A Localization Algorithm of Wireless Sensor Network Based on Statistical Uncorrelated Vector

Min Wang
Hunan Mechanical & Electrical Polytechnic, Hunan, China
wangmin_507@126.com

Abstract—For exploring wireless sensor network – a self-organized network – a new node location algorithm based on statistical uncorrelated vector (SUV) model, namely SUV location algorithm, is proposed. The algorithm, by translating the node coordinate, simplifies the solution to double center coordinate matrix, and gets the coordinate inner product matrix; then it uses statistical uncorrelated vectors to reconstruct the coordinates of the inner product matrix and remove the correlation of inner matrix of coordinates caused by the ranging error, so as to reduce the impact of ranging error on subsequent positioning accuracy. The experimental results show that the proposed algorithm does not consider the network traffic, bust still has good performance in localization. At last, it is concluded that reducing the amount of communication of sensor nodes is beneficial to prolong the service life of the sensor nodes, thus increasing the lifetime of the whole network.

Keywords—wireless sensor, SUV, inner product matrix

1 Introduction

In the localization process of wireless sensor networks, it is generally assumed that the neighboring nodes can communicate with each other, and can measure the distance between each other. Usually, the smaller the ranging error, the higher the positioning accuracy, so it is always expected that through improving the hardware structure of the node or increasing the transmission power of node, it can reduce the ranging error, but the ranging method with high precision will cause the increase of node hardware cost and energy consumption. For example, TOA ranging technology requires strict time synchronization, TDOA ranging technology requires the two signal transmission equipment, and AOA ranging technology requires multiple signal receiving devices. Increasing the transmitting power of nodes will shorten the service life of wireless sensor nodes, and all these do not follow the design principles of wireless sensor network nodes with low cost and low power consumption. In RSSI, TOA, TDOA, and AOA four kinds of common ranging technologies, positioning mechanism hardware cost based on RSSI has the lowest cost and it is most easily implemented, so most popular among researchers at home and abroad, but relatively, the ranging error is the maximum [1]. The basic principle of RSSI ranging is to convert

the signal intensity attenuation into the signal propagation distance. Since that the wireless signal, with the increase of transmission distance, the signal intensity will be reduced by a certain law. However, the propagation characteristics of signals are related to environmental parameters, and the signal attenuation rule is affected by the temperature, humidity, obstacles, communication mode, multipath effect and other factors [2-3], so it is difficult to establish a unified signal transmission attenuation model. In addition, the RF antenna height is also an important factor affecting the signal transmission attenuation model. The accuracy of the localization algorithm based on RSSI is not only related to the ranging accuracy based on the signal intensity, but also related to the characteristics of the localization algorithm. Many scholars have put forward a lot of RSSI positioning algorithms. For example, the famous RADAR positioning system, multi-point positioning method [4], the maximum a posteriori probability location algorithm based on the model and the map [5].

This paper is devoted to the study of location problem based on RSSI ranging. In order to reduce the influence of ranging errors on the positioning accuracy, a location algorithm based on statistical uncorrelated vector set (SUV) is proposed. The algorithm, by coordinate transformation, simplifies the solution equation of the center coordinate matrix, uses statistical uncorrelated vectors to reconstruct the coordinates of the inner product matrix and reduce the ranging noise interference, and directly uses the inner product matrix to calculate the coordinates of the nodes. The experimental results showed that the SUV location algorithm can effectively improve the positioning accuracy in the case of large ranging error, especially suitable for wireless sensor network node localization based on low cost hardware.

2 SUV location algorithm

2.1 Centralized SUV localization algorithm

A centralized location algorithm is an algorithm for focusing the desired information to a central node or a computer, and locating a single or all of the nodes. Suppose the original coordinate of N nodes $\{X_1, X_2, \ldots, X_N\}$ is $X_i'\left(X_i' \in R^d\right)$, and the set $N = \{1,2\},\ldots N\}$ represents a collection of sensor nodes. All nodes are transitional, so X_1' is at the origin of coordinate, the translation of $a = (a_1, a_2,\ldots,a_d)$, then the new coordinate is $X_i\left(X_i \in R^d\right)$, in which X_1' is an arbitrary node in the reference nodes, and then the coordinate matrix of reference node after transition is $X_N = [X_1, X_2,\ldots,X_N]^T$. The 2-2 distance between all nodes forms a measurement matrix $D_N = \left(\tilde{d}_{ij}\right)_{N \times N}$, where the measurement distance between the nodes is:

$$\tilde{d}_{ij} = N\left(d\left(x_i, x_j\right), \sigma\right) \qquad i, \quad j \in N \tag{1}$$

