Wireless Sensor Network Multi-Hop Positioning Algorithm Based on Continuous Regression

https://doi.org/10.3991/ijoe.v12i10.6199

Weimin Han[1], Shijun Li[2,4], Weidong Li[3]
[1] Hunan Vocational Institute of Technology, Xiangtan, China;
[2] Hunan Institute of Engineering; Hunan Provincial Key Laboratory of Wind Generator and Its Control Xiangtan, China
[3] Powerchina Zhongnan Engineering Co.Ltd, Changsha, China
[4] Hunan province Cooperative Innovation Center for Wind Power Equipment and Energy Conversion, Xiangtan, China

Abstract—**The traditional multi-hop positioning algorithm is easily affected by the network anisotropy, thus resulting in unstable positioning performance. The wireless sensor network multi-hop positioning algorithm based on continuous regression is put forwarded in the paper to address this problem. By utilizing the continuous regression model, the mapping relationship between the hop count and Euclidean distance is constructed so as to transform the positioning process model into regression prediction. Theoretical analysis and simulation results show that the improved algorithm improves the positioning accuracy, and avoids the influence of the network topology anisotropy on the performance of the algorithm. The algorithm requires little expenditure and few parameters so it can be adapted to wireless sensor networks with irregular nodes distribution, and can be of great engineering application value.**

Index Terms—**wireless sensor network, multi-hop positioning algorithm, continuous regression, ML-CR node.**

I. INTRODUCTION

Wireless sensor networks (WSN) have been widely used in large-scale data acquisition and environmental monitoring because of its low cost, strong adaptability and other attributes. However, about 80% of the application is related to the node position among the numerous application areas. Therefore, developing a large scale positioning system with low-cost, high precision and suitability for practical application is very much needed in WSN.

The range-free positioning method is often used for positioning of large scale wireless sensor networks. Most of range-free positioning methods obtain the hop count between nodes through the connection between the nodes and estimate the location of unknown nodes by averaging hop distance and hop count. Because the method is easily implemented and has relatively low hardware requirements, it is quite suitable to meet the requirements of positioning of large scale wireless sensor networks. However, the positioning accuracy of range-free multi-hop positioning technology is greatly affected by the network topology. In an isotropic network with high node density and even distribution, the ideal positioning results can be achieved [1]. However, under the condition of network environment anisotropy such as uneven and sparse distribution of nods, the positioning effect is very poor.

Multi-hop positioning is more suitable for application in the actual environment and has lower computational complexity, higher positioning accuracy and better adaptability. As for the defects of multi-hop positioning, on the basis of the common multi-hop positioning method and partial least squares regression algorithm, the multi-hop positioning algorithm based on continuous regression is proposed in this paper[2]. The ML-CR method takes the continuous regression model as the optimal relational model between actual distance and node hop count, and it predicts the distance between an unknown node and a beacon node.

II. CHARACTERISTICS OF WIRELESS SENSOR NETWORK

Compared with a traditional network, the wireless sensor network has the following characteristics:

(1) The sensor network has a large quantity of nodes and high density.

With MEMS technology, a sensor network node is micro. The communication and sensing radius of a node is limited to between ten meters to tens of meters, and most of the nodes are in sleep mode so as to prolong network lifetime [3]. Therefore, a redundancy sensor node will be deployed to generally guarantee network reliability. Node quantity and density of a sensor network are several orders of magnitude higher than those of an Adcock network, which can be plagued by a series of problems such as signal conflict, selection of information transmission path and cooperative work of nodes. A sensor network can operate under adverse environmental conditions, and nodes in the network may be invalidated due to various reasons. In order to guarantee normal operation of the network, a sensor network must be equipped with a certain fault-tolerant capability, and sensor nodes may have a certain failure rate [4].

(3) Node energy, calculation capacity and storage capacity are all limited

As a sensor node is micro and powered by batteries with limited energy and that are difficult to supplement or replace. Power consumption of a sensor node is one of the key constraints during the design of the whole sensor network node [5]. It determines the working life of the network. Furthermore, calculation and storage capacities of sensor nodes are limited, making them unable to conduct complicated operations, and matured protocol and algorithm in the traditional hutment network are too expensive for a sensor network. Therefore, a simple and effective protocol and algorithm must be redesigned.

(4) Varying network topology structure

As a sensor network has its own characteristics, the sensor node switches between operation mode and sleep mode, and the sensor node may be invalid due to various reasons at any given time, or a new sensor node may enter the network to increase robustness of the network. The topological structure of a sensor network may change, and this challenges validity for various algorithms (routing algorithm and link quality control protocol) in the network. In additional, if the node can move, it will also cause topological changes of network [6].

