Journal of Computers Vol. 31 No. 4, 2020, pp. 187-197
doi:10.3966/199115992020083104014

A Multi-hop Localization Through Model Selection for Irregular Networks

Ru-Lin Dou[1*], Xu-Ming Fang[1], Ya-Nan Liu[1], Xiao-Hui Mo[2]

[1] School of Software Engineering, Jinling Institute of Technology, Nanjing, China
{drl, xmfang, yanan.liu}@jit.edu.cn

[2] School of Computer Engineering, Jinling Institute of Technology, Nanjing, China
mxh407@jit.edu.cn

Received 10 March 2020; Revised 6 April 2020; Accepted 6 May 2020

Abstract. Identifying the location of nodes in wireless networks is a major challenge because its accuracy impacts the efficiency of location-aware protocols and applications. This paper presents a novel localization scheme called Multi-hop Localization through Model Selection (MLMS) which can significantly improve localization accuracy in different irregular networks. The proposed scheme contains three steps: data collection, establishing the skeleton model, and location estimation. In the data collection step, the proximity information of the irregular network is collected. In the establishment of the skeleton model step, the skeleton model among the practical distances and the proximity among nodes is constructed using regression which is supervised by Bayesian Information Criterion (BIC). Specifically, the skeleton model when the BIC value is the smallest is deemed to be optimal. In the location estimation step, any unknown node calculates its location in a distributed manner using the optimal skeleton model. The simulation results demonstrated that the proposed scheme can greatly reduce the estimation error and quickly achieve estimation location in networks with different irregular topologies. Simulation results show that, compared with similar algorithms recently reported, MLMS improves localization accuracy by more than 32 %.

Keywords: Bayesian information criterion, irregular networks, model selection, multi-hop localization, skeleton model

1 Introduction

The interactions between man and nature will change drastically when 5G mobile communication technology officially enters in commercialization phase and realizes inter-person, inter-thing, and person/thing information transmissions without the limit of time and space [1-3]. The location information of mobile devices is the premise and foundation of developing other 5G applications, as it provides the location information of monitored events or traced objectives, and offers technical support for improving routing efficiency of the network, optimizing network coverage, realizing topology control, etc. According to the previous researches, wireless localization can be divided into single-hop localization and multi-hop localization [4]. The single-hop localization relies on direct communications between the node to be localized and the reference node (a node with known location) and a typical example is satellite localization. However, mankind spends more than 80% of time in complex and closed spaces (e.g., buildings in cities, forests, etc.) [5]. In such environments, satellites are difficult to directly connect with mobile terminals, and relay nodes are needed to assist them in information exchange between satellites and the mobile terminals. Localization that relies on relay nodes for information exchange is defined as multi-hop localization.

In multi-hop localization, the location of the node to be localized relative to the reference node (also known as the anchor node) is obtained first, and then the absolute location information is obtained

* Corresponding Author

through exchanging information with wireless network and collaboration. Multi-hop localization can be further divided into range-based and range-free [4]. Range-based multi-hop localization requires installation of hardware devices on nodes, then the specific location is found out after figuring out estimated distance between the unknown nodes and the anchors through a wireless measurement technology (such as radio signal strength (RSS), time of arrival (ToA), and time difference of arrival (TDoA) [4]). Range-free multi-hop localization relies on connectivity information between nodes (e.g., hop count) to estimate distances between nodes and requires no installation of hardware devices. The method of using hop count to describe distances between nodes is easy to implement and has low requirements on hardware. It is very suitable for localization in large-scale scenarios. Over the past years, various multi-hop localization methods have been proposed [8-14]. However, most of previous methods [8-9] are based on assumptions that the network topology is regular and no obstacles are present in the area. In other words, no large holes are present in the network topology. In practical networks, however, the presence of obstacles leads to failed node communications, resulting in holes in the network topology. Hence, the network topology is usually irregular in practical cases. Assuming two-node pairs (N_1/N_3 pair and N_2/N_4 pair) with similar Euclidean distances ($d_{N_1N_2} \simeq d_{N_3N_4}$, see Fig. 1) are present in the network, and there are huge differences ($h_{N_1N_2} \gg h_{N_3N_4}$) in the hop distance of information transmission in their routes due to the presence of obstacles. If some unknown node adopts these two node pairs for location estimation, the estimation results with a substantial difference will be obtained.

