A New Multi-Hop Localization Based on ℓ_2 Constraint Least Square for Anisotropic Networks

Bo Hu[1], Sainan Zhang[1], Zhenyan Yuan[2]

[1] Center of information construction and management, Nanjing Normal University of Special Education, China
[2] Command & Control Engineering College, Army Engineering University, China
hb@njts.edu.cn, zsn@njts.edu.cn, 543060738@qq.com

Abstract

Multi-hop localization is a common method for wireless networks. However, when this network is anisotropic, the performance of multi-hop localization is greatly reduced due to deviation of the relationship between hop counts and physical distance. In order to improve the performance of multi-hop localization in anisotropic networks, this paper uses the ℓ_2 constraint least square to build a mapping relationship that represents the anisotropy of a network based on anchors. During transformations between the hops and the physical distances, the proposed algorithm can prevent over-fitting through the constraint of some space. The proposed algorithm has strong adaptability to the complex deployment environment; it overcomes the shortcoming of the traditional algorithm which applies only to the isotropic network. We also compare our method with several related methods, and the results show that our proposed method is more efficient than others in different topological networks. Furthermore, high accuracy can be obtained by this method without setting complex parameters.

Keywords: Multi-hop localization, Anisotropic networks, ℓ_2 constraint least square, Over-fitting

1 Introduction

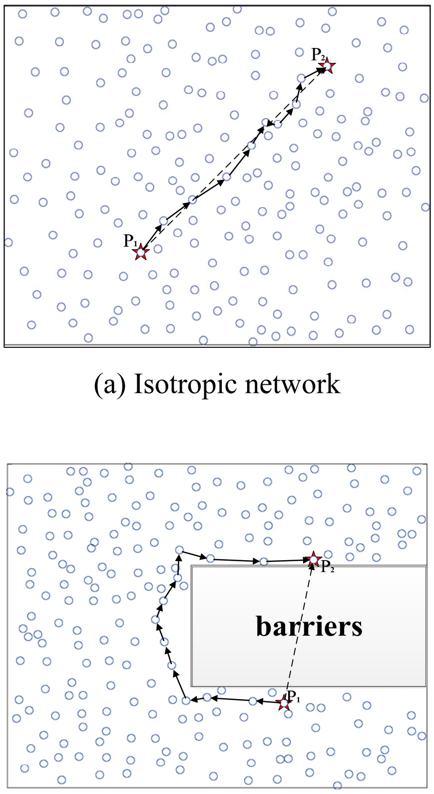
With the fast development of modem microelectronics technology, as a new approach to obtain information, wireless communication has also experienced fast development, and it has been broadly used in various fields such as military [1], disability rehabilitation [2], navigation for blind people [3] and environmental protection [4]. Different from previous communication methods, wireless communication is a data-centered communication technology, which can realize information transmission between "people and object" and "object and object" at any time and at any location [5]. Among various wireless communication applications, wireless node localization is one important supporting technology, and it can not only provide the location information of monitored event or tracked target, but also provide technical support to improve the routing efficiency, optimize network coverage and realize topology control.

Localization refers to determination of the targets' location in a specific coordinate system based on certain approach or method; while for wireless localization, the wireless communication technology is used to determine the coordinate location of the target [6]. According to whether it requires measuring the actual distance between nodes during the localization process, the localization algorithm includes range-based localization algorithm and range-free localization algorithm [7]. According to the results of previous research, we can see that the range-based wireless localization technique generally require complicated hardware equipment, so their application in a large-scale network is severely restricted [8]. Due to consideration of cost and network scale, the range-free wireless localization method has the advantages on the aspects of cost and technology, especially the range-free localization method based on hops. The range-free localization method based on hops is also called the "multi-hop range-free localization method", which requires simple network protocol with numerous system expandability. However, this kind of method generally assumes horizon communication between wireless nodes and even distribution of nodes, but it has ignored that in actual network environment, there might be various problems such as irregular deployment of nodes, uneven distribution of nodes and possible obstacles. These problems will cause hops between nodes to change according to the change of communication direction, and that the network might present anisotropy [9].

Take Figure 1 for example, we use $d_{S_1S_2}$ (dotted line) to represent the physical distance between nodes S_1 and S_2 in a certain network, and use $h_{S_1S_2}$ (arrowhead straight line) to represent approximate length of the shortest path between nodes S_1 and S_2. When the

*Corresponding Author: Bo Hu; E-mail: hb@njts.edu.cn
DOI: 10.3966/160792642019102006021

nodes are distributed in the network evenly and densely, $d_{S_1S_2} \propto h_{S_1S_2}$, and this kind of network is also called the isotropic network; however, when this area has an irregular deployment of nodes, uneven distribution of nodes or obstacles that might affect signal transmission, the network would have an anisotropic problem, which will result in $d_{S_1S_2} \ll h_{S_1S_2}$. Therefore, we could intuitively understand that when the nodes have even distribution in the network, the physical distance between nodes presents a proportional relation with hops; when there is anisotropic problem in the networks, the proportional relation between physical distance and hops would not be tenable anymore.



(a) Isotropic network

(b) Anisotropic network

Figure 1. Nodes distribution in wireless network

According to the above analysis, the anisotropic network is much more complicated than the isotropic network. Therefore, the localization of nodes in anisotropic network is significantly more difficult than that in an ideal environment. As shown in Figure 2, if there is any error in the estimated length of the shortest path between unknown nodes and anchor nodes, it will generate significant influence on the location estimation result. According to the problem of multi-hop localization in anisotropic network, we propose a multi-hop localization algorithm that not only applies to the ideal environment, but also applies to the anisotropic network, and this method has relatively low computation complexity and high localization accuracy, *i.e.*, ML- ℓ_2 CLS (Multi-hop Localization though ℓ_2 Constrained Least Squares). The ML- ℓ_2 CLS method tries to improve the performance of localization algorithm through partial spatial constraint of network, so that it can adapt to different network environment.

