

## Water Quality Monitoring Network Localization Method Based on RSSI Ranging

CHANG Bo<sup>1, a</sup>, ZHANG Xin-rong<sup>2, 3, b</sup>, XU Bao-guo<sup>3, c</sup>

<sup>1</sup>Faculty of Electronic Information Engineering, HuaiYin Institute of Technology, huaiyan 223003, China

<sup>2</sup>Faculty of Automation Engineering, HuaiYin Institute of Technology, huaiyan 223003, China

<sup>3</sup>School of Internet of Things Engineering, JiangNan University, wuxi 214122, China

ammm33534@sohu.com, bnn33@163.com, xbg@jiangnan.edu.cn

**Keywords:** Water quality monitoring; Wireless sensor network (WSN); RSSI ranging; WLS algorithms; Positioning

**Abstract.** In order to understand the water environment quality and improve the production efficiency, a wireless sensor network localization algorithm about water quality monitoring based on RSSI ranging is proposed. Firstly the dynamic path loss index is obtained for ranging, then the measurement range is corrected by the relative error coefficient of the reference node, finally, the weighted least square (WLS) algorithm is used to estimate the node coordinates. By simulation and localization experiments, it shows that this method takes full consideration of the influence of the location error and the number of reference nodes on the positioning accuracy, which can meet the requirements of complex network environment and the location of the monitoring system. Compared with the LS algorithm, the proposed localization algorithm has higher accuracy and lower computational complexity.

### Introduction

Sensor Networks Wireless (WSN) is able to self-organize and rapidly expand, through its internal numerous sensor nodes to perceive corporately and collect monitoring information and sends the processing results to the terminal equipment [1]. WSN has the characteristics of high fault tolerance and long working life, which makes it have a wide application prospect in the field of environmental water quality monitoring in the lake river, and attracts many scholars to carry out the relevant application research [2]. Location problem is a key problem in the water quality monitoring of wireless network. Global Positioning System (GPS) is limited by the cost and deployment environment, and cannot be implemented, so locating research is necessary. The solution is to estimate the localization algorithm using wireless communication based on the location of the anchor node [3].

The positioning algorithm is divided into two categories, that is, range-based and range-free algorithm [4]. The time of arrival (TOA) and angle of arrival (AOA) ranging algorithm can obtain higher accuracy by distance or angle measurement, but the cost is higher. RSSI (Signal Strength Indicator Received) algorithm is used to measure the received signal strength, and to estimate the distance between nodes using the wireless communication mode. The power consumption and the cost of RSSI are lower, but RF signal propagation modeling is influenced by environmental factors, so appropriate measures should be taken to reduce the ranging error [5].

Some scholars have proposed the method for calculating the parameters of the path loss model for on-line real time correction on the basis of the study of signal propagation modeling and no ranging [6]. The RSSI ranging algorithm was improved by analyzing the signal propagation model [7], so the location accuracy of the monitoring nodes was improved. The dynamic ranging method based on RSSI and LQI is proposed by using the piecewise linear approximation method to obtain the function curves of RSSI and LQI [8]. The reference [9] is based on the Bayes estimation method, using the grid algorithm to reduce the computational complexity of the positioning algorithm, and

improve the efficiency of positioning algorithm. The combined ranging algorithm is also emphasized combined with multiple ranging methods. On the basis of the observation of TOA and RSSI, a kind of joint positioning algorithm is proposed to improve the positioning accuracy [10]. The above methods have some practical value, but their disadvantage are that it is more complex and more operational, which is not suitable for the use of water quality monitoring system.

In this paper, we designed a location algorithm, in which the path loss factor in the current monitoring area is obtained through the periodic measurement of RSSI, and the relative error of the measurement distance and the actual distance between reference nodes are obtained. The error correction of the ranging is greatly reduced the bad influence of the signal propagation modeling. In the positioning stage, the node position is solved by using the Taylor series least squares iterative algorithm, which can reduce the computation complexity of node coordinates and consider the prior information of the observed data, so it can improve the positioning accuracy.

