Research Article

Toward Collinearity-Avoidable Localization for Wireless Sensor Network

Xiaoyong Yan,[1,2] Aiguo Song,[2] Jimin Yu,[1] and Zhong Yang[1]

[1]*School of Intelligence Science and Control Engineering, Jinling Institute of Technology, Nanjing 211169, China*
[2]*School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China*

Correspondence should be addressed to Aiguo Song; a.g.song@seu.edu.cn

In accordance with the collinearity problem during computation caused by the beacon nodes used for location estimation which are close to be in the same line or same plane, two solutions are proposed in this paper: the geometric analytical localization algorithm based on positioning units and the localization algorithm based on the multivariate analysis method. The geometric analytical localization algorithm based on positioning units analyzes the topology quality of positioning units used to estimate location and provides quantitative criteria based on that; the localization algorithm based on the multivariate analysis method uses the multivariate analysis method to filter and integrate the beacon nodes coordinate matrixes during the process of location estimation. Both methods can avoid low estimation accuracy and instability caused by multicollinearity.

1. Introduction

Wireless sensor network (WSN) [1–3] refers to a sort of wireless network comprised of large amounts of static or mobile sensor network nodes in forms of self-organization and multihop. The aim of such network is to cooperatively detect, process, and transmit targets' monitoring information within the coverage area of the network, as well as report the information to users. As a new pattern of obtaining information, it possesses many advantages such as low cost, easy deployment, self-organization, and flexibility, so it has been widely applied in various domains, such as national defense and military affairs [4], environment inspection [5], traffic management [6], long-distance control of dangerous region [7], and so on. WSN has shown its significance and capability in application.

In many application problems related to sensor network, location information of nodes is of great importance to the monitoring activity of the whole network, which plays a critical role in many applications. Monitoring data without nodes' location information is often of no use. 80% of information provided by sensor nodes to users related with the monitored area is connected with location [8].

Generally, the localization process can be roughly divided into two phases [9, 10]: distance measurement phase and location estimation phase. In distance measurement phase, sensor nodes communicate with neighbors to estimate distance between pairs of devices. In location estimation phase, a localization method is used based on previous estimated distance, and sensor nodes can finally estimate their physical locations in the form of coordinates. "Zero error" is the eternal pursuit of localization algorithm. Owing to the limited computing capacity of sensor and complexity in the network environment, each stage would generate some errors that have significant influence on the final coordinate estimation. Therefore, the final estimated locations of the unknown nodes are mainly affected by the distance measurement between the nodes and the relative location of reference nodes. Most researchers carried out studies for the measurement accuracy [11–13] and have achieved some results, especially in recent years, for further development of robust estimation, making use of this advanced technology to the localization accuracy and algorithm design of localization mechanism. However, when it is used as the reference nodes for location estimation, in other words, when the topological shape between beacon nodes is collinear or approximately collinear, that is, there

is multicollinearity [14, 15], the localization accuracy of surrounding unknown nodes is poor, which can even reduce the localization accuracy of the whole monitoring area. At present, most localization methods conduct research in accordance with the ranging error during the localization process, and seldom consideration had been given to the impact of beacon nodes on localization accuracy.

The research object of this paper is the impact of the relative location between beacon nodes on localization accuracy, and the discussion is divided into two parts: the first part starts from analysis of the topology quality of beacon nodes; firstly, the multicollinearity problem caused by poor localization units under two-dimensional and three-dimensional localization environment will be analyzed. Then, the shape of poor localization units will be provided, which will also be quantified, the quality determination standard will be provided, and last but not the least, the impact of multicollinearity problem caused by collinear or approximately collinear localization units on the localization accuracy will be solved; the second part starts from the coordinate matrix of localization units; the dimensionality reduction method in multivariate analysis is used to reconstruct the beacon nodes used in location estimation, and by eliminating the data with low SNR (signal to noise ratio), the noise is reduced and the data with multicollinearity is eliminated.

The rest of the paper is organized as follows. In Section 2, we analyze two-dimensional and three-dimensional localization unit topological qualities. In Section 3, we formulate a series of newly developed localization algorithms based on the geometry analysis of the localization unit and describe these WSN localization algorithms. In Section 4, we make use of multivariate analysis to analyze the localization unit and formulate our localization algorithm based on multivariate analysis. In Section 5, we give justification of the applicability and effectiveness of our approaches in WSN location. Section 6 gives the conclusion.

2. Topological Analysis of Beacon Nodes

In general, the more beacon nodes are chosen by the unknown node, the more accurate the estimated location is [9, 16]. But, in fact, the topology of beacon nodes and the topological structure formed between beacon nodes and the unknown node will greatly affect unknown node's estimation result. Localization algorithm in the two-dimensional space requires a minimum of three reference nodes within the whole operational field [16]; however, in a three-dimensional space, due to increase of one dimensionality, it requires at least four beacon nodes to estimate the location of an unknown node [16]. Without loss of generality, a localization unit (LU) is defined as a beacon node group which can determine at least one unknown node and directly affect the final localization result. In a two-dimensional plane, when the distance between each beacon node and the unknown node in a LU is calculated, the trilateration or multilateration can be used to determine the location of the unknown node. However, during calculation of the distance between the unknown node and beacon node, there generally exists certain error, which makes the three circles fail to meet at
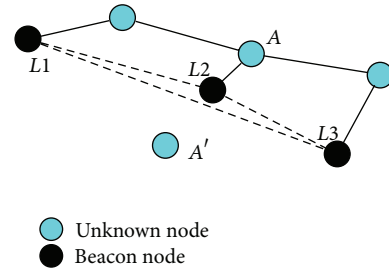

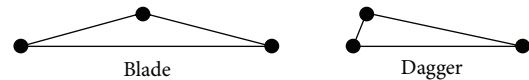Figure 1: Collinearity of beacon nodes.

Figure 2: The poor-quality of LU in two-dimensional space.

one point during the trilateral positioning, so the estimation method should be used to determine the location of unknown node. When three beacon nodes spread approximated on a straight line, that is, the three nodes are of collinearity, the location of unknown nodes may not be able to be estimated with ordinary least-squares estimation method, and the error rate can be up to 200% [17, 18]. See Figure 1. The distance between node A and each beacon node $L1$, $L2$, and $L3$ has been given. As the three beacon nodes are almost on a straight line, that is, the three nodes are approximately of collinearity, the location of unknown node A may be A or A'. As for this, A's practical location cannot be figured out under such circumstance.

Because in a two-dimensional space, the three beacon nodes that constitute the LU form a triangle, there are two kinds of collinearity phenomena: when the three points of triangle are in a straight line, the location process has complete collinearity, which is rare in reality; more often, approximate collinearity tends to occur; the LU with approximate collinearity has at least one small angle, and the formed triangle has a big aspect ratio. In other words, the triangle has at least one small angle, and three vertices of triangle are approximately collinear. It is also easy for us to know there are two types of triangle (as shown in Figure 2): one type has no short edge and is called blade; the other type has one short edge and is called dagger [19, 20].

During three-dimensional location estimation, if a certain beacon node is known, and the distance between the undetermined unknown node and this node can be observed, then this undetermined track is a sphere. In order to determine the location of the undetermined node, its distances to at least four known nodes should be determined; create four positioning spheres with the four known nodes as the centers of sphere and with the observed four distances as radiuses, two spheres can intersect at one space curve, and four spheres can intersect at one point. Because it requires at least four beacon nodes to conduct three-dimensional localization, so a tetrahedron determines the LU of three-dimensional localization. Similar to the two-dimensional space, the topology quality of tetrahedron also affects the localization accuracy
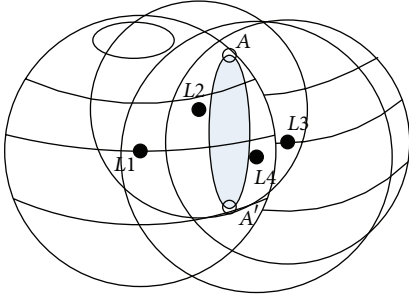
FIGURE 3: Four coplanar beacons in three-dimensional space.

of unknown node with it as the reference. Similarly, the distance error is inevitable in an actual environment, which causes that the four spheres in the localization not necessarily have an intersection point. When the relative locations of four beacon nodes are approximately coplanar, four spheres have two intersection points, so it is difficult to estimate the location of unknown node. As shown in Figure 3, if the geometric distributions of beacon nodes $L1$, $L2$, $L3$, and $L4$ are completely coplanar, when using the traditional location estimation method, unknown node's estimated coordinate may be A or A', in which case node A's physical coordinate cannot be estimated and the errors also will be up to 200%. If the deployment environment noise is large, the estimated location of the unknown node A will be far away from its true location.

It is generally believed that tetrahedron is the expansion of triangle in the three-dimensional space, and therefore, there are two situations in which collinearity phenomenon occurs: when the volume of tetrahedron is close to zero, it is equal to occurrence of approximately collinear phenomenon; when the volume of tetrahedron is zero, it is equal to occurrence of complete collinear phenomenon. The researchers found that when tetrahedral volume is equal to zero or tends to zero, the composition of the tetrahedron triangle always contains one or more triangles with a large aspect ratio [19, 21]. Cheng et al. [19, 22] made detailed study on the tetrahedrons, proposing nine kinds of poor-quality tetrahedron whose structures are shown in Figure 4.