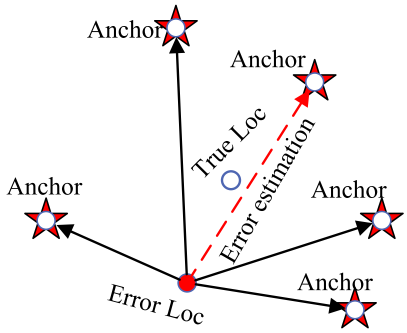


Figure 2. Schematic diagram of location error caused by distance error

2 Related Work

DV-Hop is the most famous multi-hop range-free localization algorithm, which was proposed by Niculescu et al. [10] based on the DV algorithm and the localization principle of GPS, and this is a distributive multi-hop range-free localization algorithm that conducts computation hop by hop. The localization process of DV-Hop algorithm mainly consists of three stages,

In Stage I, each node computes the minimum hops from other connecting nodes in the network;

In Stage II, each anchor node estimates the average hop distance based on its physical distance and hops from other reference nodes and broadcasts it to the network, and the estimated average hop distance can be expressed with Formula 1.

$$c_i = \frac{\sum_{i \neq j} d_{ij}}{\sum_{i \neq j} h_{ij}} \quad (1)$$

In which, d_{ij} refers to the physical distance between anchor node i and anchor node j, and h_{ij} is corresponding minimum hops. The unknown node only records the first average hop distance it receives, and then combines the minimum hops between various anchor nodes to compute the distance between them.

In Stage III, according to the recorded estimated distance from each anchor node, the unknown node will estimate its own location via the multilateration method.

In the isotropic network, the DV-Hop algorithm can be used to convert hops between nodes into physical distance in a linear way, but in the anisotropic network, this method to convert hops to physical distance won't work. However, if hops are still used to replace actual physical distance, it will cause sharp decline in the performance and estimation accuracy of algorithm.

When a network is anisotropic, hops between nodes may not match physical distance well. To address this issue, Lim and Hou [11] from UIUC proposed a modified approach PDM (Proximity Distance Map) that estimates the distance between nodes. First of all, the PDM localization algorithm identifies the physical distance and hops between anchor nodes with matrixes respectively; then, it uses TSVD (Truncated Singular Value Decomposition) to obtain the mapping model between two matrixes; finally, with the help of least square method, the mapping model is used to estimate the physical distance between the unknown node and anchor nodes. Compared to previous multi-hop localization algorithms, PDM can more actually establish the relationship between hops and physical distance, and effectively explore related implicit information behind data, such as the network topology structure and correlation, and as a result, it can effectively improve the localization performance. However, the PDM method has ignored the conversion of order of magnitude between hops and physical distance, which tends to result in fluctuation of localization performance of the algorithm under different density and different communication radius of nodes.

Inspired by the PDM method, Lee et al. [12] proposed two range-free multi-hop localization algorithms were based on support vector regression, *i.e.*, LSVR (Localization through Support Vector Regression) and LMSVR (Localization through Multi-dimensional Support Vector Regression). By introducing the kernel function, these two algorithms convert localization problem to kernel regression. However, depending on the well-known Occam's razor principle [13], a high number of model parameters tend to generate the over-fitting problem. During actual applications, the LSVR and MSVR algorithms are seriously affected by the over-fitting problem, and its localization performance is even poorer than that of the early PDM method. Then, in 2015 and 2016, Yan *et al.* used the KRR (kernel ridge regression) [14] and KPLS (kernel partial least squares) [15] methods to replace the multi-parameter SVR method, and the localization accuracy was much improved. However, the KRR and KPLS methods are also based on nonlinear method, and the algorithm has high requirement for hardware.

One of the simple and efficient schemes is weighted least square multi-hop localization [16]. The idea is to reduce localization errors by giving the optimal weights to each estimated distances. Each node estimates distance to an anchor by its hop count to the anchor. One drawback of this approach is that the algorithm requires distribution variance of nodes is equal. However, this is hard to do in a real environment. Recently, selective 3-Anchor DV-hop approach is presented in [17]. In this approach, every unknown node estimates its location by picking only three best anchor nodes. Zhao et al. [18] proposed the Locally Weighted Linear Regression DV-hop (LWLR-DV-hop) method on the basis of selective 3-Anchor DV-hop. LWLR-DV-hop approach employs local weighted least square to further reduce the deviation in hop count and physical distance conversion. However, some common nodes in the network may have little anchor nodes in close range, so they have to choose a longer distance anchor nodes. In addition, more distant anchor nodes contain more accumulated errors, which reduces the localization accuracy of the common nodes.

In this paper, by referring to the PDM, LSVR and MSVR methods, we propose an efficient multi-hop localization method that can not only adapt to complicated operation environment, but also control the complexity of the model. The algorithm limits selection of anchor nodes within a certain scope with the ℓ_2CLS method to reduce the complexity of model, which can prevent the over-fitting problem. In the meantime, the algorithm also conducts standard processing of hops and physical distance, which can prevent the influence of localization scale on the algorithm performance.

3 Localization Problem Formulation

The nodes localization problem of anisotropic network can be specifically described as:

Assume there are $m+n$ nodes $S=\{S_1,S_2,\cdots,S_m,\cdots,S_{m+n}\}$ in a $d(d=2\,or\,d=3)$-dimension network, in which, the first m nodes $\boldsymbol{V}_m\triangleq\{S_i\,|\,i=1,2,\cdots,m\}$ are anchor nodes, and the rest n nodes are unknown nodes $\boldsymbol{V}_n\triangleq\{S_j\,|\,j=m+1,m+2,\cdots,m+n\}$. For a 2-dimension network, the coordinate of node S_α can be expressed as:

$$Cor(S_\alpha)=(x_\alpha,y_\alpha)^{\mathrm{T}}\quad for\,\alpha=1,2,\cdots,m,\cdots,m+n \quad \textbf{(2)}$$

The $m+n$ nodes in the d –dimension network space can be abstracted into an undirected graph $G=(V_m\cup V_n,E)$, in which, E represents the edge set between all node pairs that can communicate with each other. If and only if nodes $\alpha,\beta\in V_m\cup V_n$ in undirected graph G are neighbors will there exist an edge $E_{\alpha,\beta}$ between nodes α,β. For the anchor nodes, because their coordinates are known in advance, the physical distance between any two anchor nodes can be obtained, and their physical distance can be expressed as:

$$d(S_i,S_k)=\|Cor(S_i)-Cor(S_k)\| \\ =\sqrt{(x_i-x_k)^2+(x_i-x_k)^2}\in\mathbb{R}^2 \\ S_i,S_k\in \boldsymbol{V}_m \quad (3)$$

Assume the distance between anchor nodes and unknown nodes is expressed with $\delta(S_i,S_j)$. Therefore, we can sum up the range-free multi-hop localization problem in anisotropic network as: the process to recover the coordinates of all unknown nodes in V_n under the constraints of the set of all anchor nodes V_m and $\delta(S_k,S_l)$. However, during multi-hop localization, the physical distance between unknown node and anchor nodes is unknown, and only the hops between them can be obtained, *i.e.*, $h(S_i,S_j)$ is used to express the hops between unknown nodes and anchor nodes. Assume there is a mapping function $f:h(S_i,S_j)\rightarrow\delta(S_i,S_j)$, and the multi-hop localization problem can be defined as: the process to recover the coordinates of all unknown nodes in V_n under the constraints of the set of all anchor nodes V_m and mapping functions f.