### Monitoring Node Ranging Based on RSSI

**RSSI Ranging Model.** In the lake and river environment, there are different species and different sizes of fishes and other intensive facilities, and there is uneven distribution. Therefore, there are factors such as multipath, diffraction and barriers making the RF signal propagation model become more complex. In this paper, we use a simplified lognormal distribution model. RSSI values can be expressed as:

$$P_R(d) = P + G - P_L(d) \quad (1)$$

where  $P_R(d)$  (unit:  $dBm$ ) is the signal intensity after a distance of  $d$ ,  $P$  is a node transmit power,  $G$  is a node antenna gain,  $P_L(d)$  is the signal strength loss after a distance of  $d$ . RF communication modeling can also be written as the following expression:

$$P_R(d) = P_R(d_0) - 10n \lg(d/d_0) + X_\sigma \quad (2)$$

where  $n$  is the path loss factor;  $d_0$  is the reference distance, which value is  $1m$  normally;  $P_R(d_0)$  is the received signal strength at  $d_0$  after the signal is transmitted;  $X_\sigma$  is the Gaussian random variables which mean is 0.  $P_R(d_0)$  is the signal intensity after a distance of  $d_0$ , calculated by Eq. 1 as follows:

$$P_R(d_0) = P + G - P_L(d_0) \quad (3)$$

The path loss  $P_L(d_0)$  after the reference distance  $d_0$  is:

$$P_L(d_0) = -10 \lg \left[ \frac{G_t G_r \lambda^2}{(4\pi)^2 d_0^2 L} \right] \quad (4)$$

where  $G_t$  (unit:  $dB_i$ ) is the transmit antenna gain of a node;  $G_r$  (unit:  $dB_i$ ) is the receiving antenna gain;  $L$  is the system loss coefficient;  $\lambda$  is the RF signal wavelength.

**Dynamic Acquisition of Path Loss Exponent.** The received signal strength RSSI between the pending nodes and the reference nodes is measured periodically, and then the path loss exponent  $n$  is calculated by using the RF wireless signal propagation model. By Eq. 2, it can be obtained:

$$P_R(d) = PX - 10n \lg(d/d_0) \quad (5)$$

where,  $PX = P_R(d_0) + X_\sigma$ . If the reference node  $R_3$  receives the RSSI value of other two reference nodes,  $R_1$  and  $R_2$ , there are the following formulas:

$$P_R(d_1) = PX - 10n \lg(d_1/d_0), \quad P_R(d_2) = PX - 10n \lg(d_2/d_0) \quad (6)$$

where  $P_R(d_1)$  is the value between  $R_3$  and  $R_1$ ,  $P_R(d_2)$  is the value between  $R_3$  and  $R_2$ ,  $d_1$  is the actual distance between  $R_1$  and  $R_3$ ,  $d_2$  is the actual distance between  $R_2$  and  $R_3$ . When  $d_0 = 1m$ , then:

$$P_R(d_1) - P_R(d_2) = 10n \lg(d_2/d_1) \quad (7)$$

From Eq. 7, we know that the path loss exponent  $n$  depends on  $P_R(d_1)$ ,  $P_R(d_2)$  and  $d_1$ ,  $d_2$ . This method not only enhances the adaptability of the RSSI ranging algorithm to the monitoring environment, but also improves the positioning accuracy of the nodes.

**Acquisition of Accurate Ranging Value.** In the actual location, the error of RSSI measurement results in inaccuracy for the distance between nodes calculated by Eq. 2. One solution is to obtain the relative ranging error of RSSI by comparing the measurement distance and the actual distance between the reference nodes, and then to correct error as locating the monitoring node to improve the accuracy of the measurement.

