

# Support Vector Machine Based Range-Free Localization Algorithm in Wireless Sensor Network

Tao Tang<sup>1(✉)</sup>, Haicheng Liu<sup>1</sup>, Haiyan Song<sup>1</sup>, and Bao Peng<sup>2</sup>

<sup>1</sup> College of Electrical and Information Engineering,  
Heilongjiang Institute of Technology, Harbin 150070, Heilongjiang, China  
Tt8854@126.com

<sup>2</sup> School of Electronic and Communication,  
Shenzhen Institute of Information Technology, Shenzhen 518172, China

**Abstract.** Localization method is critical issues in Wireless Sensor Network (WSN) system. The existing node localization algorithms, especially range-based algorithms, did not consider the distances measured error and this may result in severe location errors that degrade the WSN performance. In this paper, a new algorithm called Support Vector Machine based Range-free localization (RFSVM) algorithm in WSN is proposed. This algorithm introduced a new matrix called transmit matrix which maps the relationship between the hops and distance. And use the SVM model to estimate the position of unknown nodes. This algorithm does not need any addition hardware, and the experiments shows that it can lead to the localization accuracy character good enough.

**Keywords:** Wireless Sensor Networks · Localization · Range-free · Support Vector Machine

## 1 Introduction

One of the critical issues in Wireless Sensor Network (WSN) research is to determine the physical positions of nodes [1]. This is because: information from the sensors is useful only if node location information is also available; additionally, some routing protocols use position to determine viable routes [2, 3]. The process of determining the position of WSN nodes is known as localization. The objective of localization is to determine the virtual or physical coordinates of each node in the network [4]. Along with the employing in military activities such as reconnaissance, surveillance [5], and target acquisition [6], environmental activities [7], or civil engineering such as structural health measurement [8], the localization system becomes a significant base in WSN.

A simple way to ensure this is to equip every node in the WSN with GPS. Due to economical issues, only a small subset of the network can affordably be equipped with GPS; such nodes are called anchor nodes, or beacons. An automatic localization process is required for the rest of the nodes (unknown nodes) in the network, using the anchor nodes as reference points. This process is commonly known as location

discovery (LD) [9]. The existing node localization algorithms can be divided into Range-based and Range-free algorithms.

Range-based employed distance measurements between sensor nodes through a received signal strength indicator (RSSI), time of arrival (TOA), or time difference of arrival (TDOA), the angle can be measured by the angle of arrival (AOA) [10] and et al. However, the distances measured among the nodes in a WSN usually contain errors. These errors range from slight errors to large ones, and are difficult to characterize by a standard model such as Gaussian. These distance measurement errors may result in severe location errors that degrade the WSN performance, and therefore should be taken into account during the LD. Range-free algorithms do not need absolute range information. Typical range-free algorithms including centroid, DV-Hop [11], amorphous [12], APIT [13], etc. The authors in [14] proposed a coarse grained range-free algorithm to lower the uncertainty of nodes positions using radio connectivity constraints. In [15], the authors used geometry method to determine the sensor node location based on the cross point of the two chords in a circle. The range-free algorithms are more economical, cost-effective, and feasible for the large-scale WSN. However, the Range-free always have accuracy problem.

