

# A Local Localization Algorithm based on WSN

Tian Zhang\*

*Changchun Normal University, Changchun, 130032, China*

---

## Abstract

After studying the topological structure of neighboring nodes in the WSN, we present a local localization (LLA) algorithm by combining the ideas of principal manifold learning and nonlinear dimension algorithm. This algorithm is particularly suitable for determining the relative locations of sensor nodes in the large-scale and low-density WSNs, where the low connectivity between nodes and the large ranging error between long-distance nodes usually make accurate localization quite difficult. In this algorithm, based on the pair-wise distance between each node and its neighbor nodes within a certain communication range, the local geometry of the global structure is firstly obtained by constructing a local subspace for each node, and those subspaces are then aligned to give the internal global coordinates of all nodes. Combined with the global structure and the anchor node information, we can finally calculate the absolute coordinates of all unknown nodes in the least squares algorithm.

*Keywords:* wireless sensor network; neighbour nodes; localization algorithm

(Submitted on November 4, 2017; Revised on December 1, 2017; Accepted on December 19, 2017)

© 2018 Totem Publisher, Inc. All rights reserved.

---

## 1. Introduction

Wireless sensor network (WSNs) is a new information acquisition and processing technology. It makes a wide application prospect in the fields of environmental monitoring, target tracking, intelligent traffic and intrusion detection.

In these applications, most of them get to know the location information of sensor nodes in the WSNs when they get the monitoring information. Obtaining sensor node location information is of great importance. Only the monitoring data of the location information can be used to specify the location of the event and to achieve the target location and tracking.

Position distribution of acquisition sensor nodes can improve the efficiency of network routing, so as to realize the automatic configuration load balancing and network topology of the network coverage, improve the quality of the whole network, can also provide named network space. So, one of the key technologies of WSNs is to design a scientific and effective localization algorithm.

The essence of WSNs node positioning is to explore the positional relationship between nodes by means of the relationship between distribution characteristics of all nodes in the network and metrical information [5,7,15], and then locate the coordinate of unknown nodes or determine the coordinate of nodes through the manifold of node distribution. In the field of machine learning, dimensionality reduction refers to mapping data points in high dimensional space to low space by a certain mapping technique; the essence of it is to realize the conversion from high dimensional space to the low one by finding a kind of mapping relation according to the distribution feature of all separate data points in a data set and the spatial relative information among those data points. Data points after conversion keep as many of their features as possible and remove noise in the high dimension space [6,13,14]. If nodes in WSNs are regarded as separately distributed data points and the metrical information relation among neighboring nodes as the spatial relative information of such data points, then the WSNs node positioning problem is transformed to the problem of data dimensionality in the machine learning field [1,12]. Technology for reducing dimension itself can eliminate redundant information and noisy information in original high dimensional data

---

\* Corresponding author.

E-mail address: zhangtianzt2017@sina.com

and can affect the intrinsic structural characteristics of data. Hence, WSNs positioning algorithm based on dimension reduction technology can tolerate measuring noise to a certain degree. Partial algorithms can even suppress measurement noise and make it especially adaptive to the node positioning problem in the case of big measuring error.

Many researchers have conducted related studies. Jenkins [8] applied the most typical dimension reducing technique in machine learning - Principal for Node Positioning in WSNs. However, since PCA dimension reducing method is a linear way of reducing dimensionality, when there is big measuring noise, positioning precision is rather lower. Since MDS is a linear dimension decreasing approach, when nonlinear input is not high, MDS-MAP positioning algorithm can achieve higher accuracy of positioning. But, when metrical information (distance or hop) is of high non-linearity, for instance, RSSI based measuring technology or nodes of lower density in a certain range, measuring error can sometimes reach up to 50%; MDS-MAP positioning algorithm has poor precision and sometimes it even can fail [4,9]. Patwari et al introduced the idea of manifold learning to WSNs node positioning and successfully realized the WSNs positioning based on Hessian manifold learning algorithm. Those algorithms assume the location of any node can be linearly expressed by the position of its neighboring node; or assume neighbor nodes exist at the same one manifold. Then, use it as the foundation of estimating the coordinates of nodes. Although methods based on the manifold learning idea can get higher positioning precision, they require setup of parameters, with limited practicability [10,11]. Besides, Juang et al solved the non-isotropic WSNs positioning issue in the approach of eigen decomposition [2,3].

With regards to the big error of positioning algorithm caused by then range measuring problem in a big area and lower density, we propose a local locating algorithm (LLA) based on MDS algorithm and the idea of manifold learning algorithm. Usually in the same manner of measuring distance e.g. RSSI ranging technology, the closer the distance between nodes, the higher the precision of measuring we will get. Based on that criterion, firstly, we use neighboring nodes in certain communication ranges to construct a local subspace; then, we use all local subspaces to form a global coordinate system; finally, with global structure and anchor node information, we use the least square method to map out the absolute coordinate of all nodes in the network.

## 2. Problem Description

Suppose WSNs have  $N$  nodes  $\{X_1, \dots, X_N\}$ , which are deployed in  $d$ -dimensional space of a big area. Sensor nodes have limited range of communication but can only realize single hop ranging with neighbors in a certain distance.  $N = \{1, 2, \dots, N\}$  is a collection of sensor nodes;  $x_i \in R^d$  means coordinate of the  $x_i$  node. Without loss of generality, define the number of anchor nodes as  $M$  ( $M \ll N$ ). The coordinates of them can all be acquired during system initialization; the number of unknown nodes is  $N-M$ ; the distance between nodes is obtained by ranging technology. Normally, in the same way of range measuring, the closer the distance nodes are, the higher the accuracy of measuring will achieve.