The actual distance from the reference node  $R_i(x_i, y_i)$  to the other reference node is marked as  $r_{ik}$ ,  $i = 1, 2, \dots, n$ ,  $k \neq i$ , where  $n$  is the number of reference nodes in the network. The measurement distance from  $R_i$  to other reference nodes is marked as  $d_{ik}$ , then the average relative error between the actual distance and the measurement distance for the reference node  $R_i$  is as follow:

$$\mu_i = \frac{1}{n-1} \sum_{k=1, k \neq i}^n \frac{r_{ik} - d_{ik}}{d_{ik}} \quad (8)$$

where,  $\mu_i$  reflects the accuracy of the measurement distance for the reference node  $R_i$ , called the relative error coefficient.

After the node receives a broadcast message from the reference node, according to the value of  $P_R(d)$ , the measurement distance  $d$  is calculated by Eq. 2, then based on the  $\mu_i$  value of each reference node, the measurement range is corrected by the Eq. 9.

$$d_{ui}^c = d_{ui}(1 + \mu_i) \quad (9)$$

where,  $d_{ui}$  is the measurement distance between the monitoring node and the reference node  $R_i$ ,  $d_{ui}^c$  is the correction distance between the monitoring node and the reference node  $R_i$ .

### Monitoring Node Localization Algorithm

For the monitoring nodes in the water quality monitoring network, there is a distance equation:

$$d_i^2 = (x - x_i)^2 + (y - y_i)^2 \quad (10)$$

Supposing that the initial coordinate is  $(x_0, y_0)$ , we assume:

$$f(x, y) = \sqrt{(x - x_i)^2 + (y - y_i)^2} \quad (11)$$

At the point of  $(x_0, y_0)$ , by Taylor series expansion to the Eq. 11, and neglect of the two orders and more weight, we have:

$$f(x, y) = f(x_0, y_0) + f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y \quad (12)$$

Eq. 12 can also be written as:

$$\hat{d}_i = d_{i0} + a_x \Delta x + a_y \Delta y \quad (13)$$

where,  $\hat{d}_i = \sqrt{(x - x_i)^2 + (y - y_i)^2}$ ,  $d_{i0} = \sqrt{(x_0 - x_i)^2 + (y_0 - y_i)^2}$

$$a_x = f_x(x_0, y_0), \quad a_y = f_y(x_0, y_0)$$

Considering the error, Eq. 13 can be abbreviated as:

$$h = G\Delta + \varepsilon(d_2) \quad (14)$$

where,  $\Delta = [\Delta x \quad \Delta y]^T$ ,  $h = [\hat{d}_1 - \hat{d}_{10}, \dots, \hat{d}_N - \hat{d}_{N0}]^T$

$$G = \begin{bmatrix} \frac{x_0 - x_1}{\hat{d}_{10}} & \frac{y_0 - y_1}{\hat{d}_{10}} \\ \vdots & \vdots \\ \frac{x_0 - x_N}{\hat{d}_{N0}} & \frac{y_0 - y_N}{\hat{d}_{N0}} \end{bmatrix}$$

The weighted least square (WLS) algorithm is used in the Eq. 14,  $\Delta$  can be estimated as:

$$\Delta = (G^T W G)^{-1} G^T W h \quad (15)$$

where,  $W$  is the covariance matrix of measurement error, that is:  $W = \text{cov}^{-1}(\varepsilon)$ , then the updated parameter vector can be obtained:

$$\theta(k+1) = \theta(k) + \Delta \quad (16)$$

### Simulation and Experimental

In order to reflect the influence of the number of anchor nodes, the density and the communication radius on the positioning error, the average position error of the nodes is the main evaluation criteria. The average value of the results obtained from the simulation results of the 100 times under the same parameters is used as the final result. The location error of the nodes in the network is defined as  $E_{ai}$ .

$$E_{ai} = \frac{\sqrt{(x_{ci} - x_i)^2 + (y_{ci} - y_i)^2}}{R} \quad (17)$$

where,  $R$  is the communication radius,  $p_i = [x_{ci} \ y_{ci}]^T$  is the final estimation position of the node  $i$ ,  $z_i = [x_i \ y_i]^T$  is the true position of the node  $i$ , and the average position error is  $E_a$ .

$$E_a = \frac{\sum_{i=1}^N \sqrt{(x_{ci} - x_i)^2 + (y_{ci} - y_i)^2}}{NR} \quad (18)$$

where,  $i=1, 2, \dots, N$ ,  $N$  is the number of unknown nodes in the network.