(5) Data-centered

Attention is generally paid to the value of a certain observation index in a certain area in the sensor network rather than specific observation of data changes in a certain node. For example, in the experiment, the sensor network will know "temperature in northeast corner of the detection area" rather than "temperature value detected by node No.8" [7]. This is the data-centered characteristic of a sensor network. Data transmitted by a traditional network is linked to the physical address of the node. Data-centered characteristics require that the sensor network break away from the addressing process of a traditional network so as to conduct information fusion quickly and extract useful information and send it directly to users [8].

As a wireless sensor network is distributed randomly and intensively, it suits the adverse environment of a battlefield. In a military field, a wireless sensor network can be used to gather intelligence about the enemy and monitor deployment of partners [9]. It can also be used to monitor equipment, materials and the battlefield itself to assess the threat of an attack of biological and/or chemical weapons. The military can place a large numbers of inexpensive micro sensor nodes in the area of interest using airplanes so as to monitor changes in the surrounding environment in real time through these sensor nodes, and they can also send monitoring data to the command's monitoring center through a satellite channel or ground base station [10]. Conveyance and collection of information has become a vital aspect in modern warfare, and prompt information acquisition and response time are of great importance on the battlefield. A sensor network can provide prompt and accurate information for command, and it is of great importance to enhance national defense and offensive military capability [11].

III. POSITIONING ALGORITHM

The positioning algorithm is generally divided into two stages: distance measurement and positioning calculation. In the first stage, in the self-organization process of the whole network, each node will collect an RSSI value from itself to the adjacent node of one hop. When the collected quantity reaches the requirement, the algorithm will conduct data filtering, calculate its mean value and save the final result. After the network organization is completed, the beacon node will broadcast positioning data frames, containing beacon node ID, coordinate information, data frame life cycle, hop count and RSSI accumulated value. The hop count and RSSI accumulated value will be initialized to 0. When an unknown node receives the positioning data frame, it will refer to the local history table to verify whether the frame has been received. If it does not receive the frame, it will save the data frame directly. If it receives the frame, it will compare the data frame with the RSSI accumulated value stored in the local table to determine whether the RSSI accumulated value of the newly received data frame is smaller than the local value [12]. If it is smaller, the unknown node will save the coordinate and the RSSI accumulated value of the data frame and will add one to the hop count and subtract one from the life cycle. It will also total the RSSI value and the accumulated value from the sending node to the receiving node at the same time and will store the value into the accumulated value and then transmit the newly handled data frame [13]. If the frame is larger than the local value, the frame will be discarded. If the life cycle is 0, data frame forwarding will be stopped. In the second stage, when the unknown node receives coordinates of three beacon nodes or above, it can conduct positioning calculation with trilateration or maximum likelihood estimate (MLE) to determine the coordinate of the unknown node [14]. The life cycle value of the data frame can be changed according to different network sales so as to promote network coverage of positioning. However, the accuracy of distance measurement will decline [15].

A. ML-CR node positioning model

If there is a sensor network $S=\{S_1,S_2,\dots,S_{m+n}\}$ in a two-dimensional space and it contains m beacon nodes and an unknown nodes, the node coordinate can be expressed in Equation (1).

$$\mathrm{pos}\left(S_p\right)=\left(x_p,y_p\right)^T \qquad p=1,\mathrm{L}\ ,m+n \tag{1}$$

where positions of m beacon nodes $S_i \in B$ are known and positions of the remaining n nodes $S_i \in U$ are unknown and $B=\{S_i|i=1,2,\dots,m\}$, $U=\{S_j|j=m+1,\dots,m+n\}$. Minimum hop count and distance collected by the node are represented with two groups of vector data sets respectively. $h(S_i,S_j)\in H=\{0,1,2,\dots\}$ is used to represent the number of branches from node S_i to node S_j and Euclidean distance from node S_i to node S_j can be expressed with Equation (2).

$$\begin{aligned} d\left(s_i,s_j\right) &= \left\| pos\left(s_i\right)-pos\left(s_j\right)\right\| \\ &= \sqrt{\left(x_i-x_j\right)^2+\left(y_i-y_j\right)^2} \in R^2 \end{aligned} \tag{2}$$

