Research on Centroid Localization Algorithm in Wireless Sensor Networks

Pingfang Hu¹, Bo Zhang²

¹ Department, Computer Science Department, Wuhan Donghu University, Wuhan City, Hu Bei Province, 430212, China
² Department, Computer Science Department, Wuhan Donghu University, Wuhan City, Hu Bei Province, 430212, China

*Corresponding author’s e-mail: zhangbo@wdu.edu.cn

Abstract. Localization of nodes is the key technology for application of wireless sensor network. The positioning accuracy of centroid localization algorithm completely depends on the size and the distribution density of anchor nodes. Anchor nodes are small and the density of anchor nodes are distributed randomly in actual application, so the positioning accuracy of centroid localization algorithm is relatively low. In order to improve the accuracy of nodes, a method which is centroid localization algorithm of wireless sensor network is proposed based on received signal strength indicator. In the algorithm, firstly, point-to-point signal intensity is computed by RSSI and converts its inverse distance as the centroid algorithm. Then the weights of centroid localization algorithm are used for calculating coordinates of the unknown node. Through the simulation algorithm for testing and analysis, the simulation results show that, compared with the traditional algorithm, the positioning error of the centroid localization algorithm is obviously reduced and the node location accuracy is improved. It is an effective localization algorithm for application of wireless sensor networks.

1. Introduction

Wireless sensor network (WSN) monitors, perceives and collects the network information online in a self-organizing and cooperative way, and then transmits it to the user terminal. Wireless sensor network is low-cost, easy to arrange and flexible in network structure. It has a very broad application prospect in military and national defence, environmental monitoring, disaster relief, medical and health, industrial and agricultural control and other important fields. In the process of wireless sensor network application, it is necessary to continuously obtain relevant information and analysis it, so as to determine the source of data. For example, in military and national defence applications, it is not enough for wireless sensor network to only obtain the information of "enemy aircraft happened". Only when we get the data containing location information such as "where to find the enemy aircraft", can our army make corresponding countermeasures, which needs to be realized through the positioning technology of wireless sensor network. Therefore, location technology is the basis of wireless sensor network applications, and is an important research direction in wireless sensor network research.

In the application process of wireless sensor network, node localization is the first problem to be solved, which has attracted great attention of scholars at home and abroad. At present, according to whether the distance between nodes needs to be measured in the localization process, localization algorithms can be divided into two categories: range based localization and range free localization [3] [4].
Received signal strength indicator (RSSI) is the wireless signal strength value sent by the neighbour node from the end of the node, which can be used to determine the relative distance between nodes. It does not need additional hardware support, and can be measured by the sensor node itself. Aiming at the shortage of current localization technology in wireless sensor networks, a node localization method based on RSSI is proposed. This method combines RSSI technology with centroid algorithm, uses RSSI value to judge the relative distance between nodes, collects the distance information between all nodes, takes RSSI value as the weight of centroid algorithm, and then obtains the predicted position of unknown nodes through centroid algorithm. The simulation results show that the positioning accuracy of the algorithm is significantly higher than that of the original centroid positioning algorithm.

2. The basic principle of node location

In wireless sensor networks, according to whether the location of the node itself is known, the unknown node of the sensor node and the beacon node are separated. The unknown node indicates that the location of the node in the wireless sensor network needs to be located, while the beacon node, also known as the anchor node, indicates that its location has been determined. Beacon node obtains its own accurate position through manual deployment or global positioning system (GPS) positioning equipment, which is used as the reference node for unknown node positioning. Considering the limitation of network cost and energy consumption, the proportion of anchor nodes in sensor nodes is very small. The anchor node and the unknown node are deployed in the sensing area of the wireless sensor, as shown in Figure 1.