In the isotropic network, the distance between nodes can be approximately expressed by hops. However, in the anisotropic network, the node's sensing ability is not only related to distance, but also to the communication direction between nodes. In recent years, a popular research subject of multi-hop localization is to build the hops-physical distance conversion model based on machine learning method. Inspired by that method, through constraint of partial space in anisotropic network, we can obtain relatively accurate hops-physical distance conversion models to improve the localization performance. In this paper, we propose the ML-ℓ_2CLS method, and its entire process is showed in Figure 3.

Step 1(Data collection): After the program starts operation, any anchor node $S_i\in\boldsymbol{V}_m$ in the network exchanges hops information with the remained nodes $S_\alpha\in\boldsymbol{V}_m\cup\boldsymbol{V}_n$, and in the meantime, it will send its location information $Cor(S_i)$ to other nodes in the network.

Step 2 (Building model): Any anchor node $S_i\in\boldsymbol{V}_m$ uses ℓ_2CLS to build its mapping model $\ell_2\text{CLS}_i$, and broadcasts it to other nodes in the network. Finally, m mapping models $(\ell_2\text{CLS}_1,\ell_2\text{CLS}_2,\cdots,\ell_2\text{CLS}_m)$ are sent to all nodes in the network.

Step 3 (Location estimation): The unknown node $S_k\in\boldsymbol{V}_n$ in the network uses the mapping model obtained in Step 2 to estimate its physical distances from all anchor nodes; finally, the coordinate of unknown nodes are estimated by the multilateration method in a distributed manner.

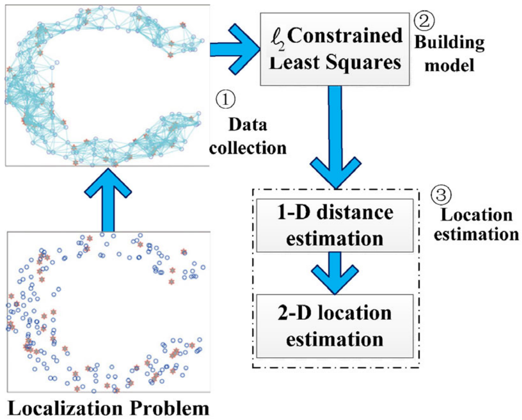


Figure 3. The framework of ML-ℓ_2CLS algorithm. Firstly, the algorithm collects by distance vector protocol. After that, in the building model phase, the mapping model is constructed by ℓ_2CLS, using data consist of the known hop-counts and physical distances. Finally, the physical distances of the unknown node are predicted by the learned mapping model and the location of unknown nodes are estimated by multilateration method.

During the operation process of algorithm, each node adopts the flooding broadcast information to compute hops in the network, so the communication overhead of ML-ℓ_2CLS is $O(n^2)$, in which, n refers to the number of nodes in the network.

3.1 Data Collection

At the data collection stage, the actual communication process is as follows: each anchor node $S_i\in\boldsymbol{V}_m$ broadcasts a data packet to all other nodes, and in this way, each node would know the shortest hops from other nodes in the network.

Assume $\boldsymbol{h}_i=[h_{i,1},\cdots,h_{i,m}]^{\mathrm{T}}$ is the hops vector between anchor nodes $S_i\in\boldsymbol{V}_m$ and other anchor nodes in the network, in which, $h_{i,k}=h(S_i,S_k),S_i,S_k\in\boldsymbol{V}_m$, and $h_{i,i}=0$. The hops between all anchor nodes can be expressed with matrix $\boldsymbol{H}$,

$$\boldsymbol{H}=[\boldsymbol{h}_1,\cdots,\boldsymbol{h}_m] \quad (4)$$

Similarly, assume $\boldsymbol{d}_i=[d_{i,1},\cdots,d_{i,m}]^{\mathrm{T}}$ is the physical distance vector between anchor node $S_i\in\boldsymbol{V}_m$ and other anchor nodes, in which, $d_{i,k}=d(S_i,S_k),S_i,S_k\in\boldsymbol{V}_m$, and $d_{i,i}=0$. The distance information between all anchor nodes can be expressed as,

$$\boldsymbol{D}=[\boldsymbol{d}_i,\cdots,\boldsymbol{d}_m] \quad (5)$$

Assume $\boldsymbol{h}_j$ is the hops vector between j th unknown node $S_j\in\boldsymbol{V}_n$ and all other anchor nodes $S_k\in\boldsymbol{V}_m$, and $\boldsymbol{h}_l$ can be expressed as,

$$\boldsymbol{h}_l=\left[h_{l,1},\cdots,h_{l,m}\right]^{\mathrm{T}} \tag{6}$$

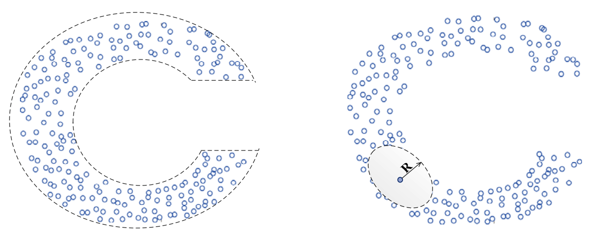
In which, $h_{k,l}=h\left(S_k,S_l\right), S_k\in \boldsymbol{V}_m \ and\ S_l\in \boldsymbol{V}_n$.