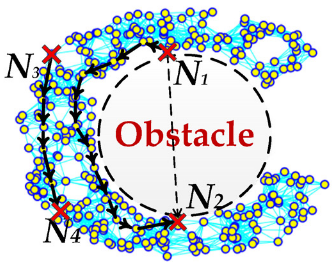


Fig. 1. Irregular network

Vural et al. [6] conducted a series of experiments and found that distances corresponding to different hop counts are homogeneously distributed. If both anchors and unknown nodes are present in the network, hop counts and distances are in one-to-one correspondence. Based on the previous studies and practical distributions of anchors and unknown nodes, we propose a novel Multi-hop Localization through Model Selection (called MLMS). The propose MLMS algorithm is composed of three steps: the data collection step, establishing the skeleton model step, and location estimation step. First, the shortest-path (hop-counts) of the given network is collected. Second, the relationships between practical distances and proximity among nodes are modeled using a regression approach which is supervised by the Bayesian Information Criterion (BIC). Finally, each node achieves its location by the multilateration approach in a distributed manner. To guarantee trade-off model accuracy and estimation accuracy during establishing the skeleton model, Bayesian Information Criterion (BIC) [7] was employed for model selection in this paper. In this way, the correlations of variables were effectively considered, efficiency and accuracy of localization were guaranteed, and the proposed algorithm applies to various network environments.

The rest of the paper is organized as follows: In section 2, we review some of the existing multi-hop algorithms. We present the network model and propose our scheme in section 3. In section 4, we discuss simulation setup and parameters and analyze the simulation results of the proposed methods with those of previous methods via relevant analysis using plots and tables. We conclude the paper in section 5.

2 Literature Review

The DV-hop localization algorithm [8] is currently the most widely recognized multi-hop localization method. In DV-hop, the location information of anchors was spread over the entire network via broadcast and the average per-hop distance in the network was computed based on the hop counts and the practical distance between different anchors. Then, the product of the average per-hop distance and hop count between the anchor and the unknown node is viewed as their estimated distance. Finally, multilateration is employed to estimate the location of the unknown node. Despite its high efficiency and low cost, the DV-hop algorithm has not been widely applied as it doesn't consider the effects of irregular topology.

Another approach is Selective 3-Anchor DV-hop (Selective-3) [9]. The idea is to choose the best 3 anchors to compute the coordinates of an unknown node. Specifically, three anchors are selected according to the connectivity between nodes and used them to estimate the location of the unknown node. However, the Selective-3 is not convincing as its method of obtaining the optimal 3 anchor nodes comes from a special case, rather than a complete theoretical derivation. Indeed, simulations based on codes provided in that author's doctoral dissertation revealed that the estimation accuracy of the Selective-3 algorithm is superior to the ordinary DV-hop algorithm only. Meanwhile, the Selective-3 algorithm is limited by complicated calculations in the selection of anchors, resulting in long calculation time. Hence, the Selective-3 algorithm is not an ideal option for real-time calculations.

Xiao et al. [10]. proposed the Reliable Anchor-based Localization (RAL) algorithm. In RAL, the expected progress in the amorphous method is analyzed and a lookup containing the correlation of hop count and minimum hop length is designed [11]. In this lookup, different hop counts are set to match different expected progress. In the process of localization, the RAL algorithm eliminates the unreliable anchors (with large accumulated errors) and only chooses reliable anchors to carry out localization by searching and comparing the elements in the lookup. If reliable anchors near the node to be estimated are insufficient, the RAL algorithm may employ unknown nodes that have been localized as new anchors for localization. Yan et al. [12] considered that irregular networks caused heteroscedasticity problems during distance estimation, and then they proposed an optimal weighted DV-hop method (OWDV-hop). In this approach, localization results are ensured to be optimal by using optimal weight functions to eliminate heteroscedasticity and using a geometric constraint algorithm to correct the outlier estimation location. Recently, the Accurate Analytical-based Multi-hop Localization (AAML) is presented in [13]. In AAML, distances between any two nodes are estimated using the local distribution with nodes linked in pairs, then the preliminary estimation of unknown nodes is conducted by using a weighted hyperbolic equation, and at last, anomalous location estimations are corrected with the help of weighted Taylor series and geometric constraint algorithm. Nevertheless, the AAML algorithm heavily depends on the distribution density of nodes. In other words, the AAML algorithm can only deliver good results at high density of nodes.

It has one of the research hotspots in recent years that improve localization performance using machine learning techniques. Machine learning-based localization approaches utilize the correspondence between the distribution character of the anchors and metrical information and correlate and build a skeleton model to estimate the locations of the unknown nodes. Compared with the conventional methods, machine learning-based methods can effectively clarify the information behind data such as network topology and node correlation thus, effectively improving the localization efficiency. Lee et al. [14]. proposed LSVR and MSVR, which are two Support Vector Regression (SVR)-based multi-hop range-free localization methods. In LSVR and MSVR, the relation between hop count and practical distance is non-linear. Distances between nodes are predicted using SVR which is based on kernel technique. According to the principle of Occam's razor [15], the more the complicated model, the higher is the chance of overfitting. Hence, SVR-based localization methods may be exposed to overfitting with the increased quantity of anchors, resulting in degraded localization accuracy. Additionally, SVR-based localization methods are readily exposed to complicated calculations and the kernel parameters in the kernel functions are manually set, resulting in poor applicability of SVR-based localization methods.

