

An Improved DV-Hop Localization Algorithm for Wireless Sensor Networks

Chengping Peng^a, Guangcong Liu^b

School of Computer, Guangdong University of Technology, Guangzhou 510000, China

^apechpi@163.com, ^bliugc@gdut.edu.cn

Keywords: Wireless sensors networks, Localization, DBSCAN, Collinearity, DV-Hop.

Abstract. Considering the topology of real distance of the anchor node and the real trilateration composed of the unknown positioning node, an improved DV-Hop localization algorithm for wireless sensor networks was proposed. Firstly, unknown node will be measured by the trilateration-based triangle. Any of the three anchor nodes which meet collinear degree within a predetermined range, will be divided into one group. According to multiple groups of location estimation, finally by using clustering analysis, the result of the location could be optimized by selecting the largest cluster, removing the noise points and boundary points and leaving the core point with small errors. In the end, the unknown node estimation of location will be obtained.

1. Introduction

Wireless Sensor Networks (WSNs) ^[1] consist of a large number of sensor nodes, each sensor can have the ability to sense object, process data, and communicate with other nodes. The sensor node collects the surrounding physical information which is closely related with the location information. Only when the collected physical information combined with their own position information, the collected information can be meaningful ^[2]. Obviously, positioning technology in the sensor network is very important. In the process of positioning, the positioning algorithm is divided into two categories according to whether it is necessary to carry out distance measurement: positioning algorithm based on distance measurement and positioning algorithm without distance ^[3]. The positioning algorithm based on distance measurement uses the node's own hardware to obtain the signal intensity, azimuth angle or response time to realize the distance measurement between points.

In data mining, a large number of data is often used for clustering analysis, analysis of data distribution, characteristics exploring for each class, so that a specific set of collections on a more accurate centralized analysis can be conducted to find useful information. However, there are many clustering algorithms, one of which is classical DBSCAN clustering algorithm ^[7-9]. It is a density-based clustering based on a high-density connectivity region detecting. It can classify multiple classes of data. The similarity between the elements is the largest, the similarity between the elements is the smallest, and it can discover the clusters of arbitrary shapes in the environment with noise, and has strong clustering ability. Therefore, in the image pattern recognition, business intelligence application, Web search and many other fields have a wide range of applications ^[10].

In this paper, the density-based clustering analysis method is applied to the wireless sensor network location. Combining the DBSCAN clustering method with the DV-Hop localization algorithm, a DV-Hop localization algorithm based on DBSCAN clustering method with high positioning accuracy is proposed. DB DV-Hop algorithm, which uses the trilateral measurement method to obtain a set of multiple positioning results by using the anchor nodes, and then cluster analysis of a large number of positioning results to get the coordinates of the unknown node.

2. DV-Hop algorithm and triangulation method

2.1 Three Stages of the DV-Hop algorithm positioning process

The first stage: This stage is called distance vector exchange: each anchor node broadcasts its own location information packet to the surrounding neighbor node, including hop number information, anchor node number and its coordinate position. In the flooding broadcast, through the distance vector exchange protocol, the network of each anchor node can obtain the information of the minimum number of hops of all other anchor nodes.

The second stage: broadcast the correction value stage. Through the first stage, the minimum number of hops information obtained, and then according to the actual distance between the anchor nodes, you can find the average distance of each anchor node, (1) And then sends it as a correction value to the network through broadcast. The unknown node of the whole network only accepts the correction value of the nearest anchor node. Finally, the unknown node calculates the estimated distance from each anchor node by the correction value, (2).

$$HopSize_i = \sum_{i \neq j} \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} / \sum_{i \neq j} h_{ij} \quad (1)$$

$$d_{ij} = h_{ij} \times HopSize_i \quad (2)$$

The third stage: position calculation phase. When the unknown node obtained three or more anchor nodes to estimate the distance information, the trilateral measurement method or maximum likelihood estimation method would be conducted to calculate the unknown node coordinate.

2.2 Trilateral measurement method

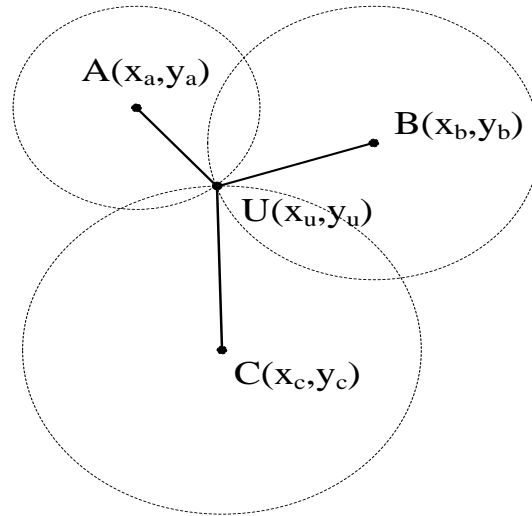


Fig.1 The sketch map of trilateration

As shown in Fig. 1, the three sides AU, BU and CU and three anchor nodes A (x_a, y_a), B (x_b, y_b) and C (x_c, y_c) are known,