$D=[d_{i1},d_{i2},\dots,d_m]$. Therefore, the positioning problem can be formulated as:

$$\begin{aligned} &Estimate \quad pos\left(S_j\right) \\ &Given \quad pos\left(S_i\right), d\left(S_i,S_k\right), and \quad h\left(S_i,S_j\right) \end{aligned} \tag{3}$$

As there is a mapping relation between the minimum hop count and the actual distance of beacon nodes, namely:

$$D=H\beta+e \tag{4}$$

Where $\beta=\{\beta_1,\beta_2,\dots,\beta_{m+n}\}$ is regression coefficient and e is random error. Equation (4) shows that there are serious multiple correlations in the variables in the equation, or the quantity of sample points in the equation is less than the quantity of the variables [16]. If Equation (4) is forcibly calculated in this circumstance, the estimate will be invalid. In addition, the accuracy of estimate value β is related to both input and output variables. Input and output together decide prediction direction.

In order to gain the optimal linear relationship between hop count and actual distance, optimal estimate value $\overline{\square}$ of β is needed for the equation. Therefore, $\|e\|^2\rightarrow\min$ is necessary. At this time, $H^T H\overline{\square}=H^T D$. Equation (4) is expressed below with continuous regression model:

$$D=T_h\zeta+e \tag{5}$$

where T_h is compositional vector matrix and h is quantity of hidden variables. $T_k=HW_k$ and $w_h=\{w_1,w_2,\dots,w_h\}$. Selection criteria for continuous regression weight vector $w_i(i=1,2,\dots,h)$ is as follows:

$$w_i = \left\{ Cov\left(H\alpha, \mathrm{D}\right)^2 Var\left(H\alpha\right)^{\frac{\delta}{1-\delta}-1} \right\} \to \max \quad (6)$$

This meets the following constraint condition. For any $j<i$, there is:

$$\|w_i\| = 1 \qquad Cov\left(Hw_i, Hw_j\right) = 0 \quad (7)$$

In Equation (6): $0<\delta<1$, *Dov* represents the sample covariance estimate and *Var* represents the sample variance estimate. The final purpose of the continuous regression model is to gain regression coefficient β. Parameter ξ in modular form (5) is estimated by total least square. There is:

$$\zeta_{\delta,\ h} = \left(T_h^T T_h\right)^{-1} T_h^T D \quad (8)$$

For the given δ and h, there is

$$\overline{D}_{\delta,h} = HW_h \zeta_{\delta,h} \quad (9)$$

Regression coefficient β will be determined by Equation (10)

$$\beta_{\delta,\ h} = W_h \left(T_h^T T_h\right)^{-1} T_h^T D \quad (10)$$

After estimated item β is substituted into Equation (4), the prediction model will be obtained. After the hop count from an unknown node to a known node is input into the equation, the corresponding estimated distance will be obtained. The least square method will be used to estimate unknown node coordinate according to known node coordinate and the estimated distance to obtain the estimated coordinate.

B. ML-CR algorithm

The ML-CR algorithm proposed in the paper is comprised of two stages: training and positioning.

(1) Training stage

In order to eliminate the dimension difference between hop count and Euclidean distance units, the algorithm conducts standardization treatment for *H* and *D* (including minus mean value and initial standard deviation). It is assumed that $\overline{H}$ and $\overline{D}$ are matrices of *H* and *D* after standardization respectively. After $\overline{H}$ and $\overline{D}$ undergo continuous regression, the cross inspection method will be used to gain the optimal parameters δ and *h* and corresponding regression coefficient $\overline{\beta}_{\delta}$,*h*. Therefore:

$$\beta_{\delta,\ h} = W_h \left(T_h^T T_h\right)^{-1} T_h^T D = \left(\left(\overline{H}W_h\right)^T \left(\overline{H}W_h\right)\right)^{-1} \left(\overline{H}W_h\right)^T \overline{D} \quad (11)$$

Now, m training models can be obtained from the sensor network. $\beta_{train}=\{\beta_1,\beta_2,\ldots,\beta_m\}$ and they will be broadcast to each node *Sp* in the network[17].

(2) Positioning stage

Each unknown node S_j uses its hop count matrix H_{test} to beacon node and previous training model β_{train} to predict its Euclidean distance D_{pred} from the unknown node, namely:

$$D_{pred} = \overline{H}_{test}\beta_{train} + repmat\left(\overline{h}, n\right) \quad (12)$$

Where $\overline{H}_{test}$ is matrix of H_{test} after standardization treatment; $\overline{h}$ is mean value of *H* and *repmat*($\overline{h}$,*n*) is accumulation of n rows of $\overline{h}$. H_{test}

C. Network model parameter

For the positioning algorithm simulation of the whole wireless sensor network, the following network parameters must be defined:

(1) Sensor node coordinate system

As wireless sensor network nodes are distributed in a two-dimensional plane in the experiment, the two-dimensional coordinate system is also be used during simulation and the expressed coordinate will be considered the absolute coordinate. For example, coordinates of node x and y are 2m and 6m respectively.