3.2 Building Model

During the building model stage, any anchor node $S_i\in \boldsymbol{V}_m$ builds a hops-physical distance mapping model. According to the literature [19], there is a mapping relation $d\propto \boldsymbol{w}^{\mathrm{T}}\boldsymbol{h}$ between the physical distance and hops. After anchor node S_i collects m pairs of hops and physical distance $\left(\boldsymbol{h}_k,d_{i,k}\right)$, it will build the mapping model f_i ,

$$f_i\left(\boldsymbol{h}_j\right)=\theta_i^{\mathrm{T}}\boldsymbol{h}_j \tag{7}$$

In order to prevent inconsistent order of magnitudes during hops-physical distance conversion with the change of factors such as network scale and node communication radius, we have conducted standard processing of data pair $\left(\boldsymbol{h}_k,d_{i,k}\right)$. According to Formula 7, we can see that parameter $\left\{\theta_i\right\}_{i=1}^{m}$ can be set freely, so any one anchor node uses all anchor nodes to build the mapping model, as shown in Figure 4 (a).



(a) Hops space of ordinary multi-hop localization
(b) Hops space of Subspace-Constrained

Figure 4. Hops space

In order to avoid the over-fitting problem during convert hops to physical distance, we constrain the parameters within a certain scale, *i.e.*,

$$\min_{\theta} J\left(\theta\right) \quad subject\ to\ \left\|\theta\right\|\leq R^2 \tag{8}$$

In which, $\min_{\theta} J\left(\theta\right)=\frac{1}{2}\sum_{i=1}^{m}\left(f_i\left(\boldsymbol{h}_j\right)-d_{i,j}\right)^2$; R refers to the radius, as shown in Figure 4(b). For Formula 8, we can adopt the ℓ_2 constrained least squares, and the mapping model can be solved within certain radius scope.

Through the Lagrange duality method, Formula 8 can be transferred to Formula 9, *i.e.*

$$\max_{\lambda}\min_{\theta}\left[J\left(\theta\right)+\frac{\lambda}{2}\left(\left\|\theta\right\|^2-R\right)\right] \quad subject\ to\ \lambda\geq 0 \tag{9}$$

In which, λ is the Lagrange multiplier, and we use $\lambda/2$ here to divide out 2 generated by computation of partial differential related to θ . If the Lagrange multiplier λ of Lagrange duality method is decided by the radius R of circle, which is directly specified, the solution $\hat{\theta}$ to the least square learning method constrained by ℓ_2 can be obtained through the following formula,

$$\hat{\theta}=\arg\min_{\theta}\left[J\left(\theta\right)+\frac{\lambda}{2}\left\|\theta\right\|^2\right] \tag{10}$$

In the literature [20], they proved that the algorithm would have optimum performance when there are 5 to 15 hops between nodes. Therefore, in this paper, we assume the anchor node only use radius $\boldsymbol{R}$ to build the mapping model for neighbor anchor nodes within four hops.

In Formula 10, the first item $J\left(\theta\right)$ represents the fitting degree of training samples, and the minimum value can be obtained by combining the second item $\frac{\lambda}{2}\left\|\theta\right\|^2$ to prevent over-fitting of training samples. For the objective function of Formula 10, set the partial differential of parameter θ as 0, and the solution $\hat{\theta}$ to the least square learning method constrained by ℓ_2 can be obtained through Formula 11,

$$\hat{\theta}=\left(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}+\lambda\boldsymbol{I}\right)^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{d} \tag{11}$$

In which, $\boldsymbol{I}$ is the identity matrix. By adding the hop matrix $\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}$ with $\lambda\boldsymbol{I}$, it can improve the algorithm performance to prevent over-fitting, which can further solve the inverse matrix more stably. If consideration is given to design the SVD (singular value decomposition) of hops matrix $\boldsymbol{H}$, *i.e.*,

$$\boldsymbol{H}=\sum_{i=1}^{m}\kappa_i\psi_i\varphi_i^{\mathrm{T}} \tag{12}$$

In which, κ_i , ψ_i and φ_i are called the singular value, left singular vector and right singular vector respectively. All singular values are negative, and the singular vectors satisfy orthogonality, *i.e.*,

$$\psi_i^{\mathrm{T}}\psi_k=\begin{cases}1\left(i=k\right)\\0\left(i\neq k\right)\end{cases}$$

$$\varphi_i^{\mathrm{T}}\varphi_k=\begin{cases}1\left(i=k\right)\\0\left(i\neq k\right)\end{cases}$$

At this point, the solution $\hat{\theta}$ to the ℓ_2 constraint least squares, or the hops-physical distance mapping model can be expressed as,

$$\hat{\theta} = \sum_{k=1}^{m} \frac{\kappa_k}{\kappa_k^2 + \lambda} \psi_k^{\mathrm{T}} \boldsymbol{d} \varphi_k \tag{13}$$

Through constraint of ℓ_2 CLS, the over-fitting problem during building model process can be alleviated in a certain degree. However, according to Formula 11, we find that the accuracy of mapping model depends on the Lagrange multiplier λ, and different accuracies of mapping model can be obtained by choosing different λ values. Common methods used to determine λ include GCV (Generalized Cross-Validation), L-corner, etc. We utilized the GCV method in this paper. The GCV method is an intuitive method based on the posterior estimation information, which is not dependent on the prior information and assumption. The λ value is determined based on the criterion of minimum mean value, and its computation criterion is,

$$GCV(\lambda) = \frac{\frac{1}{m}\left\| \left(\boldsymbol{I} - \boldsymbol{H}\left(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H} + \lambda \boldsymbol{I} \right)^{-1} \boldsymbol{H}^{\mathrm{T}} \right) \boldsymbol{d} \right\|^2}{\left[\frac{1}{m} trace\left(\boldsymbol{I} - \boldsymbol{H}\left(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H} + \lambda \boldsymbol{I} \right)^{-1} \boldsymbol{H}^{\mathrm{T}} \right) \right]^2} \tag{14}$$

By obtaining the minimum value of Formula 14, the optimal λ value can be determined.

