and then stabilize as the communication radius $R$ increases. When $R$ is less than 30 m, the ALA of the proposed MDV and MDV-EC decreases with the increase of $R$, and when $R$ is greater than 30 m, the two algorithms have less variation. This is because when the communication radius is small, the connectivity of the network is poor and the average positioning accuracy of all nodes is large. When $R$ increases, the connectivity of the network is improved, thus enhancing the accuracy of distance estimation and localization accuracy. As $R$ continues to increase, the communication coverage of the sensor nodes in the network becomes too large and the nodes are affected by the deployment area boundaries during distance estimation. While higher network connectivity contributes to the accuracy of distance estimation, larger $R$
also makes distance estimation more negatively influenced by boundaries, which overall leads to a flat ALA.

Figure 10 compares the localization accuracy of the four algorithms in different signal noise environments. As can be seen from the figure, the ALA of the four algorithms increases with the increase of signal noise, and the best accuracy in all experimental scenarios is the proposed MDV-EC algorithm, while the worst accuracy is DV-Hop. When the signal noise in the network environment is 0, i.e., the communication coverage area of every node is a standard circle, the ALA of the proposed MDV-EC algorithm is slightly better than the DV-RND algorithm, both around 0.1, while the ALA of the MDV algorithm without the error correction is about 0.12. With the increase of signal noise in the network, the ALA value of DV-RND is gradually larger than MDV, while the ALA of the proposed MDV-EC algorithm is always the smallest. Therefore, the proposed MDV-EC has stronger environmental robustness and good localization performance in different noisy environments. In addition, the proposed error correction method can further reduce the node localization errors in different noisy environments.

Figure 11 illustrates the random distribution of 200 sensor nodes, including 20 anchor nodes. The blue circles represent unknown nodes and the red stars represent anchor nodes. Set communication radius \( R = 20 \text{ m} \), other parameters are preset values. The data of Figs. 12 and 13 are obtained in this node deployment of Fig. 11.

Figure 12 compares the error vectors of the MDV algorithm without the error correction scheme and the MDV-EC algorithm with the error correction scheme. The green line represents the error vector of MDV, the black line is the error vector of MDV-EC, and the red line segments on the anchor nodes represent the LEV of the respective anchor nodes. It can be seen that there is a strong similarity in the magnitude and direction of the localization error between neighboring nodes. The main reason for this phenomenon is the similar node distribution of neighboring nodes in the distance estimation, which leads to strong similarity in their ranging and localization errors. Furthermore, according to Fig. 12, the error vector length of MDV-EC of most nodes (black line segment) is shorter than MDV (green line segment), which indicates the localization error of the proposed MDV-EC is smaller than MDV, and the proposed error correction method
is effective in reducing the localization error of nodes by exploring the local similarity of localization errors.

Figure 13 compares the cumulative distribution curves (CDF) of localization errors for the four algorithms. As shown in the figure, the CDF curve of the proposed MDV-EC is always at the highest position, followed by MDV, DV-RND and DV-Hop algorithms. Among them, the MDV-EC algorithm has 90% of the node localization errors within 10 m, and the maximum value of the localization error is about 15 m, which is better than the performance of the MDV algorithm. While about 75% of the nodes in DV-RND have errors within 10 m, and the number of node localization errors within 10 m in DV-Hops algorithm accounts for about 52%. Therefore, the proposed MDV-EC algorithm outperforms MDV, DV-RND, and DV-Hops in terms of localization accuracy.

6 Conclusion

In order to enhance the localization accuracy for WSNs, a multi-hop distance-vector localization algorithm named MDV-EC is proposed in this paper. The proposed MDV-EC estimates the distance between neighboring nodes based on the neighbor information and corrects the estimated distance with reference to the local path correction coefficient, rather than corrects the estimated distance from a global point view. By investigating the local similarity characteristics of the localization error in the distance vector localization algorithm, an error correction method is applied to further reduce the localization error. The error correction method uses the estimated and true positions of the anchor nodes to calculate the localization error vector for each anchor node. Then correct the initial estimated position by referring the localization error vector of nearby anchor node. The simulation results demonstrate that the proposed MDV-EC algorithm outperformed DV-Hop and DV-RND in terms of localization accuracy. And the proposed error correction method is effective to reduce the localization error. In the future research, we will focus on exploring the quantization relationship between the local error similarity and the location error vector and investigating a better error correction method to reduce the node localization error.
Acknowledgements This work was supported by the Key Research and Development Projects of Hunan Province (2018SK2055); Ministry of Emergency Management of the People’s Republic of China Key Technologies for Safety Production and Major Accidents Prevention and Control (Hunan-0001-2018AQ); Hunan Provincial Natural Science Foundation of China (2021JJ50093).