(2) Node distribution density

Node distribution density P refers to the number of nodes in a given unit area. If there are n sensor nodes in a place with an area of S, the node distribution density of that place is P=S/n.

(3) Network connectivity

The number of nodes that can communicate directly with other nodes in the network is called network connectivity; namely, the number of neighboring nodes with a single hop of a node. The value of network connectivity is based on node density and radio frequency communication radius R of the node. An increase in node density or radio frequency communication radius of a node will increase network connectivity.

(4) Beacon node density

Beacon node density refers to ratio between the number of beacon nodes in the network and total node number.

(5) Measurement method for distances between nodes

In actual positioning, distances between nodes are measured by the RSSI method and there is a measurement error. The estimation model is used in the simulation experiment to estimate distances between nodes in this paper.

Apart from the basic parameters above, the following problems will be resolved for the establishment of the sensor network and related simulations:

(1) Area s of node distribution area, number of node n and communication radius R for the radio frequency module of a node will be preset, as these parameters determine node density and network parameters such as sensor network connectivity and beacon node density.

(2) Position coordinates of unknown nodes are generated randomly and special consideration is needed for beacon nodes. The trilateration-based positioning algorithm used in this paper requires the presence of at least three beacon nodes during positioning of unknown nodes. The positions of four beacon nodes are generated according to this requirement and their positions are fixed and distributed around the network. Other beacon nodes generated randomly are distributed among these four beacon nodes.

(3) If the estimated distance between nodes B and A is less than the radio frequency communication radius R of node A, then node B is determined to be a single-hop adjacent node of node A.

As the wireless sensor network is distributed randomly and intensively, it is suitable for the adverse environment of a battlefield. In the military field, a wireless sensor network can be used to gather intelligence about the enemy and monitor the deployment of partners. It can also be used to monitor equipment, materials and the battlefield itself to ascertain any threat of an attack using biological and/or chemical weapons. The military can place a large quantity of inexpensive micro sensor nodes in the area-of-interest from airplanes so as to monitor any changes in the surrounding environment in real time through these sensor nodes. They can also send monitoring data to the command monitoring center through satellite channels or a ground base station.

(4) Each sensor node is calculated in parallel in practical application and several positioning coordinate values will be sent to a certain aggregation node at the same time. Therefore, each node is calculated according to the sequence in the global positioning module. The performance of the simulation positioning algorithm is measured by the following indexes.

IV. SIMULATION AND ANALYSIS

A. *Algorithm complexity analysis*

The complexity of ML-CR method is comprised of communication and computation. Similar to the communication process such as distance vector- light-flooding algorithm, fuzzy algorithm, close distance-mapping algorithm, positioning algorithm based on support vector regression and ridge regression algorithm, each node needs the flooding method to calculate hop count between nodes. Therefore, their communication overheads are the same. Communication overheads of six different methods are about $o(n^2m)$. N is the number of nodes and m is the number of beacon nodes. After each unknown node determines its skip distance from the beacon node, DV-hop and Amorphous method will use the least square method to estimate the unknown position, which requires computation overhead of $o(m^3)$. The PDM method cuts off singular value decomposition to dispose of data, which requires computation overhead of $o(m^3)$, and then it will operate with least squares. Therefore, computation overhead of the PDM method is higher than that of the DV-hop and Amorphous methods. Prompt information acquisition and response in modern warfare are increasingly of great importance, and sensor networks can fill this need.

LMVR adopts the regression method based on support vector machine, and a quadratic programming problem must be resolved in the SVM solution procedure. Its computation complexity is between $o(m^2)\approx o(m^3)$ generally according to the optimization method. However, as regression based on traditional SVM is multi-input single-output, for m training samples, the computation cost of its modeling is $o(m^3)\approx o(m^4)$. In addition, the selection of nuclear parameter, penalty coefficient and insensitive loss function width also require computation cost, while achieving the optimum is difficult. RR adopts the ridge regression method which is similar to least squares. Therefore, after the ridge parameter is selected, computation overhead is $o(m^3)$. If the method in reference [2] is adopted, the computation overhead can be reduced to $o(m^2\log m)$. The computation overhead of the ML-CR method proposed in this paper is similar to that of the RP method. There will be additional overhead after optimal parameter δ and h are obtained and the total computation overhead is $o(m^3)$. Table 1 lists communication and computation overhead complexity of six types of positioning algorithms.

