Hierarchical self-localization of underwater wireless sensor network nodes

ZHANG Hua †, LIU Yu-liang
School of Electromechanical Engineering, Zhejiang Ocean University, Zhoushan 316000, P. R. China

Received 3 May 2012; received in revised form 1 July 2012

Abstract: The follow-up application of underwater wireless sensor network is influenced by accuracy of self-localization of nodes. The self-localization of nodes is discussed in this paper. First of all, nodes of underwater wireless sensor network are classified into several levels according to the accuracy of position of nodes and the levels are from the first to the fifth in accordance with accuracy of nodes from high to low respectively. Secondly, the level of anchor nodes can be known by those unknown nodes from the information given by the anchor nodes themselves, At the same time the unknown nodes are able to be located in the area controlled by the first level of anchor nodes that are as the aggregation. Then the positioning algorithm is designed correspondingly in accordance with the accuracy level of nodes. Finally, the positioning algorithm is simulated and analyzed. The result shows that the unknown nodes can be located effectively by hierarchical control.

Keywords: hierarchical self-localization; underwater wireless sensor network; accuracy; classification; density

CLC number: TP212.9               Document code: A

1 Introduction

The self-localization technology of nodes is a prerequisite for the application of wireless sensor network. The accuracy of self-localization is related to the validity of data collected by sensor nodes, and it has no meaning without precise location of the monitoring message [1-3]. The technology of wireless sensor networks is using in the marine environment with the continued expansion of the depth and the application of wireless sensor networks [4-6]. Studies on underwater sensor networks use sonar as a means of communication because the radio is not suitable to the marine environment on the depth and breadth [4-6]. It needs the underwater wireless sensor networks to continue normal work and the nodes to collect data of underwater wireless sensor networks at different depths in the ocean including the chemical composition, characteristics of temperature, salinity, depth, metal ion concentration, and acoustic Doppler profile of the currents in the seawater. There are two types of underwater wireless sensor network layout. Firstly, each sensor node in the seawater is connected with one buoy on the water. And the second, sensor nodes are anchored at the bottom of the seabed pulled to the ocean surface by the buoy (as shown in Fig. 1) [9-11].

Ref. [11] described the application prospects and facing challenges of the underwater acoustic sensor network, and discussed on the short-range acoustic communication and MAC protocols. Ref. [12] considered that the AMHL was the most suitable way for the application of track-oriented UWSN by using the PLACE self-localization method on multi-hop. There are a lot of things to do before the application to the actual situation because the simulation environment used in the model is ideal without influences on the parameters independent of external [12-15]. The overall
accuracy of self-localization will be improved by increasing the control flag in the node communication header, and the different algorithm will be taken by different accuracy levels with the self-localization of unknown nodes which distinguish the anchor classification level automatically in this paper.

2 Design of anchor node’s classification

The accumulated error of self-localization of nodes has not been solved in previous literature although the relationship between the number of anchor nodes and the accuracy of node self-localization had been demonstrated. There must have been a certain number of anchor nodes as to self-localization of unknown nodes [11-13]. Different anchor nodes will be dealt with different approach according to their classification in order to reduce the accumulated error in the self-localization of the unknown nodes.

The first level of anchor nodes They can be used as the master node of a group, and other nodes’ information can be given by the base station directly or other methods.

The second level of anchor nodes The information can be given by the first level of anchor nodes directly.

The third level of anchor nodes The information of unknown nodes can be obtained from the second level of anchor nodes that the accuracy of self-localization of this kind of nodes is lower than the second one’s.

The fourth level of anchor nodes The information of unknown nodes can be obtained from the second level of anchor nodes or the third level of anchor nodes that the accuracy of self-localization of this kind of nodes is lower than the third one’s.

The fifth level of anchor nodes This kind of unknown nodes is distributed in the margin of which self-localization is much more difficult because the error accumulated is large. Such nodes are not discussed as the focus.

One known node can not be located well without the node spatial information and its computing speed is impacted by the information storage format. The information table of node is shown in Table 1.