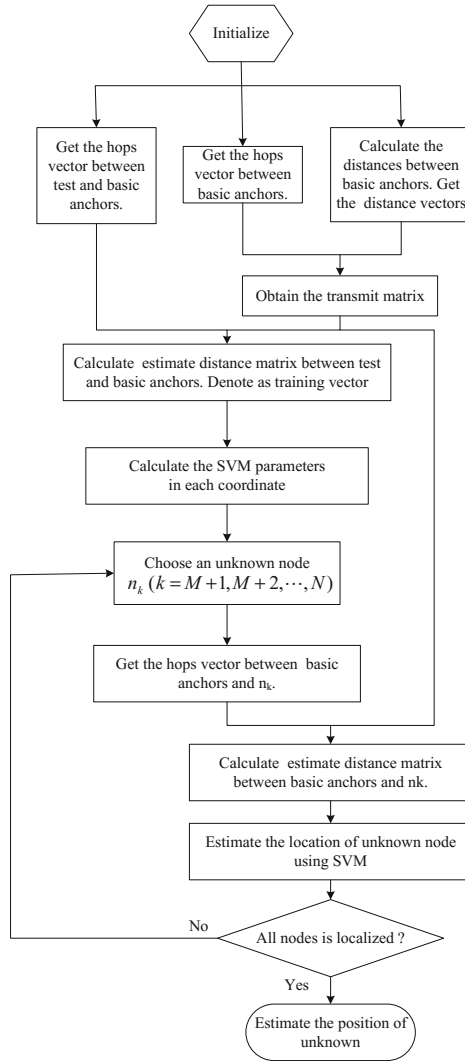
In this paper, a new algorithm called Support Vector Machine based Range-free localization (RFSVM) algorithm in WSN is proposed. This algorithm first introduced a new matrix called transmit matrix which maps the relationship between the hops and distance. Then use the transmit matrix to calculate the training vector data and test vector. Finally, establish the SVM model and estimate the position of unknown nodes.

## 2 Network Model and Assumptions

Suppose that we have  $M + N$  nodes,  $\{n_1, n_2, \dots, n_{M-1}, n_M, n_{M+1}, \dots, n_{M+N}\}$ , random deployed in a two-dimension space  $[0, D] \times [0, D]$ . Among them,  $M$  nodes  $\{n_1, n_2, \dots, n_M\}$  are anchor nodes, the locations of which are supposed to be known. The anchors are randomly static deployed in WSN with density  $\rho_L$ . Other nodes  $\{n_{M+1}, n_{M+2}, \dots, n_{M+N}\}$  are regarded as unknown nodes, and they are randomly deployed with a density  $\rho_S, \rho_S \gg \rho_L$ . We assume the communication radius of each node is  $R$ .

## 3 The Design of RFSVM Algorithm

The localization process of RFSVM algorithm can be divided into three steps, which are transmit matrix obtained, SVM [16] model build and localization. The WSN is initialized and transmit matrix is got in the first step. Then, the localization system get the SVM model through the training vector date which is calculated by the transmit matrix. Finally, the unknown nodes justify which area is it in using SVM method. The whole process of RFSVM algorithm is shown in Fig. 1.



**Fig. 1.** The process of RFSVM algorithm

### 3.1 Transmit Matrix Obtained

First of all, the anchors are separated into two parts: basic anchor  $n_i(i = 1, 2, \dots, K)$  and test anchor  $n_j(j = K + 1, K + 2, \dots, M)$ . Let  $\mathbf{H}$  be the hops vector between the  $i$ th basic anchor node and the other basic anchor nodes, therefore

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \vdots \\ \mathbf{h}_K \end{bmatrix} = \begin{bmatrix} h_{1,1} & h_{2,1} & \cdots & h_{K,1} \\ h_{1,2} & h_{2,2} & \cdots & h_{K,2} \\ \vdots & \vdots & \ddots & \vdots \\ h_{1,K} & h_{2,K} & \cdots & h_{K,K} \end{bmatrix} \in \mathbf{Z}^{K \times K}$$

The distance vector between them can be calculated as

$$\mathbf{D} = \begin{bmatrix} d_{1,1} & d_{2,1} & \cdots & d_{K,1} \\ d_{1,2} & d_{2,2} & \cdots & d_{K,2} \\ \vdots & \vdots & \ddots & \vdots \\ d_{1,K} & d_{2,K} & \cdots & d_{K,K} \end{bmatrix} \in \mathbf{Z}^{K \times K}$$