3. Localization Method Based on the Geometrical Analysis of LU

3.1. Geometrical Analysis of Two-Dimensional LU. In order to solve the impact of collinearity phenomenon on localization accuracy in the two-dimensional space, the researchers have proposed multiple solutions based on analysis of LU in different scenarios. In accordance with the fact that the LU has a triangular form in the two-dimensional space, Poggi and Mazzini [23] proposed the concept of collinearity (also called degree of collinearity, DC). They used the smallest value of the three heights of triangle as the DC's parameter of a triangle and used it to measure the topology quality of LU; the more the three beacon nodes that constitute the LU are close to be collinear, the lower its DC is, and otherwise the higher the DC is. Their experimental result also shows

that the more the LU are close to be collinear (i.e., low DC), the higher the location error of unknown node is, which might even be nonlocalized; the more the LU are close to equilateral triangle, the higher the localization accuracy of unknown nodes is. Later, Wu et al. [24, 25] proposed another standard of DC, that is, the biggest cosine value of the interior angles of triangle. Similarly, for the DC by using the method proposed by them the lower the value is, the worse the estimation result is; the higher the DC is, that is, the more the LU are close to equilateral triangle, the better the positioning result is. The literature [25] also provides another definition of DC: assuming the longest side length of the triangle formed by three random points in the plane is $l_{\max}$, the corresponding height of this length is $h_{\min}$, the ratio $2\sqrt{3}/3$ between $h_{\min}$ and $l_{\max}$ is defined as the collinearity of this triangle, and when the three nodes are collinear, the collinearity is 0. In this way, the value range of DC is $[0, 1]$, and the smaller the collinearity is, the closer to be collinear these three nodes are. The several DC determination methods mentioned above are actually used to measure the quality of triangle unit. Many years ago, researchers [20] had studied in detail the measurement criterion for the quality of triangle unit and provided various scientific judgment and quality evaluation methods from different perspectives. They believed that the measurement criterion for the quality of triangle positioning unit should satisfy the following principle: the translation, rotation, inverse, reflection, and uniform scaling of triangle unit should change its measured value; when and only when the triangle is an equilateral triangle, use the biggest measured value; when the triangle area is close to zero, its measured value is also close to zero. Based on the standard mentioned above, the researchers provided various methods to determine the topology quality of triangle: the smallest angle measurement method; the longest and shortest side measurement method; area-side length measurement method; inner and external radius measurement method; inner radius-shortest side measurement method; shortest height-longest side measurement method. The quality determination method mentioned above has the following definition formula:

(1) the smallest angle measurement method

$$q_{\alpha\min} = \frac{3\alpha_{\min}}{\pi}, \tag{1}$$

(2) the longest and shortest side measurement method

$$q_{Ll} = \frac{l_{\min}}{l_{\max}}, \tag{2}$$

(3) area-side length measurement method

$$q_{ALS} = \frac{4\sqrt{3}A}{l_1^2 + l_2^2 + l_3^2}, \tag{3}$$

(4) inner and external radius measurement method
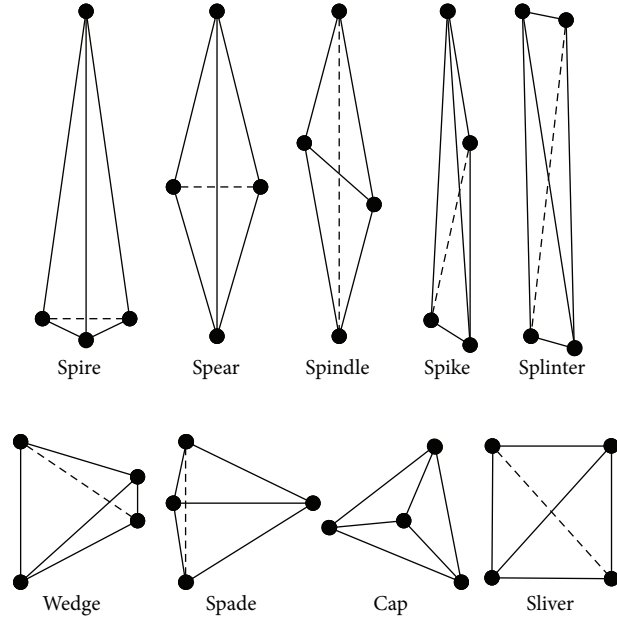
$$q_{Rr} = \frac{2r}{R}, \tag{4}$$

FIGURE 4: The poor-quality of LU in three-dimensional space.

(5) inner radius-shortest side measurement method

$$q_{Lr} = \frac{2\sqrt{3}r}{l_{\max}},$$ (5)

(6) shortest height-longest side measurement method

$$q_{Lh} = \frac{2h_{\min}}{\sqrt{3}l_{\max}},$$ (6)

where $\alpha_{\min}$ is the smallest inner angle; $l_{\min}$ and $l_{\max}$ are the length of the shortest and longest edge, respectively; l_1, l_2, and l_3 are the length of the three sides of the triangle; A is the area of the triangular element; r is the inradius of a triangle; R is the circumradius of a triangle; and $h_{\min}$ is the minimum height of the triangle.

Literatures also demonstrate that the formulas above are equivalent [20, 26]; the formulas all tend to zero in the case that the triangles mesh's area tends to zero; metric formula value tends to one in the case that the triangles mesh tends to an equilateral triangle mesh.

3.2. 2D Localization Algorithm Based on Determination of the Geometrical Shape of LU. The unknown node obtains its "distance" to the beacon nodes through various methods such as RSSI, ToA, TDoA, and AoA, relative distance or skip distance, and after it has communicated with more than three surrounding beacon nodes, the unknown node can use trilateration or multilateration to estimate the unknown node. Because the quality of LU has a huge impact on the final estimation result, during the estimation process, the quality of LU should be determined, and the six criterions mentioned in the previous section are approximately equivalent. Therefore, by refereeing to the DC determination criterion, the concept

of degree of multicollinearity (DM) is proposed in this paper as the measurement criterion for LU quality and then develops its corresponding novel location algorithm in the two-dimensional space, called two-dimensional location estimation-shape analysis (2D LE-SA). Assume there are n nodes in total in the monitoring area; their actual coordinate is $\{\mathbf{x}_i\}_{i=1}^n$, the first m are beacon nodes, and their coordinates are known; after obtaining the distance matrix $\mathbf{D}$ between nodes, see Algorithm 1 for the location estimation method of unknown node.

3.3. Geometrical Analysis of Three-Dimensional LU. For location estimation in the three-dimensional space, Zhou et al. [27, 28] proposed the optimized selection principle of beacon nodes, and by establishing the error area of four beacon nodes, it could improve the localization accuracy and provide better localization service through certain distribution of the four beacon nodes. However, during the solving process, the tangent plane was used to replace the sphere for signal propagation, which caused the idea that the signal could not reach certain part in the built model, and the obtained solution needs further discussion. In addition, there is another three-dimensional localization algorithm based on elevation-type reference nodes [29]. By installing each beacon node is equipped with an omnidirectional antenna that can be raised up and lowered down when needed. Using its movable antenna makes it possible for the beacon nodes to send signal at different heights, and after determining the height of unknown node, the method of projection is used to project the unknown node to the two-dimensional plane and the trilateration method is used to obtain its two-dimensional coordinate. This method uses the lifting equipment to obtain the vertical coordinate of node; after obtaining the projection, the two-dimensional coordinate relation can be skillfully

---

Input: **D**: distance matrix between nodes; $\{x_1, x_2, \ldots, x_m\}(m \geq 3)$: coordinate of beacon nodes.
Output: $\{\hat{x}_{m+1}, \hat{x}_{m+2}, \ldots, \hat{x}_n\}$: estimated location of the non-beacons.
   (1) Divide the beacon nodes collected by the unknown node into a series of LU groups in accordance with their IDs and
       by obtaining the combinatorial number, and calculate the DM value of each subgroup. Any one of Formulas (1)–(6)
       can be used to calculate the DM value.
   (2) Compare the DM value of each positioning unit with the set DM threshold value, eliminate the subgroups with poor
       unit quality (low measured value), only keep the subgroups with good quality, and record the DM value of the kept
       positioning units and corresponding estimated locations obtained by using trilateration or multilateration.
       It can be believed that the bigger the DM value is, the better the quality of positioning unit, and it
       has bigger contribution to the accuracy of final positioning result. Assume there is such a multicollinearity
       weight after the positioning units with poor quality have been eliminated, and its expression is as the following:
       $$W_i = \frac{DM_i}{\sum_i DM_i}$$
   (3) At last, multiply the obtained weight with the estimated location of corresponding positioning unit, and add the
       corresponding products to obtain the final estimated location.

ALGORITHM 1: 2D LE-SA.

---

obtained, so the algorithm has low complexity, but the nodes have a high deployment cost, and the application scope is small.

In the three-dimensional space, it requires at least four beacon nodes to form the positioning unit; these four nodes form a tetrahedron, and the tetrahedron can be regarded as the expansion of triangle in the three-dimensional space, so they have a certain connection. It is generally believed that tetrahedron mesh's quality criteria include the following: the metric will not change in the case of tetrahedron mesh cells' translation, rotation, reflection, and equal scaling; the metric unit reaches the maximum in the case of a regular tetrahedron and tends to zero in the case that its volume tends to zero. Based on the criteria, researchers have proposed many criteria for measurement of which the most common ones include the minimum solid angle θ, radius ratio ρ, coefficient Q, and coefficient γ. They are, respectively, defined as follows [19, 21, 26]:

(1) the minimum solid angle θ

$$\theta = \min(\theta_1, \theta_2, \theta_3, \theta_4), \tag{7}$$

where θ_1 is given by $\sin(\theta_1/2) = 12V/(\prod_{2 \leq j < k \leq 4}[(l_{1i} + l_{1j})^2 - l_{ij}^2])^{0.5}$. $\theta_2, \theta_3, \theta_4$ can be obtained by rotation of indices;

(2) radius ratio ρ

$$\rho = \frac{3r}{R}, \tag{8}$$

where r and R are the inradius and circumradius of the tetrahedron mesh, respectively;

(3) coefficient Q

$$Q = C_d \frac{V}{\left[\sum_{1 \leq i < j \leq 4} l_{ij}\right]^3}, \tag{9}$$

where the coefficient $C_d = 1832.8208$ is applied so that the highest value of Q (for equilateral element) is equal to 1;

(4) coefficient γ

$$\gamma = \frac{72\sqrt{3}V}{\left[\sum_{1 \leq i < j \leq 4} l_{ij}^2\right]^{1.5}}. \tag{10}$$

In the above expression, V denotes the volume of tetrahedron mesh with vertexes P_1, P_2, P_3, P_4, l_{ij} representing the length of the edge joining P_i and P_j.