In (1), $d(x_i, x_j) = \sqrt{\sum_{\lambda=1}^{d}(x_{i\lambda} - x_{j\lambda})^2}$ is the Euclidean distance between the sensor node X_i and X_j. In order to get the inner product between the coordinates of the nodes, introduce the coordinate inner product matrix $B_N = \left(d_{ij}\right)_{N \times N}$, and b_{ij} is:

$$b_{ij} = \left(\tilde{d}_{1i}^2 + \tilde{d}_{1j}^2 - \tilde{d}_{ij}^2\right)/2 \tag{2}$$

Because of the error of the measurement value $\tilde{d}_{ij}$, the correlation of the data that is only correlated in the d dimension is extended to the high dimension, and the eigenvalue decomposition of B_N is performed, and the number of eigenvalues greater than 0 will be greater than d. If only the first d eigenvalues and their corresponding eigenvectors are chosen to form the relative coordinate, the error will be larger.

In wireless sensor networks, because the B_N is usually singular, there is at least one eigenvalue of 0 for the B_N feature decomposition, which leads to the in-existence of the inverse matrix of the eigenvalue matrix. In the following, the eigenvalue decomposition of B_N can be expressed as:

$$B_N = \left(V_1, V_2\right)\begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix}\begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix} \tag{3}$$

Define $B_N' = V_1 D_1 V_1^T$, where D_1 is a diagonal matrix composed of r eigenvalue greater than 0, V_1 is the matrix composed of the eigenvectors corresponding to the r eigenvalue greater than 0, D_2 is the diagonal matrix constituted by the rest N+1-r eigenvalues, and V_2 is the matrix constituted by the eigenvectors corresponding to the eigenvalues of D_2.

The coordinates of unknown nodes can be calculated by $X_N x_k^T = b_k'$:

$$x_k^T = \left(X_N^T X_N\right)^{-1} X_N^T b_k' \tag{4}$$

In (4), $b_k' = \begin{bmatrix} b_{1,\,k}' \\ b_{2,\,k}' \\ \cdots \\ b_{m,\,k}' \end{bmatrix}$, $k = \{M+1, M+2, ..., N\}$, then the absolute coordinate of the unknown node is:

$$x'_k = x_k - a \tag{5}$$

2.2 Distributed SUV localization algorithm

The distributed localization algorithm only needs to communicate with the neighbor nodes during the positioning process and the location process is completed in the node. In the distributed SUV localization algorithm, only the coordinates of anchor nodes are used, and the distance information between unknown nodes and anchor nodes is specified, to make a distributed location of the specified unknown nodes.

It is assumed that in the wireless sensor, there are K anchor nodes {X1, X2,... Xk}, the original coordinate is $X'_i (X'_i \in R^d)$, and the set K={1,2,... K} represents a collection of K anchor nodes. Preform the coordinate translation of all anchor nodes, to make X'_1 located in the origin of coordinates, and the translation is a=(a₁,a₂,...,a_d), then re-set the new coordinate $X_i (X_i \in R^d)$, in which X'_1 is an arbitrary node of anchor node, and then the anchor node coordinate matrix after the translation is $X = [X_1, X_2,..., X_K]^T$. Because the anchor node coordinate information is known, the distance between anchor nodes is determined. The 2-2 distance between the anchor nodes forms a measurement matrix $D_K = (d_{ij})_{K \times K}$, in which $d_{ij} = \sqrt{\sum_{\lambda=1}^{d} (x_{i\lambda} - x_{j\lambda})^2}$ is the Euclidean distance between nodes, $i, j \in K$.

Similarly, in order to get the product relationship between the node coordinates, introduce the coordinate inner product matrix of coordinates $B_K = (b_{ij})_{K \times K}$, in which b_{ij} is:

$$b_{ij} = (d_{1i}^2 + d_{1j}^2 - d_{ij}^2)/2 \tag{6}$$

In distributed location system, the unknown nodes, through two ways, get the parameter a, D_K and B_K. One is to directly store the three parameter in the unknown nodes, the other is through a wireless signal to seed to the unknown nodes. In wireless sensor networks, for arbitrary unknown node K+1, we can the measure distance $\tilde{d}_{1,(k+1)}, \tilde{d}_{2,(k+1)},..., \tilde{d}_{k,(k+1)}$ between it and all the other anchor nodes by the ranging technology. We assume that the unknown nodes can communicate with all the anchor nodes.