Suppose communicating radius of all nodes in the network is  $CD$ ; each node can connect with others through one or more hops and each node can send, receive and measure distance information of one hop between two nodes. When node  $X_i$  locates within communicating radius  $CD$  of node  $X_j$ , the Euclidean distance between two nodes is defined as below:

$$d(x_i, x_j) = \sqrt{\sum_{\lambda=1}^d (x_{i\lambda} - x_{j\lambda})^2} \quad (1)$$

Considering measuring error grows along with increasing distance and is affected by lots of factors such as blocks, communication model, and environment. Here, we simulate measuring distance between two nodes within  $CD$  in the manner of increasing noise. During the simulation, the following equation is applied to measure the distance between nodes.

$$d(x_i, x_j) = \begin{cases} (N(d(x_i, x_j), \sigma)) \\ 0 \end{cases} \quad (2)$$

The locating algorithm based on measuring technology faces a lot of difficulties. Since WSNs are of big size and have plentiful nodes, it's better to keep the hardware cost of node as low as possible; moreover, the simpler the hardware structure, the better they will adapt. Likewise, for the reason of cost and application environment, the number of anchor nodes that know

their positioning information in the network cannot be more. The cost anchor node is usually 10 times that of unknown nodes. So, more anchor nodes mean bigger network deployment cost and that the positioning algorithm must reduce its dependence on the number of anchor nodes. Measuring technique relies directly on the hardware of sensor nodes and due to complications of the working environment and diversity of sensor nodes, the ranging technology can lead to big errors in actual use, which results in the differences between measuring distance and the actual data between two nodes.

When there is a blockage between node A and node D, they can't communicate directly, i.e. it's not possible to complete direct measurement of distance, but indirect measurement through multiple hops or choosing the best path is. Assume completing distance measuring in the way of multiple hops; the distance between node A and D is equal to the sum of the distance between A and B, B and C and C and D. Obviously, the measuring distance is far bigger than the Euclidean distance between A and D. The existence of block changes the topological structure of network and the network connectivity. In essence, the existence of block makes convex region where there are no covering loopholes change to non-convex region where there are covering loopholes, bringing enormous error to the measuring technology. Therefore, for the existence of block, the positioning precision of some locating algorithms based on distance measurement is apparently affected. In the actual deployment area of WSNs, the influence of block is inevitable. So, the proposed locating algorithm will focus on the solution of measuring distance based on one hop between neighboring nodes and the use of coordinates of anchor nodes to estimate that of unknown nodes. It is shown in Figure 1.

### 3. Local Localization Algorithm

Give sample collection  $\{X_1, \dots, X_N\}$  of locating sensor nodes, where N is number of nodes;  $x_i \in R^d$  is the coordinate of the  $X_i (1, 2, \dots, N)$  node, and  $x_i$  is on the origin of coordinate i.e.  $x_i = 0$ . Without loss of generality, set number of anchor nodes as M ( $M \ll N$ ). In order to get neighbour nodes of one node, use network graph G to describe the topological structure of WSNs, where each vertex represents one node in the network. If node  $X_i$  and  $X_j$  are in certain communicating range CD, let one edge exist between  $X_i$  and  $X_j$  in the graph G. Suppose G's weight matrix is H; when one edge exists between vertex  $X_i$  and  $X_j$ ,  $H_{ij} = 1$ ; otherwise,  $H_{ij} = 0$ ; then weight matrix is a symmetrical sparse one.

$$H_{ij} = \begin{cases} 1, \tilde{d}_{i,j} \leq CD \\ 0, \tilde{d}_{i,j} > CD \end{cases} \quad (3)$$

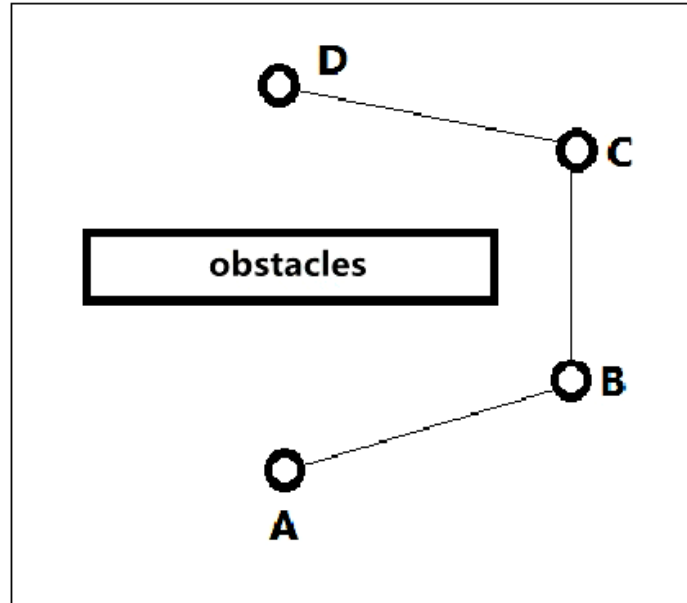


Figure 1. Influence of obstacles on the ranging technology

For the  $i$ th row in the weight matrix, if there are  $k$  elements with value 1, it means that there are  $k$  neighbouring nodes in communicating range  $R$  of node  $X_i$ . Redefine node  $X_i$ 's  $k$  neighbouring nodes as  $\{X_{i1}, \dots, X_{ik}\}$ , which includes node  $X_i$  itself; then, the distance matrix formed by neighbouring nodes of node  $X_i$  is  $D^{(i)} = [\tilde{d}_{pq}]$ , where  $\tilde{d}_{pq}$  is one hop measuring distance between  $X_{ip}$  and  $X_{iq}$ ; next, by following the equation, we can get the inner product matrix  $B^{(i)} = [b_{pq}]$  of coordinate:

$$b_{pq} = (d_{1p}^2 + d_{2p}^2 - d_{pp}^2) / 2 \quad (4)$$

Coordinate the inner product matrix here is a symmetric positive semi-definite matrix. In order to get the local coordinates, coordinate the inner product matrix eigenvalue decomposition:

$$B^{(i)} = V_i A_i V_i \quad (5)$$

Perform eigenvalue decomposition of  $R$ ; choose  $d$  groups of feature vectors to which the biggest  $d$  feature values correspond as the global relative coordinate in the  $d$ -dimensional space. After global relative coordinate is acquired, with the absolute coordinate of  $n$  anchor nodes, through the conversion of coordinate, cover all global relative coordinates to the global absolute coordinate. In the 2-dimensional node coordinate positioning problem, at least three anchor nodes are required that are not on the same straight line in order to realize the transformation of coordinate. For that, the formula is:

$$x_i = CT_i + b \quad (6)$$

So far, the whole LLA positioning algorithm is completed:

Setp1: Extracting local information. Calculate the weight matrix  $G$  of the network graph  $H$ ; and then for each node  $X_i$ .

Setp2: Reconstruction of Global combinatorial matrix. Calculation of global combinatorial matrix by iterative accumulation  $R$

Setp3: Calculate global relative coordinates. Eigenvalue decomposition of global combinatorial matrix, and select the largest  $d$  eigenvalues corresponding to the  $d$  set of feature vectors as the global relative coordinates in  $d$ -dimensional space;

Setp4: Coordinate transformation. Combined with enough anchor nodes, the least squares method is used to convert all global relative coordinates into global absolute coordinates.

In short, the LLA positioning process includes the following steps:

The weights are used to determine the number of neighbour nodes, and the local subspace is generated. Then, all local subspaces are combined into a global coordinate relative space including all nodes. Then, the coordinates of the unknown nodes can be located by the method of least square method, which is based on the coordinate of the anchor nodes.

#### 4. Experimental Analysis and Results

The performance of LLA localization algorithm is analysed and evaluated by simulation. Choose Matlab as the simulation platform, as all the algorithms are written in the Matlab language. In the simulation experiments, the sensor nodes are deployed in the region of  $100 * 100 \text{ m}^2$ . We use the performance of three factors to evaluate the algorithm: the communication distance of the sensor nodes (CD), the number of sensor nodes (NN) and the distance error between nodes (sigma). In Figure 2, sensor nodes are deployed in the form of regular deployment and random deployment.

In the process of simulation, the distance between nodes is measured by using the Noisy Disk model:

$$d(x_i, x_j) = \begin{cases} (N(d(x_i, x_j), \sigma)) \\ 0 \end{cases} \quad (7)$$