Then, let us think about a linear matrix called transmit matrix  $\mathbf{T}$  which optimally maps the hops vector  $\mathbf{H}$  to the geographical distance  $\mathbf{D}$ ,

$$\mathbf{T} = \begin{bmatrix} \mathbf{t}_1 \\ \mathbf{t}_2 \\ \vdots \\ \mathbf{t}_k \end{bmatrix} = \begin{bmatrix} t_{1,1} & t_{2,1} & \cdots & t_{K,1} \\ t_{1,2} & t_{2,2} & \cdots & t_{K,2} \\ \vdots & \vdots & \ddots & \vdots \\ t_{1,K} & t_{2,K} & \cdots & t_{K,K} \end{bmatrix} \in \mathbf{Z}^{K \times K}$$

where  $t_{i,j}$  represents the effect of proximity to the  $j$ th basic anchor node on the geographic distance to the  $i$ th basic anchor node. We derive the transition by minimizing the following square error for  $i = 1, 2, \dots, K$

$$e_i = \sum_{k=1}^K (d_{i,k} - \mathbf{t}_i \mathbf{h}_k)^2 = \|\mathbf{d}_i^T - \mathbf{t}_i \mathbf{H}\| \quad (1)$$

Finally, our goal is to minimize (1). Therefore, the transmit matrix  $\mathbf{T}$  could be obtained through this equals.

### 3.2 SVM Model Constitution

In this step, each dimension of WSN coordinate value is separated into  $M$  pieces for  $M$  classes. So, in the two-dimension space  $[0, D] \times [0, D]$ , we have  $M$  classes  $\{cx_1, cx_2, \dots, cx_M\}$  in  $x$  dimension and  $M$  classes  $\{cy_1, cy_2, \dots, cy_M\}$  in  $y$  dimension. Then, each anchor  $n_j$  ( $j = K + 1, K + 2, \dots, M$ ) computed the distance to all other anchor nodes  $n_i$  ( $i = 1, 2, \dots, K$ ) using the transmit matrix  $\mathbf{T}$ . Supposed that the hops between  $n_j$  and  $n_i$  is  $h_{ij}$ , therefore

$$\mathbf{H} = \begin{bmatrix} h_{K+1,1} & h_{K+2,1} & \cdots & h_{M,1} \\ h_{K+1,2} & h_{K+2,2} & \cdots & h_{M,2} \\ \vdots & \vdots & \ddots & \vdots \\ h_{K+1,K} & h_{K+2,K} & \cdots & h_{M,K} \end{bmatrix} \in \mathbb{Z}^{K \times (M-K)}$$

And the estimate distance matrix between  $n_j$  and  $n_i$  is

$$\mathbf{D}' = \mathbf{T} \times \mathbf{H} = \begin{bmatrix} d'_{K+1,1} & d'_{K+2,1} & \cdots & d'_{M,1} \\ d'_{K+1,2} & d'_{K+2,2} & \cdots & d'_{M,2} \\ \vdots & \vdots & \ddots & \vdots \\ d'_{K+1,K} & d'_{K+2,K} & \cdots & d'_{M,K} \end{bmatrix} = [\mathbf{D}'_{K+1} \quad \mathbf{D}'_{K+2} \quad \cdots \quad \mathbf{D}'_M] \in \mathbb{Z}^{K \times (M-K)} \quad (3)$$

Where,  $\mathbf{D}'_j = [d'_{j1} \ d'_{j2} \ \cdots \ d'_{jk}]^T$  ( $j = K+1, K+2, \dots, M$ ). Thus, a distance vector  $[d'_{j1}, d'_{j2}, \dots, d'_{jk}]$  is constructed in each anchor  $n_j$ ,  $d'_{ji}$  distance between test anchor  $n_j$  and basic anchor  $n_i$ .

Because of the position acknowledgement in test anchor  $n_j$ , the location classes of it known as  $(cx_j, cy_j)$ . Then, transform the anchor nodes  $n_j$  distance vector  $[d'_{j1}, d'_{j2}, \dots, d'_{jk}]$  into the training data which are  $\{[d'_{j1}, d'_{j2}, \dots, d'_{jk}], cx_j\}$  in  $x$  dimension and  $\{[d'_{j1}, d'_{j2}, \dots, d'_{jk}], cy_j\}$  in  $y$  dimension. Therefore, two SVMs will be built with a Gauss kernel function as formula (4)

$$K(x, x_i) = \exp\left\{-\frac{|x - x_i|^2}{\sigma^2}\right\} \quad (4)$$

In each SVM, arbitrary test data  $x$  can be classified by

$$f(x) = \text{sgn}\left\{\sum_{i=1}^l \alpha_i^* y_i K(x_i, x) + b^*\right\} \quad (5)$$