Literatures [30] also demonstrate that the formulas above are equivalent; the formulas all tend to zero in the case that the tetrahedron mesh's volume tends to zero; metric formula value tends to one in the case that the tetrahedrons mesh tends to a regular tetrahedron mesh.

3.4. 3D Localization Algorithm Based on Determination of the Geometrical Shape of LU. Similar to positioning in the two-dimensional environment, in the three-dimensional monitoring area, the unknown node obtains more than four surrounding beacon nodes and its distances to these beacon nodes to conduct localization. Due to the impact of the LU formed by beacon nodes on location estimation, during the localization process, it requires quality determination of these LU. The four criterions mentioned in the above section are approximately equivalent, and we can choose any one out of Formulas (7)–(10) as the criterion to measure the quality of three-dimensional tetrahedron and then develop its corresponding novel location algorithm in the three-dimensional space, called three-dimensional location estimation-shape analysis (3D LE-SA), and see Algorithm 2 for the detailed procedure.

4. Localization Algorithm Based on Multivariate Analysis

4.1. Multicollinearity Problem during the Localization Process. In accordance with the literature, we know that the equation set of the distance between unknown node and beacon node can be transformed into the form of $\mathbf{Ax} = \mathbf{b}$

Input: **D**: distance matrix between nodes; $\{x_1, x_2, \ldots, x_m\}$ ($m \geq 4$): coordinate of beacon nodes.
Output: $\{\widehat{x}_{m+1}, \widehat{x}_{m+2}, \ldots, \widehat{x}_n\}$: estimated location of the non-beacons.
  (1) Divide the beacon nodes collected by the unknown node into a series of LU groups in accordance with their IDs and
      by obtaining the combinatorial number, and calculate the DM value of each subgroup. Any one of Formulas (7)–(10)
      can be used to calculate the DM value.
  (2) Compare the DM value of each positioning unit with the set DM threshold value, eliminate the subgroups with poor
      unit quality (low measured value), only keep the subgroups with good quality, and record the DM value of the kept
      positioning units and corresponding estimated locations obtained by using trilateration or multilateration.
      It can be believed that the bigger the DM value is, the better the quality of positioning unit, and it has bigger
      contribution to the accuracy of final positioning result. Assume there is such a multicollinearity weight after
      the positioning units with poor quality have been eliminated, and its expression is as the following:
      $$W_i = \frac{DM_i}{\sum_i DM_i}$$
  (3) At last, multiply the obtained weight with the estimated location of corresponding positioning unit, and add the
      corresponding products to obtain the final estimated location.

ALGORITHM 2: 3D LE-SA.

[31, 32]. Because the deployment environment has various interference sources, noise within the node, and rounding off caused by quantification of signal, error exists in the distance measurement, and the actual equation set generally exists in the form of $\mathbf{Ax} = \mathbf{b} + \xi$, in which ξ refers to the error. In order to obtain the optimal solution of location estimation, and also considering the convenience of computation, the square of error is generally used as the criterion, and in order to obtain the optimal solution, calculate the partial derivative of loss equation and set it as zero; that is,

$$\frac{\partial \|\xi\|^2}{\partial \mathbf{x}} = -2\mathbf{A}^T \mathbf{b} + 2\mathbf{A}^T \mathbf{Ax} = 0. \tag{11}$$

The formula (11) can be recast as

$$\mathbf{A}^T \mathbf{b} = \mathbf{A}^T \mathbf{Ax}. \tag{12}$$

If the beacon nodes that constitute the LU are not in a straight line, that is, square matrix $\mathbf{A}^T\mathbf{A}$ is reversible, then, for the equation, the common least square method can be used to obtain the estimated coordinate of unknown node: $\widehat{\mathbf{x}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$. If the beacon nodes that constitute the LU are or are approximately in a straight line, at this moment, there will be multicollinearity phenomenon in the estimation; if forced implementation of least square method continues, it will cause instability to the estimated value, and under severe situation, the multicollinearity might even cause abnormality in the signal of estimated value, which makes the estimated result lose all its meanings. When the LU are completely collinear, matrix $(\mathbf{A}^T\mathbf{A})^{-1}$ does not exist, which makes it impossible to use the least square method to estimate the location of unknown node; when the LU are approximately collinear, $|\mathbf{A}^T\mathbf{A}| \approx 0$, it results in big diagonal element of matrix $(\mathbf{A}^T\mathbf{A})^{-1}$, which increases the variance of the parameter estimated value, and the estimated value becomes invalid.

The concept of multicollinearity in multivariate analysis was firstly proposed by Frisch in 1934 [33], and its initial meaning is that some independent variables in the regression model are linear dependent, and for the location estimation algorithm, it means at least two columns in $\mathbf{A}$ matrix have liner relation; that is, columns $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ in matrix $\mathbf{A}$ have relational expression:

$$k_1\mathbf{a}_1 + k_2\mathbf{a}_2 + \cdots + k_n\mathbf{a}_n = 0, \tag{13}$$

where not all of the constants $k_1, k_2, \ldots, k_n$ are zero. Obviously, the linear relation of columns in matrix $\mathbf{A}$ has caused the abnormality of matrix $\mathbf{A}^T\mathbf{A}$, which makes the algorithm of location estimation completely invalid. However, in actual application, this kind of situation is rare. Under most circumstances, certain data columns in matrix $\mathbf{A}$ can be approximately expressed by other data columns, not completely; in other words, columns $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ in matrix $\mathbf{A}$ have

$$k_1\mathbf{a}_1 + k_2\mathbf{a}_2 + \cdots + k_n\mathbf{a}_n + \xi = 0, \tag{14}$$

where ξ is a stochastic error. At this moment, it can be called nearly collinearity, and complete collinearity and nearly collinearity together are called multicollinearity. If there is multicollinearity problem, it is not treated and location estimation continues. Although sometimes nearly collinearity can be used to calculate the location of unknown node, it will increase the variance of estimated value; the estimated value is unstable, its confidence interval is increased, and the estimation accuracy is reduced, and under severe situation, it might cause the estimated location and actual location to have a mirror-image relation along the straight line formed by beacon nodes.

4.2. Detection and Remedies for Multicollinearity Problem. During the location estimation process, because the positioning units have a collinear or approximately collinear geometrical relationship, it results in the fact that the columns in matrix $\mathbf{A}$ constituted by LU also have a collinear or approximately collinear relation. These collinear or nearly collinear relations result in an unstable model during the estimation computation, and under severe circumstance, it

may even affect the accuracy of location estimation. Methods like variance inflation factor (VIF), condition index (CI), and variance proportions (VP) are usually used to diagnose multicollinearity [34–36]. In accordance with the literature [36], if VIF > 10, it is generally believed that the model has a strong multicollinearity relation; if the CI is between 10 and 30, there is weak multicollinearity relation, if it is between 30 and 100, there is medium multicollinearity relation, and if is bigger than 100, there is strong multicollinearity relation; among the big CI, the variable subset consisting of independent variables with a variance proportion bigger than 0.5 is regarded as related variable set.

At present, in theory and practical engineering applications, there are also various methods that can be used to overcome the impact of multicollinearity. Researchers have proposed various detection methods and remedial measures, but different methods have different effects in engineering application. Ridge regression (RR) and principal component regression (PCR) are the most common remedial measures for multicollinearity problem [35, 37, 38] used by researchers. RR was proposed by Ae [38] in 1962, by introducing offset k (also known as "ridge parameter"), and the estimated unbiasedness is sacrificed for significant decrease of variance in the estimated value, in order to realize the final purpose of increasing estimation accuracy and stability. For the estimation model $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{b}$, after introducing the ridge parameter k, a new estimation model $(\mathbf{A}^T\mathbf{A} + k\mathbf{I})\mathbf{x} = \mathbf{A}^T\mathbf{b}$ can be obtained. Due to introduction of the ridge parameter k, the location estimation is no longer unbiased, but the multicollinearity problem is solved, which reduces the variance in the estimated value, and in the meantime, the estimation becomes stable. RR is easy and feasible, and in a certain degree, it overcomes the impact of multicollinearity on the estimated value, so it has been widely applied in the engineering practices. The key of RR is how to choose appropriate ridge parameter k, and ridge parameter k does not have specific meaning, which causes that the selection of k is too subjective. In accordance with Formula (14), we can see that most multicollinearity is caused by noise, while ridge regression reduces variance in the estimated value only by adding k, and the RR method retains all variables, so the ridge regression method does not apply to the scenario with severe noises.