![Figure 1 anchor node and unknown node graph](image)

In the area of wireless sensor network, there are a large number of sensor nodes randomly deployed in it. If the manual positioning method is used, it will be a huge workload and unrealistic in large-scale wireless sensor network. Another way to directly obtain the node position is to equip each node with GPS. In theory, it can complete the sensor node positioning through GPS. However, due to the high cost of GPS equipment, high energy consumption and line of sight, it is too expensive to install GPS equipment on all nodes. At present, the main solution is to first use GPS positioning or pre place a small number of anchor nodes, and take these anchor nodes as reference nodes, and then use the node positioning algorithm to obtain the location information of the node by sending and receiving data with unknown nodes. In the range free node localization algorithm, the centroid localization algorithm is easy to implement and has less communication overhead, which has become the most widely used method in wireless sensor network localization. However, the positioning accuracy of the localization algorithm has a great relationship with the density and distribution of anchor nodes. If the anchor nodes are dense and distributed more evenly, the positioning accuracy will be higher, otherwise the positioning accuracy will be quite poor, which leads to the low positioning accuracy. The application is limited. In the ranging based location algorithm, the ranging technology based on the strength of RSSI has low cost, low energy consumption, and few hardware equipment, which attracts the attention of many scholars. However, the ranging error is large, which leads to the low positioning accuracy.
In this paper, RSSI ranging technology and centroid localization algorithm are combined to make use of their advantages to accurately locate wireless sensor nodes.

3. Centroid localization algorithm for Wireless Sensor Networks

3.1. RSSI node ranging

Because the signal strength will decrease in the process of signal propagation, the principle of RSSI is to send node broadcast packets. For the signal with transmission power value, the receiving node converts the transmission power difference into the distance between nodes according to the specific signal propagation model by comparing the received signal power and transmission power. Therefore, the closer the distance between nodes is, the larger the RSSI value is, and the farther the distance between nodes is, the smaller the RSSI value is. When RSSI is used to measure the absolute distance between nodes, each data frame received by the wireless communication transceiver module will get a RSSI value reflecting the change of signal strength in time, and then the propagation loss can be converted into the distance between nodes by using the path loss model. In this paper, the path loss model is called shadowing model. The shadowing model is obtained through the combination of analysis and experience. It includes two parts. The first part is the path loss model. $p_r(d)$ is used to represent the average energy received when the distance is $d$. In fact, a distance close to the center, $d_0$, is used as the reference. Compared with $p_r(d_0)$, $p_r(d)$ can be expressed as:

$$\frac{p_r(d)}{p_r(d_0)} = \left(\frac{d}{d_0}\right)^\beta$$

(1)

Where, $p_r(d_0)$ represents the average distance of $d_0$, and $\beta$ is the path attenuation factor. Generally, the empirical value is obtained from the site measurement. The more obstacles there are, the larger the relative value is. Therefore, with the increase of distance, the average energy received will decline faster and faster.

Generally, the path loss model is measured in dB, Therefore, the expression of path loss model is obtained by transforming equation (1):

$$\log \left(\frac{p_r(d)}{p_r(d_0)}\right) = -10\beta \log \left(\frac{d}{d_0}\right)$$

(2)

The second part of shadowing model represents the change of received energy when the distance is fixed, which is a lognormal random variable. Therefore, a complete shading model is expressed as follows:

$$\log \left(\frac{p_r(d)}{p_r(d_0)}\right) = -10\beta \log \left(\frac{d}{d_0}\right) + X_{dB}$$

(3)

Where, $X_{dB}$ is a Gaussian distribution variable with zero mean.

In this way, the signal strength of each unknown node when receiving the anchor signal is as follows:

$$\text{RSSI} = \text{transmit power} + \text{antenna gain path loss}$$

(4)

3.2. Centroid location algorithm

Centroid localization algorithm is an outdoor localization algorithm based on connectivity and independent of distance. The basic idea of centroid localization algorithm is: the unknown nodes first determine which anchor nodes are in their own network communication range, then take these anchor nodes as the centroid of the polygon, that is, the geometric center of the polygon, and finally estimate the position according to the centroid. In the process of practical application, the anchor node periodically broadcasts a signal of the anchor node's own ID and location information to the neighbour nodes. When the unknown node receives the information, it stores it. When the number of beacon signals of an anchor node exceeds the preset threshold, the node considers that it is connected with
the anchor node, and the node determines its position as the centroid of the polygon composed of all the connected anchor nodes. The specific description is as follows:

Suppose that in wireless sensor networks, according to certain rules with \((R_1, R_2, \ldots, R_n)\), n anchor nodes are deployed in a certain area. The anchor node sends S messages to the surrounding when its period is T, then the positioning time \(t\) is:

\[
t = T \times (S + 1 - \varepsilon)
\]  

Where, \(\varepsilon\) is a random number between 0 and 1.