Because $b_{ij} = b_{ji}$, the coordinate inner product matrix B_{K+1} can be represented by the following formula:

$$B_{k+1} = \begin{bmatrix} & & & b_{1,(k+1)} \\ & B_k & & ... \\ & & & b_{k,(k+1)} \\ b_{1,(k+1)} & ... & b_{k,(k+1)} & b_{(k+1),(k+1)} \end{bmatrix}$$

(7)

Similar to the centralized SUV localization algorithm, B_{K+1} is usually singular, and there is at least one eigenvalue of 0 when B_{K+1} is conducted with characteristics decomposition, which will cause that the inverse matrix of eigenvalue matrix does not exist. In the following, the eigenvalue decomposition of Bn can be expressed as:

$$B_{k+1} = (V_1, V_2) \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix}$$

(8)

In (8), D_1 is the diagonal matrix composed of r eigenvalues more than 0, V_1 is the matrix composed of the eigenvectors corresponding to r eigenvalues more than 0, D_2 refers to the diagonal matrix composed of the rest of K+1-r eigenvalues, and V_2 is the matrix composed of the eigenvectors corresponding to the eigenvalues of D_2.

Reconstruct the coordinate inner product matrix:

$$B_{k+1}^{'} = V_1 D_1 V_1^T$$

(9)

Because $B_k = (b_{ij})_{k \times k} = X_k X_k^T$, we can get:

$$X_k x_{(k+1)}^T = b_{(k+1)}^{'}$$

(10)

Then the unknown node coordinate $x_{(k+1)}^T$ can be calculated:

$$x_{(k+1)}^T = (X_k^T X_k)^{-1} X_k^T b_{(k+1)}^{'}$$

(11)

In (11), $b_{k+1}^{'} = (b_{1,(k+1)}^{'}, b_{2,(k+1)}^{'}, ..., b_{(k+1),(k+1)}^{'})^T$.

At last, through the coordinate transition, we can get the absolute coordinate of the unknown nodes is:

$$x_{(k+1)}^{'} = x_{(k+1)} - a$$

(12)

3 Experimental simulation

The distance measurement technology based on RSSI, by the strength of the RSSI signal received, makes use of signal propagation attenuation model to obtain the distance between the signal point to the receiving point. Radio frequency signal is used in wireless sensor localization. Because the RSSI is easy to be affected by the environment and other factors, there is no fixed propagation error model. It is necessary to determine the model parameters according to the specific environment for on-site measurement.

Select the CC2430 chip as the main chip of the experimental node. CC2430 is the first 2.4 GHz RF system single chip in line with ZigBee technology [6], which adopts the CC2420 chip architecture, and integrates ZigBee RF front-end, memory and micro controller on a single chip. The current loss of the chip in working is 27mA, and the current loss is less than 27mA or 25mA in the receiving or transmitting mode. Hardware supports for CSMA/CA function, digital RSSI/LQI and powerful DMA function. Each node can receive / transmit frequency signal; a node sends a signal, and the other node receives the signal sent and measures the RSSI. The parameters of the RSSI signal transmission error model are estimated in the lab.

3.1 RSSI signal transmission attenuation model [7].

P_{ij} is defined as the signal strength (mW) of node j received by node i, and assuming that $P_{ij}=P_{ji}$, P_{ij} is a log normal random variable, and $P_{ij}(dBm)=10\log_{10}P_{ij}$ obeys the Gauss distribution:

$$P_{ij}(dBm) \sim N\left(\overline{P_{ij}}(dBm), \sigma_{dB}^2\right) \tag{13}$$

$$\overline{P_{ij}}(dBm) = P_0(dBm) - 10n_p \, log_{10}\left(d_{ij} / d_0\right) \tag{14}$$

$\overline{P_{ij}}(dBm)$ refers to the energy average value (DB MW), σ_{dB} is the standard deviation, $d_{ij} = \sqrt{\sum_{\lambda=1}^{d}\left(x_{i\lambda} - x_{j\lambda}\right)^2}$ is the Euclidean distance between the node i and the node j, $P_0(dBm)$ indicates the energy received when the node has the distance of d_0, and n_p suggests the path loss coefficient related to environment.