According to statistics and the maximum entropy principle, the information in signal data set generally refers to the variation of data in this set, while the variation can be measured by the total variances; the bigger the variance is, the more information is contained in the data and the smaller variance the noise has, and the signal to noise ratio is actually the variance ratio between signal and noise [39]. Therefore, when choosing the data that can best explain the system, the values with big variances in multiple observations are actually chosen, and this kind of data is called principal components (PCs). Principal component analysis (PCA) [33, 40] is a method which uses a small amount of PCs to disclose the internal structure of multiple variables through recombination of the original data. It is generally believed that the data with a big variance is closely related to the PCs
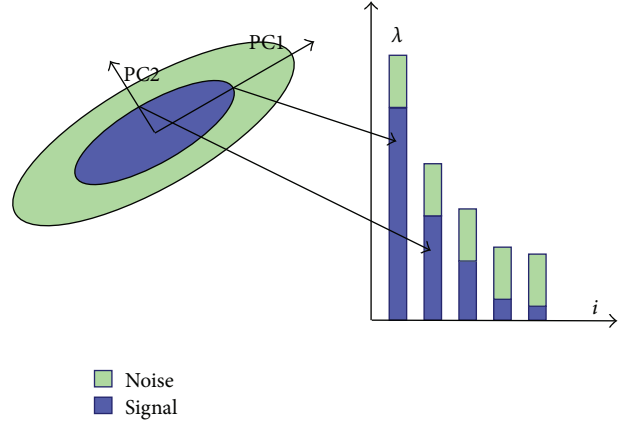

FIGURE 5: Two views of the "directional" information versus the "unidirectional" noise.

with a big eigenvalue, while other data with a small variance has a strong connection to the PCs with a small eigenvalue. Therefore, different PCs have different effects and impacts on location estimation, and the location accuracy is not in direct proportion to the number of PCs, so it will help increase the stability and accuracy of model by choosing the PCs which can better explain the estimated value to estimate and analyze the data. PCA could transform the original data with a high correlation into mutually independent or irrelevant data, and the data with the biggest signal to noise ratio occurs in the first PCs, and as the eigenvalue becomes smaller, the signal to noise ratio of the data contained in its corresponding eigenvector also becomes smaller. Figure 5 shows the result after such transformation [41].

Through PCA computation of data, only the first several dimensions of PCs are kept, in which not only the scale of original data matrix is compressed, but each obtained new variable which is the linear combination and comprehensive result of the original variables, and it has certain realistic significance. Among the vectors with relatively concentrated noise and small eigenvalue, by eliminating these data with a small signal to noise ratio, the redundancy and noise can be eliminated, and in the meantime, the multicollinearity between variables can also be eliminated. Massy proposed PCR in 1965 [33] based on the idea of PCA, which uses the PCA to retain low-order PCs, ignore high-order PCs, and then run least squares to regression analysis.

4.3. Localization Algorithm Based on Multivariate Analysis.
Due to the multicollinearity problem between the coordinate matrices of beacon nodes, it causes matrix $\mathbf{A}^T\mathbf{A}$ irreversible or the fact that $\mathbf{A}^T\mathbf{A}$ cannot used for node estimation. Therefore, we can use the multicollinearity detection method to determine whether the positioning data has multicollinearity phenomenon; then, the PCA method from PCR method can be used to reconstruct matrix detection. The part with eigenvalue of zero, close to zero, or with a very small eigenvalue will be eliminated (only the part with a cumulative variance contribution rate bigger than 90% is kept), and at last, the location is estimated. Because during the computation

Input: **D**: distance matrix between nodes. $\{x_1, x_2, \ldots, x_m\}$ $(m \geq 3)$: coordinate of beacon nodes.
Output: $\{\widehat{x}_{m+1}, \widehat{x}_{m+2}, \ldots, \widehat{x}_n\}$: estimated location of the non-beacons.
    (1) Conduct standardization treatment to matrix **A**.
    (2) To matrix **A** after standardization treatment, use PCA to extract PCs and the score vector. (see Formula (15)).
    (3) Use conditional index to determine whether there is multicollinearity problem. If there is, eliminate corresponding
        PCs with a small characteristic root in accordance with the cumulative variance contribution rate.
    (4) Use the left PCs and PCR, and obtain the final location estimation through Formula (18).

ALGORITHM 3: LE-PCR.

process, the impact of correlation has been considered when choosing data, it has ensured the model's estimability. In the meantime, based on assurance of accuracy, certain data with insignificant impact on the system (noise data) will be abandoned to reduce the model's order and significantly reduce the calculated amount.

This paper uses PCA to conduct feature extraction to matrix **A**, and the obtained first d components form a matrix to replace the original matrix **A** to conduct multivariate analysis. Although part of the data is lost, the accuracy and stability of estimation are increased.

After standardization of matrix **A**, break it down to the sum of d exterior products of vectors; that is,

$$\mathbf{A} = t_1 p_1^T + t_2 p_2^T + \cdots + t_d p_d^T, \tag{15}$$

where t refers to the score vector; p is PCs. Formula (15) can also be expressed as

$$\mathbf{A} = \mathbf{T}\mathbf{P}^T. \tag{16}$$

It is easy to know that, in matrix **T**, each vector has a mutually orthogonal relation with each other; in matrix **P**, each sector is also orthogonal with each other, and each vector length is one. In accordance with above description, it is not difficult to obtain that

$$t_i = \mathbf{A} p_i. \tag{17}$$

Therefore, we can obtain the following conclusion: each score vector is actually the projection of matrix **A** in the direction of its corresponding PCs vector.

In this way, we can obtain the final location estimation as

$$\widehat{\mathbf{x}} = \mathbf{P} \left(\mathbf{T}^T \mathbf{T} \right)^{-1} \mathbf{T}^T \mathbf{b}. \tag{18}$$

The complete steps of PCR-based location estimation (location estimation-PCR, LE-PCR) are described in Algorithm 3.

5. Simulation and Experiments

The wireless sensor network has the characteristic of a big scale. It might require deployment of hundreds or even thousands of nodes in order to verify a localization algorithm, and it is impossible to realize a real network of such a big scale under current experimental conditions. In addition, in order to determine the quality of a localization algorithm, it also

requires verifying its adaptability under different scenarios; sometimes it might even require adjusting the parameter of algorithm under the same scenario, and these are difficult to realize under current experimental conditions. Therefore, during research of the localization algorithm for large-scale nodes, the method of software simulation is usually used to evaluate the quality of localization algorithm.

The algorithm referred to in this paper is mainly in accordance with the impact of the relation between beacon nodes on the localization accuracy, and the impact of the distance measurement between nodes on the location estimation accuracy is not a main issue considered in this section. Based on that, this section adopts DV-Hop based on the range-free localization algorithm between beacon nodes as the carrier to verify the concept proposed in this section. In addition, there are many technical standards to measure a localization algorithm, while this paper mainly studies the impact of LU on the positioning performance, so the performance parameter of ALE (average localization error) is used to examine the algorithm performance.

ALE is mainly verified to evaluate localization accuracy, and it is described as follows:

$$\text{ALE} = \frac{\sum_{i=1}^{n} \sqrt{\left(\widehat{x}_i - x_i \right)^2 + \left(\widehat{y}_i - y_i \right)^2}}{n \times R} \times 100\%. \tag{19}$$

In the formula, $(\widehat{x}_i, \widehat{y}_i)$ represents the estimated coordinate location of the ith node, (x_i, y_i) represents the actual coordinate location of the ith node, (x_i, y_i) represents the number of the unknown nodes, and R represents the communication radius. It can be seen from the above formula that ALE refers to the ratio of the average error of the Euclidean distance from the estimation location of all nodes to the real location in the area to the communication radius. ALE can reflect the stability of the localization algorithm and the positioning accuracy; when the communication radius of the node is given, if the average localization error of the node is smaller, then the positioning accuracy of the algorithm is higher, and vice versa.

First of all, this section briefly introduces the DV-Hop localization algorithm; then, two-dimensional and three-dimensional DV-Hop algorithms are used to verify the algorithms based on geometrical analysis of two-dimensional and three-dimensional LU, respectively; because the localization processes of multivariate analysis in the two-dimensional and three-dimensional spaces are too similar, in the final part of this section, only two-dimensional DV-Hop algorithm is used to verify the idea of PCR algorithm.
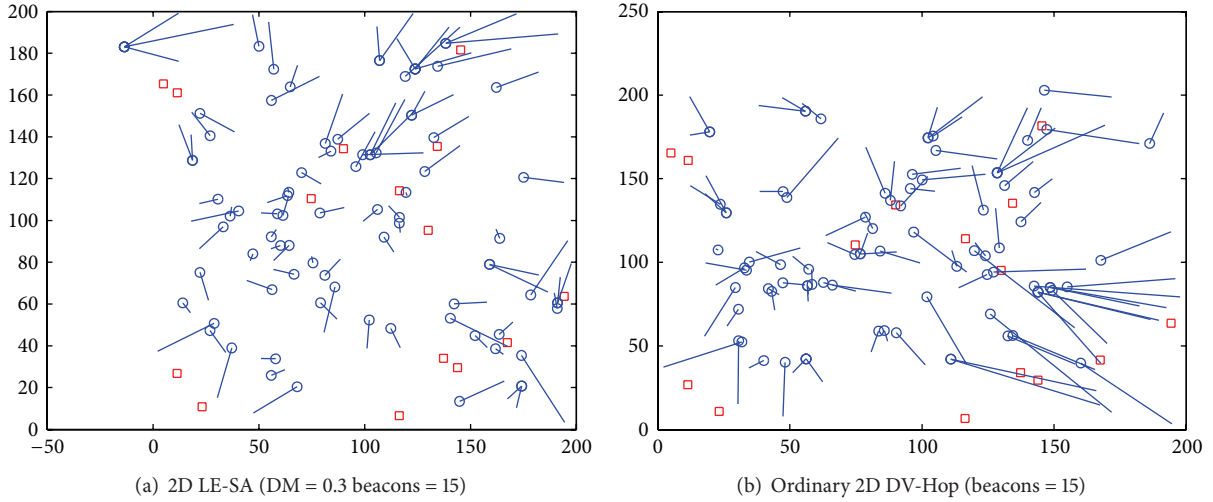
(a) 2D LE-SA (DM = 0.3 beacons = 15)

(b) Ordinary 2D DV-Hop (beacons = 15)

FIGURE 6: Localization results in 2D environment.