3 Design of location algorithm

The density of the anchor nodes affects the ultimate accuracy of nodes self-localization from the self-localization process. It may be positioned to one unknown node by the second level of anchor nodes or by the third level of anchor nodes, and may more likely be co-located by the second and the third levels of anchor nodes from the possibility of the self-localization accuracy of anchor nodes. Three anchor nodes are discussed as subject to investigate the problem in a region with a certain number of unknown nodes and anchor nodes by dispenser (Fig. 2).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Network-id</th>
<th>Level</th>
<th>Anchor</th>
<th>Local-id</th>
<th>Position</th>
<th>Accuracy</th>
<th>Algorithm</th>
<th>Fcs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Describe</td>
<td>Network Identification</td>
<td>Level of anchor nodes</td>
<td>ID of anchor node</td>
<td>Local identity</td>
<td>Location coordinates</td>
<td>Location accuracy</td>
<td>Algorithm Selected</td>
<td>Parity bit</td>
</tr>
</tbody>
</table>

Fig. 1 Distribution of underwater sensor networks
3.1 The second level of anchor nodes around the unknown nodes only

Assuming the three anchor nodes (A, B, and C) are the second level of anchor nodes around the unknown node O (Fig. 2). The coordinates of the three anchor nodes are given by the first level of anchor nodes so that the accuracy of the location of anchor nodes is much high. The distance between the unknown node to the three nodes of the second level of anchor nodes is \( d_1, d_2, \) and \( d_3 \), respectively. The coordinate of the unknown node is the geometric center of the triangle formed three anchor nodes of the second level of anchor nodes without the influence of the external environment, but the fact is that the coordinate of the centroid can not be used as the real coordinate of unknown node. The error is exiting between the measuring distance and the coordinates.

\[
\rho_i = \frac{1}{n} \sum_{i=1}^{n} \left( \sqrt{(x-x_i)^2 + (y-y_i)^2} - d_i \right). \tag{1}
\]

The error of node’s location coordinate can be expressed as \( o_i = o_j + \rho_i \), and the adjustment of the loop implementation is completed when the error is less than a quantitative.

3.2 The third level of anchor nodes around unknown nodes only

Assuming the three anchor nodes (A, B, and C) are the third level of anchor nodes around the unknown node O (Fig. 2). The coordinates of the third level of anchor nodes are located by the second level of anchor nodes, and the accuracy of coordinate of the second level is much higher than that of the third one because of iterative error. The estimate coordinate of one unknown node is \( O(x_{est}, y_{est}) \) and the centroid coordinate is \( O(x_w, y_w) \). The anchor nodes build a center of mass coordinates communicating with the unknown node. The error exits between the estimate coordinate and the centroid coordinate.

\[
x_{est} = x_w + \Delta x; y_{est} = y_w + \Delta y, \tag{2}
\]

where \( \Delta x \) and \( \Delta y \) denote the error between the real coordinate and the centroid coordinate of the node respectively. \( f(x, y) \) is the Euclidean distance between the unknown node and anchor nodes who communicate with the unknown node and

\[
f(x, y) = \sqrt{(x-x_i)^2 + (y-y_i)^2}; i = 1, 2, 3, \ldots, n. \tag{3}
\]

The substance of the Taylor series is the solution of the tangent approximation to solve equation. Eq. (3) is transformed by the Taylor series and the expansion of the Taylor series is shown in Eq. (4). Eq. (4) is Lagrange remainder.

\[
f(x, y) = f(x_{est} + \Delta x, y_{est} + \Delta y) = f(x_{est}, y_{est}) + \left( f_x, f_y \right)_x (\Delta x) + \left( f_x, f_y \right)_y (\Delta y) + R;
\]

\[
R = \frac{1}{2!} \left( \begin{array}{c} \Delta x, \Delta y \\ f_x^{(2)}, f_y^{(2)} \end{array} \right) \left( f_x, f_y \right)_x (\Delta x) (\Delta y) = (x_{est} + \frac{\partial}{\partial x} x_{est} + \frac{\partial}{\partial y} y_{est}) (\Delta x) (\Delta y);
\]

\[
X^* = (x_{est} + \Delta x, y_{est} + \Delta y). \tag{4}
\]

There is approximation error to a certain extent where there is a high-end remainder of Taylor series with the high-end discarding. Assuming the distance between anchor nodes and unknown nodes is \( d_i, i = 1, 2, 3, \ldots, n \) to reduce the error of the coordinates, respectively. The modified formula is as follows with further constraining the distance range of unknown coordinates.

\[
f(x, y) \approx f(x, y) \bigg|_{x_{est}} + \left( f_x, f_y \right)_x (\Delta x).
\]

Eq. (5) is expanded with multiple Taylor series in
the real coordinates of unknown node. The first-order partial derivative is seek in Eq. (5) for \( x \) and \( y \) respectively.

\[
f_x = \frac{x - x_i}{f_{x_i}}, \quad f_y = \frac{y - y_i}{f_{y_i}}. \tag{6}
\]