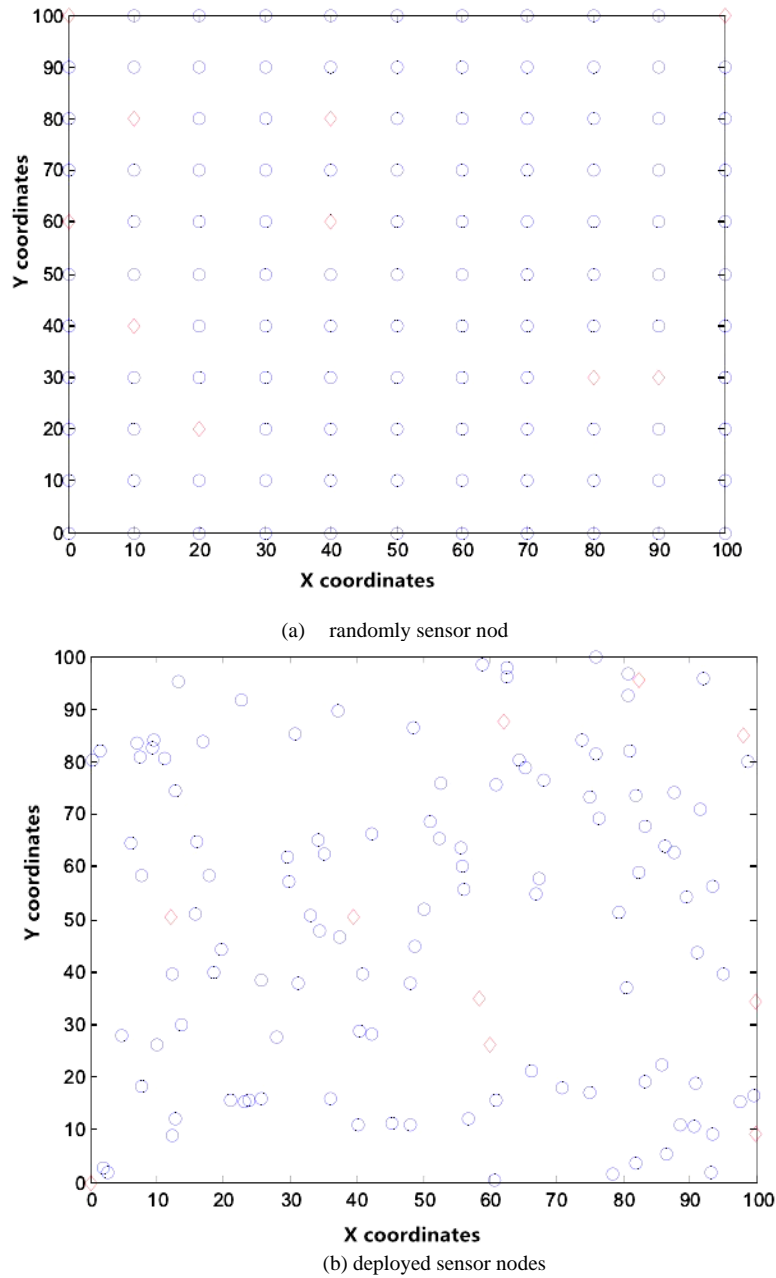
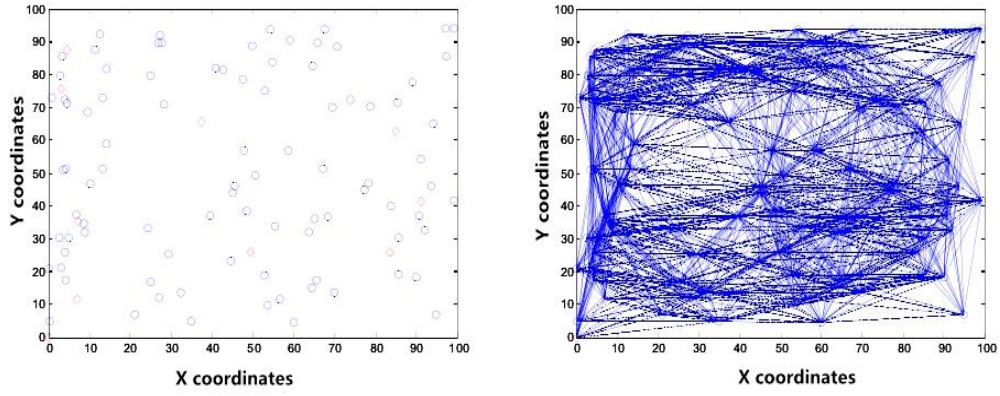
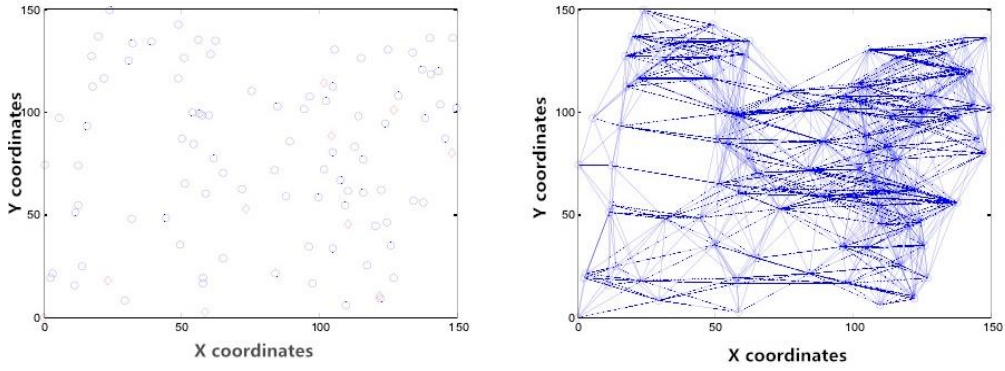


Figure 2. Regularly and randomly deployed sensor nodes

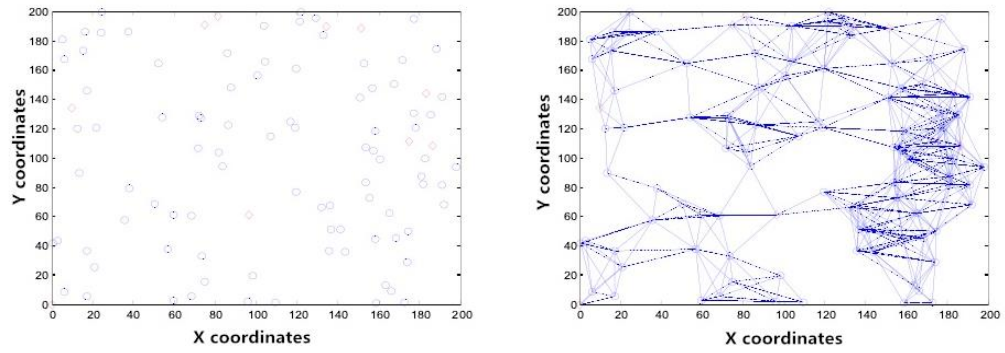
Usually, the communicating ability of sensor nodes is relatively stable, i.e. communicating distance of sensor nodes is fixed. However, the deployment area of WSNs will vary according to different actual applications. Yet, full coverage of networks can be achieved by increasing the number of deployed nodes, which no doubt raises the deployment cost of networks. Figure 3 presents node deployment and network connectivity, which are achieved by randomly deploying 100 sensor nodes (where 10 is anchored nodes) in the area of  $100 \times 100 \text{ m}^2$ ,  $150 \times 150 \text{ m}^2$ ,  $200 \times 200 \text{ m}^2$  and  $250 \times 250 \text{ m}^2$ . Since CD is fixed, the node's connectivity decreases. In the area of  $250 \times 250 \text{ m}^2$ , isolated nodes are seen on the upper right corner. In that case, the positioning problem of node will be directly influenced. Figure 4 gives the average connectivity of the network and average number of neighbouring anchor nodes in the network. In the area of  $100 \times 100 \text{ m}^2$ ,  $150 \times 150 \text{ m}^2$ ,  $200 \times 200 \text{ m}^2$  and  $250 \times 250$ , the average connectivity of network is 30.62, 18.46, 9.5 and 6.26; the average number of neighbouring anchor nodes is 2.9, 1.79, 0.87 and 0.58.