Use CC2430 production nodes to receive and transmit RF signal, and measure the RSSI value to verify the above statistical model. To obtain the above model, three parameters must be determined: d_0, $P_0(dBm)$ and n_p. Usually, take $d_0=1$ m, $P_0(dBm)$ value is the RSSI value measured when the distance between the node X_i and the node X_2 is $d_{ij}=1$m. When d_0 and $P_0(dBm)$ are determined, the node X_i and the node X_2 are placed in different positions, measuring the RSSI value for several times, substituting in (14) to calculate for several times, and the average value of n_p is obtained. So far, the three parameters of the RSSI signal attenuation model can be obtained.

In Figure 1, the dot represents the average value of RSSI signal when sensor nodes are in different distance. In this experiment, 10 RSSI values are measured at each position, and then averaged. According to the measured value, we can get RSSI signal transmission attenuation model parameters: $P_0(dBm)\approx-45$, $n_p\approx3$. Two parameters are substituted in (14), we can get RSSI- distance data model as shown in the smooth curve in Figure 1.
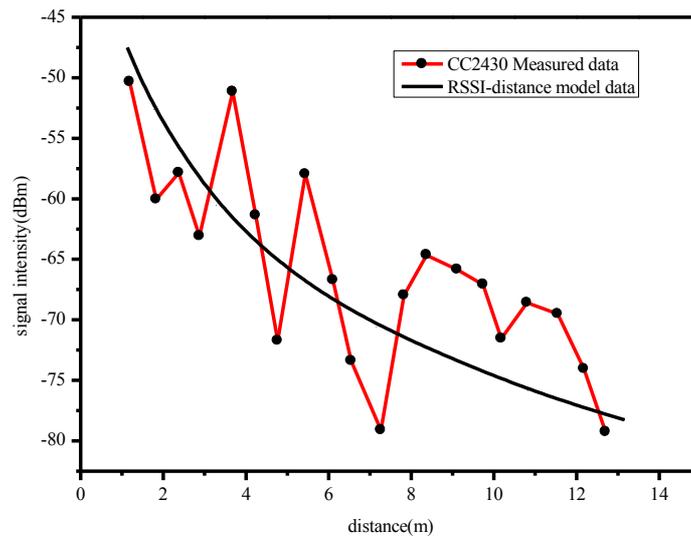
Fig. 1. Relationship between RSSI signal intensity and distance

It is found that, for two fixed sensor nodes with a certain distance, not only RSSI will fluctuate, but the farther away from each other, the larger the fluctuation frequency and amplitude. Figure 2 shows the distribution of the signal intensity with time when the transceiver is in the distance of 1, 2, and 3 meters, each position has a total of 105 sets of signals, and each signal is separated by 2s. We can see that the RSSI value fluctuates around the mean in a fixed position, and the greater the distance, the stronger the fluctuating. Figure 1 and Figure 2 show that the above model is in line with the actual situation.

3.2 Experimental results of SUV positioning based on RSSI

After the RSSI signal transmission attenuation model is obtained, the performance of the distributed SUV localization algorithm based on RSSI ranging is evaluated. Measurement environment is also selected in the laboratory. Eight reference nodes are placed around the laboratory of 10 x $10m^2$, and the equipment and furniture are normally placed indoors. The unknown node sends the signal, all reference nodes receive the signal, and the measured signal's RSSI value is reported to a gateway through the wireless signal, and transmitted to the PC server by the gateway to carry on the localization computation.
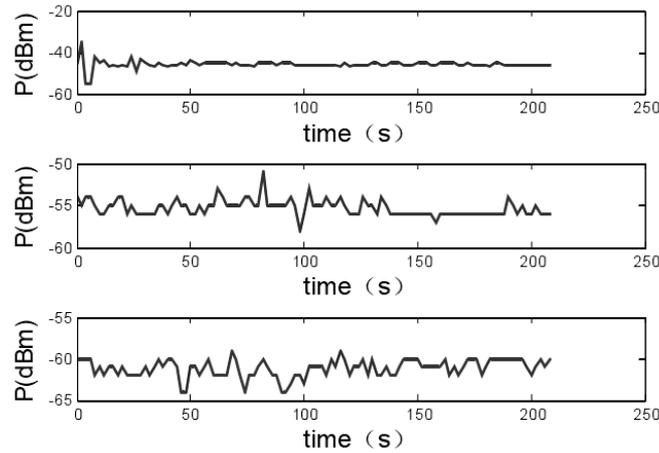
Fig. 2. RSSI signal intensity fluctuations with time of two nodes at the distance of 1, 2, 3 meters