3.3 Location Estimation

Each unknown node starts location estimation after receiving m hops-physical distance mapping models and the hops vector from unknown nodes to anchor nodes $\left\{ \left(\hat{\theta}_1, \boldsymbol{h}_1 \right), \left(\hat{\theta}_2, \boldsymbol{h}_2 \right), \cdots, \left(\hat{\theta}_m, \boldsymbol{h}_m \right) \right\}$. Let

$$\hat{\boldsymbol{d}}_j = \left[\hat{d}_{j,1} \cdots \hat{d}_{j,m} \right]^{\mathrm{T}} \tag{15}$$

is the estimated distance between unknown nodes $S_j \in \boldsymbol{V}_n$ and m anchor nodes $\{S_i\}_{i=1}^{m} \subset \boldsymbol{V}_m$, in which, $\hat{d}_{i,j}$ is the estimated distance between $S_j \in \boldsymbol{V}_n$ and $S_i \in \boldsymbol{V}_m$. The estimated distance vector $\hat{\boldsymbol{d}}_j$ can be obtained through the following method,

$$\begin{aligned} \hat{\boldsymbol{d}}_j &= \left[\hat{d}_{j,1} \cdots \hat{d}_{j,m} \right]^{\mathrm{T}} = \left[f_1\left(\boldsymbol{h}_j \right) \cdots f_m\left(\boldsymbol{h}_j \right) \right]^{\mathrm{T}} \\ &= \left[\hat{\theta}_1 \boldsymbol{h}_j \cdots \hat{\theta}_m \boldsymbol{h}_j \right]^{\mathrm{T}} \end{aligned} \tag{16}$$

In the network, after the unknown node completes estimation of its distances to all anchor nodes, it would use the multilateration estimation method for position estimation.

4 Algorithm Simulation and Performance Analysis

In order to verify the performance of ML-ℓ_2CLS, we designed a series of simulation experiments on the MATLAB platform, and compared its performance with the performance of similar localization algorithms, such as DV-hop, PDM and LSVR. In the experiment, we mainly investigated the influence of various indices on the localization performance, such as the anisotropic factors, proportion of anchor nodes and node density. Therefore, in the experiment, we would not focus on factors such as the node communication cost, node life cycle and message passing method. The experiment scenario used the alphabetic network topology (C-shaped, S-shaped, Z-shaped) commonly adopted by similar algorithms, and for specific network parameters, see Table 1.

Table 1. Network parameters

Nodes number	400
Anchor nodes ratio	Between 30,40,50,60,70,80, at 10 intervals
Network topology	Within the range of 500*500, topologies are C-shaped, S-shaped and Z-shaped
Radio radius	The communication radius of nodes is 50, 60, 70 and 80, respectively.

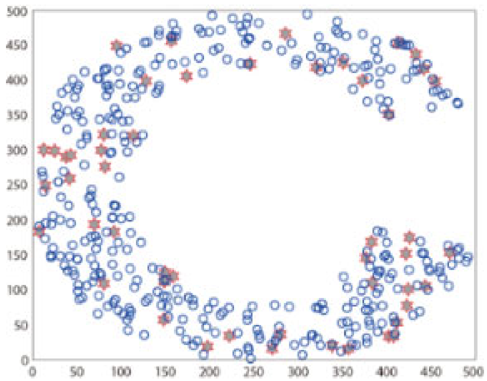
To be fair, we use root mean squares (RMS) as the criterion for localization accuracy, and RMS can be expressed as,

$$RMS = \sqrt{1/N_t \sum_{i=1}^{N_t} \left(\left(\hat{x}_i - x_i \right)^2 + \left(\hat{y}_i - y_i \right)^2 \right)} \tag{17}$$

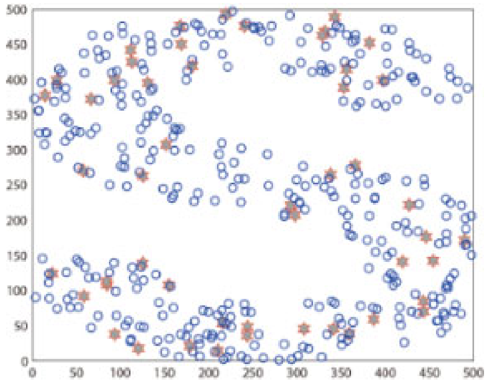
In which, $(\hat{x}_i, \hat{y}_i)$ refers to the estimated coordinate position of i -th node, and (x_i, y_i) is the actual coordinate position of i -th node; N_t refers to the number of nodes that can be located.

4.1 Anisotropic Factors

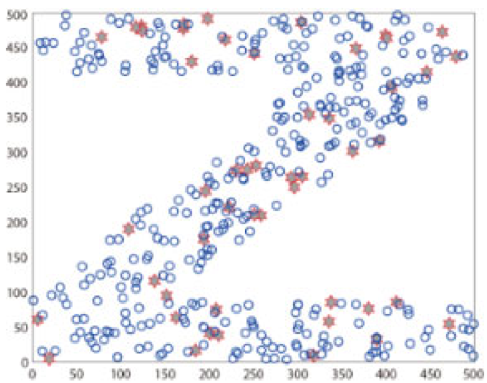
The anisotropic problem can be caused by various reasons, such as irregular deployment of nodes, uneven distribution of nodes and barriers. The C-shaped, S-shaped, and Z-shaped network topology can reflect the anisotropic problem, so we assume the nodes are evenly distributed in the C-shaped, S-shaped, and Z-shaped networks. Figure 5 shows that there are 400 nodes in the network, 12.5% of which are anchor nodes (represented by hexagrams), and the rest nodes are unknown nodes (represented by circles).