$$\begin{cases} (x_a - x_u)^2 + (y_a - y_u)^2 = AU^2 \\ (x_b - x_u)^2 + (y_b - y_u)^2 = BU^2 \\ (x_c - x_u)^2 + (y_c - y_u)^2 = CU^2 \end{cases} \quad (3)$$

The first two equations in the above equation are subtracted from the third term, and after the deformation, the coordinate equation is used to represent the coordinate position:

$$\begin{bmatrix} x_u \\ y_u \end{bmatrix} = \begin{bmatrix} 2(x_a - x_c) & 2(y_a - y_c) \end{bmatrix}^{-1} \begin{bmatrix} x_a^2 - x_c^2 + y_a^2 - y_c^2 - AU^2 + CU^2 \\ x_b^2 - x_c^2 + y_b^2 - y_c^2 - BU^2 + CU^2 \end{bmatrix} \quad (4)$$

3. DB DV-Hop algorithm

3.1 Estimated distance calculation

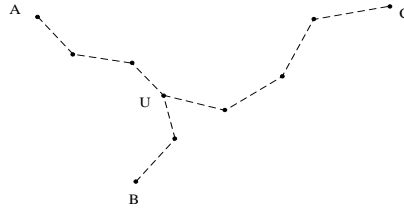


Fig.2 The sketch map of shortest path between the unknown node and anchor nodes

As shown in Fig. 2, the unknown node U is connected with the anchor nodes A, B and C, and from the near-far anchor nodes are B, A and C respectively. The average jump distance of the three anchor nodes is calculated by the formula (1). The larger the contribution of the anchor node to the U, the higher the similarity is. The higher the similarity, the higher the similarity. In this paper, we propose a method to calculate the weighting coefficient λ of the anchor node:

$$\lambda_i = \omega_i / \sum_{j=1}^m \omega_j \quad (5)$$

among them,

$$\omega_i = \exp \left(h_{\min} - h_{iu} / \sum_{j=1}^m h_{ju} \right) \quad (6)$$

Where h_{\min} is the number of hops between the anchor node i and the unknown node U, m is the number of anchor nodes involved in the location, and 3 is the number of hops. Thus, the average jump distance of the unknown node is obtained by the formula (7):

$$HopSize_u = \sum_{i=1}^m \lambda_i \cdot HopSize_i \quad (7)$$

The distance between the unknown node and the anchor node can be estimated by substituting the average jump distance in equation (2), after obtaining the average jump distance of the unknown node.

3.2 Multiple positioning for multiple set

Through the last estimated distance calculation method for the estimation of trilateral UA, UB and UC value, then using triangular trilateral AB, AC and BC, three times, using cosine theorem, $\angle AUB$, $\angle AUC$ and $\angle BUC$ to determine the sum of a triangle relations with 2π . If the sum of the triangle is equal to 2π (the actual range can meet the error range can be equal to the processing), according to the unknown node in the triangle within the positioning method to locate, and vice versa, according to the positioning node in the triangle positioning method outside the positioning.

In order to reduce the error generation, improve the positioning accuracy of the trilateral measurement method, select the "excellent" anchor nodes, anchor nodes must meet the following conditions:

1. The three anchor nodes cannot be on a straight line or an approximate straight line, that is, the collinearity of the three points is less than a given threshold k (called the collinearity upper limit k).

2. Any two nodes in the three anchor nodes can not rely on too close, especially from the unknown node.

In the case of three points in a straight line, the concept of covariance is introduced. Document[11] defines the covariance as the cosine of the smallest interior angle in the triangle consisting of three anchor nodes. Since the smallest inner angle in the triangle is $[0, \pi/3]$, Therefore, the covariance range.

4. DBSCAN clustering model

4.1 DBSCAN clustering algorithm description

1. randomly select a point object that is not accessed and does not belonging to any cluster, whether the number of dots in the detected Eps field is greater than or equal to MinPts, and if so, add them to the cluster ClassID together with the point object, And then tag the object to been accessed.

2. Select the unexposed point object from the cluster ClassID and determine if the density is less than the threshold MinPts, or less than the border point object. If not, it is the core point object and all points in the Eps field are added to the cluster ClassID. Recursively select unsearched point objects from the clustered ClassID until all the point objects in the clustered ClassID are accessed, and the clustering ClassID is fully generated.