4 Results and Discussion

Figure 3 shows the relationship between CD and the average connectivity of the network, the average number of neighbor anchor nodes, and the positioning error. With the increase of 10m of CD, the average connectivity of the nodes in the network is increased by more than 10 nodes, and with the improvement of the network connectivity, the positioning error decreases.
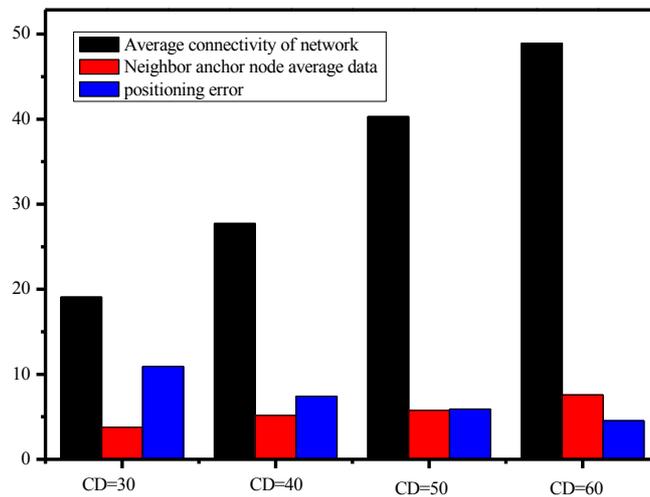
Fig. 3. The relationship between the node communication distance and the average connectivity of the network, the average data of neighbor anchor nodes and the positioning error

In order to better evaluate the performance of the centralized SUV location algorithm, in the following, from the deployment number (NN) of sensor nodes, the ranging error between nodes (σ), node communication distance (CD), we view the positioning precision of the algorithm, as the standard for the evaluation of the positioning performance of the algorithms proposed in this chapter. In the following three groups of experiments, the sensor nodes are randomly deployed in the area of 100 x 100m², and we select any six of the nodes as the anchor node of the known coordinates.

First of all, we evaluate the effect of NN on the performance of SUV localization algorithm. Figure 4 (a) is, in the case of range error σ=0.1, to continue to add sensor nodes to the network, and get the positioning error graph of SUV algorithm. It can be seen that when the NN is more than 40, the positioning error of the three kinds of CD decreases; when the NN increases from 70 to 100, the three curves become more gentle. Taking CD=50 as an example, when NN=70, the average positioning error is 3.85M; when NN=100, the average positioning error is 3.23M, reduced by 0.62m.

Figure 4 (b) gives, when σ changes, the relationship between average positioning error and NN. Similar to Figure 4 (a), the curve slope of NN between 30 to 40 is greater than the curve slope when NN is between 40 to 70, and the curve slope of NN between 40 to 70 is greater than the slope between 70 to 100. That is to say, when NN>70, under the same parameter conditions, the variation range of positioning error is no great, which provides a reference for the deployment of sensor networks. From Figure 4 (b), it can also be seen that, the influence of σ on SUV positioning is relatively large (the spacing between the three curves is obvious). In the case of NN=80, the positioning error average values when σ is 0.1, 0.2 and 0.3 were 3.26m, 6.11m, and 8.52m, respectively, and the positioning accuracy decreased significantly.
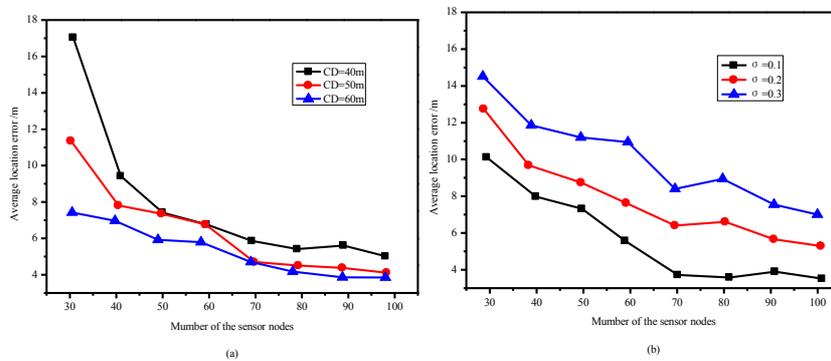
Fig. 4. (a) the influence of the number of sensor nodes (b)the impact of sensor nodes on the positioning error under different ranging distance