3 Network Model and the Proposed Algorithm

3.1 Network Model

Suppose there are n nodes deployed randomly in a flat space. They form a linked network through ad hoc network technology, in which the first m ($m \ll n$) nodes known their locations, and the rest n – m nodes are unknown nodes to be localized. Define the real coordinates of the anchor a as c_a, where

$a \in \{1,2,\cdots,m\}$; define the real location of the unknown nodes to be estimated as c_u, where $u \in \{m+1, m+2, \cdots, n\}$.. Without loss of generality, the following assumptions can be made:

(1) All nodes belong to the same type, which means they have identical communication radius r.

(2) The communication range of a node is a circle with the node as the center and r as the radius.

(3) Nodes can get the information of their adjacent nodes by sending and forwarding messages, adjacent relation model is $R_i = \{ j | j \neq i \&\& d_{ij} < r \}$. Herein, R_i refers to the adjacent node collection of node i, i and j refer to any two nodes in the network, d_{ij} refers to the Euclidean distance between two nodes, which is defined as:

$$d(i,j) = d_{ij} = \left(\sum_{i,j=1}^{n} \left(\boldsymbol{c}_i - \boldsymbol{c}_j \right)^2 \right)^{\frac{1}{2}} \tag{1}$$

If any node i and j are linked, their metric information can be obtained, which is minimum hop count, shown as $h(i,j) = h_{ij} \triangleq \{0,1,2,\cdots\}$. The minimum hop count vector of node i and other linked nodes in the network is $\boldsymbol{h}_i = \left[h_{i1}, h_{i2}, \cdots, h_{in} \right]^{\mathrm{T}}$, and the minimum hop count matrix among nodes in the whole network is $\boldsymbol{H} = \left[\boldsymbol{h}_1, \boldsymbol{h}_2, \cdots, \boldsymbol{h}_n \right]$. Thus, multi-hop localization problem can be abstracted to the process of solving estimated coordinates $\hat{\boldsymbol{c}}_u$ of the unknown node u under the constraints of the coordinates c_a of the anchor a and hop count h_{au} between the anchor a and the unknown node u. Herein, the process of distance estimation between anchors and unknown node can be abstracted to a mapping function: $f : h(a,u) \rightarrow \hat{d}(a,u)$, where $\hat{d}_{a,u}$ refers to the estimated distance between the anchor a and the unknown node u. The estimated location $\hat{\boldsymbol{c}}_u$ of the unknown node u) is expected to approach c_u.

3.2 Multi-hop Localization Through Model Selection

Node localization in the MLMS algorithm consists of three stages: data collection, model development and location estimation.

Step 1 Data collection. Hop counts and distances between nodes are collected for model development. For the entire network, the broadcast messages sent from any node $i, i \in \{1,2,\cdots,m\}$ to other nodes in the network contain its number, hop count information and coordinates (if it is an anchor). When the program runs for some time, every node in the network has got the minimum hop count with all the other nodes. Therefore, there is communication cost $O(n^2)$ in data collection phase. In this way, any anchor in the network can form a global matrix (H) with minimum hop count between anchors and a global matrix (D) with minimum distances between anchors.

Step 2 Establishing the skeleton model. Through data collection, any anchors $N_i, i \in \{1,2,\cdots,m\}$ collect other (m-1) hop count and distance data pair $\left(\boldsymbol{h}_k, d_{i,k} \right), k = 1,\cdots,m \cap k \neq i$. Assuming that the constraints for estimation of inter-node distances is a mapping function:

$$f_i\left(\boldsymbol{h}_j \right) = \boldsymbol{\theta}_i^{\mathrm{T}} \boldsymbol{h}_j + e_i \tag{2}$$

where, $\boldsymbol{\theta}_i$ refers to the estimation model for distance between anchor N_i and other (m-1) anchors. To prevent the problem of inconsistent quantity level when the scale of network, node communication radius and other factors change, and the hop count and physical distance are converting, the standardized processing is done to data pair ($\boldsymbol{h}_k$, $d_{i,k}$) in advance.

Fig. 2 shows a typical multi-hop network. As illustrated, from N_1 to N_2, node 1→2→3→4→5 needs to be passed, if $\boldsymbol{N}_1$ and $\boldsymbol{N}_2$ are source node and objective node, respectively. Therefore, there is correlation in the hop counts between nodes and vectors in the hop count matrix of the hop count/distance conversion model are very likely to have correlations. In this case, constraints are needed during model development (*i.e.*, model penalty is needed).