Hundreds of sensor nodes are deployed in a large scale wireless sensor network, and it is difficult to actualize a real network under current experiment conditions. Therefore, the MATLAB simulation method is adopted in this paper to evaluate the advantages and disadvantages of the positioning algorithm. The experiment area is set as 1000×1000. Uniform distribution and random distribution of nodes are used to analyze the influence of network topology anisotropies on algorithm performance. Three types of network topology structures are set in the experiment: C-shape, X-shape and Z-shape structures. Positioning accuracy is generally defined as the ratio between the error value and the node communication radius. For example, a positioning accuracy of 10% means that the positioning error is equal to 10% of the node communication radius.

The distributed positioning method is also used to divide the positioning area into grids, and the positioning accuracy depends on the size of the grid, such as radar and positioning method based on compressed sensing. In order to evaluate the positioning accuracy of the entire network, average positioning error is generally used. It is defined as the ratio of the average error of Euclidean distance between the estimated position and the real position of the unknown node. Indexes to evaluate the positioning error, maximum error and minimum error are considered and compared in this paper.

$$\begin{aligned} ALE &= 1\Big/ N_t R \sum_{i=1}^{N_t} \sqrt{\left(\bar{x}_i - x_i\right)^2 + \left(\bar{y}_i - y_i\right)^2} \\ MAE &= \max\left(1\Big/ N_t R \sum_{i=1}^{N_t} \sqrt{\left(\bar{x}_i - x_i\right)^2 + \left(\bar{y}_i - y_i\right)^2}\right), i = 1,\ldots N_t \\ MIE &= \min\left(1\Big/ N_t R \sum_{i=1}^{N_t} \sqrt{\left(\bar{x}_i - x_i\right)^2 + \left(\bar{y}_i - y_i\right)^2}\right), i = 1,\ldots N_t \end{aligned} \tag{13}$$

TABLE I. SIX KINDS OF LOCALIZATION ALGORITHM OF COMMUNICATION AND COMPUTING COMPLEXITY OVERHEAD

Localization algorithm	**Communication complexity**	**Computational complexity**
DV-hop	O(n2m)	O(m3)
Amorphous	O(n2m)	O(m3)
PDM	O(n2m)	O(m2)+O(m3)
LSVR	O(n2m)	O(m2)- O(m3)
RR	O(n2m)	O(m2logm)
ML-CR	O(n2m)	O(m3)

B. *Uniform node deployment*

200 nodes are deployed uniformly in the experimental area in total and the distance is set between nodes at 50 and the node communication radius R=150, 200, 250, 300 and 350. The positioning result of the C-shape wireless sensor network of the beacon node 20 is shown in Table 2.

The positioning result of the C-shape when the communication radius is R=200, and the number of beacon nodes M=30, 40, 50, 60 and 70 is shown in Table 3. The positioning results of the six algorithms in uniformly deployed networks of C-shape, X-shape and Z-shape when the communication radius is R=250 and the number of beacon nodes M=20 are shown in Table 4.

It can be seen from the simulation result that in regular node deployment, the ML-CR algorithm put forward in this paper has good positioning accuracy. When the communication radius is R=250 and the number of beacon nodes is M=20, the positioning accuracy is improved about 60% compared with the traditional multi-hop positioning algorithm DV-hop.

C. Random node deployment

500 nodes are deployed in the experimental area at random and the distance is set between nodes at 50. The positioning result of the C-shape when the number of beacon nodes is M=100 and the communication radius of node is R=150, 200, 250, 300 and 350 is shown in Table 4. The positioning result of C-shape when the communication radius is R=200 and the number of beacon nodes is M=30, 40, 50, 60 and 70 is shown in Table 5.

It can be seen from the simulation result shown in Table 5 that with random node deployment, the positioning algorithm based on the continuous regression model put forward in this paper has good positioning accuracy as well. When the communication radius of a node is R=250 and the number of beacon nodes is M=100, the positioning accuracy is improved about 63%, compared with the traditional multi-hop positioning method DV-hop.