Where  $x_i$  is support vector,  $y_i \in \{-1, 1\}$  is class label corresponded to  $x_i$ ,  $K(x, x_i)$  is kernel function,  $\alpha_i^*$  is Lagrange multiplier corresponded to  $x_i$ ,  $b^*$  is classification threshold value and  $\text{sgn}$  is symbol function.

Through calculating the training data, support vector  $x_i$ , Lagrange multiplier  $\alpha_i^*$ ; and classification threshold value  $b^*$  are achieved for two SVMs respectively.

### 3.3 Localization

Like step 2, first we estimate the distance between the unknown nodes  $n_k$  ( $k = M+1, M+2, \dots, N$ ) and the basic anchors  $n_i$  ( $i = 1, 2, \dots, K$ ). Supposed that the hops between  $n_k$  and  $n_i$  is  $h_{ik}$ , therefore

$$\mathbf{H} = [\mathbf{H}_{M+1} \quad \mathbf{H}_{M+2} \quad \cdots \quad \mathbf{H}_{M+N}] \in \mathbb{Z}^{K \times N} \quad (6)$$

And the estimate distance matrix between  $n_k S$  and  $n_i S$  is

$$\mathbf{D}' = \mathbf{T} \times \mathbf{H} = \begin{bmatrix} d'_{M+1,1} & d'_{M+2,1} & \cdots & d'_{M+N,1} \\ d'_{M+1,2} & d'_{M+2,2} & \cdots & d'_{M+N,2} \\ \vdots & \vdots & \ddots & \vdots \\ d'_{M+1,K} & d'_{M+2,K} & \cdots & d'_{M+N,K} \end{bmatrix} = [\mathbf{D}'_{M+1} \quad \mathbf{D}'_{M+2} \quad \cdots \quad \mathbf{D}'_{M+N}] \in \mathbb{Z}^{K \times N} \quad (7)$$

Where,  $\mathbf{D}'_k = [d'_{k1} \quad d'_{k2} \quad \cdots \quad d'_{kK}]^T (k = M+1, M+2, \cdots, M+N)$ .

Thus, the test vector date of  $n_k$  can be denoted as  $\{[d'_{k1}, d'_{k2}, \cdots, d'_{kK}]\}$ . By considering the distance vectors in unknown nodes as testing data  $x$ , the classes of two coordinate values are attained according to Eq. (5). If the classes of two coordinate values of node  $n_k$  are  $(cx_i, cy_j)$ , which are decided by SVMs, the node  $n_k$  is believed to locate in the cubic unit  $[(i-1)D/M, iD/M] \times [(j-1)D/M, jD/M]$ . The cubic unit centroid  $[(i-\frac{1}{2})D/M, (j-\frac{1}{2})D/M]$  is used as the estimated position of the node  $n_k$ . The maximum error of localization for node  $n_k$  is  $\sqrt{2}D/2M$ , when the SVMs perform correct classification.

The process of SVM model constitution and localization step is shown in Fig. 2.

## 4 Emulation and Analysis

In this paper, MATLAB simulation environment is used to simulate and analyze the two-dimension localization algorithm. Supposing all nodes is distributed randomly in a two-dimension space, which is 50 m x 50 m in size. The two-dimension space is divided into 25 small cubic units, which are 10 m x 10 m in size. Thus, the distance of every dimension is  $D = 50$  m, and the number of classes for every dimension is  $M = 5$ . In all of the experiments, the total number of nodes is set to  $N = 250$  (50 anchor and 200 unknown node). The communication radius is assumed to be  $R = 25$  m. Due to the random distribution of sensor nodes, the randomness of the localization result derived from the experiment can not be ignored. The experiment is repeated for 50 times on the same network parameter settings and use statistical methods to obtain accurate experimental results. All nodes will be laid in two-dimension region uniformly and randomly in each experiment. Therefore, the error of all unknown nodes location is shown in Fig. 3.