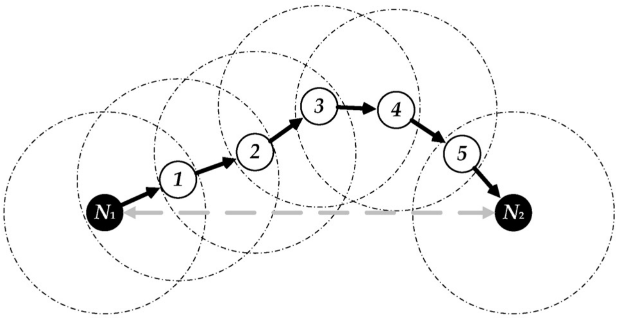


Fig. 2. Data exchange between nodes

The objective function of distance estimation model shall be:

$$\boldsymbol{L}(\lambda,\boldsymbol{\theta}_i)=\|\boldsymbol{d}-\boldsymbol{H}\boldsymbol{\theta}_i\|^2+\lambda\|\boldsymbol{\theta}_i\|^2 \tag{3}$$

where L refers to a likelihood function, λ refers to the penalty parameter.

The hop count distance mapping model is:

$$\hat{\boldsymbol{\beta}}_{\mathrm{i}}=\arg\min\left\{\|\boldsymbol{d}_i-\boldsymbol{H}\boldsymbol{\beta}_i\|^2+\lambda\|\boldsymbol{\beta}_i\|^2\right\}=\left(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}+\lambda\boldsymbol{I}\right)^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{d}_i \tag{4}$$

Herein, the effects of the penalty term are proportional to the value of λ. If λ→0, Eq 4 degrades to a general least square. As shown in Fig. 2, nodes are related to each other and λ = 0 is not valid. If λ→∞, the loss function contains the penalty term only and its minimization inevitably leads to $\boldsymbol{\theta}_i=0$. Hence, the selection of λ is the key step in establishment of accurate mapping correlation of hop count and distance. The mapping correlation model of hop count and distance was obtained by minimizing residual sum of squares of minimized hop count and physical distance of anchors. Intuitively, increasing the quantity of anchors leads to enhanced accuracy of model β. Nevertheless, increasing quantity of anchors will lead to increased system cost and exacerbated variance of β. In other words, increasing quantity of anchors leads to degradation of estimation accuracy in the following localization stage. Therefore, λ with balanced *a* can only be obtained by balancing residual and variance via model selection. Among criteria commonly used for model selection (AIC and BIC), BIC was employed for hop count-distance mapping model selection. BIC is defined as:

$$BIC=-2L+p\ln n \tag{5}$$

where, p refers to the quantity of parameters, L refers to a likelihood function and n refers to the quantity of anchors. Its single calculation complexity is O (1).

The literature [7] points out that the λ corresponding to minimum BIC was regarded as the optimized one. In the operation process, select BIC through Newton-Raphson method iteration, until it reaches minimum. The complexity of single calculating BIC is O (1).

Step 3 Location estimation. In the development of mapping model, any unknown node in the network corresponds to m optimized mapping models of hop count and distance. Once hop counts between the unknown node and anchors connected to this node are obtained, the distances between the unknown node and anchors can be estimated. Define m mapping models received by the unknown node (u) as $\left(\hat{\boldsymbol{\theta}}_1,\hat{\boldsymbol{\theta}}_2,\cdots,\hat{\boldsymbol{\theta}}_m\right)$, hop counts between u and m anchors are $\left(h_{u,1},h_{u,2},\cdots,h_{u,m}\right)$ and estimated distances between u and m anchors are:

$$\hat{\boldsymbol{d}}_u=\left[\hat{d}_{u,1}\cdots\hat{d}_{u,m}\right]^{\mathrm{T}}=\hat{f}_1(\boldsymbol{h}_u)\cdots\hat{f}_m(\boldsymbol{h}_u)=\hat{\boldsymbol{\theta}}_1^{\mathrm{T}}\boldsymbol{h}_u\cdots\hat{\boldsymbol{\theta}}_m^{\mathrm{T}}\boldsymbol{h}_u \tag{6}$$

If $m\geq 3$, its own location is estimated in a distributed manner through multi-lateral estimation. The complexity of multi-lateral estimation is O (m^3) and m is the quantity of anchors. Fig. 3 shows the flow chart of the proposed MLMS algorithm.