(a) C-shaped topology network



(b) S-shaped topology network

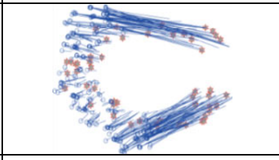
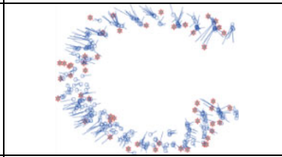
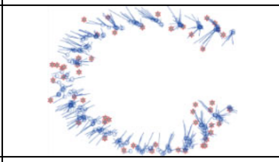
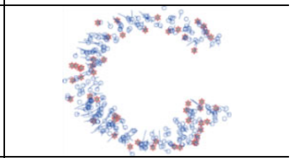
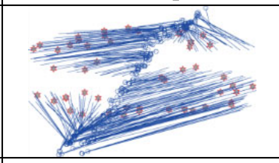
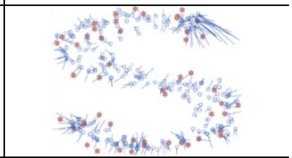
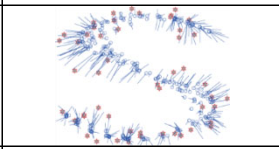
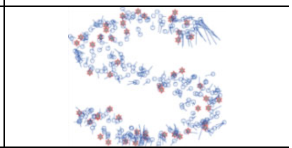
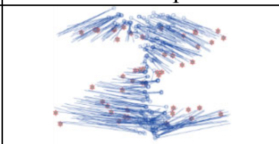
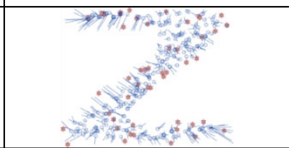
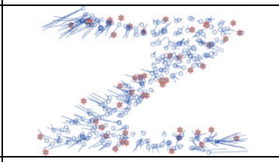
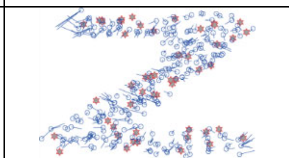


(c) Z-shaped topology network

Figure 5. Anisotropic network

Assume all nodes are homogeneous, *i.e.*, they have equal transmission radius R =60. Table 2 shows the final localization results of the four localization algorithms of DV-hop, PDM, LSVR and ML-ℓ_2CLS under the node distribution as shown in Figure 5. In it, the circle represents the estimated location of unknown nodes; the straight line connects the actual coordinate of unknown nodes with its estimated coordinate, and the longer the line is, the bigger the localization error.

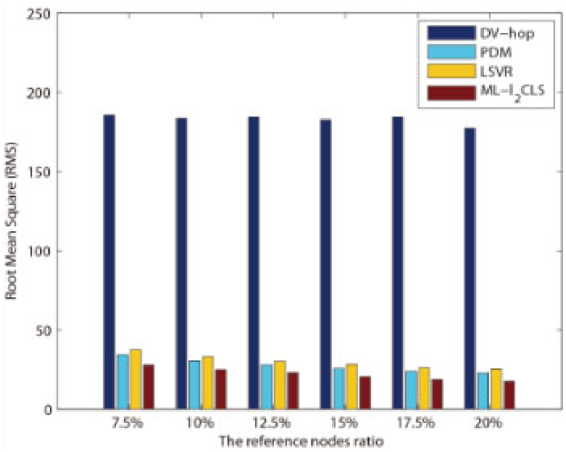
Table 2. Localization results of four algorithms under anisotropic networks

	DV-hop	PDM
C-shaped		
RMS	162.014	26.731
	LSVR	ML-ℓ_2CLS
C-shaped		
RMS	31.633	21.904
	DV-hop	PDM
S-shaped		
RMS	223.341	31.273
	LSVR	ML-ℓ_2CLS
S-shaped		
RMS	37.851	24.534
	DV-hop	PDM
Z-shaped		
RMS	124.803	28.945
	LSVR	ML-ℓ_2CLS
Z-shaped		
RMS	32.971	23.499

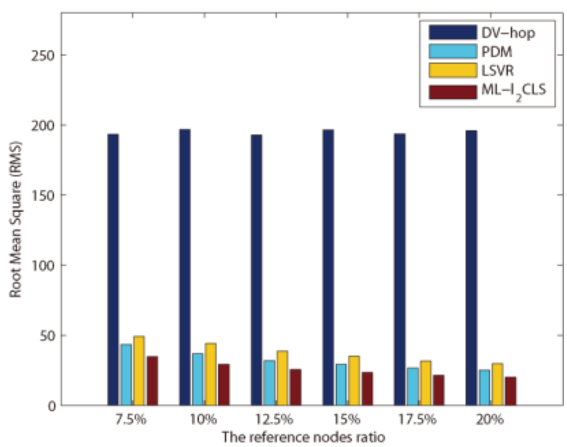
According to Table 2, the DV-hop method uses fixed hops-physical distance conversion coefficient, and in the anisotropic network, the conversion between hops and distance tends to generate deviation; although the PDM method can solve the deviation problem during hops-physical distance conversion in a certain degree, it cannot well solve the data class problem of hops and physical distance; for LSVR, the selection of the combinations of multiple parameters tends to result in over-fitting of final results (the estimated positions of unknown nodes tend to present a curve); the ML-ℓ_2 CLS method proposed in this paper has higher localization accuracy than all the above localization methods.

4.2 Anchor Nodes Ratio

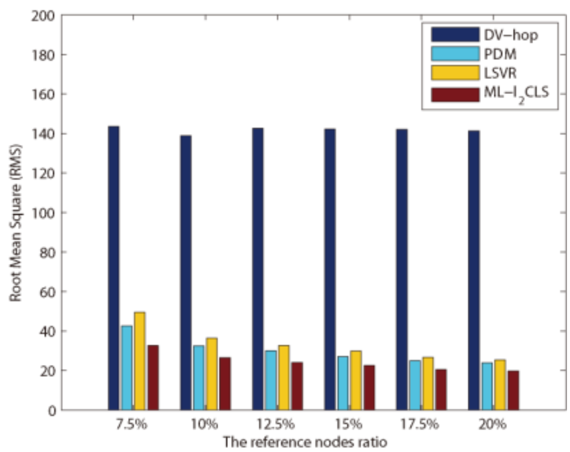
Another factor that affects the node localization accuracy is the ratio of anchor nodes. In order to reduce the one-sidedness of the results of one experiment, we conducted 100 simulations to each type of anisotropic network. Figure 6(a) to Figure 6(c) shows the histograms of RMS mean values in 100 simulation experiments under different proportions of anchor nodes. We can see that the DV-hop method cannot well adapt to localization in the anisotropic network, its RMS value is close to 200 in C-shaped and S-shaped networks and close to 150 in Z-shaped network, which presents fluctuation with the increase of anchor nodes. The LSVR and PDM methods and the ML- ℓ_2 CLS method proposed by us have higher localization accuracy than the DV-hop method, and the RMS value decreases progressively with the increase of anchor nodes. The ML- ℓ_2 CLS method proposed in our paper has the lowest RMS values in three anisotropic networks. In the C-shaped topology, the localization accuracy of ML- ℓ_2 CLS method is 88.9%, 19.8% and 26.6% higher than that of the DV-hop, PDM and LSVR methods respectively; in the S-shaped topology, the localization accuracy of ML- ℓ_2 CLS method is 86.7%, 19.8% and 32.2% higher than that of the DV-hop, PDM and LSVR methods respectively; in the Z-shaped topology, the localization accuracy of ML- ℓ_2 CLS method is 82.8%, 19.2% and 27% higher than that of the DV-hop, PDM and LSVR methods respectively.