The range-free positioning method is often used for positioning of large scale wireless sensor networks. Most of range-free positioning methods obtain the hop count between nodes through the connection between the nodes and estimating the location of unknown nodes with the average hop distance and hop count. Because the method is easy to be realized and has relatively low hardware requirements, it is quite suitable for the positioning requirements of large scale wireless sensor networks. However, the positioning accuracy of range-free multi-hop positioning technology is greatly affected by the network topology. In the isotropic network with high node density and even distribution, the ideal positioning results can be achieved. However, in the network environment, with anisotropy such as uneven and sparse distribution of nodes, the positioning effect is poor.

TABLE II.
VARIOUS ALGORITHMS POSITIONING ERROR INDEX CHANGING WITH COMMUNICATION RADIUS R (M=20)

Localization algorithm	Error indicators	R=150	R=200	R=250	R=300	R=350
DV-hop	ALE	2.3021	1.2021	1.5245	0.7415	0.4215
	MAE	1.7254	1.5212	1.2652	0.5845	0.2514
	MIE	2.0125	1.2632	1.3265	0.6525	1.0254
Amorphous	ALE	0.2351	2.3252	2.0154	2.0147	0.2658
	MAE	0.3652	2.3021	2.0214	2.5265	0.3202
	MIE	0.5241	0.2514	2.3201	2.0124	0.6252
PDM	ALE	0.5241	0.7541	1.3252	1.6585	0.3254
	MAE	0.5695	0.6215	1.0214	1.2652	1.3625
	MIE	0.5454	0.8514	1.5245	1.6251	2.0125
LSVR	ALE	0.3625	0.5021	2.3625	1.6021	0.9526
	MAE	0.8541	0.6254	2.1542	1.0485	0.3652
	MIE	0.5265	0.2154	2.0126	1.2653	0.5141
RR	ALE	0.6584	0.6584	0.6958	0.3526	0.4512
	MAE	0.5124	0.9545	0.5487	0.3145	0.5252
	MIE	0.4215	0.5241	0.5254	0.2514	0.3261
ML-CR	ALE	0.4112	0.5144	0.6585	0.3652	0.4152
	MAE	0.4854	0.6524	0.6459	0.9852	0.2012
	MIE	0.4856	0.3626	0.4852	0.3658	0.8585

TABLE III.
DIFFERENT ALGORITHMS POSITIONING ERROR INDEX CHANGING WITH COMMUNICATION RADIUS R (R=20)

Localization algorithm	Error indicators	R=150	R=200	R=250	R=300	R=350
DV-hop	ALE	1.2022	1.8545	1.5425	1.3265	2.0124
	MAE	1.3625	1.5265	1.3253	1.3255	2.2123
	MIE	1.9585	1.8452	2.0121	1.2522	0.5423
Amorphous	ALE	1.9025	1.7142	0.9525	0.6855	0.6251
	MAE	2.2012	1.2515	0.4582	0.7896	0.3625
	MIE	1.6325	1.6250	0.4956	0.8512	0.4715
PDM	ALE	0.6625	0.9525	0.2565	2.3625	0.8516
	MAE	0.7415	0.5236	0.7825	1.3658	0.3652
	MIE	0.5265	0.7415	0.9201	2.0123	0.1246
LSVR	ALE	0.5215	0.8526	0.6582	1.3625	0.4102
	MAE	0.5201	0.5265	0.4852	1.0254	0.8203
	MIE	0.6585	0.5236	0.6958	1.3625	0.5025
RR	ALE	0.4154	0.4585	0.1545	1.2012	0.3625
	MAE	0.4568	0.6585	0.1589	0.4151	0.5142
	MIE	0.4747	0.3652	1.8475	0.2514	0.3625
ML-CR	ALE	0.4150	0.4584	1.3652	0.3625	0.1545
	MAE	0.5485	0.6525	1.0256	0.4569	0.8523
	MIE	0.4525	0.4152	0.4858	0.4578	1.2523

TABLE IV.
VARIOUS ALGORITHMS POSITIONING ERROR INDEX CHANGING WITH COMMUNICATION RADIUS R (M=20)