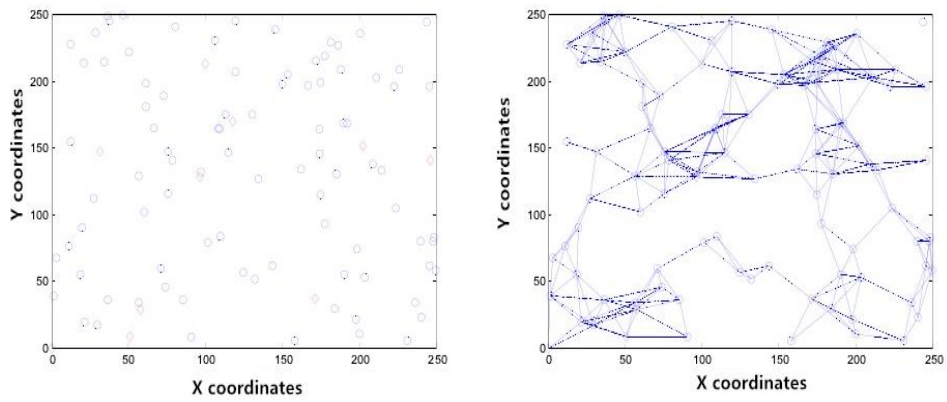
(a) Deployment of random deployment is  $100 \times 100 \text{ m}^2$



(b) Deployment of random deployment is  $150 \times 150 \text{ m}^2$



(c) Deployment of random deployment is  $200 \times 200 \text{ m}^2$



(d) Deployment of random deployment is  $250 \times 250 \text{ m}^2$

Figure 3. Randomly deployed sensor nodes and network connectivity graphs (NN=100, CD=40 m)

The communication distance of sensor nodes is not just associated with hardware cost, but also affects directly the energy consumption of nodes. Strong communication ability indicates that it consumes more. WSNs may be deployed in a vast place where it's sparsely populated. In many cases, sensor nodes are of one-time disposal. How to reduce energy consumption and extend the node's service life is realistic issues facing WSNs. For a sensor network node self-positioning technology, it's an important concern about how to realize high precision of positioning in the case of limited communicative capability of nodes. When the deployment area of the sensor network is definite with no variation in the number of deployed nodes, a change in the node's communicating distance will have a direct impact on the average connectivity of network.

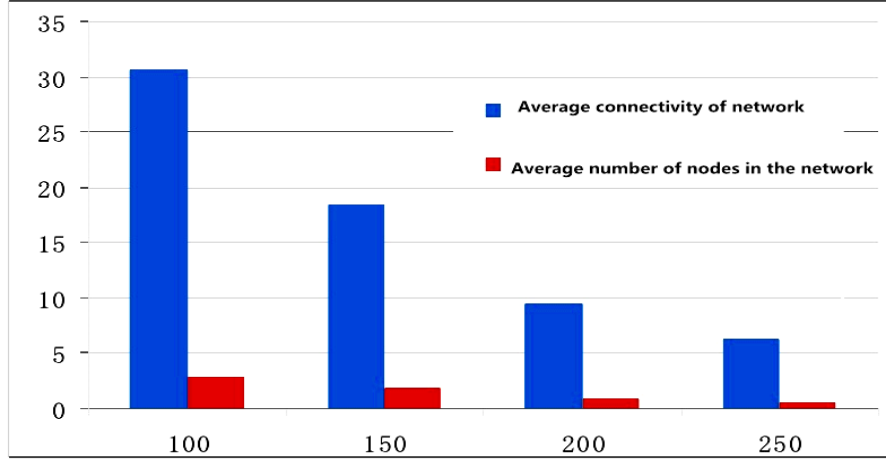


Figure 4. Average network connectivity and number of neighbour anchor nodes. (Deployment areas are  $100 \times 100 \text{ m}^2$ ,  $150 \times 150 \text{ m}^2$ ,  $200 \times 200 \text{ m}^2$ ,  $250 \times 250 \text{ m}^2$ , respectively,  $NN=100$ , anchor node number is 10,  $CD=40 \text{ m}$ )

Figure 5 shows a random deployment of 100 sensor nodes in the area of  $100 \times 100 \text{ m}^2$  and network node connectivity when the communicating distance  $CD$  is between  $10\text{m} \sim 35\text{m}$ . With a growing  $CD$ , the connectivity of nodes becomes stronger. Figure 6 is a column chart of the network's average connectivity and average number of anchor nodes in six situations, where the average connectivity of network is 3.1, 5.72, 9.68, 14.4, 19.42 and 24.68; the average quantity of the neighbouring anchor nodes is 0.32, 0.59, 1, 1.45, 1.85 and 2.29.

In the following three groups of simulation experiments, 100 sensor nodes are deployed randomly in the area of  $100 \times 100 \text{ m}^2$  and six anchor nodes are randomly chosen whose position information is known.

At first, we analyse quantitatively the performance of LLA from the aspect of a node deployment number. In certain deployment space, the number of sensor nodes directly matters network hardware cost, routing algorithm of nodes, network energy consumption and network coverage. The key is how to choose the appropriate number of nodes, complete required event monitoring or environmental parameter monitoring, and obtain the locational information of sensor nodes.