(a) Effect of anchor nodes fraction in C-shaped network



(b) Effect of anchor nodes fraction in S-shaped network



(c) Effect of anchor nodes fraction in Z-shaped network

Figure 6. Simulation results of anisotropic network with different ratio of anchors

Table 3 shows the time analysis of four multi-hop localization algorithms in above experiment environment with 50 anchor nodes, and 100 experiments were conducted.

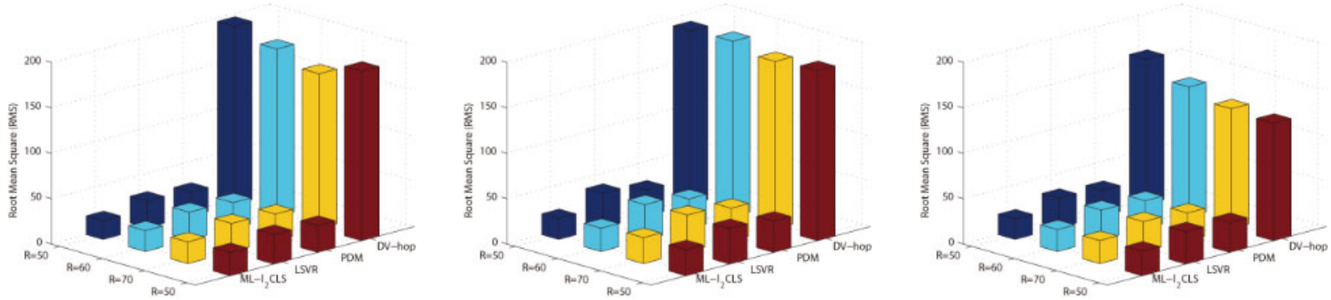
Table 3. Running time in anisotropic network

CPU TIME(m)		**DV-hop**	**PDM**	**LSVR**	**ML-ℓ_2CLS**
50anchors,350 unknown nodes (C-shaped)	**average**	1.068	1.504	2.067	1.706
	median	1.066	1.501	2.066	1.702
	Worst	1.098	1.534	2.161	1.795
	standard deviation	0.009	0.017	0.029	0.020
50anchors,350 unknown nodes (S-shaped)	**average**	1.068	1.502	2.066	1.709
	median	1.067	1.501	2.063	1.708
	Worst	1.116	1.535	2.114	1.745
	standard deviation	0.011	0.017	0.023	0.017
50anchors,350 unknown nodes (Z-shaped)	**average**	1.075	1.521	2.081	1.718
	median	1.073	1.523	2.081	1.718
	Worst	1.124	1.588	2.129	1.767
	standard deviation	0.013	0.021	0.023	0.016

When building the model, the ML-ℓ_2CLS method has the same computation complexity as the PDM method, both of which is $O(m^3)$. However, the ML-ℓ_2CLS method chooses the optimum λ value based on the GCV method, so the operation time of ML-ℓ_2CLS is more optimal than that of DV-hop method, slightly more optimal than that of PDM method, and significantly more optimal than that of LSVR method.

4.3 Node Density

In this section, we also conducted experiment to estimate the influence of node density on localization error. In general, the number of nodes' average neighbors $\rho = \frac{N_t \times R^2}{S_A}$ is used to represent the node density, in which, S_A refers to the size of node deployment area. In order to change the density of nodes in the network, we adopt the way to adjust the communication radius of the node in the experiment. When the node communication radius changes within the range of [50, 80], in the C-shaped network, the variation range of ρ is [2.6, 5.6]; in the S-shaped area, the variation range of ρ is [2.19 4.36]; in the Z-shaped area, the variation range of ρ is [2.22 4.65]. In Fig.7, the histograms show the RMS mean values by using four multi-hop localization methods in the anisotropic network under different communication radiuses, and these mean values were obtained from 100 simulation experiments. In theory, with the increase of node communication radius, the number of neighbors would increase accordingly, and the hop distance (*i.e.*, hops) of nodes would become closer to the actual distance. However, it is not the case in reality. As described in the literature [21], in the multi-hop network, the node distribution is Poisson distribution, and the increase of communication radius will not increase the localization accuracy, but results in fluctuation. According to Figure 7, we can see that although the localization accuracy presents fluctuation with the change of communication radius, no matter how the communication radius changes, the method proposed in this paper is always superior to the DV-hop, PDM and LSVR methods. High transmitting distance will increase the node transmission distance, but it involves high energy consumption, whereas although small transmitting distance requires lower energy, it will involve many hops and routes. Therefore, according to the deductions of literature [22] and partial space radius set in this paper, we set the communication radius at 60 in the experiment in Section 5.1 and 5.2.



(a) The average RMS of different communication radius in C-shaped networks.

(b) The average RMS of different communication radius in S-shaped networks.

(c) The average RMS of different communication radius in Z-shaped networks.

Figure 7. Effects on localization accuracy of nodes' communication radius

5 Conclusion

This paper proposes a range-free localization method ML-ℓ_2CLS, which can be used to build the hops-physical distance relationship model with ℓ_2CLS. First of all, the algorithm standardizes the hops and physical distance to prevent the order of magnitudes problem during the conversion process. Then, the algorithm prevents the over-fitting problem during the model building process by restricting the hops scale, and in this way, it can control the complexity and improve the generalization ability of the model. In the meantime, the algorithm has also obtained the optimal value of Lagrange multiplier with the GCV method, in this way to improve the algorithm's localization accuracy. Various experiments of anisotropic network all show that compared to similar algorithms, the ML-ℓ_2 CLS method proposed in this paper has better localization performance and stronger adaptability.

Acknowledgements

The paper is sponsored by the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (15KJB520009, 16KJD520004), Project of Modern Educational Technology of Jiangsu Province (2015-R-4430), and Philosophy and Social Science Research in Colleges and Universities in Jiangsu Province (2015SJB001).