Localization algorithm	Error indicators	R=150	R=200	R=250	R=300	R=350
DV-hop	ALE	2.0120	1.2012	1.0236	0.7415	0.4525
	MAE	2.3262	1.4520	0.3625	1.2525	1.2021
	MIE	1.7452	1.2002	1.0251	0.6252	0.6252
Amorphous	ALE	2.0325	2.0121	2.3602	1.5021	0.2012
	MAE	2.3620	2.3021	2.0325	2.3602	1.0236
	MIE	2.0212	2.1020	0.6625	0.4252	0.3625
PDM	ALE	0.3652	0.3625	0.5626	0.4582	0.1026
	MAE	0.8585	0.1542	0.6925	2.0358	0.2323
	MIE	0.4546	0.6253	0.7454	1.3602	0.2514
LSVR	ALE	0.5156	0.5212	0.7525	0.2025	0.2925
	MAE	0.7445	0.5236	0.5656	0.4858	0.2360
	MIE	0.5012	0.5145	0.6925	0.6256	0.2545
RR	ALE	0.6256	0.6202	0.6252	0.6564	0.3266
	MAE	0.1526	0.4251	0.5012	0.8545	0.1212
	MIE	0.4152	0.5514	0.5412	0.6933	0.1745
ML-CR	ALE	0.4695	0.4201	0.3232	0.3656	0.3656
	MAE	0.4825	0.5236	0.3025	0.5151	0.4125
	MIE	0.4152	0.8021	0.4152	0.3626	0.2515

TABLE V.
VARIOUS ALGORITHMS POSITIONING ERROR INDEX CHANGING WITH COMMUNICATION RADIUS R (M=20)

Localization algorithm	Error indicators	R=150	R=200	R=250	R=300	R=350
DV-hop	ALE	2.0212	0.5454	1.0212	0.7444	0.4584
	MAE	0.3625	0.6251	1.6253	1.2031	1.0255
	MIE	2.3652	0.2541	0.3652	0.4512	0.2512
Amorphous	ALE	2.0121	0.6255	1.5152	1.2021	0.6251
	MAE	2.1212	0.4152	1.2635	2.0121	1.3205
	MIE	1.0362	0.2012	2.0112	1.3021	0.6323
PDM	ALE	1.0254	0.5252	0.7141	2.031	0.3232
	MAE	0.8525	1.3602	0.8565	0.4515	0.2012
	MIE	0.6565	1.6251	0.1023	0.3653	0.2111
LSVR	ALE	0.3695	1.5121	0.3020	0.5141	0.1251
	MAE	0.4584	1.2020	0.4012	0.5454	0.3622
	MIE	0.4512	1.2515	0.3289	0.2626	0.3262
RR	ALE	0.4521	2.0321	0.3602	0.5452	0.3625
	MAE	0.4878	0.6252	0.3625	0.3665	0.4545
	MIE	0.8525	0.4522	0.5252	0.6262	0.3625
ML-CR	ALE	0.5656	0.4858	0.4152	0.3626	0.5151
	MAE	0.6454	0.4154	0.3626	0.5252	0.5252
	MIE	0.6254	0.4859	0.5151	0.1456	0.4151

V. CONCLUSION

A wireless sensor network multi-hop positioning algorithm based on continuous regression is put forward in this paper. The method defines the positioning problem as a regression problem, and it effectively eliminates the impact of network anisotropy on positioning performance. The computing complexity of the method is less than that of the traditional method. The experiment result shows that compared with the traditional DV-hop algorithm, the positioning accuracy of the method put forward in this paper has been greatly improved, but the ML-CR algorithm requires a certain sufficient number of known nodes in the training stage to improve accuracy. In addition, the model in this paper will be further optimized.

The traditional multi-hop positioning algorithm is easily affected by the network anisotropy, thus resulting in unstable positioning performance, and the wireless sensor network multi-hop positioning algorithm based on continuous regression is put forwarded in the paper to address this problem. By utilizing the continuous regression model, the mapping relationship between the hop count and Euclidean distance is constructed so as to transform the positioning process model into regression prediction. Theoretical analysis and simulation results show that the improved algorithm improves the positioning accuracy, and it avoids the influence of the network topology anisotropy on the performance of the algorithm. The algorithm requires little expenditure and few parameters so it can be adapted to wireless sensor networks with irregular nodes distribution, and can be of great engineering application value.

REFERENCES

[1] J. Zhipeng, Z. Sanguo, "Reliability evaluation of wireless sensor networks using logistic regression," *2010 WRI International Conference on Communications and Mobile Computing*, CMC 2010, Vol. 3, pp. 334-338, April 2010.