5.1. Introduction of DV-Hop Localization Algorithm. The DV-Hop localization algorithm proposed by Niculescu et al. from Rutgers University [41–43] is one of a series of distributed localization algorithms; it is a localization algorithm not related to the distance, and it smartly uses the distance vector routing and the idea of GPS localization, and this algorithm has great distributivity and expandability. Its localization principle is as follows: firstly, the minimum hop from the unknown node to the beacon node is calculated, then the average distance of each hop is estimated, then the minimum hop is multiplied with the average distance of each hop to obtain the estimated distance between the unknown node and beacon node, and at last trilateration is used to calculate the coordinate of unknown node. DV-Hop method has great distributivity and expandability, and the positioning process consists of the following three steps.

Step 1. The DV-Hop localization algorithm uses the classic distance vector exchange protocol to make all nodes in the deployment area obtain the hop of beacon nodes.

Step 2. The beacon node calculates the average distance of each hop in the network, and after obtaining the locations of other beacon nodes and distance of hop, the beacon node calculates the average distance of each hop in the network, uses it as an adjusted value, and broadcasts it to the network. The average distance of each hop can be expressed by the following formula:

$$\text{HopSize}_i = \frac{\sum_{i \neq j} \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2}}{\sum_{i \neq j} h_i}, \qquad (20)$$

where (x_i, y_i) and (x_j, y_j) refer to the coordinates of beacon nodes i and j, respectively; h_i refer to the hops of beacon node i and all other beacon nodes. When the unknown node obtains its distance to three or more beacon nodes, it can enter Step 3, that is, calculation of node location.

Step 3. Suppose an unknown node receives the flood messages from three beacons. It uses trilateration or maximum likelihood method to determine its location.

Similar to the common DV-Hop algorithm, the 3D DV-Hop algorithm also consists of three steps.

In accordance with the above description, the DV-Hop is also considered as localization algorithm based on beacon nodes, and its estimation result is related to multicollinearity between nodes in a certain degree. During the solving process, $\mathbf{A}^T\mathbf{A}$ must be reversible; if $|\mathbf{A}^T\mathbf{A}| = 0$ or $|\mathbf{A}^T\mathbf{A}| \approx 0$, the matrix has multicollinearity problem; that is, exact or approximate linear relation occurs in the columns of matrix $\mathbf{A}$, and its existence will cause bad consequence to the final localization accuracy, and when there is complete multicollinearity, the multilateral measurement might even fail. When only incomplete multicollinearity occurs, though the estimated value of location can be obtained, it is unstable. In the meantime, the variance of estimated parameter value will increase, and the increase depends on the severity of multicollinearity.

5.2. 2D LE-SA DV-Hop. In this group of simulation experiments, we suppose 100 nodes were randomly and evenly distributed in a 200 m × 200 m area, and the node communication radius is 50 m. It was assumed the number of beacon nodes increased from 10 to 20. In the meantime, the DM value increased from 0.1 to 0.7, and the step size is 0.1. To reduce the statistical variability, under the same number of beacon nodes and DM value, the reported results here are averaged over 20 repetitions.

Figure 6 shows the location result of 2D LE-SA DV-Hop and ordinary 2D DV-Hop. The squares are beacons and the circles denote the non-beacons. Each line connects a true node location and its estimation. The length of each line denotes the estimation error. We set the number of beacon nodes as 15, DM = 0.3, and plot the location result of each sensor node in **Figure 6(a)**. The ALE is about 29.1%. The final
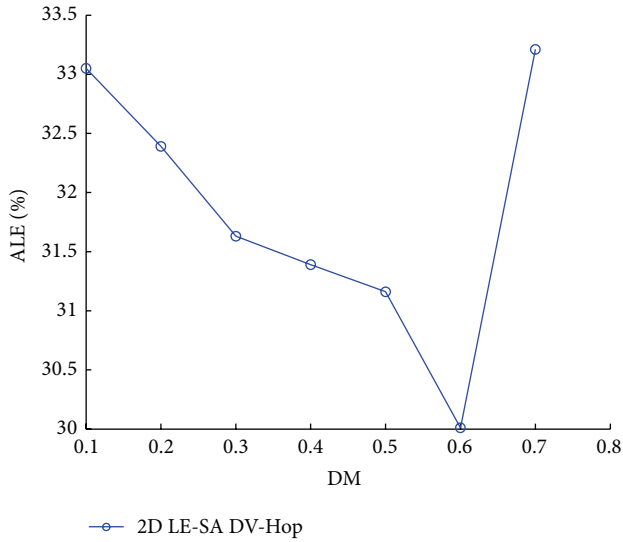
FIGURE 7: ALE on locations based upon DM in 2D environment.

FIGURE 8: ALE on locations based upon the number of beacons in 2D environment.

estimation of each sensor node of ordinary 2D DV-Hop is shown in Figure 6(b). The ALE is about 40.7%. Obviously, the length of the straight line in Figure 6(a) is shorter than that in Figure 6(b).

We present a quantitative analysis (beacons are fixed 15) of the effect of ALE and DM in Figure 7. We can see that when DM is between 0.1 and 0.6, the ALE value has monotonic decrease; when DM > 0.6, the ALE curve presents a rising trend. The reason is that the DV-Hop algorithm is a localization algorithm based on distance vector routing; it uses the hop distance between nodes to replace the linear distance between nodes, and as the hop distance increases, the error between nodes also increases.

The location estimation method based on shape analysis in this paper actually refers to that during the localization process; only the LU with a high shape quality are chosen during the localization process, and the LU with a poor shape quality are eliminated. When the LU determines that the DM value of multicollinearity is big, the beacon node near the location node (with a small hop) does not satisfy the estimation requirement; it can only choose the beacon nodes far away (with a big hop) as the reference nodes, which causes the idea that the hop distance used in localization is far longer than the actual distance, and the final estimation result will increase instead of decreasing. For this kind of situation, the researchers usually add the threshold value of hops to restrict the hop distance with big hops, but the threshold value of hops will also generate nodes that cannot be estimated in the monitoring area, which further reduces the coverage of monitoring area. Therefore, a compromise should be made to ensure the localization accuracy on the one hand, and the localization coverage on the other hand.

Figure 8 shows the change of ALE curve with the increase of beacon nodes (10–20) in the 2D LE-SA and ordinary 2D DV-Hop when DM = 0.3. In accordance with Figure 8, we can see that because shape analysis is added to the 2D LE-SA algorithm, it has inhibited multicollinearity, and the ALE
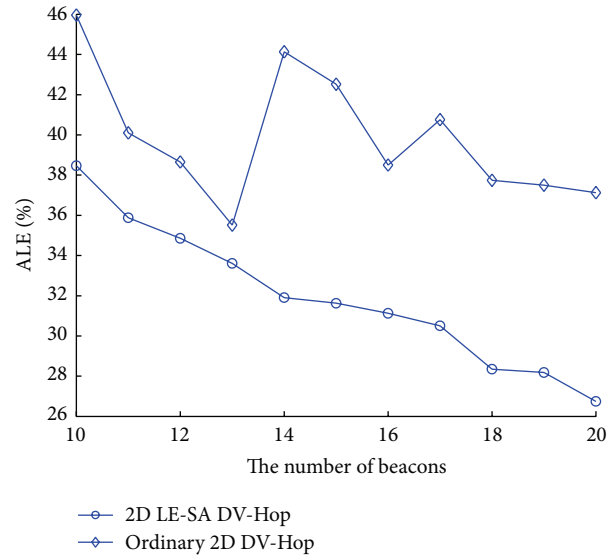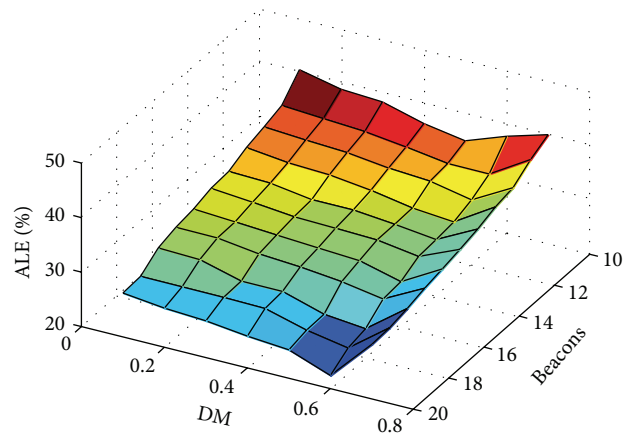

FIGURE 9: ALE on locations based upon the number of beacons and DM in 2D environment.

curve of 2D LE-SA falls with the increase of beacon nodes; while the ALE curve of ordinary 2D DV-Hop algorithm is not improved with the increase of beacon nodes, it presents up and down motion.