In the positioning time \(t\), every time \(T\), the anchor node will send a beacon packet containing its own ID and location information to its neighbour nodes. The unknown node records the number of packets sent by each anchor node. When a positioning time \(t\) is completed, the communication success rate of each anchor node is calculated:

\[
CM_i = \frac{N_{\text{receive}}(i, t)}{N_{\text{send}}(i, t)} \times 100\%
\]

Where, \(N_{\text{receive}}(i, t)\) is calculated from the ID contained in the anchor node group, \(N_{\text{receive}}\) indicates the number of packets received from the anchor node \(R_i\) in the positioning time, \(N_{\text{send}}\) indicates the total number of packets sent by the anchor node \(R_i\) in the positioning time.

Set the threshold value of communication success rate of unknown node as \(CM_{\text{th}}\), when \(CM_i > CM_{\text{th}}\), the node is considered to be in the communication range of the anchor node \(R_i\) and connected with the anchor node \(R_i\). If the coordinate information of \(n\) anchor nodes is known, the coordinates of anchor nodes are \((x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\), these nodes form a polygon with \(n-1\) edges. If the unknown node \(D\) coordinates are \((x, y)\), the centroid of the polygon is the estimated value of the physical position of the unknown node \(D\):

\[
(x, y) = \left(\frac{x_1 + x_2 + \cdots + x_n}{n}, \frac{y_1 + y_2 + \cdots + y_n}{n}\right)
\]

It can be seen from the above analysis that the centroid algorithm has a direct relationship with the density of anchor nodes in the network. When the density of anchor nodes is low, the accuracy of the unknown node position obtained by the centroid localization algorithm is too low. If the RSSI ranging algorithm is directly used to calculate the unknown node position, the positioning accuracy is not ideal due to various factors, so we can combine the characteristics of the two The RSSI information is introduced into the centroid localization algorithm, and the RSSI value received by the node is used as the weight of the centroid to calculate the weight, so as to obtain higher accuracy.

3.3. Centroid location algorithm based on RSSI

In view of the defects of traditional centroid location algorithm, RSSI information is introduced into its location, which can play the role of auxiliary information. The centroid of the polygon is analysed. Based on the centroid localization algorithm, the weight of each node is added to reflect the influence of different anchor nodes on the unknown points. If the distance between the anchor node and the unknown node is closer, then according to the RSSI value, the greater the proportion of the corresponding node coordinates is, so the coordinate calculation of the unknown node \(D\) becomes:

\[
\begin{align*}
    x_{\text{est}} &= \frac{1}{n} \times \left(\frac{x_1}{RSSI(1)} + \frac{x_2}{RSSI(2)} + \cdots + \frac{x_n}{RSSI(n)}\right) \\
    y_{\text{est}} &= \frac{1}{n} \times \left(\frac{y_1}{RSSI(1)} + \frac{y_2}{RSSI(2)} + \cdots + \frac{y_n}{RSSI(n)}\right)
\end{align*}
\]

Therefore, the traditional weighted centroid algorithm makes full use of the RSSI data information, so the prediction accuracy of the centroid localization algorithm is improved. The specific steps of the algorithm are as follows:

1. Anchor nodes periodically broadcast their own information to neighbour nodes, which is composed of their own location information and node ID;
(2) After receiving the information, the unknown node only records the RSSI value of the same anchor node. If an anchor node receives multiple RSSI values, it needs to calculate its RSSI value and take the average RSSI.

(3) When the number of anchor node information received by the unknown node exceeds the predicted threshold K, it will not receive new information. Then, the unknown nodes are sorted according to the RSSI value from strong to weak, and the mapping and three sets of RSSI value and node to anchor node distance are established.

Anchor node set: \( B_{set} = (a_1, a_2, \ldots, a_k) \); distance mapping set from anchor node to unknown node: \( D_{set} = (d_1, d_2, \ldots, d_k) \); anchor node location set:

\[ P_{set} = \{(x_1,y_1),(x_2,y_2),\ldots,(x_k,y_k)\}; \]

(4) The RSSI value of unknown node is calculated by formula (4), and the RSSI value is converted into distance \( d \), from large to small, the distance \( D \) is divided into three groups, namely: \( \{d_1,d_2,d_3\}, \{d_4,d_5,d_6\}, \{d_7,d_8,d_9\}; \)

(5) Use formula (8) to calculate the three groups respectively, and get the coordinates of three locating points, so as to form a triangle;

(6) Taking the centroid of triangle composed of locating points as the location coordinates of unknown nodes \((x_m,y_m)\);

(7) The location error of unknown nodes is calculated:

\[
\text{ERROR} = \sqrt{(x_m - x)^2 + (y_m - y)^2}
\]

Where, \((x_m,y_m)\) is the coordinates of the unknown node, and \((x, y)\) is the real location of the unknown node. The flow chart of centroid location algorithm based on RSSI.