References

[1] M. P. Đurišić, Z. Tafa, G. Dimić, V. Milutinović, A Survey of Military Applications of Wireless Sensor Networks, *Proc. 2012 Mediterranean Conference on Embedded Computing*, Bar, Montenegro, 2012, pp. 1-4.

[2] O. Sohaib, H. Lu, W. Hussain, Internet of Things (IoT) in E-commerce: For People with Disabilities, *Proc. 2017 12th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, Siem Reap, Cambodia, 2017, pp. 419-423.

[3] K. Duarte, J. Cec´ılio, J. S. Silva, P. Furtado, Information and Assisted Navigation System for Blind People, *Proceedings of the 8th International Conference on Sensing Technology*, Liverpool, UK, 2014, pp. 470-473.

[4] D. Gao, H. Lin, F. Zhang, Y. Liu, Maximum Network Throughput Based on Distributed Algorithm for Rechargeable Wireless Sensor Networks, *Adhoc & Sensor Wireless Networks*, Vol. 35. No. 3-4, pp. 193-215, January, 2017.

[5] L. Atzori, A. Iera, G. Morabito, From "Smart Objects" to "Social Objects": The Next Evolutionary Step of the Internet of Things, *IEEE Communications Magazine*, Vol. 52, No. 1. pp. 97-105, January, 2014.

[6] Z. Yang, C. Wu, Y. Liu, *Location-based Computing: Localization and Localizability of Wireless Networks*, Tsinghua University Press, 2014.

[7] S. Zhang, X. Liu, J. Wang, J. Cao, G. Min, Accurate Range-Free Localization for Anisotropic Wireless Sensor Networks, *ACM Transactions on Sensor Networks*, Vol. 11. No. 3. Article No. 51, pp.1-27, May, 2015.

[8] D. Zhang, Z. Fang, H. Sun, J. Cao, HTCRL: A Range-Free Location Algorithm Based on Homothetic Triangle Cyclic Refinement in Wireless Sensor Networks, *Information*, Vol. 8, No. 2, pp. 1-15, June, 2017.

[9] B. Xiao, L. Chen, Q. Xiao, M. Li, Reliable Anchor-Based Sensor Localization in Irregular Areas, *IEEE Transactions on Mobile Computing*, Vol. 9, No.1, pp. 60-72, January, 2010.

[10] D. Niculescu, B. Nath, DV Based Positioning in Ad Hoc Networks, *Telecommunication Systems*, Vol. 22, No. 1-4, pp. 267-280, January, 2003.

[11] H. Lim, J. C. Hou, Distributed Localization for Anisotropic Sensor Networks, *ACM Transactions on Sensor Networks*, Vol. 5, No. 2, pp. 1-26, March, 2009.

[12] J. Lee, W. Chung, E. Kim, A New Kernelized Approach to Wireless Sensor Network Localization, *Information Sciences*, Vol. 243, No. 10, pp. 20-38, September, 2013.

[13] P. Domingos, The Role of Occam's Razor in Knowledge Discovery, *Data Mining and Knowledge Discovery*, Vol. 3, No. 4, pp. 409-425, December, 1999.

[14] X. Yan, A. Song, Z. Yang, W. Yang, An Improved Multihop-Based Localization Algorithm for Wireless Sensor Network Using Learning Approach, *Computers & Electrical Engineering*, Vol. 48, No. C. pp. 247-257, November, 2015.

[15] X. Yan, Z. Yang, A. Song, W. Yang, Y. Liu, R. Zhu, A Novel Multihop Range-Free Localization Based on Kernel Learning Approach for the Internet of Things, *Wireless Personal Communications*, Vol. 87, No. 1, pp. 269-292, March, 2016.

[16] C. Jo, C. Lee, Multilateration Method Based on The Variance of Estimated Distance in Range-free Localization, *Electronics Letters*, Vol. 52, No. 12, pp. 1078-1080, June, 2016.

[17] L. Gui, T. Val, A. Wei, R. Dalce, Improvement of Range-free Localization Technology by a Novel DV-Hop Protocol in Wireless Sensor Networks, *Ad Hoc Networks*, Vol. 24, pp. 55-73, January, 2015.

[18] W. Zhao, S. Su, F. Shao, Improved DV-Hop Algorithm Using Locally Weighted Linear Regression in Anisotropic Wireless Sensor Networks, *Wireless Personal Communications*, Vol. 98, No. 4, pp. 3335-3353, February, 2018.

[19] W. Zheng, X. Yan, W. Zhao, C. Qian, A Large-Scale Multi-Hop Localization Algorithm Based on Regularized Extreme Learning for Wireless Networks, *Sensors*, Vol. 17, No. 12, pp. 1-18, December, 2017.

[20] F. Shahzad, T. R. Sheltami, E. M. Shakshuki, DV-maxHop: A Fast and Accurate Range-Free Localization Algorithm for Anisotropic Wireless Networks, *IEEE Transactions on Mobile Computing*, Vol. 16, No. 9, pp. 2494-2505, September, 2017.

[21] H. Takagi, L. Kleinrock, Optimal Transmission Ranges for Randomly Distributed Packet Radio Terminals, *IEEE Transactions on Communications*, Vol. 32, No. 3, pp. 246-

257, March, 1984.
[22] L. Kleinrock, J. Silvester, Optimum Transmission Radii for Packet Radio Networks or Why Six Is a Magic Number, *Proceedings of the IEEE National Telecommunications Conference*, Birmingham, Ala, 1978, pp. 431-435.

Biographies

Bo Hu received the B.S. degree in school of computer science and technology from Jiangsu normal university in 2001. He received M.S. degree in college of computer and information from Hohai University in 2012. He is the director and senior engineer at the center of information construction and management, Nanjing normal university of special education. His current research interests lie in the areas of wireless sensor networks, information accessibility, and the internet of things.

Sainan Zhang received the M.S. degree in educational technology professional in 2013 from Nanjing normal university. She is currently an engineer with the center of information construction and management, Nanjing normal university of special education. Her current research interests include cloud computing, virtualization technology and internet of things.

Zhenyan Yuan graduated from Nanjing College of political science of PLA in 2006, and worked in China satellite maritime measurement and control center during the year 2006 to 2011, and work in military data and information engineering teaching and research office, in command & control engineering college of army engineering university from 2012 to the present. Her current research interests include cloud computing, big data and Internet of Things.