Secondly, we compare the effect of measurement error on the performance of a node localization algorithm. Set CD=50m, in the case of NN of 100, 70, and 40, continuously increase σ. The average localization error results of SUV localization algorithm in these parameters are shown in Figure 5 (a). It can be seen that with the increase of ranging error, the average positioning error becomes larger. When NN is

100 or 70, the change amplitude of the curve is slower than when NN is 40; in the three cases, when σ increased from 5% to 40%, the positioning error was increased by 7.50 m, 6.63 m, and 10.92 m; it also verified the relationship between the number of node and the positioning error in the last group experiment. Then, set the NN=100 in the region, and in the case of CD of 40 m, 50 m, and 60 m, calculate the positioning error by changing σ, and the results are shown in Figure 5 (b). It can be seen that the positioning error increases with the increase of σ; when σ increased from 0.05 to 0.40, the positioning error in three kinds of CD were increased by 8.10 m, 7.50 m, 5.03 m, respectively.
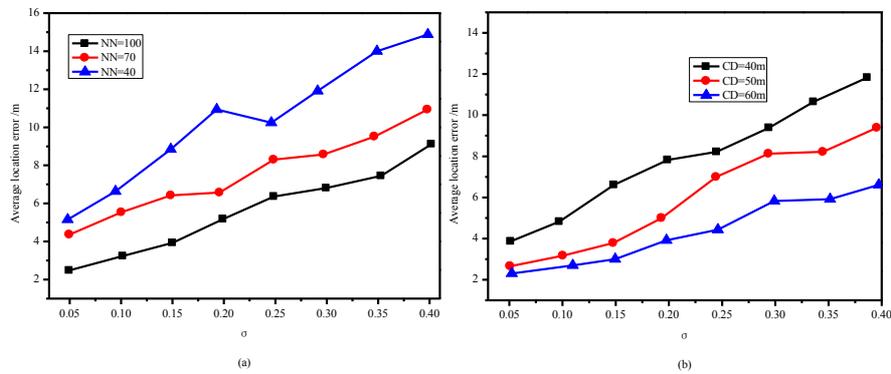
Fig. 5. (a)the total number of different nodes (b) the influence of sensor node ranging error on positioning error under different node communication distance

Finally, we evaluate the effect of CD on the performance of SUV localization algorithm. Set 100 sensor nodes randomly distributed in the region, and in the case of σ of 0.1, 0.2, and 0.3, continuously increase CD and observe the change of SUV positioning algorithm average positioning error, the results as shown in Figure 6 (a). It can be seen that the curve for σ as 0.3 is higher about 4 m than the other two curves; and when the CD is less than 40 m, the position error that the curve gives is more than 10 m; the CD is 30 m, the positioning error is up to 15.58 m. Then, set σ of 0.1, deploy 100, 70, and 40 sensor nodes, respectively. Similarly, by increasing the CD, we observe the change of positioning error, the results as shown in Figure 6 (b). It can be seen that the positioning error is 5.39 m for the curve when the number of nodes is 40, even at CD=60 m, and the overall positioning error is relatively large. For the curve with the number of nodes of 70, after CD>40m, the curve change is relatively stable; the positioning error of CD=60 m is less about 1.12 than that when CD=40 m. For the curve with the number of nodes of 100, after CD>50 m, the average positioning error is less than 3 m.
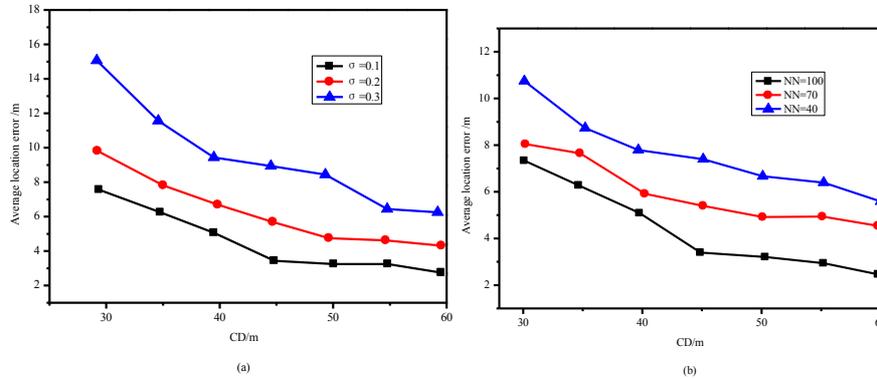
Fig. 6. (a) different ranging errors (b)the influence of the communication distance of the sensor nodes on the positioning error under different total number of nodes