The unknown node \( O(x_{\text{est}}, y_{\text{est}}) \) is expanded by the Taylor series in Eq. (3) as

\[
f(x, y) \approx f(x, y)\bigg|_{x_{\text{est}}} + \left(f_x, f_y\right)_{x_{\text{est}}} \left(\Delta x, \Delta y\right). \tag{7}
\]

Eq. (6) are shown as follows in the right side.

\[
f(x, y)\bigg|_{x_{\text{est}}} = \sqrt{(x_{\text{est}} - x_i)^2 + (y_{\text{est}} - y_i)^2} - d_i,
\]

\[
\left(f_x, f_y\right)_{x_{\text{est}}} \left(\Delta x, \Delta y\right) = f_x * \Delta x + f_y * \Delta y\bigg|_{x_{\text{est}}} = \frac{x_{\text{est}} - x_i}{f_{x_i}} \Delta x + \frac{y_{\text{est}} - y_i}{f_{y_i}} \Delta y. \tag{8}
\]

Eq. (8) can be obtained from the known information.

3.3 The second and the third levels of anchor nodes around unknown nodes

3.3.1 More second level anchor nodes and less third level ones

There is one unknown node that is located by some anchor nodes. Some of the anchor nodes are the second level of anchor nodes with high accuracy in their position and others are the third level of anchor nodes with low accuracy in their position when the number of anchor nodes are not enough to locate all the unknown nodes in a dispenser wireless sensor networks. It is assumed that \( A \) and \( B \) are the second level of anchor nodes and \( C \) is the third level one in a model of three anchor nodes. There are two circles taking \( A \) and \( B \) as the center and their radius are the communication distance between the anchor node and the unknown node respectively. The actual location of the unknown node is in the cross-regional where is in the internal graphic constituted by anchor nodes with high precision as shown in Fig. 3.

[Diagram: Fig. 3 The number of the second level of anchor nodes is more than that of the third one]

The anchor node \( C \) is the third level of anchor nodes. Therefore, its accuracy of coordinate is not refined, so that the coordinates of the third level should be effectively corrected. It is difficult that position coordinates are fixed directly in fact thus the communication distance are considered to fix in this paper because the communication distance between nodes are the same impacted in the same environment. The coordinates of \( A \) and \( B \) can be known and the distance between them also can be measured. It equals 1 without measuring error. The coefficient \( k \) is denoted as the radio between the measuring distance and the calculating distance of nodes.

\[
k = \frac{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}}{d_{i2}}. \tag{9}
\]

The shadow zone in Fig. 3 may be the area of the unknown nodes. There exits error between the revised measure distances and the calculate one which can be described as

\[
n = \sqrt{(x - x_1)^2 + (y - y_2)^2} - kd_i. \tag{10}
\]

It is the best coordinate of the unknown node when the value of Eq. (10) is the least.

3.3.2 Less second level anchor nodes and more third level ones

There is one unknown node that is located by some anchor nodes. Some of the anchor nodes are the second level of anchor nodes with high accuracy in their position and others are the third level of anchor nodes with low accuracy in their position when the number of anchor nodes are less to locate all the unknown nodes
in a dispenser wireless sensor networks. The accuracy of self-localization of unknown nodes is not high due to the error transmission when one unknown node is located by three anchor nodes only. There are two second level anchor nodes and one third level node among the three nodes. There are more than one of the second level of anchor nodes to locate around one unknown node and two of the second one are selected, and the coordinates of them are expressed as \((x_{m1}, y_{m1}), (x_{m2}, y_{m2})\). The measuring distance between the two nodes is \(d_{m1,m2}\). The coefficient \(k_m\) is denoted as the ratio between the measuring distance and the calculating distance of nodes.

\[
k_m = \frac{\sqrt{(x_{m1} - x_{m2})^2 + (y_{m1} - y_{m2})^2}}{d_{m1,m2}}.
\]  

(11)

The coordinates of the third level of anchor nodes around the unknown node needs to be amended while that of the second one does not need, and the other coordinates of the third level of anchor nodes are expressed as \((x_{na1}, y_{na1}), \ldots (x_{nak}, y_{nak})\), \((l < n)\), respectively. The coefficient \(v_m\) is denoted as the ratio between the measuring distance and the calculating distance of nodes.

\[
v_m = \sum_{j=1}^{l} \left( \sqrt{(x - x_j)^2 + (y - y_j)^2} - d_j \right) + \sum_{j=l+1}^{n} \left( \sqrt{(x - x_j)^2 + (y - y_j)^2} - k_m d_j \right).
\]  

(12)

It is the best coordinate of the unknown node when the value of Eq. (12) is the least.