Figure 9 shows the resulting ALE as a function of DM and the number of beacon nodes in the 2D LE-SA. Obviously, after DM > 0.6, the monotonic decrease trend of ALE changes into increase; however, when the DM value is fixed, ALE presents decrease with the increase of beacon nodes. Therefore, it shows that by setting DM threshold value in the localization algorithm, this can help eliminate the impact of multicollinearity and increase the stability of algorithm and accuracy of estimation. However, if the set threshold of DM is too high, it will reduce the reference nodes used for localization, which will further reduce the algorithm performance and coverage. Therefore, pretest should be conducted before setting DM value in a specific area.
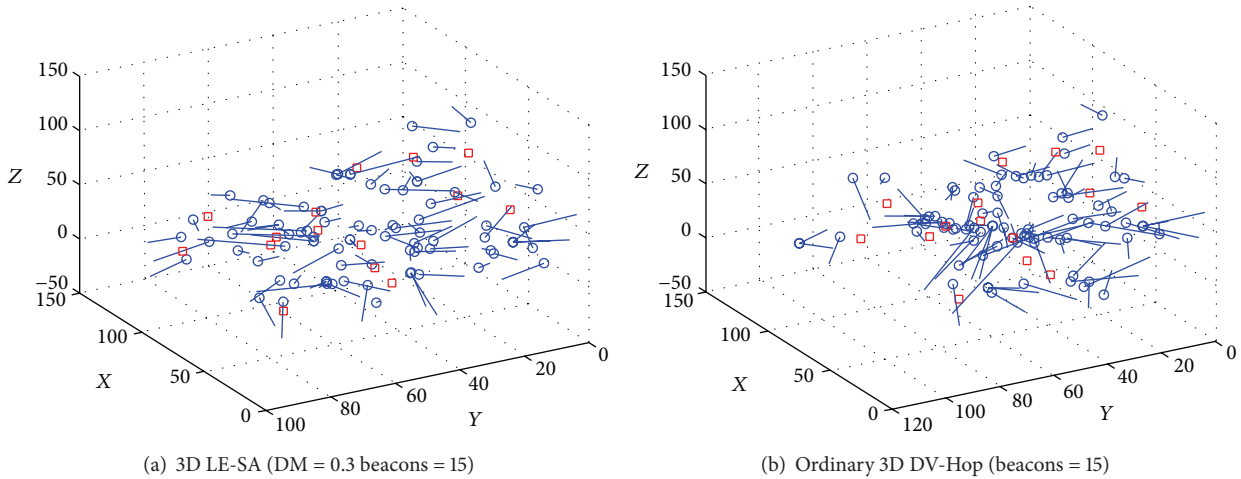
(a) 3D LE-SA (DM = 0.3 beacons = 15)

(b) Ordinary 3D DV-Hop (beacons = 15)

FIGURE 10: Localization results in 3D environment.

5.3. 3D LE-SA DV-Hop. During actual application, it is impossible to place the nodes in a purely two-dimensional plane, and they are generally in a three-dimensional scenario, such as under the water, on the slope, and in a space. Therefore, for the shape analysis and localization method mentioned in this paper, the criterion to determine a three-dimensional shape is different from that to determine a two-dimensional shape, and it needs experiment to verify.

In the simulation experiments, it was assumed that 100 nodes were randomly and evenly distributed in a three-dimensional environment of 100 m × 100 m × 100 m. We set the communication radius of nodes 50 m and the number of beacon nodes increased from 10 to 20; in the meantime, the DM value increased from 0 to 0.6. All of the reported simulation results are the average over 50 trials.

Figure 10 shows the localization result when the number of beacon nodes is 15, where the DM value of 3D LE-SA DV-Hop is 0.3. Figure 10(a) shows the location results of 3D LE-SA DV-Hop where ALE is 31.9% while the ALE of ordinary 3D DV-Hop ALE is 42.6% in Figure 10(b).

We also present a quantitative analysis of the effects of DM in the 3D LE-SA DV-Hop algorithm when there are 15 beacon nodes in Figure 11. We can see that when DM ≤ 0.3, the ALE value has monotonic decrease; when DM > 0.3, the ALE curve presents a rising trend. The reason is that similar to the two-dimensional scenario, a big DM value has restricted selection of surrounding reference beacon nodes, and due to a big error between the beacon nodes far away and the unknown node, it increases the error of the final estimation result instead of reducing it. The difference is that the three-dimensional space causes the nodes to become more "sparse"; when DM is bigger than 0.3, it will cause change to ALE, and the change is more significant; when DM = 0.5, ALE is close to 90%; when DM = 0.6, ALE is even bigger than 120%. The threshold value of hops is added in order to maintain the localization accuracy, and after the beacon nodes with big hops have been limited, it will also cause the decrease of coverage.
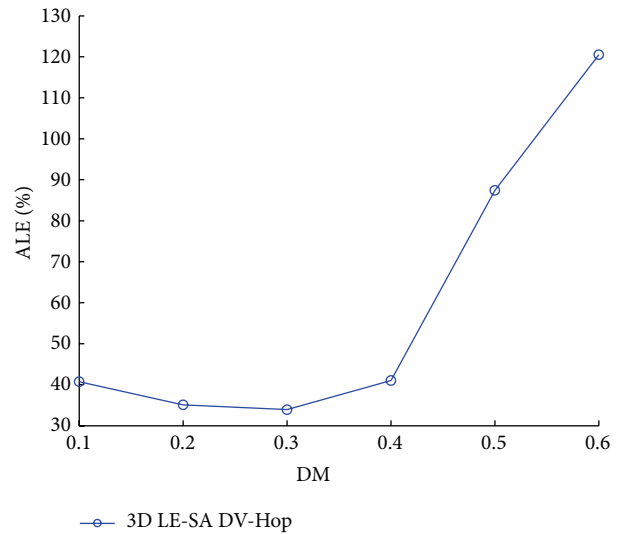

FIGURE 11: ALE on locations based upon DM in 3D environment.

Figure 12 shows the change of ALE curve with the number of beacon nodes varying from 10 to 20. Similarly, the common ordinary 3D DV-Hop algorithm also fails to address the impact of multicollinearity, which makes the ALE curve present up and down motion. Through the 3D LE-SADV-Hop with DM, the impact of multicollinearity problem is avoided, its ALE curve falls with increase of beacon nodes, and its accuracy is better than the ordinary algorithm.

5.4. LE-PCR DV-Hop. Because the localization method based on multivariate analysis only conducts computation processing to the coordinate matrix of beacon nodes, the two-dimensional and three-dimensional methods are very similar; therefore, the third group of experiments in this section only compares the situations in the two-dimensional environment. In this set of experiments, 100 sensors are randomly placed in a 200 m by 200 m square region, and the communication radius is 50 m. Due to the fact that the

FIGURE 12: ALE on locations based upon the number of beacons in 3D environment.

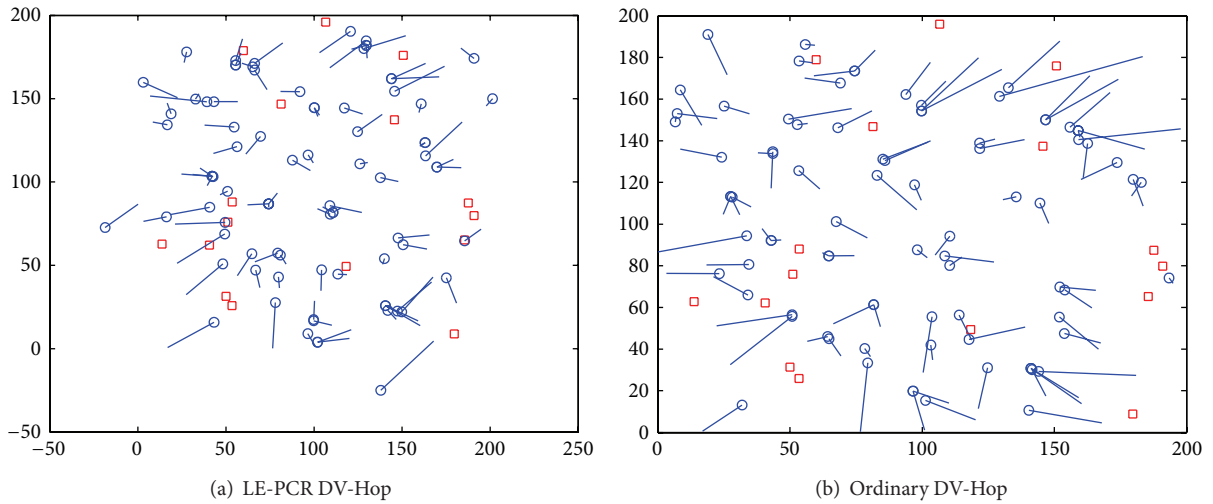(a) LE-PCR DV-Hop

(b) Ordinary DV-Hop

FIGURE 13: Localization results in 2D environment.

beacon nodes are collinear or approximately collinear, this causes the idea that matrix $\mathbf{A}^T\mathbf{A}$ is irreversible or that matrix $\mathbf{A}^T\mathbf{A}$ cannot be used in node estimation. We can use the PCA method to reconstruct matrix $\mathbf{A}$ and use CI to determine whether there is multicollinearity problem; the part with an eigenvalue of zero, close to zero, or with a very small eigenvalue will be eliminated (only the part with a cumulative variance contribution rate bigger than 90% is kept), and in this way, in the reobtained data, there is no collinear part, and some of noise is eliminated as well.

Similar to the first group of experiments, we show location results of each sensor node in Figure 13. We set 16 beacon nodes were randomly distributed in a monitoring area, cumulative variance contribution rate bigger than 90%, and the final solution of LE-PCR DV-Hop in Figure 13(a). The ALE is about 27.5%. The squares are the beacons, and the circles denote the non-beacons. Each line connects a

true sensor location and its estimation. The final estimated location of ordinary DV-Hop is shown in Figure 13(b). The ALE is about 35.8%.

Because the PCR-based method conducts recombination and screening to the coordinate information of beacon nodes, in other words, during the location estimation process, as much coordinate information of beacon nodes as possible will be maintained, and the information not important to location estimation will be eliminated. In addition, the impact of correlation has been considered when choosing data, so it has ensured the estimability of location estimation process. In the meantime, based on assurance of estimation accuracy, it can reduce the estimation model's order and significantly reduce the computational complexity.