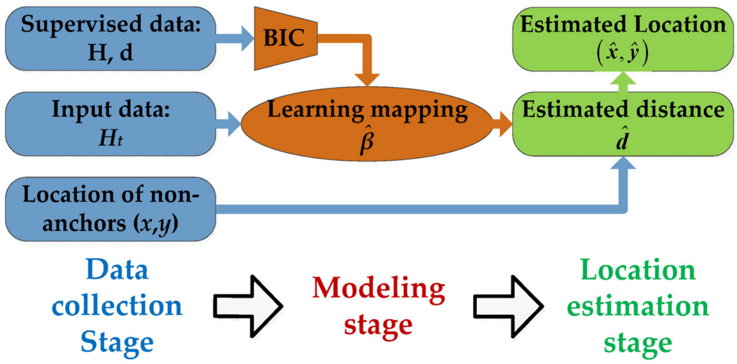


Fig. 3. The framework of MLMS. The framework of MLMS

The data is collected during the network initialization phase. After that, the mapping is firstly trained by BIC, using supervised data consist of the known hop-counts and physical distances. At last, in the location estimation phase, the physical distances of the unknown node are predicted by the learned mapping and estimate location using multilateral and anchor node coordinates.

4 Simulations Results

This section analyzes and evaluates efficiency of the MLMS algorithm through simulations. Comparison of the MLMS algorithm, RAL [10], Selective-3 [9], OW DV-hop [12] and LSVR [14] was done. Kernel parameters needed to be set for LSVR. These however, were manually set in the original paper, which made the algorithm lack self-adaptive ability, the self-adaptive setting of kernel parameter method [16] was adopted in this paper.

Simulations were realized by MATLAB and nodes are assumed to be randomly and homogeneously distributed in the irregular network topology. For each experimental condition, simulations were executed for 80 times and the results were statistically analyzed. To be fair, root mean square(RMS) error [17]was adopted as the basis for evaluation of average localization error. RMS is defined as:

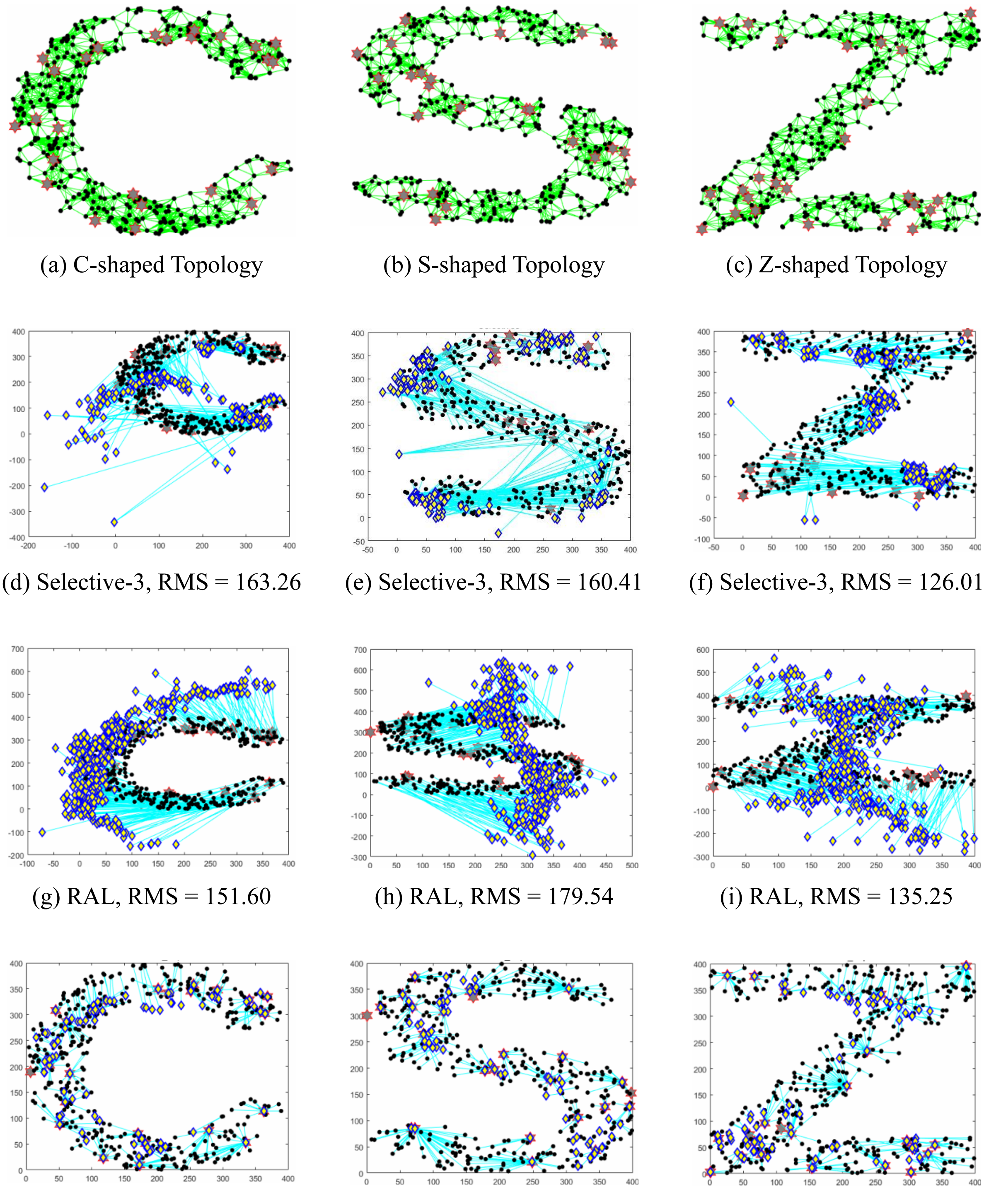
$$RMS = \sqrt{1/N_t \sum_{i=1}^{N_t}\left((\hat{x}_i - x_i)^2 + (\hat{y}_i - y_i)^2\right)} \quad \textbf{(7)}$$