5 Conclusion

The algorithm greatly simplifies the solution of center coordinate matrix and gets the coordinate inner product matrix, and uses statistical uncorrelated vectors to reconstruct the coordinates inner product matrix and the correlation of the coordinate inner product matrix caused by removing the ranging error. After the appropriate transformation, the SUV localization algorithm can realize the wireless sensor node centralized localization and the distributed localization two kinds of plans. We use TI company's CC2430 as the main chip, design and produce relevant wireless sensor nodes. Based on these hardware nodes, a lot of tests are carried out to analyze the influence of the RSSI signal transmission attenuation model and the actual environment on the RSSI signal. The proposed algorithm does not consider the network traffic. In the process of positioning, reducing the amount of communication of sensor nodes is beneficial to prolong the service life of the sensor nodes, thus increasing the lifetime of the whole network. Under the premise of ensuring the positioning accuracy, it is of great significance to reduce the network traffic positioning technology as much as possible.

6 References

[1] Andreyev, P., Grishko, A., & Yurkov, N. (2016). The temperature influence on the propagation characteristics of the signals in the printed conductors. International Conference on Modern Problems of Radio Engineering. Telecommunications and Computer Science (pp.376-378).

[2] Da-li, Z., Jing, S., Hao, J., & Ying, Y. (2016). Pseudo Base Location for Mobile Terminal with Abnormal Dynamic Access. International Journal of Future Generation Communication and Networking, 9(11): 253-262. https://doi.org/10.14257/ijfgcn.2016.9.11.23

[3] Guo, H., Zhang, H. (2017). Development of double-pair double difference earthquake location algorithm for improving earthquake locations. Geophysical Journal International, 208(1): 333-348. https://doi.org/10.1093/gji/ggw397

[4] Hui, J., Yang, Y., Hui, Y., & Luo, L. (2016). Research on Identify Matching of Object and Location Algorithm Based on Binocular Vision. Journal of Computational and Theoretical Nanoscience, 13(3): 2006-2013. https://doi.org/10.1166/jctn.2016.5147

[5] Jia, K., Li, M., Bi, T., & Yang, Q. (2016). A voltage resonance-based single-ended online fault location algorithm for DC distribution networks. Science China Technological Sciences, 59(5): 721-729. https://doi.org/10.1007/s11431-016-6033-2

[6] Wang, H., Tong, L., Yu, L., & Ben, H. (2015). The research of facial features localization based on posterior probability deformable model. IEEE International Conference on Mechatronics and Automation (pp.2392-2396). IEEE. https://doi.org/10.1109/icma.2015.7237861

[7] Xu, Q. (2014). Design and development of a compact flexure-based, precision positioning system with centimeter range. IEEE Transactions on Industrial Electronics, 61(2): 893-903. https://doi.org/10.1109/TIE.2013.2257139

[8] Yang, B., Lei, Y., & Yan, B. (2016). Distributed multi-human location algorithm using naive Bayes classifier for a binary pyroelectric infrared sensor tracking system. IEEE Sensors Journal, 16(1): 216-223. https://doi.org/10.1109/JSEN.2015.2477540

[9] Zhang, S., Gao, H., & Song, Y. (2016). A New Fault-Location Algorithm for Extra-High-Voltage Mixed Lines Based on Phase Characteristics of the Hyperbolic Tangent Function. IEEE Transactions on Power Delivery, 31(3): 1203-1212. https://doi.org/10.1109/TPWRD.2015.2461678

[10] Zhang, Y., Liang, J., & Wang, P. (2016, August). Mutual impedance parameter modeling and accurate location algorithm of angled space crossed transmission lines. In Electricity Distribution (CICED), 2016 China International Conference on (pp. 1-6). IEEE. https://doi.org/10.1109/ciced.2016.7576202

7   Author

Min Wang is with Hunan Mechanical & Electrical Polytechnic, Hunan, China (wangmin_507@126.com).