Figure 14 shows the change of ALE curve as the number of beacon nodes is gradually increased from 10 to 20 under multiple deployments of the algorithm under the same
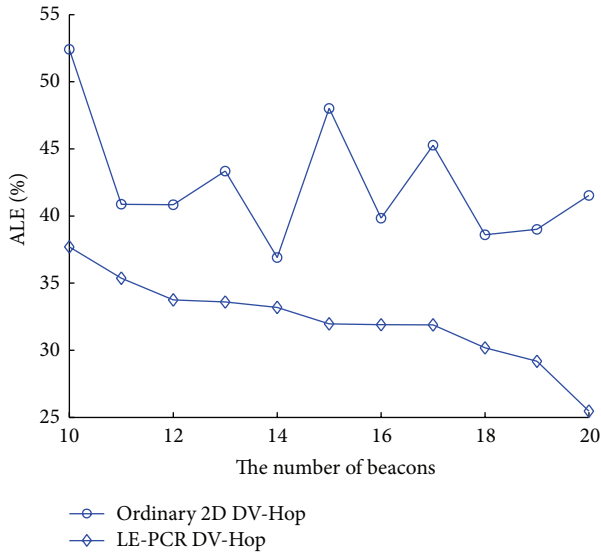
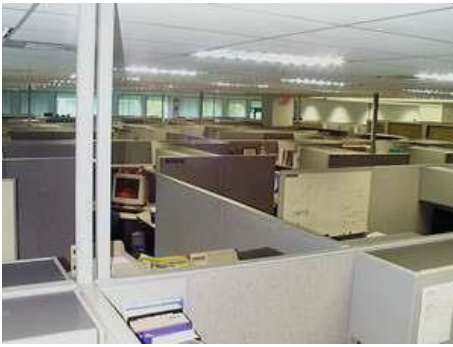FIGURE 14: ALE on locations based upon the number of beacons in 2D environment.

FIGURE 15: Actual collection area.

scenario (50 times, use average value of ALE). In the actual environment, because the impact of multicollinearity and noise is inevitable, in accordance with Figure 14, we can see that, for the ordinary DV-Hop algorithm (due to the fact that the nodes are randomly redeployed in each experiment, the ALE in Figures 8 and 14 are not the same), ALE does not decrease with the increase of beacon nodes, and the ALE curve presents up and down motion; while in the PCR-based localization algorithm, through reconstruction of the beacon location data, the useful information, the multicollinearity problem, and noise are rearranged, and by setting a certain threshold value (cumulative variance contribution rate), part of the multicollinearity data and noise can be eliminated so that the ALE curve will fall with the increase of beacon nodes. In addition, the ALE of improved method, which is all lower than 40%, is significantly lower than that of common algorithm.

5.5. Performance Evaluation Based on Actually Measured Data. In this group of experiments, we use actually measured data set provided by the SPAN lab. As shown in Figure 15,

the network consists of 44 sensor nodes that are deployed in a rectangular office area of $12 \times 14 \, \mathrm{m}^2$. We randomly choose 4 to 13 nodes as the beacon nodes and make the node communication radius 5 m. In the experiment, the DM value is set to 0.3, and cumulative variance contribution rate is 90%.

Figure 16 shows the localization results of three algorithms under the circumstance that the number of beacons is 9, in which Figure 16(a) shows nodes deployment; Figure 16(b) is localization result of ordinary DV-Hop method, ALE = 36.32%; Figure 16(c) is localization result of 2D LE-SA method, ALE = 31.1%; Figure 16(d) shows localization result of 2D LE-PCR method, ALE = 30.6%. From Figure 9, it can be seen that localization results of 2D LE-SA and 2D LE-PCR method are close, and their localization performance is better than ordinary DV-Hop method.

Figure 17 plots curves of ALE of repeated experiments by three localization algorithms varying with the quantity of beacon nodes in SPAN lab. It is easy to find ALE of ordinary DV-Hop fluctuating as the strongest, and accuracy is the lowest; however, SA-based and PCR-based methods can obtain more stable and higher precision results. Comparing with the ordinary DV-hop algorithm, SA-based and PCR-based algorithms considered the multicollinearity factor of beacons, which makes them obtain better accurate localization results.

6. Conclusion

In this paper, we analyze the problem caused by multicollinearity during the localization computation process. Firstly, we give two kinds of poor-quality of 2-dimensional LU and nine kinds of poor-quality of 3-dimensional LU, respectively. Secondly, we give the corresponding six triangle and four tetrahedron judgment formula. Finally, we employ PCR algorithm, which is dimensionality reduction method, to conduct recombination and extraction to the coordinate matrix of beacon nodes and make use of beacons data to estimate the location of unknown node.

The method based on geometrical analysis of LU is direct. By setting the threshold value, the LU smaller than this threshold value will be excluded. Because the location estimation process only uses the LU with high quality, so it can promote localization accuracy, and the algorithm is stable. However, because some of the LU are excluded from the location estimation process, the number of nonestimated nodes is increased in certain area. In addition, the biggest threshold value of DM should be selected in accordance with the distribution area. The PCR-based localization method extracts PCs in the coordinate matrix, and because there is no correlation between PCs, the impact of multicollinearity problem is avoided. In addition, after abandoning certain PCs that contain noise, the overall localization accuracy is increased, the calculated amount is reduced, it does not need to select threshold value such as DM, and it only requires setting the cumulative variance contribution rate. However, the PCR-based method is a biased estimation method, and certain estimation accuracy will inevitably be lost.

(a) Node distribution

(b) Ordinary DV-Hop

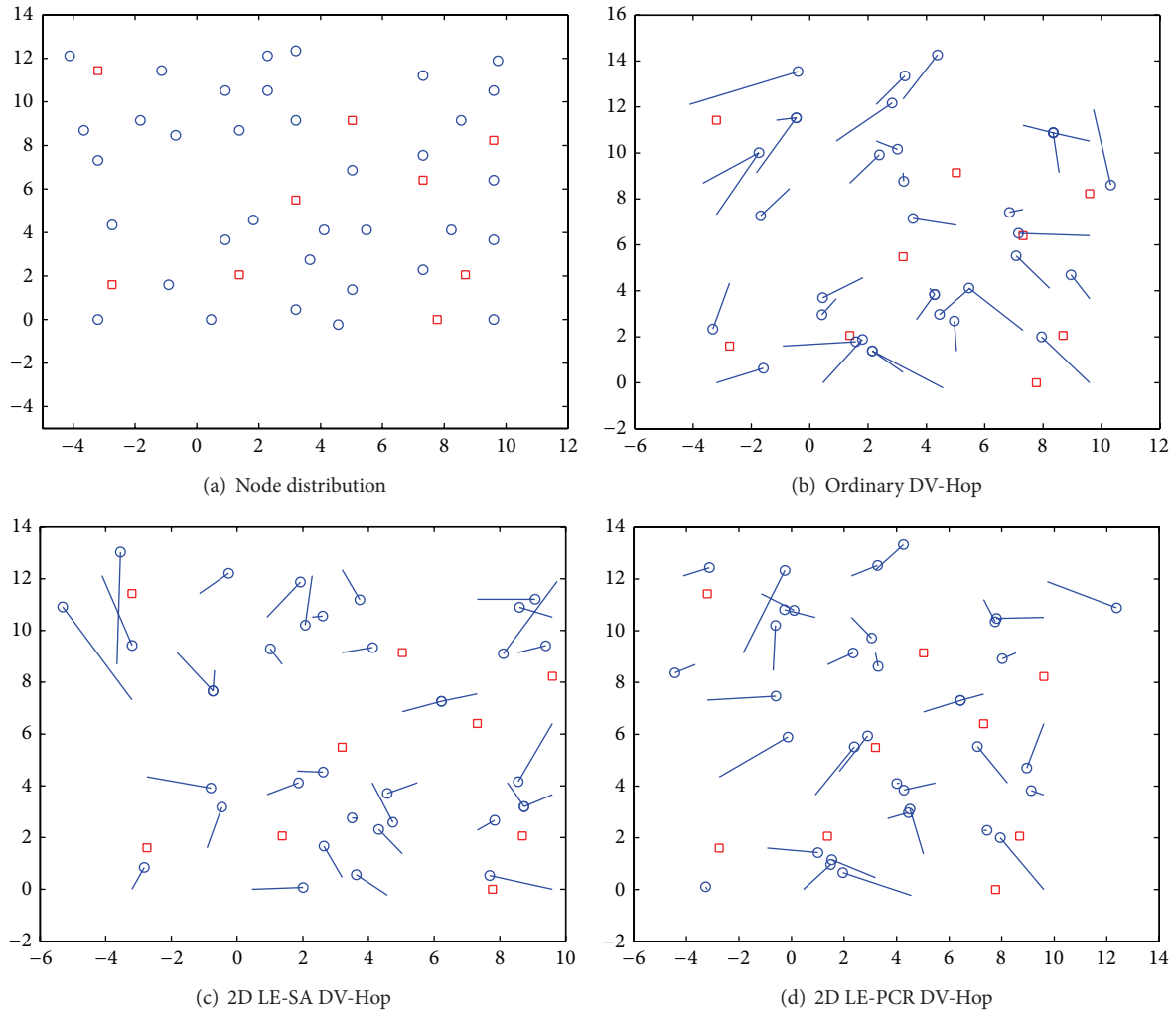(c) 2D LE-SA DV-Hop

(d) 2D LE-PCR DV-Hop

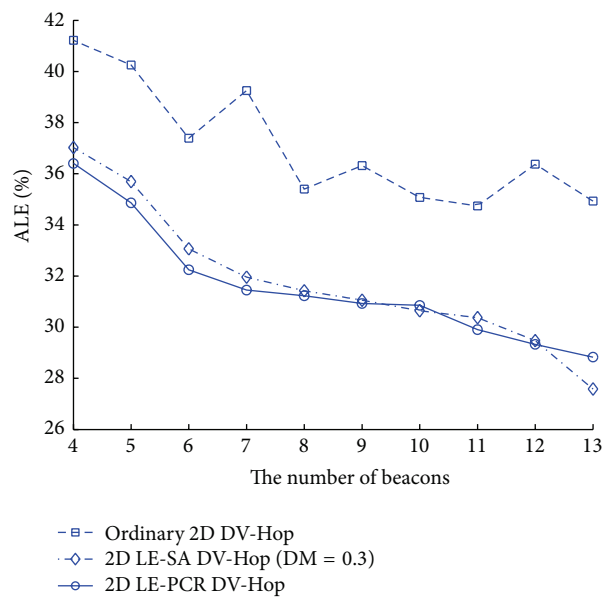FIGURE 16: Location estimates with actual measured data.