Conflict of interest There is no conflict of interest exists in the submission of this manuscript, and all authors have approved the manuscript that is enclosed.

References

1. Praveen Kumar, D., Amgoth, T., & Annavarapu, C. S. R. (2019). Machine learning algorithms for wireless sensor networks: A survey. Information Fusion, 49, 1–25. https://doi.org/10.1016/j.inffus.2018.09.013
2. Zhang, K., Zhang, G., Yu, X., Hu, S., & Li, M. (2022). Clustering the sensor networks based on energy-aware affinity propagation. Computer Networks. https://doi.org/10.1016/j.comnet.2022.108853
3. Osamy, W., Khedir, A. M., Salim, A., Ali, A. I. A., & El-Sawy, A. A. (2022). Coverage, deployment and localization challenges in wireless sensor networks based on artificial intelligence techniques: A review. IEEE Access, 10, 30232–30257. https://doi.org/10.1109/access.2022.3156729
4. Liu, X., Han, F., Ji, W., Liu, Y., & Xie, Y. (2020). A novel range-free localization scheme based on anchor pairs condition decision in wireless sensor networks. IEEE Transactions on Communications. 68(12), 7882–7895. https://doi.org/10.1109/comm.2020.3020553
5. Kumar, P., Chaturvedi, A., & Kulkarni, M. (2012). Geographical location based hierarchical routing strategy for wireless sensor networks. In International Conference on Devices. IEEE.
6. Liu, J., Wang, Z., Yao, M., & Qiu, Z. (2015). VN-APIT: Virtual nodes-based range-free APIT localization scheme for WSN. Wireless Networks, 22(3), 867–878. https://doi.org/10.1007/s11276-015-1007-z
7. Luomala, J., & Hakala, I. (2019). Analysis and evaluation of adaptive RSSI-based ranging in outdoor wireless sensor networks. Ad Hoc Networks, 87, 100–112. https://doi.org/10.1016/j.adhoc.2018.10.004
8. Zhang, H., Wang, Z., & Gulliver, T. A. (2017). Two-stage weighted centroid localization for large-scale wireless sensor networks in ambient intelligence environment. Journal of Ambient Intelligence and Humanized Computing, 9(3), 617–627. https://doi.org/10.1007/s12652-017-0458-8
9. Nemer, I., Sheltami, T., Shakshuki, E., Elkhail, A. A., & Adam, M. (2020). Performance evaluation of range-free localization algorithms for wireless sensor networks. Personal and Ubiquitous Computing, 25(1), 177–203. https://doi.org/10.1007/s00779-020-01370-x
10. Niculescu, D., & Nath, B. (2003). DV based positioning in Ad hoc networks. Telecommunication Systems, 22, 267–280. https://doi.org/10.1023/A:1023403232340
11. Shahzad, F., Sheltami, T. R., & Shakshuki, E. M. (2017). DV-maxHop: A fast and accurate range-free localization algorithm for anisotropic wireless networks. IEEE Transactions on Mobile Computing, 16, 2494–2505.
12. Huang, X. (2020). Multi-node topology location model of smart city based on Internet of Things. Computer Communications, 152, 282–295. https://doi.org/10.1016/j.comcom.2020.01.052
13. Liu, G., Qian, Z., & Wang, X. (2019). An improved DV-hop localization algorithm based on hop distances correction. China Communications, 16(6), 200–214.
14. Cao, Y., & Wang, Z. (2019). Improved DV-hop localization algorithm based on dynamic anchor node set for wireless sensor networks. IEEE Access, 7, 124876–124890. https://doi.org/10.1109/access.2019.2938558
15. Chen, T., Hou, S., & Sun, L. (2022). An enhanced DV-hop positioning scheme based on spring model and reliable beacon node set. Computer Networks. https://doi.org/10.1016/j.comnet.2022.108926
16. Zhang, K., Zhang, G., Yu, X., & Hu, S. (2021). Boundary-based anchor selection method for WSN node localization. Arabian Journal for Science and Engineering, 46(4), 3779–3792. https://doi.org/10.1007/s13369-020-05286-9