where, $(\hat{x}_i, \hat{y}_i)$ refers to the estimated coordinates of the i-th node, (x_i, y_i) refers to the practical coordinates of the i-th node and N_t refers to the quantity of nodes to be localized.

Since network irregularity, anchors quantity, execution efficiency are the key parameters for evaluation of localization algorithm efficiency, efficiency evaluation and statistical analysis were conducted for these parameters.

4.1 Effects of Irregular Network Topology

Fig. 4 depicts typical network topological shapes for evaluation of irregular network localization efficiency. Table 1 shows the simulation parameters in this section.



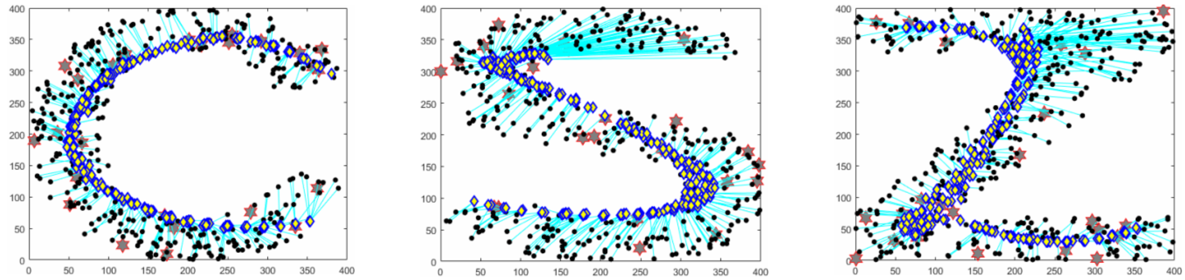
(a) C-shaped Topology (b) S-shaped Topology (c) Z-shaped Topology

(d) Selective-3, RMS = 163.26 (e) Selective-3, RMS = 160.41 (f) Selective-3, RMS = 126.01

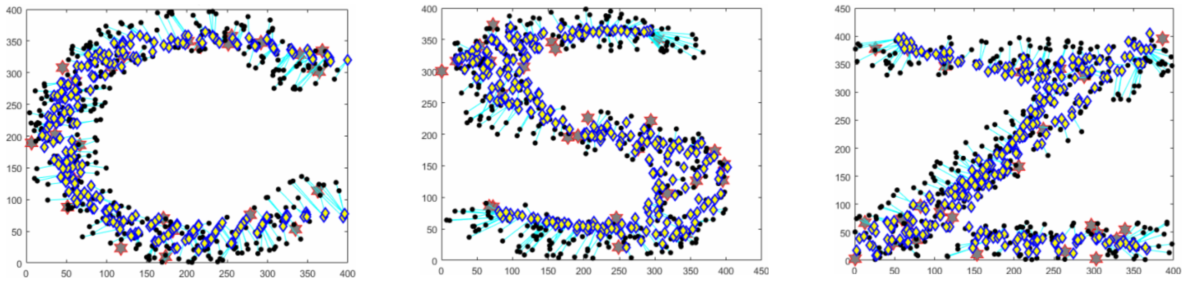
(g) RAL, RMS = 151.60 (h) RAL, RMS = 179.54 (i) RAL, RMS = 135.25

(j) OW DV-hop, RMS = 36.32 (k) OW DV-hop, RMS = 42.01 (l) OW DV-hop, RMS = 42.36

Fig. 4. The location estimation results for C-shaped, S-shaped and Z-shaped topologies by different algorithms



(m) LSVR, RMS = 37.39 (n) LSVR, RMS = 78.19 (o) LSVR, RMS = 60.27



(p) MLMS, RMS = 24.64 (q) MLMS, RMS = 26.77 (r) MLMS, RMS = 25.12

Fig. 4. The location estimation results for C-shaped, S-shaped and Z-shaped topologies by different algorithms (continue)

Table 1. Simulation parameters

Parameters	Value and description
Distribution areas and topological shapes of nodes	400*400, C-shaped, Z-shaped, S-shaped
Communication model and communication radius	Regular, R=40
Quantity and distribution of nodes	400, random, homogeneously distributed
Quantity of anchors	25

In Fig. 4(d) to Fig. 4(r) describes separately the localization results where the MLMS algorithm and four algorithms of same type exhibit same distribution under different network topologies. Herein, round shape indicates the practical location of the unknown nodes, diamond shape indicates estimated location of the unknown nodes, hexagram indicates anchors, real coordinates and their estimated coordinates of unknown nodes are linked with straight lines. The longer the lines the larger is the localization error.