**Influence of Different Ranging Errors on the Positioning Accuracy.** The simulation environment is set to the rectangle area of  $100m \times 100m$ . The algorithm A represents the algorithm of this paper, the algorithm B is the Taylor series expansion algorithm based on the real coordinates, the algorithm C represents the LS basic algorithm. The number of reference nodes is  $n=20$ , the number of nodes is 400. The influence of different ranging errors on the positioning accuracy is tested; the simulation results are shown in Fig. 1.

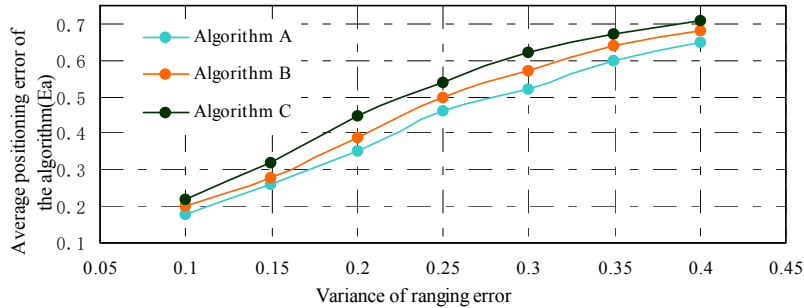


Fig. 1 Influence of different ranging errors on the positioning accuracy

From Fig. 1, the algorithm C is very sensitive to the range error, and with the increase of the distance measurement error, the positioning accuracy of the algorithm is obviously decreased. Algorithm B reduces the bad influence of ranging error on the positioning accuracy, but the effect is not obvious. The algorithm A can greatly reduce the bad influence of the ranging error, and thus has a higher positioning accuracy. With the increase of the error variance, the positioning accuracy of the three algorithms are gradually decreased, but the algorithm A's positioning accuracy has been higher than the algorithm B and C. The main reason is that the positioning error is mainly caused by the error of the ranging error caused by the unknown node positioning calculation error. Because the algorithm A uses the measurement range correction, the accuracy of the positioning is more obvious.

**Relationship between Number of Anchor Nodes and Location Accuracy.** The simulation environment is set to the rectangle area, and the 100 nodes are randomly distributed in the region of  $100m \times 100m$ , the node's communication radius is  $40m$ . The simulation results are shown in Fig. 2.

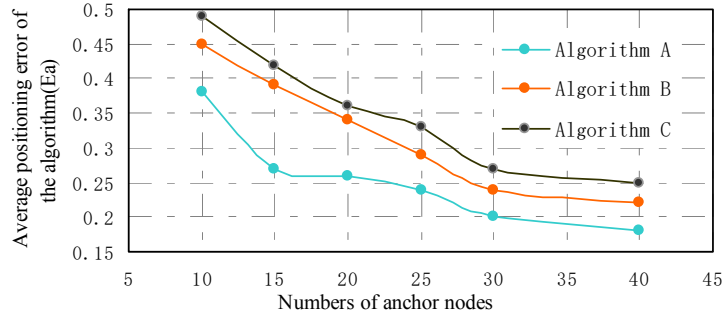


Fig. 2. Relationship between number of anchor nodes and location accuracy

The results from Fig. 2 show that the algorithm *B* and *C* are less than the number of reference nodes. The main reason is that the information used to calculate distance and position is reduced in the network, so the distance error between the unknown node and the anchor nodes become large, when the reference node ratio is lower. The algorithm *A* can effectively correct the measurement distance using the correction factor of multiple reference nodes, so it can reduce the positioning error caused by less reference nodes.