4. simulation study

4.1. simulation environment
The simulation is carried out in a hypothetical free space environment. The wireless sensor nodes are randomly deployed in a 100 m × 100 m square area with free space. There are 200 unknown nodes and 40 anchor nodes in the area. The location of the anchor node is known. The coordinates of the unknown node are determined by the C language random generating function. The location is shown in Figure 2, where the anchor node is marked with black Note that others are unknown nodes. Hardware environment: CPU: Core Duo 2. Memory: 2G, hard disk: 160g, operating system: Windows 10. All the experiments are programmed in C language and implemented on UNIX platform.

4.2. Parameter setting and evaluation
In the process of RSSI ranging, due to the influence of air reflection, object blocking, multipath and climate, Gaussian noise \(X_{db} \) with mean value of 0 and variance of 10, is the path attenuation factor \(\beta\), taking 5. In order to verify the effectiveness of the localization algorithm, there is no improved centroid localization algorithm in the reference model. The evaluation index of the model is average positioning error.

![Figure 2 deployment of wireless sensor nodes](image)
4.3. Results and analysis

the average positioning error of the two algorithms in the network under different number of anchor nodes. It can be seen that the average positioning error of the two algorithms decreases with the increase of the number of anchor nodes. However, under the same conditions, the average positioning error of this algorithm is less than the original centroid positioning algorithm, because with the increase of anchor node density, the positioning accuracy is higher and higher, and the positioning algorithm in this paper introduces the RSSI information, so the accuracy is improved more, and the error is smaller.

Results shows the average positioning errors of the two algorithms show a downward trend with the increase of the number of nodes. Under the same conditions, the average positioning error of the improved text location algorithm is far less than that of the original centroid location algorithm. This is because the number of nodes is increased, the distribution of nodes is more uniform and reasonable. At the same time, the improved location algorithm in this paper gets more information and eliminates the impact on the average location error.

5. conclusion

 Wireless sensor network nodes are deployed in an open environment, the location of most of the nodes can not be determined in advance, which is uncontrollable. But in the specific application process, we need to obtain the geographic location information of the nodes, so that we can know the exact location of the information source. Therefore, the node positioning technology plays a very important role in the practical application of wireless sensor network Use. The traditional centroid algorithm is widely used in the localization of wireless sensor nodes because of its simple computational complexity, but it does not consider the RSSI information, so the positioning accuracy is not high. Therefore, based on the analysis of the sensor centroid localization algorithm, this paper proposes a weighted centroid localization algorithm based on RSSI. The simulation results show that compared with the traditional centroid localization algorithm, the positioning accuracy of this algorithm is greatly improved, and the positioning error is correspondingly reduced, which provides an effective positioning method for large-scale wireless sensor networks.

Acknowledgments

This paper was co-supported by Hubei Key Laboratory of Big Data in Science and Technology in 2019, “Research and application of regional science and technology based on big data of hot spring industry in Dawan District of Guangdong, Hong Kong and Macao” under grant 20KF011014.

References

[1] Ma, Z. C., Sun, Y. N., Mei, T. Overview of wireless sensor networks [J]. Acta communication Sinica, 2014, 25 (4): 114-124.
[2] Bulusu, N., Heidemann, J., Esrin, D. GPS-less Low Cost Outdoor Localization for Very Small Devices[J].IEEE Personal Communications,2010,7(5):28-34.
[3] Wang, F. B., Shi,L., Ren,F.Y. Self localization system and algorithm in wireless sensor networks [J]. Acta Sinica Sinica, 2015, 16 (5): 1148-1157.
[4] Su, J., Wan,J.W., Wang,Y. Research on relative localization algorithm in wireless sensor networks [J]. Journal of sensor technology, 2017, 20 (12): 2695-2700.