3. Go to the first step to continue executing the algorithm repeatedly to find other clusters; if a point object does not belonging to any cluster and has not accessed, it is treated as a noise point object and marked as accessed.

4. If all the point objects in the point object are processed and there are no unexposed point objects, the algorithm ends.

4.2 DBSCAN positioning optimization

1. In order to reduce the computational complexity, the parameter MinPts in the algorithm is set to 4 [14], and the radius of the domain is the same as that of the unknown node. Eps:

$$Eps = \sqrt{MinPts \cdot prod(\max(o) - \min(o)) \cdot gamma(2)/m \cdot \pi} \quad (8)$$

2. The DBSCAN algorithm randomly selects an unattached point S from the set of data point objects, detects whether the number of point objects contained in the field of S is less than MinPts, and if not less than MinPts, point S is the core point, and S the class number is marked as ClassID, and then find all points about Eps and MinPts that are reachable from S, and mark their class numbers as ClassID. If less than MinPts in the field of S, the point S is temporarily marked as a noise point.

3. Search for all points on Eps and MinPts that are reachable from S, loop through the core points of the next cluster, and continually expand the number of core points in the cluster until no new core points are added, and finally form a cluster class named ClassID. If there are unattached data objects in the data object collection, go to step 2 until all point objects are accessed.

4. Find the cluster with classID with the largest number of data points in multiple clusters, and remove the boundary points from them, leaving only the core points in the cluster and the average of all the core points then position of the unknown node finally estimated.

5. Summary

Based on the DBSCAN clustering algorithm in the field of data mining, this paper explores and studies the positioning process of DV-Hop algorithm. Considering that the real distance between anchor nodes can be used as a known condition, it has not been fully utilized. Finally, the DBSCAN clustering algorithm is used to effectively filter the estimated position, so as to obtain a more accurate coordinate of the unknown node, and then the DBSCAN clustering algorithm is used to select the estimated position. The experimental results show that the DB DV-Hop algorithm proposed in this paper has higher positioning accuracy than the DV-Hop algorithm and is not affected by the individual positioning error. However, the algorithm has the advantages of communication and computation between nodes, resulting in large energy consumption of the nodes, thus reduced the life of the network. Therefore, the next step will be to study how to control the traffic and balance the energy consumption between the nodes while maintaining the high positioning accuracy..

References

- [1] Sun Limin, Li Jianzhong, Chen Yu, etc .. Wireless sensor network [M]. Tsinghua University Press. 2005.5: 148-154.

- [2] Angelo Cenedese, Giulia Ortolan, Marco Bertinato. Low-Density Wireless Sensor Networks for Localization and Tracking in Critical Environments [J]. IEEE Transactions on Vehicular Technology, 2010, 59(6) :2951-2962.
- [3] Mert B, Liu M, Shen W M, et al. Localization in cooperative wireless sensors networks: A review [C]. Proceedings of the 2009 13th International Conference on Computer Supported Cooperative Work in Design, 2009: 438-443.
- [4] Bulusu N, Heidemann J, Estrin D. GPS-Less low cost outdoor localization for very small devices [J]. IEEE Personal Communication Magazine, 2000, 7(5) : 28-34.
- [5] He T, Huang C, Blum B M, et al. Range-Free localization schemes for large scale sensor networks [C] //Proc of the 9th Annual Int'l Conf on Mobile Computing and Networking. San Diego:ACM Press, 2003, 81-95.
- [6] Niculescu D. Nath B. DV-based positioning in ad hoc networks [J]. Telecommunication Systems, 2003, 22 (1): 267~280.
- [7] Martin Ester, Hans-Peter Kriegel, Jorg Sander, et al. A density-based algorithm for discovering clusterings in large spatial databases with noise [A]. //Proceedings of the 2th international conference on knowledge discovery and data mining[C], (KDD-96).
- [8] YU Ya-fei,ZHOU Ai-wu. An Improved Algorithm of DBSCAN[J]. Computer Technology and Development, 2011, 21(2): 30-33.
- [9] ZHU Xuanzhang. Location method based on DBSCAN in wireless sensor networks[J]. Computer Engineering and Applications, 2013, 49(11): 80-83.
- [10] ZHANG Li-jie. Stable saturation density of DBSCAN algorithm[J].Application Research of Computers,2014,31(7): 1972-1975.
- [11]WU Ling-fei, MENG Qing-hu,LIANG Hua-wei.A Collinearity-Based Localization Algorithm for Wireless Sensor Networks[J],2009,22(5): 722-727.