**Influence of the Communication Radius of Test Nodes on the Positioning Accuracy.** The simulation environment is set to the rectangle area, and the 100 nodes are randomly distributed in the region of  $100m \times 100m$ , the number of reference nodes is 15. The simulation results are shown in Fig. 3.

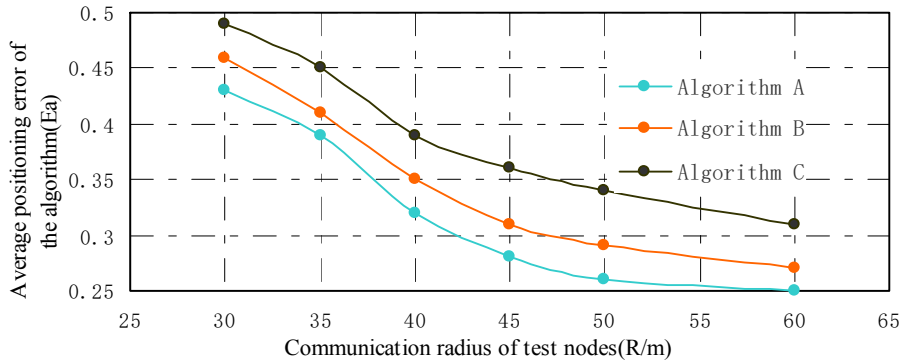


Fig. 3 Influence of the communication radius of test nodes

The results from Fig. 3 show that as the distance between the nodes increases, the reference nodes in the communication range of the monitoring node are gradually increased, so the information between the monitoring nodes and reference nodes increases. Because the more distances between anchor nodes are used to correct the distance from the unknown node to the anchor node, so the positioning accuracy is also improved. Under the same conditions, the positioning accuracy of the algorithm *A* is higher than that of *B* and *C*.

**Experiment and Analysis.** In the laboratory (about  $15m \times 15m$ ), a small wireless sensor network test system is built using CC2530 nodes. The system has 6 anchor nodes, which are deployed in the test area evenly. In addition, the 10 unknown nodes and one sink node are deployed in the selected position, and the whole system includes one console. The distance between the nodes is  $15m$ , the height from a node to the ground is about  $0.5m$ , and the data is transmitted every 20 seconds. The mean value of the 50 measurements is taken as the experimental result. According to the experimental data, the experimental results are shown in Table. 1.

Table. 1 Experimental Result

Number of nodes to be located	Node actual position	Location after measurement	Distance of two position	$E_{ai}$
01	(3.0, 3.0)	(4.2, 3.5)	1.30	0.09
02	(3.0, 9.0)	(3.4, 7.1)	1.96	0.13
03	(6.0, 6.0)	(5.3, 8.1)	2.21	0.15
04	(6.0, 12.0)	(6.6, 10.7)	1.43	0.10
05	(9.0, 3.0)	(7.4, 3.2)	1.61	0.11
06	(9.0, 6.0)	(8.5, 7.6)	1.68	0.11
07	(12.0, 3.0)	(10.4, 4.1)	1.94	0.13
08	(12.0, 6.0)	(11.1, 7.9)	2.10	0.14
09	(12.0, 9.0)	(10.8, 9.3)	1.24	0.08
10	(15.0, 6.0)	(12.7, 5.2)	2.44	0.16

By Table 1, the mean value of the algorithm A in the actual test environment is 0.120. Under the same conditions, the average positioning error of the simulation experiment is 0.113, and the positioning accuracy under actual test conditions is lower than that of the simulation. The main reason is that the RF signal between the test nodes is affected by the wall, the equipment and the barrier in the actual test environment, which leads to an increase in the error of RSSI measurement, and the positioning accuracy impaired.