[2] C. Aditi, V. Palaniandavar, "An efficient statistical approach for time synchronization in wireless sensor networks," *International*

Journal of Communication Systems, Vol. 29, pp. 722-733, March 2016. http://dx.doi.org/10.1002/dac.2944

[3] J.B. Predd, S.R. Kulkarni, H.V. Poor, "Distributed kernel regr ession: An algorithm for training collaboratively," *2006 IEEE Information Theory Workshop*, ITW 2006, pp. 332-336. http://dx.doi.org/10.1109/ITW.2006.1633840

[4] C. Jongyoo, "Development and evaluation of an ambulatory stress monitor based on wearable sensors," *IEEE Transactions on Information Technology in Biomedicine*, Vol. 16, pp. 279-286, March 2012. http://dx.doi.org/10.1109/TITB.2011.2169804

[5] W. Shuai, Z.Y. Sheng. "The simulation of mobile robots pursuit collision problem in wireless sensor network," *Proceedings of the 27th Chinese Control Conference*, pp. 190-193, April 2008.

[6] O.E.K. Aktouf, I. Parissis, "SMART service for fault diagnosis in wireless sensor networks," *Proceedings - 6th International Conference on Next Generation Mobile Applications*. NGMAST 2012, pp. 211-216.

[7] I. Aydin, M. Karaköse, "Wireless sensor network based fault diagnosis approaches," *2013 21st Signal Processing and Communications Applications Conference*, SIU 2013, pp. 201-211.

[8] S.H. Chang, "A causal model method for fault diagnosis in wireless sensor networks," *Proceedings-10th IEEE International Conference on Computer and Information Technology*, CIT-2010, pp. 155-162. http://dx.doi.org/10.1109/cit.2010.65

[9] C.L. He, "Technology research on the fault diagnosis of wireless sensor network system," *Advanced Materials Research*, vol. 84, pp. 442-445, April 2014.

[10] X.H. Jin, T.W.S. Chow, "Kuiper test and autoregressive model-based approach for wireless sensor network fault diagnosis," *Wireless Networks*, vol. 21, pp. 829-839, May 2015. http://dx.doi.org/10.1007/s11276-014-0820-0

[11] P.M. Khilar, "Intermittent fault diagnosis in wireless sensor networks," *Proceedings-10th International Conference on Information Technology*, ICIT 2007, pp. 145-147. http://dx.doi.org/10.1109/icit.2007.15

[12] Q. Li, "Wireless sensor network fault diagnosis method of optimization research and simulation," *Applied Mechanics and Materials*, vol. 347, pp. 955-959, May 2009.

[13] R.F. Liu, "Fault diagnosis of wireless sensor based on ACO-RBF neural network," *Proceedings-2010 3rd IEEE International Conference on Computer Science and Information Technology*, ICCSIT 2010, pp. 248-251.

[14] T.Y. Li, "Fault diagnosis of auxiliaries in power plants based on wireless sensor networks with vibration transducer," *Proceedings-2010 2nd IEEE International Conference on Network Infrastructure and Digital Content*, IC-NIDC 2010, pp. 732-736. http://dx.doi.org/10.1109/icnidc.2010.5657877

[15] B. Tang, B. Deng, L. Deng, "Mechanical fault diagnosis method based on multi-level fusion in wireless sensor networks," Zhendong Ceshi Yu Zhenduan/Journal of Vibration, *Measurement and Diagnosis*, vol. 36, pp. 92-96, April 2016.

[16] B. Steffen, L. Cyril, "Rollout Algorithms for Wireless Sensor Network-Assisted Target Search," *IEEE Sensors Journal*, Vol. 15, pp. 3835-3845, July 2015. http://dx.doi.org/10.1109/JSEN.2015.2393893

[17] K. Sanjeev, V. Poor, "Regression in sensor networks: Training distributively with alternating projections," *Proceedings of SPIE - The International Society for Optical Engineering*, Vol. 5910, pp. 1-15, April 2005.

AUTHORS

Weimin Han is an associate professor with Department of Wind engineering, Hunan Vocational institute of Technology, Xiangtan 411000, China. Her research interests include wind power engineering, information engineering and automation. (e-mail: B12090018@hnu.edu.cn).

Shijun Li is a doctoral student of College of Electrical and Information Engineering, Hunan University, Xiangtan 411101, China. He is mainly engaged in the study of computer control technology. He is also with Hunan province Cooperative Innovation Center for Wind Power Equipment and Energy Conversion, Xiangtan 411101, China (email: lishijun820404@qq.com).

Weidong Li is a master of Powerchina Zhongnan Engineering Co. Ltd., Changsha 410014, China. He is mainly engaged in communication technology research. (e-mail: Lsj181627@qq.com).

Submitted 03 September 2016 Published as resubmitted by the authors 17 October 2016.