17. Han, D., Yu, Y., Li, K. C., & de Mello, R. F. (2020). Enhancing the sensor node localization algorithm based on improved DV-hop and DE algorithms in wireless sensor networks. Sensors (Basel). https://doi.org/10.3390/s20203543
18. Singh, P., Mittal, N., & Singh, P. (2022). A novel hybrid range-free approach to locate sensor nodes in 3D WSN using GWO-FA algorithm. Telecommunication Systems. https://doi.org/10.1007/s11235-022-00888-0
19. Yang, J., Cai, Y., Tang, D., & Liu, Z. (2019). A novel centralized range-free static node localization algorithm with memetic algorithm and levy flight. Sensors (Basel). https://doi.org/10.3390/s19143242
20. Cui, H., Liang, Y., Zhou, C., & Cao, N. (2018). Localization of large-scale wireless sensor networks using niching particle swarm optimization and reliable anchor selection. Wireless Communications and Mobile Computing, 2018, 1–18. https://doi.org/10.11055/2018/2473675
21. Wang, Y., Wang, X., Wang, D., & Agraval, D. P. (2009). Range-free localization using expected hop progress in wireless sensor networks. IEEE Transactions on Parallel and Distributed Systems, 20, 1540–1552. https://doi.org/10.1109/TPDS.2008.239
22. Wu, G., Wang, S., Wang, B., Dong, Y., & Yan, S. (2012). A novel range-free localization based on regulated neighborhood distance for wireless ad hoc and sensor networks. Computer Networks, 56(16), 3581–3593. https://doi.org/10.1016/j.comnet.2012.07.007
23. Yun, W., Xiaodong, W., Demin, W., & Agraval, D. P. (2009). Range-free localization using expected hop progress in wireless sensor networks. IEEE Transactions on Parallel and Distributed Systems, 20(10), 1540–1552. https://doi.org/10.1109/tpds.2008.239
24. Wen, W., Wen, X., Yuan, L., & Xu, H. (2018). Range-free localization using expected hop progress in anisotropic wireless sensor networks. EURASIP Journal on Wireless Communications and Networking. https://doi.org/10.1186/s13638-018-1326-8
25. Zaidi, S., El Assaf, A., Alles, S., & Kandil, N. (2016). Accurate range-free localization in multi-hop wireless sensor networks. IEEE Transactions on Communications, 64, 3886–3900. https://doi.org/10.1109/TCOMM.2016.2590436

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

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.
Ke Zhang is now pursuing a Ph.D. degree in School of Safety Science and Emergency management, Wuhan University technology, Wuhan, China. His research interests including wireless sensor networks localization and routing protocol.

Guang Zhang is a professor in the School of Safety Science and Emergency management, Wuhan University technology, Wuhan, China. He received a Ph.D. degree from the University of Chinese Academy of Sciences, Wuhan, China in 1994. His research interests including geological disaster theory and prevention, smart safety monitoring systems.

Xiuwu Yu is currently a professor with the School of Resource & Environment and Safety Engineering, University of South China, Hengyang, China. He received a Ph.D. degree from Wuhan University of Technology, Wuhan, China, in 2012. He has published more than 10 technical papers in international journals and conferences. His research interests including smart safety monitoring systems, wireless sensor networks, and communication signal processing.

Shaohua Hu is an associate professor in the School of Safety Science and Emergency management, Wuhan University technology, Wuhan, China. He received a Ph.D. degree from Wuhan University in 2015. His research interest including smart safety monitoring systems.

Youcui Yuan is currently pursuing a master's degree in School of Safety Science and Emergency management, Wuhan University of Technology, Wuhan China. Her research area is industrial safety.