As shown in Fig. 3, the estimate errors of unknown nodes are less than 5 m when the communication range is 25 m. In the other words, the accuracy of the RFSVM algorithm is more than 80%. And if we set more anchors in the networks, we will get more accuracy of the nodes localizations.

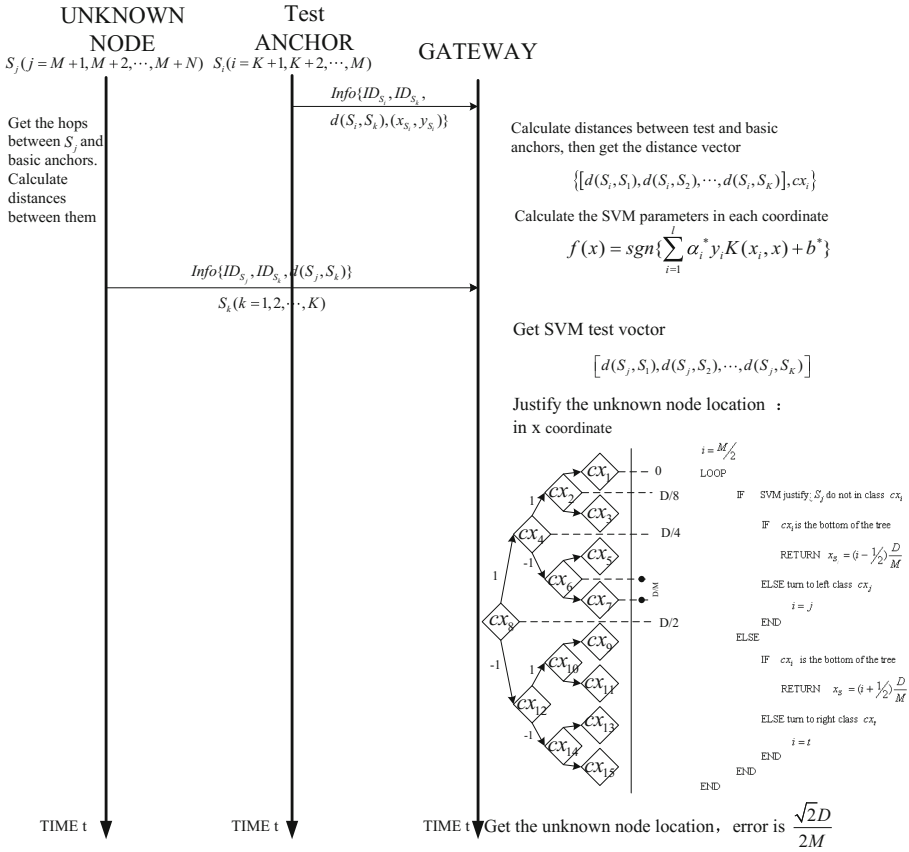


Fig. 2. The process of SVM model constitution and localization

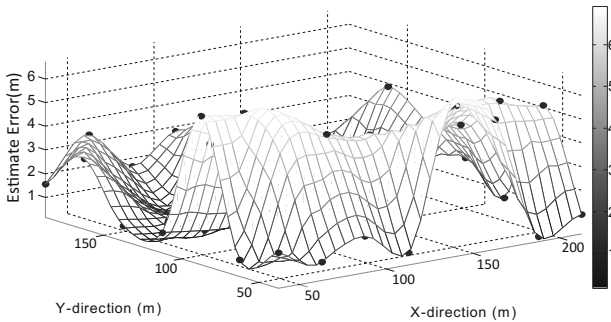


Fig. 3. Unknown nodes location error

## 5 Conclusions

In this paper, a new algorithm called Support Vector Machine based Range-free localization (RFSVM) algorithm in WSN is proposed. This algorithm introduced a new matrix called transmit matrix which maps the relationship between the hops and distance. Then use the transmit matrix and SVM method to get the unknown nodes position. The whole algorithm does not need any addition hardware such as RSSI, antennas et al. The experiment shows that the accuracy of the RFSVM algorithm is more than 80%. Along with the anchor number increasing, the location of nodes will be more accuracy.