FIGURE 17: ALE on locations based upon the number of beacons in office area.

Our method can be used for range-free localization, and it can also be used for range localization; furthermore, it can be used for tracking and locating moving targets.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] S. Khan, A.-S. K. Pathan, and N. A. Alrajeh, *Wireless Sensor Networks Current Status and Future Trends*, CRC Press, London, UK, 2013.

[2] I. M. M. E. Emary and S. Ramakrishnan, *Wireless Sensor Networks: From Theory to Applications*, CRC Press, London, UK, 2013.

[3] I. F. Akyildiz and M. C. Vuran, *Wireless Sensor Networks*, John Wiley & Sons, 2010.

[4] A. A. Hady, S. M. Abd El-Kader, and H. S. Eissa, "Intelligent sleeping mechanism for wireless sensor networks," *Egyptian Informatics Journal*, vol. 14, no. 2, pp. 109–115, 2013.

[5] Y. Zhang, W. Yang, D. Han et al., "An integrated environment monitoring system for underground coal mines—wireless Sensor Network subsystem with multi-parameter monitoring," *Sensors*, vol. 14, no. 2, pp. 13149–13170, 2014.

[6] A. Pascale, M. Nicoli, F. Deflorio, B. Dalla Chiara, and U. Spagnolini, "Wireless sensor networks for traffic management and road safety," *IET Intelligent Transport Systems*, vol. 6, no. 1, pp. 67–77, 2012.

[7] J. Heidemann, M. Stojanovic, and M. Zorzi, "Underwater sensor networks: applications, advances and challenges," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 370, no. 1958, pp. 158–175, 2012.

[8] G. Cabri, L. Leonardi, M. Mamei, and F. Zambonelli, "Location dependent services for mobile users," *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, vol. 33, no. 6, pp. 667–681, 2003.

[9] Y. Liu and Z. Yang, *Localization, and Localizability Location-Awareness Technology for Wireless Networks*, Springer, 2011.

[10] C. Wu, Z. Yang, Y. Liu, and W. Xi, "WILL: wireless indoor localization without site survey," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 4, pp. 839–848, 2013.

[11] G. Blumrosen, B. Hod, T. Anker, D. Dolev, and B. Rubinsky, "Enhancing RSSI-based tracking accuracy in wireless sensor networks," *ACM Transactions on Sensor Networks*, vol. 9, no. 3, article 29, 2013.

[12] J. Zhao, W. Xi, Y. He et al., "Localization of wireless sensor networks in the wild: pursuit of ranging quality," *IEEE/ACM Transactions on Networking*, vol. 21, no. 1, pp. 311–323, 2013.

[13] J. Gribben and A. Boukerche, "Location error estimation in wireless ad hoc networks," *Ad Hoc Networks*, vol. 13, pp. 504–515, 2014.

[14] K. Bu, Q. Xiao, Z. Sun, and B. Xiao, "Toward collinearity-aware and conflict-friendly localization for wireless sensor networks," *Computer Communications*, vol. 35, no. 13, pp. 1549–1560, 2012.

[15] D. Kotzor and W. Utschick, "Kernel methods for ill-posed range-based localization problems," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 4151–4162, 2012.

[16] G. Han, H. Xu, T. Q. Duong, J. Jiang, and T. Hara, "Localization algorithms of Wireless Sensor Networks: a survey," *Telecommunication Systems*, vol. 52, no. 4, pp. 2419–2436, 2013.

[17] H. A. de Oliveira, A. Boukerche, E. F. Nakamura, and A. A. F. Loureiro, "An efficient directed localization recursion protocol for wireless sensor networks," *IEEE Transactions on Computers*, vol. 58, no. 5, pp. 677–691, 2009.

[18] S. Tian, X. Zhang, X. Wang, P. Sun, and H. Zhang, "A selective anchor node localization algorithm for wireless sensor networks," in *Proceedings of the 2nd International Conference on Convergent Information Technology (ICCIT '07)*, pp. 358–362, Gyeongju-si, Republic of Korea, November 2007.

[19] S.-W. Cheng, T. K. Dey, H. Edelsbrunner, M. A. Facello, and S.-H. Teng, "Sliver exudation," *Journal of the ACM*, vol. 47, no. 5, pp. 883–904, 2000.

[20] J. Sarrate, J. Palau, and A. Huerta, "Numerical representation of the quality measures of triangles and triangular meshes," *Communications in Numerical Methods in Engineering with Biomedical Applications*, vol. 19, no. 7, pp. 551–561, 2003.

[21] N. Chunge, L. Jianfei, and S. Shuli, "Study on quality measures for tetrahedral mesh," *Chinese Journal of Computational Mechanics*, vol. 20, no. 5, pp. 579–582, 2003.

[22] J. R. Shewchuk, *Lecture Notes on Delaunay Mesh Generation*, University of California, Berkeley, Calif, USA, 2012.

[23] C. Poggi and G. Mazzini, "Collinearity for sensor network localization," in *Proceedings of the 58th IEEE Vehicular Technology Conference (VTC '03)*, pp. 3040–3044, 2003.

[24] L. F. Wu, M. Q. H. Meng, and H. W. Liang, "A Collinearity-Based Localization Algorithm for wireless sensor networks," *Chinese Journal of Sensors and Actuators*, vol. 22, no. 5, pp. 723–727, 2009.

[25] L. Ke-Zhong, W. Shu, H. Fu-Ping, and X. Chang-Chun, "An Improved DV-hop localization algorithm for wireless sensor networks," *Information and Control*, vol. 35, no. 6, pp. 787–792, 2006.

[26] S.-W. Cheng, T. K. Dey, and J. Shewchuk, *Delaunay Mesh Generation*, Chapman & Hall, CRC, 2012.

[27] Zhouyan, *The Optimizing Selection and Error Analysis of Position Reference Nodes in Smart Space*, Northeastern University, Shenyang, China, 2008.

[28] Y. Zhou, H. Zhao, J. Zhang, and H.-C. Li, "Location error analysis of pervasive computing," *Acta Electronica Sinica*, vol. 37, no. 2, pp. 382–386, 2009.

[29] H. S. AbdelSalam and S. Olariu, "A 3d-localization and terrain modeling technique for wireless sensor networks," in *Proceedings of the 2nd ACM International Workshop on Foundations of Wireless Ad Hoc and Sensor Networking and Computing (FOWANC '09)*, pp. 37–45, May 2009.

[30] S. Sun and J. Liu, "An efficient optimization procedure for tetrahedral meshes by chaos search algorithm," *Journal of Computer Science and Technology*, vol. 18, no. 6, pp. 796–803, 2003.

[31] J. Zheng, Y. E. Sun, Y. Huang, Y. Wang, and Y. Xiao, "Error analysis of range-based localisation algorithms in wireless sensor networks," *International Journal of Sensor Networks*, vol. 12, no. 2, pp. 78–88, 2012.

[32] Q. Qingji, S. Xuanjing, and C. Haipeng, "An improved node localization algorithm based on DV-Hop for wireless sensor networks," *Computer Science and Information Systems*, vol. 8, no. 4, pp. 953–972, 2011.

[33] K. Månsson, G. Shukur, and P. Sjölander, "A new ridge regression causality test in the presence of multicollinearity," *Communications in Statistics: Theory and Methods*, vol. 43, no. 2, pp. 235–248, 2014.

[34] C. F. Dormann, J. Elith, S. Bacher et al., "Collinearity: a review of methods to deal with it and a simulation study evaluating their performance," *Ecography*, vol. 36, no. 1, pp. 27–46, 2013.

[35] A. Wibowo and M. I. Desa, "Kernel based regression and genetic algorithms for estimating cutting conditions of surface roughness in end milling machining process," *Expert Systems with Applications*, vol. 39, no. 14, pp. 11634–11641, 2012.

[36] R. A. P. Dias, J. Petrini, J. B. S. Ferraz et al., "Multicollinearity in genetic effects for weaning weight in a beef cattle composite population," *Livestock Science*, vol. 142, no. 1–3, pp. 188–194, 2011.

[37] S.-M. Huang and J.-F. Yang, "Improved principal component regression for face recognition under illumination variations," *IEEE Signal Processing Letters*, vol. 19, no. 4, pp. 179–182, 2012.

[38] H. Ae, "Application of ridge analysis to regression problems," *Chemical Engineering Progress*, vol. 58, pp. 54–59, 1962.

[39] I. K. Fodor, *A Survey of Dimension Reduction Techniques*, 2002.

[40] S.-H. Fang and T.-N. Lin, "Principal component localization in indoor WLAN environments," *IEEE Transactions on Mobile Computing*, vol. 11, no. 1, pp. 100–110, 2012.

[41] H. Hyötyniemi, *Multivariate Regression—Techniques and Tools*, Helsinki University of Technology, Control Engineering Laboratory, Helsinki, Finland, 2001.

[42] D. Niculescu and B. Nath, "DV based positioning in Ad Hoc networks," *Telecommunication Systems*, vol. 22, no. 1–4, pp. 267–280, 2003.

[43] D. Ma, M. J. Er, B. Wang, and H. B. Lim, "Range-free wireless sensor networks localization based on hop-count quantization," *Telecommunication Systems*, vol. 50, no. 3, pp. 199–213, 2012.