Set measuring error  $\sigma = 10\%$  between nodes. When the communicating distance  $CD$  between nodes is respectively 20m, 30m and 40m, let the number of sensor node grow from 20 to 100; the generated locating error result is shown in Figure 7(a). Set  $CD=40\text{m}$ . When  $\sigma$  is respectively 10%, 20%, and 30%, make the number of sensor node grow from 20 to 100; the generated locating error result is seen in Figure 7 (b). It's noted that either  $CD$  or  $\sigma$  is fixed, and increasing the number of sensor nodes will help improve the precision of locating. In Figure 5, when node quantity is bigger than 60, the locating results in three different  $CD$ s do not vary a lot, i.e. if  $\sigma$  is fixed and  $CD$  is also fixed, when the number of nodes reaches 60, it can meet certain locating error requirement. In that case, if two more nodes are deployed, it will result in unnecessary waste and hoist the communicating amount of network while shortening network service life. In Figure 6,  $\sigma$  has a big impact on LLA. Take for example, when the number of nodes is 60,  $\sigma=20\%$ , locating error increases about 200%, then  $\sigma=10\%$ ,  $\sigma=30\%$ , locating error increases to 450% then  $\sigma=10\%$ . Besides, only when  $\sigma=10\%$ , and node number is over 40, then the locating error tends to keep stable. LLA being easily affected by  $\sigma$  lies of the algorithm depending on the local subspace, which is formed by neighbouring nodes in one hop communicating range. So, if  $\sigma$  is too big in one hop range of communication, it will heavily affect the locating error.

Secondly, we analyse the effect of  $\sigma$  on LLA. When different measuring technology is applied,  $\sigma$  varies. The above experiment suggests that  $\sigma$  has a bigger influence on the positioning precision of LLA. Below, we quantitatively analyse the



magnitude of impact. Suppose communicating distance  $CD=40m$  between all sensor nodes and the number of anchor nodes is 6. When the total number of sensor nodes is  $NN=100$ , change the measuring error  $\sigma$  in one hop range from 5% to 40%. The resultant locating error is listed in Figure 8(a). Make  $NN=100$  and the number of anchor nodes 6. When  $CD=20m$ , 30m, and 40m, let  $\sigma$  grow from 5% to 40%, the generative locating error result is put in Figure 8(b).