As illustrated, the Selective-3 algorithm and the RAL algorithm exhibited poor resistance to irregular network topology, although distance estimation by the RAL algorithm had a certain degree of rationality compared with the Selective-3 algorithm. Therefore, there was a similarity between estimated locations and the network topology. The idea of distance estimation of Selective-3 came from some exceptions. Although its localization accuracy was higher than RAL, the estimated locations were chaotic. OW DV-hop and LSVR greatly improved the localization accuracy in the irregular network. LSVR treated the estimation of inter-node distances as regression prediction, but there were too many parameters in SVR, which led to over-training and overfitting of mapping the models, making the estimated location of unknown nodes on one of the curves in the network. OW DV-hop significantly improved the localization accuracy in irregular networks by weakening the heteroscedasticity problem resulted from the error accumulation. However, it sets anomalous estimated location as the closest anchors, resulting in high dependence on the number of anchors. The proposed MLMS algorithm has not only avoided the problems mentioned above but greatly improved localization accuracy (RMS was significantly lower than other algorithms). Compared with the Selective-3, RAL, OW DV-hop, and LSVR, the localization accuracy of proposed MLMS has respectively increased by about 84.91%, 83.75%, 32.17% and 34.11% in C-shaped network. Compared with the Selective-3, RAL, OW DV-hop, and LSVR, the localization accuracy of proposed MLMS has respectively increased by about 83.31%, 85.09%, 36.28% and 65.77%

in S-shaped network. Compared with the Selective-3, RAL, OW DV-hop, and LSVR, the localization accuracy of proposed MLMS has respectively increased by about 80.07%, 81.43%, 40.70% and 58.33% in Z-shaped network.

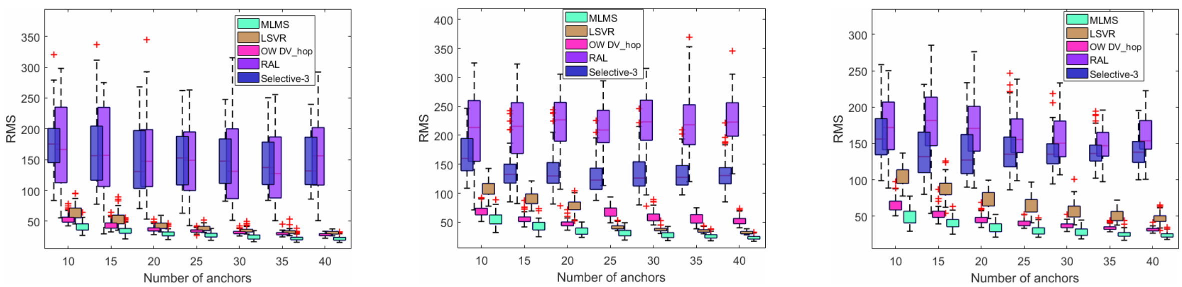
4.2 Effects of Quantity of Anchors

In RAL, OW DV-hop, LSVR and the proposed MLMS, anchors are needed while estimating the general distance per hop of nodes and also to assist the estimated location of the unknown nodes in the final location estimation. Therefore, anchors' quantity has vital influence on the efficiency of localization algorithm. This section examines influence of quantity of anchors on average localization error by setting different anchors quantity. Table 2 shows the simulation parameters in this section.

Table 2. Simulation parameters

Parameters	Value and description
Distribution areas and topological shapes of nodes	400*400, C-shaped, Z-shaped, S-shaped
Communication model and communication radius	Regular, R=40
Quantity and distribution of nodes	400, random, homogeneously distributed
Quantity of anchors	10, 15, …, 40(step sizes adopt 5)

Fig. 5 shows RMS distribution of different quantity of anchors described with boxplot after 80 experiments. It could be seen that when anchors quantity increased, the RMS median of OW DV-hop, LSVR and the proposed MLMS algorithm decreased monotonically, and among them the accuracy of MLMS was optimal and the localization efficiency was more stable (span of boxplot was small). When the quantity of anchors was different, RMS distribution of Selective-3 and RAL was wide and the RMS median did not decrease when the anchors quantity increased. Therefore, it can be concluded that Selective-3 and RAL are not the location estimation suitable for irregular networks.