The 16 sets of test data are analyzed by using the algorithm proposed in this paper. The positioning results are the mean value of the results obtained from the 50 experiments. The positioning error of the test data is shown in Fig. 4.

In the Fig. 4, the maximum value is 0.35, the minimum value is 0.18, and the mean value is 0.27. As the 6 anchor nodes are deployed in the laboratory test area, the location error of the regional edge is larger, and the overall positioning effect is better. If some more reference nodes are deployed at the edge of the test area, the location effect will be further improved.

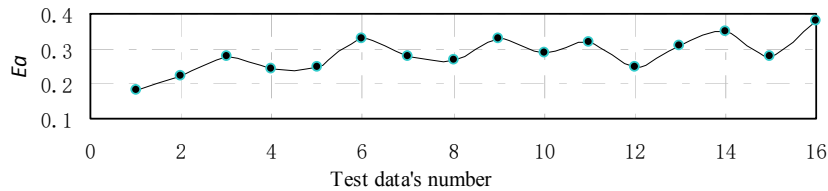


Fig. 4 Positioning error of the this paper's algorithm

## Conclusion

In this paper, an algorithm about water quality monitoring network location based on RSSI ranging is proposed. Firstly, RF wireless signal propagation model is used to obtain the path loss factor in the current monitoring region, then the ranging is corrected based on relative error, and in the positioning stage, the linear processing method of Taylor series least squares iterative algorithm is used to solve the position coordinates of the monitoring nodes. Experiments have shown that the proposed algorithm can meet the requirements of the network environment, and the positioning of the water quality monitoring system. Compared with the LS algorithm, the mean value of positioning error is 2.1425m and 2.9251m, and the average running time is 0.2372s and 1.9163s respectively, which shows that the proposed localization algorithm has higher accuracy and lower computational complexity.

## Acknowledgement

This work was supported by the National Natural Science Foundation of China (Grant No.30971689, No.21276111), We would like to thank the anonymous reviewers for their perspicacious comments.

## References

- [1] Zeng X Z, Liu G, Zheng D P. “Study and development of a field information acquisition system based on wireless technique”[C]∥ Actual Tasks on Agricultural Engineering, Opatija, CROATIA, 2006: 371~377.
- [2] ZHAO Xiao-qiang. “Intelligent Environmental System for Water Remote Monitoring”[J]. Computer Engineering, 2010, 36(17): 93-97.
- [3] Munoz D, Bouchereau F, Vargas C, “Position Location Techniques and Applications[M]. Burlington”, USA, Academic Press, 2009: 17—21..
- [4] Amitangshu P . “Localition algorithms in wireless sensor networks:current approaches and future challenges”[J]. Network Protocols and Algorithms, 2010, 2(2):45—74.
- [5] Xu Lei, Shi Weiren. “Stepwise refinement location algorithm for wireless sensor network”[J]. Chinese Journal of Scientific Instrument, 2008, 29(2): 314—319. (in Chinese with English abstract).
- [6] G. Mao, B. D. O. Anderson. “Path loss exponent estimation for wireless sensor network localization” [J]. Computer Networks 2007, 51(10): 2467-2483.
- [7] Wang Wei, Chen Dai. “RSSI positioning algorithm based on range correction and position correction” [J]. Computer engineering and design,2011,32(2): 409-412.
- [8] Zhang Jieyin, Sun Maojie. “Dynamic distance estimation algorithm based on LQI and RSSI” [J]. Electronic measurement technology,2007,30(2): 142-145.
- [9] C. Morelli, M\_ Nicoli. “Particle filters for RSS-based localization in wireless sensor networks: an experimental study”[C]. IEEE International Conference on Acoustics, Speech and Signal Processing, Toulouse, 2006: 957-960.
- [10] N. Patwari, A. O. Hero. “ Relative location estimation in wireless sensor networks” [J]. IEEE Transactions on Signal Processing, 2003, 51(8): 2137-2148.