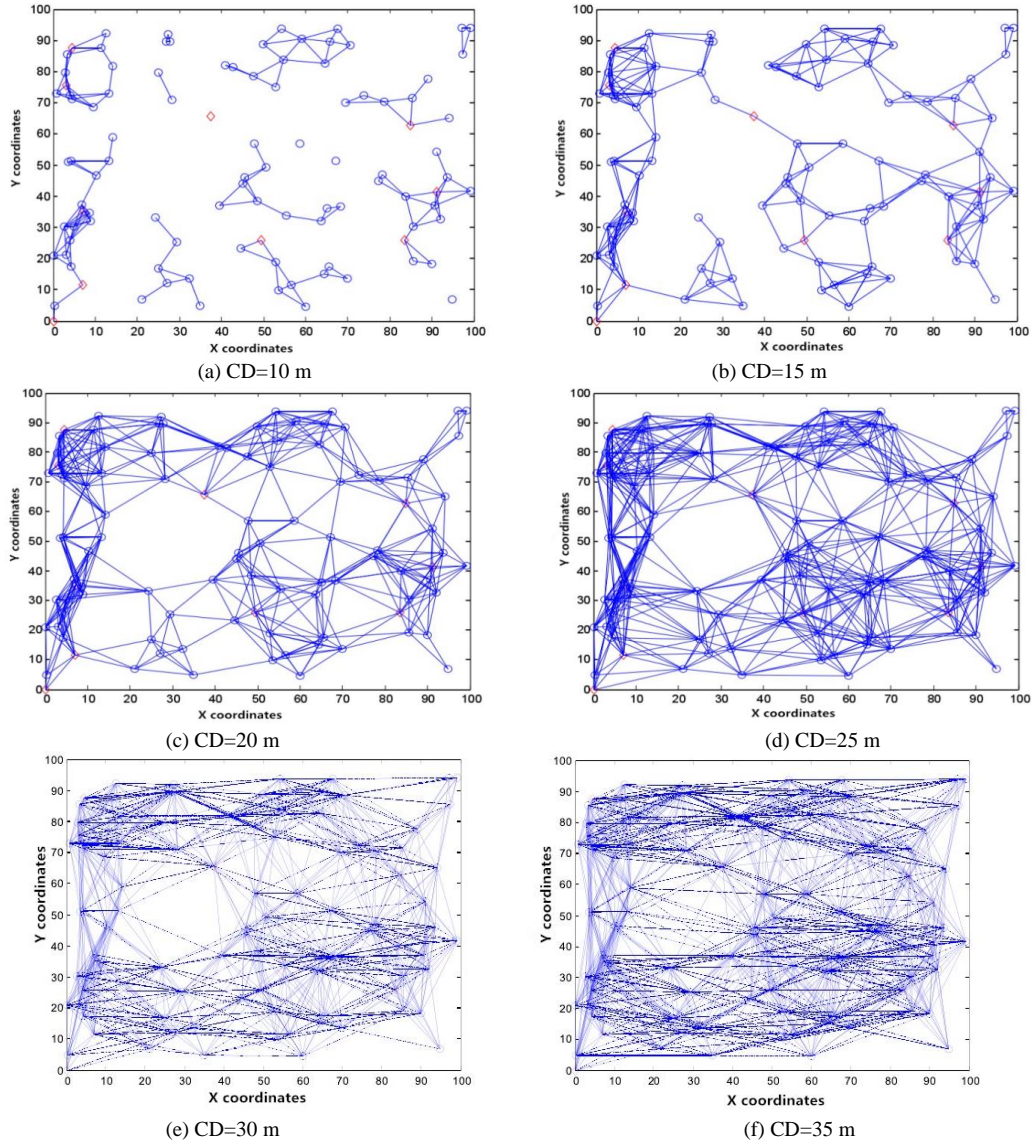


Figure 5. Network connectivity graphs when  $CD$  changes from 10 m to 30 m

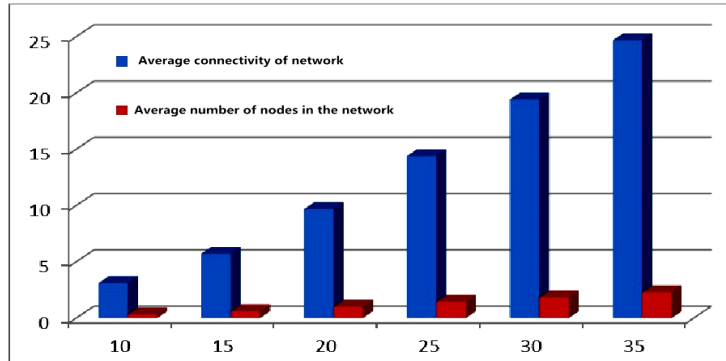


Figure 6. Average network connectivity and the number of neighbour anchor nodes (Deployment area is  $100\times100\text{ m}^2$ ,  $NN=100$ , anchor node number is 10)



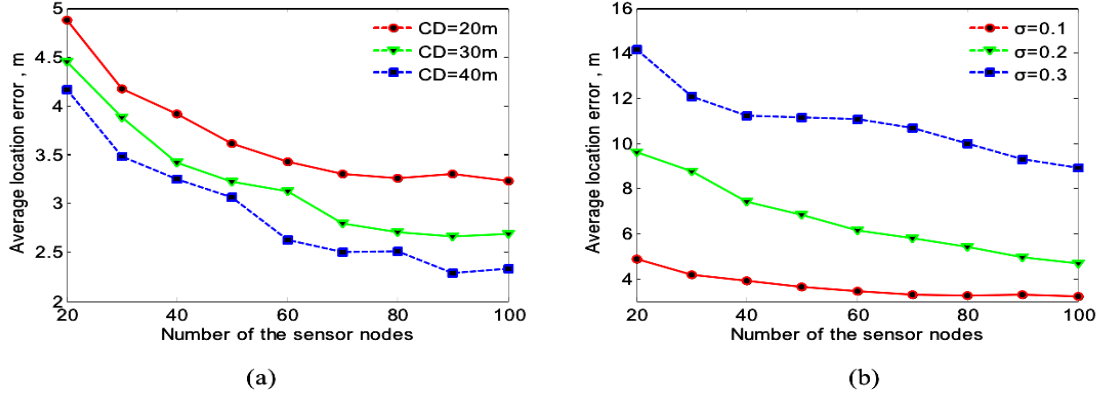


Figure 7. Average location error of LSA localization algorithm as a function of the sensor node number at  $\sigma=10\%$  (a) and CD=40 m (b)

At last, we discuss the relationship between CD and locating error by LLA method. From the above experiment, we judge that CD has a steady impact on LLA. Next, we do quantitative analysis of their relationship. CD not only matters to hardware cost of nodes, but also directly affects the consuming degree of sensor nodes' battery, affecting the service life. For instance, the locating algorithm based on RSSI shows that the propagation range of RSSI signal is affected by the node's sending power. How to set effective sending power that can satisfy node positioning requirement and reasonably prolong the service life of nodes is an issue taken into comprehensive consideration.

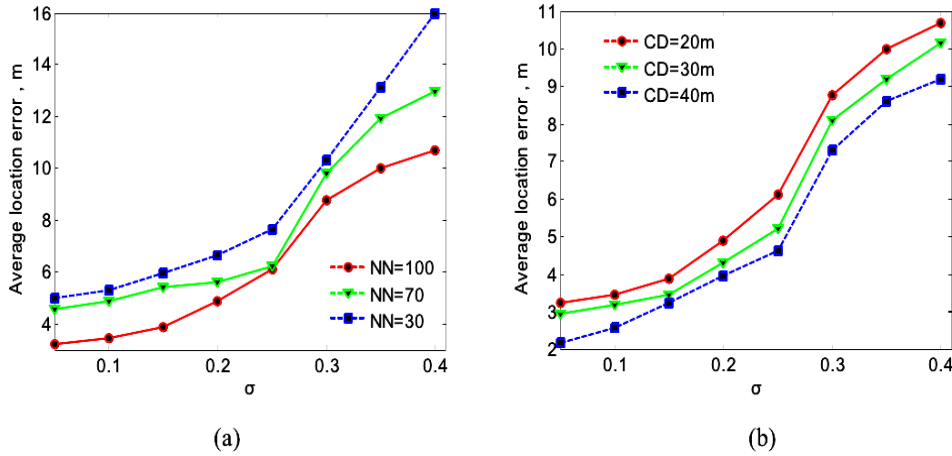


Figure 8. Average location error of LSA localization algorithm as a function of the sensor node ranging error at CD=40 m (a) and NN=100 (b)