**Acknowledgments.** This paper is by Research Fund for the Doctoral Program of HLJIT under Grant No. 2014BJ10; University Nursing Program for Young Scholars with Creative Talents in Heilongjiang Province UNPYSCT-201501.

Also supported by The Basic Research Plan in Shenzhen City under Grant No. JCYJ20130401100512995.

## References

1. Culler, D., Estrin, D., Srivastava, M.: Overview of sensor networks. *IEEE Comput.* **37**(8), 41–49 (2004)
2. Li, J., Jannotti, J., De Couto, D., Karger, D., Morris, R.: A scalable location service for geographic ad-hoc routing. In: *Proceedings of the 6th ACM International Conference on Mobile Computing and Networking (MobiCom 2000)*, pp. 120–130 (2000)
3. Amouris, K., Papavassiliou, S., Li, M.: A position-based multi-zone routing protocol for wide area mobile ad-hoc networks. In: *Proceedings of the Vehicular Technology Conference, Houston, TX* (2010)
4. Bäck, T.: *Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms*. Oxford University Press, New York (1996)
5. Lédeczi, Á., Nádas, A., Völgyesi, P., Balogh, G., Kusy, B., Sallai, J., Pap, G., Dóra, S., Molnár, K., Maróti, M., Simon, G.: Countersniper system for urban warfare. *ACM Trans. Sen. Netw.* **1**(2), 153–177 (2015)
6. Nemeroff, J., Garcia, L., Hampel, D., DiPierro, S.: Application of sensor network communications. In: *Military Communications Conference, 2001, MILCOM 2001, Communications for Network-Centric Operations: Creating the Information Force*, vol. 1, pp. 336–341. *IEEE* (2011)
7. Mladineo, N., Knezic, S.: Optimisation of forest fire sensor network using GIS technology. In: *Proceedings of the 22nd International Conference on Information Technology Interfaces, ITI 13–16 June 2000*, pp. 391–396 (2000)
8. Xu, N., Rangwala, S., Chintalapudi, K.K., Ganesan, D., Broad, A., Govindan, R., Estrin, D.: A wireless sensor network for structural monitoring. In: *SenSys 2004: Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems*, pp. 13–24. *ACM*, New York (2004)
9. Molina, G., Alba, E.: Location discovery in Wireless Sensor Networks using metaheuristics. *Appl. Soft Comput.* **11**, 1223–1240 (2011)



10. Mao, G., Fidan, B., Anderson, B.D.O.: Wireless sensor network localization techniques. *Comput. Netw.* **51**(10), 2529–2553 (2007)
11. Bulusu, N., Heidemann, J., Estrin, D.: GPS-less low cost outdoor localization for very small devices. *IEEE Personal Commun. Mag.* **7**(5), 28–34 (2010)
12. Nagpal, R.: Organizing a global coordinate system from local information on an amorphous computer. Technical Report AI Memo 1666, MIT Artificial Intelligence Laboratory (1999)
13. He, T., Huang, C.D., Blum, B.M.: Range-free localization schemes for large scale sensor networks. In: *Proceedings of the 9th Annual International Conference on Mobile Computing and Networking*, pp. 81–95. ACM, San Diego (2013)
14. Galstyan, A., Krishnamachari, B., Lerman, K., Patten, S.: Distributed online localization in sensor networks using a moving target. In: *Proceedings of the 3rd International Symposium on Information Processing in Sensor Networks*. ACM, 61–70, Berkeley (2004)
15. Su, K.F., Ou, C.H., Jiau, H.C.: Localization with mobile anchor points in wireless sensor networks. *IEEE Trans. Veh. Technol.* **54**(3), 1187–1197 (2015)
16. Tran, D.A., Nguyen, T.: Support vector classification strategies for localization in sensor networks. In: *Proceedings of the Communications and Electronics*, Hanoi (2006)