3.3.3 Steps of the algorithm

**Step 1** The information is released by nodes of the first level of anchor nodes to their control region after initialization of nodes. The nodes are classified automatically when the information received by all nodes.

**Step 2** The nodes of the second level can be located by the first level of anchor nodes firstly; they are unknown nodes in the state of waiting which can not be located directly.

**Step 3** The information is received by unknown nodes from the surrounding anchor nodes and the level of nodes can be classified by unknown nodes from the information received.

**Step 4** The position algorithm can be choose automatically by unknown nodes according to the level of anchor nodes received, and the information of the level of anchor nodes are spread.

**Step 5** The message of unknown nodes themselves is queried by the first level of anchor nodes. If the message is not known by all unknown nodes, the step is transferred to Step 2, otherwise the program is terminated.

4 Results and analysis

There are 30 beacon nodes and 10 unknown nodes arranged in the space according to the loop positioning algorithm of classification in Win7 computer under matlab7.0.1 process simulation.

Fig. 4 shows the distribution of nodes in different situations. Distribution of anchor nodes with high density and low density is shown respectively in Figs. 4a and 4b. Many unknown nodes are self-located around by anchor nodes in Fig. 4c and others are not. It explained that unknown nodes can be all located in the situation of high density of anchor nodes where the anchor nodes are as the second level of anchor nodes in Fig. 4a. And only some of unknown nodes can be located in the situation in low density of anchor nodes in Fig. 4c. Some unknown nodes turn into anchor nodes after the coordinates are known by themselves with different classification of the second, the third and the forth levels of anchor nodes in Fig. 4c.

The relationship between average connectivity with mean error and the error is decreasing with average connectivity increasing (Fig. 5). The relationship between average connectivity with density of anchor nodes and the average connectivity is increasing with density of anchor nodes increasing (Fig. 6). The average connectivity is low when density of anchor nodes is lower than 0.4 while density of anchor nodes is higher than 0.5, and the average connectivity is increasing sharply.

Fig. 7 shows that the coverage probability of nodes is increasing with the increase of the communication range.

Fig. 8 shows that the error of self-localization of nodes is decreasing with the increase of the communication radius in the effective communication range.
mainly around the unknown node, and the “High-density” one indicates that the density of anchor nodes is higher mainly around the unknown node.

**Legend:**
- * is the unknown node;
- ◆ is the anchor node;
- ▲ is the second level of anchor nodes;
- ✷ is the third level of anchor nodes; and
- ○ is the forth level of anchor nodes

**Fig. 4** Distribution of nodes in different situation where a) is a high density of anchor nodes, b) is a low density of anchor nodes and c) is self-localization of unknown nodes with low density of anchor nodes

**Fig. 5** Connectivity with error

**Fig. 6** Density of nodes with connectivity

**Fig. 7** Communication radius with coverage of anchor nodes

**Fig. 8** Position error and communication radius
Fig. 9 shows that the error of self-localization of nodes is becoming smaller with the increase of density of the anchor nodes. Line 2 indicates that the anchor nodes are the second level of anchor nodes mainly around the unknown node and Line 1 indicates that the anchor nodes are the third level of anchor nodes mainly around the unknown node.

5 Hierarchical positioning scheme for UWSN

The nodes of underwater sensor networks can be distributed as Fig. 1. We can see that anchored uw-sensors had been arranged and the nodes can send message to else nodes or surface sink. Self-localization of unknown nodes of wireless sensor network had been expounded in Refs. [1] to [11]. The nodes of wireless sensor network had been divided into many level anchor nodes in this paper based on previous literature. Self-localization of unknown nodes of UWSN also had been expounded in Refs. [12] to [15]. From those literatures we can see that distribution, working, structure of nodes of wireless sensor network and that of nodes of underwater wireless sensor network are the same. For example, node self-location scheme for underwater wireless sensor network had been proved in Ref. [13]. Mechanisms of the two type networks are the same including network topology, but means of communication are not. Therefore, the hierarchical positioning scheme of wireless sensor network can be used in underwater wireless sensor network.

6 Conclusion

The self-localization of underwater wireless sensor network nodes is discussed in this paper and different location algorithm is taken in accordance with the different accuracy of the node by the accuracy classification of anchor nodes. The simulation results show that the self-localization of nodes can be achieved by using the hierarchical localization algorithm. The communication radiuses between the anchor nodes and unknown nodes are increasing in accordance with the increased accuracy of anchor nodes and the decreased error of unknown nodes, and the error of unknown nodes is decreasing in accordance with the increased density of the anchor nodes. A new method is provided to further application for underwater wireless sensor networks of this algorithm.

References