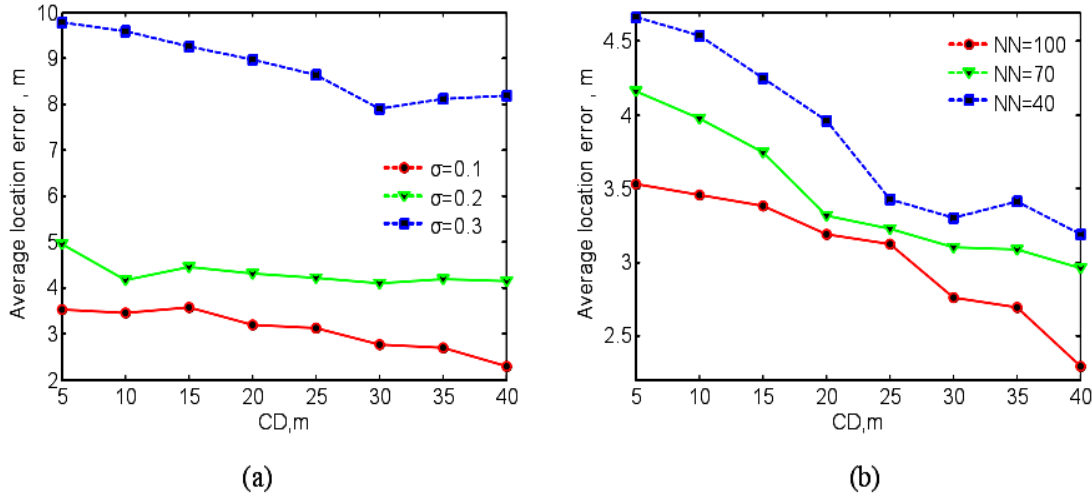


Figure 9. Average location error of LSA localization algorithm as a function of the sensor node communication distance at NN=100 (a) and  $\sigma=10\%$  (b)

Set  $NN=100$ , where the number of anchor nodes is 6, in the case of  $\sigma$  is 10%, 20%, 30% respectively, CD is from 5m to 40m to get the positioning error of the results. It is shown in Figure 9 (a).

Set  $\sigma=10\%$ , anchor node number is 6, in the case of  $NN=100$ ,  $NN=70$  and  $NN=40$ , CD is from 5m to 40m to get the positioning error of the results. It is shown in Figure 9 (b).

## 5. Conclusions

In conclusion, the paper verified the performance of the LLA method in a quantitative sense of the number of sensor nodes, communicating distance between nodes, and measuring error. It reveals that when measuring error is 10%, in the case of a small number of sensor nodes, LLA keeps a higher precision of locating. Measuring error has a big impact on the positioning precision of LLA method, especially when such aforesaid error is above 25%. The positioning accuracy of LLA decreases. Comparatively, communicating distance between sensor nodes maintains a stable influence on LLA. With the stronger communicative ability of nodes, the locating precision enhances steadily. LLA utilizes neighboring nodes in one hop range to form a subspace, then uses all sub spaces to merge into a whole space of the overall network for positioning. So, the measuring error in one hop range affects LLA a lot. Normally, the closer the distance between nodes, the higher the precision the measuring technology will achieve. So, ranging error within one hop is relatively small. Above all, the LLA method is more applicable for self-positioning of WSNs nodes of extensive range and lower density.

Through simulation experiments, the quantitative analysis of three important factors – location error between node number and the deployment of sensor nodes, communication distance of the sensor nodes on the performance of the LLA algorithm – were compared to MDS-MAP and ISOMAP algorithms. The results prove the validity and applicability of the LLA algorithm.

## Acknowledgements

This work was supported by provincial industry innovation special funds in Jilin Province (2016C078) and science and technology research project in Jilin Provincial department of education (2014269)

## References

1. Wanming Chen, "Localization Algorithm Based on Spring Model for Wireless Sensor Networks", *University of Science & Technology China*, 2009.
2. H Dai, G Chen, " Localisation Algorithm for Large-scale and Low-density Wireless Sensor Networks" ,*IET Electronics Letters*, vol.45, no.15, pp.881-883,2011
3. Juang, Ji, "Solving the Heterogeneous Positioning Problem Via Decomposition", *IET Electronics Letters*, vol.44, no.6, pp.432–433, 2016.
4. Tenenbaum J, Silva V and Langford J, " A Global Geometric Framework for Nonlinear Dimensionality Reduction", *Science*, vol.290, no.5500, pp.2319-2323,2000.
5. Fei Liu, "Localization Algorithm based on Control Strategy for Wireless Sensor Networks" , *Nanjing University of Posts and Telecommunications*, 2015
6. Fei Li, "Localization Algorithm Based on Range Finding" , *University of Electronic Science and Technology of China*, 2013
7. Zongxing Li, "Design and Research of Indoor Location Algorithm Based on RSSI and Geomagnetic Field Feature Fusion" ,*Nanjing University of Science and Technology*, 2015
8. Jenkins O. " Localization From Pairwise Distance Relationships Using Kernel PCR. Center for Robotics and Embedded Systems", *University of Southern California*, Tech. Rep. CRES-03-010, 2004.
9. Samko O, Marshall A, "Selection of the Optimal Parameter Value for the Isomap Algorithm" , *Pattern Recognition Letters*, vol.7 ,no.9 , pp.968-979,2006
10. Neal Patwari, Alfred O.Hero, "Adaptive Neighborhoods for Manifold Learning-based Sensor Localization" , *In: IEEE 6th Workshop on Signal Processing Advances in Wireless Communications*, vol.6, pp. 1098-1102,2015
11. Neal Patwari, and Alfred O.Hero, " Manifold Learning Algorithms for Localizations in Wireless Sensor Networks" , *In: Proceedings of Acoustics, Speech, and Signal Processing (ICASSP'04)*, vol.3 ,pp. 857-60 ,2014.
12. Huibo Qi, "Wireless Sensor Network Localization Algorithm Based on Mobile Anchor Node" , *Taiyuan University of Technology*, 2010
13. Renyi Su, " Localization Algorithm Based on ZigBee and Its Realization Research" , *Anhui University*, 2013
14. Weiguang Shi, "Research on Indoor Positioning Algorithm Based on Radio Frequency Identification Technology", *Tianjin University*, 2012.
15. Hua Zhu, "Research and Implementation of Algorithm Based on Wireless Signal Positioning", *Beijing University of Posts and Telecommunications*, 2014

**Tian Zhang** received his B.S degree from Northeast Normal University and received his M.S degree from Jilin University. He is a Lecturer in Changchun Normal University. His research interests include computer graphics and image, software engineering.