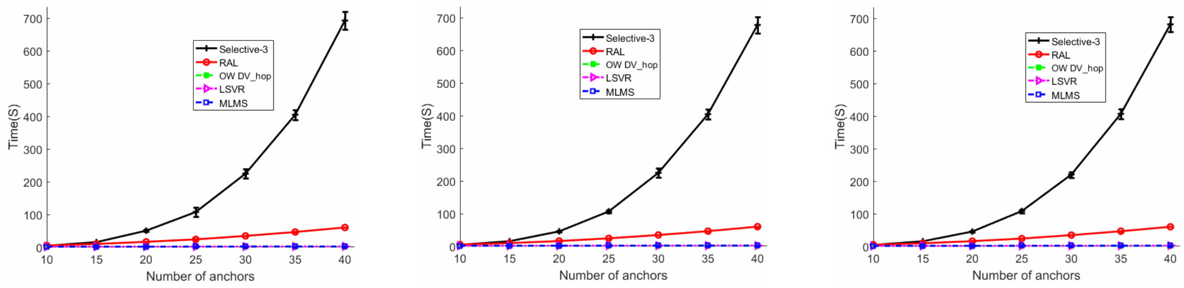


(a) C-shaped topology (b) S-shaped topology (c) Z-shaped topology

Fig. 5. Distributions of estimation errors of different algorithms with different quantities of anchors in

4.3 Execution Time

In this group of experiment, execution time distribution of Selective-3, RAL, OW DV-hop, LSVR and MLMS in 80 experiments in different anchors quantity and different network topological environment were compared. Experiments were arranged to be completed on DELL PC Server with Intel Xeon CPU E5-2630 v4 2.20 GHz, 32G RAM, Windows Server 2008 R2. All algorithms parameters of the simulation experiment are the same as those in Sect. 4.2.Each algorithm was run multiple times and the final statistical information was described with error bar graph. Fig. 6 shows the distributions of running times of different algorithms with different quantities of anchors in (a) C-shaped topology, (b) S-shaped topology, (c) Z-shaped topology. It is observed that Selective-3 adopted method of permutation and combination to select anchors and its calculation time increased rapidly as anchors' quantity increased. When anchors reached 40, the average calculation time of Selective-3 in C-shaped, S-shaped, Z-shaped topology were 692.39, 677.56, 681.43, respectively, which were far larger than other four algorithms of the same type. On the other hand, the proposed MLMS algorithm calculation time was better than RAL, close to OW DV-hop and LSVR, and was slightly more than the latter two algorithms.



(a) C-shaped topology (b) S-shaped topology (c) Z-shaped topology

Fig. 6. Distributions of running times of different algorithms with different quantities of anchors in

Table 3 provides an overview of the multi-hop localization algorithms in terms of accuracy, anchors amount and cost. All algorithms have their own characteristics and suitable applications.

Table 3. List of various localization algorithms for performance

Localization algorithm	Accuracy	Anchors amount	Cost
Selective-3	Low	Low	High
RAL	Low	Median	Low
OW DV-hop	High	Median	Low
LSVR	High	Median	High
MLMS	High	Median	Median

5 Conclusions

In this paper, we proposed an enhancement to multi-hop localization to compute localization estimation for irregular networks. The localization problem was formulated as a regression problem, and the solutions were proposed using the skeleton model which is supervised by the Bayesian Information Criterion (BIC). Only simple shortest-path (hop-counts) among the nodes was used as measurements, and each node estimated its own location in a distributed manner. This proposed algorithm relies on training of a skeleton model based on the correlations of distances and hop counts between anchors to effectively compensate heterogeneity of networks. During training, corrected BIC was introduced to evaluate the merits of from hop-count to distance conversion model, which improved the generalization capacity of model prediction and then increased localization accuracy. Simulations were conducted to compare the proposed algorithm with the previous algorithms under a variety of conditions including different irregular networks, various anchor populations and localization time. Compared with the previous algorithms, it was observed that the proposed algorithm exhibit better performance in terms of the localization accuracy and its stability.

Although the proposed algorithm can achieve higher localization performance, the proposed algorithm still has its disadvantage, namely sensitivity to the number of anchors. This will be the main direction of our future work. Besides, we only consider the two-dimensional localization in this paper, how to apply the proposed method to the three-dimensional spatial localization is also the future research direction.

Acknowledgements

This work was supported in part by the NSFC under Grants 61902163, the Major Project of the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (Grant No. 17KJD520003) and the Doctoral Scientific Research Startup Foundation of Jinling Institute of Technology (No. JIT-B-201639).

References

[1] B. Hu, S. Zhang, Z. Yuan, A new multi-hop localization based on l_2 constraint least square for anisotropic networks, Journal of Internet Technology 20(6)(2019) 1913-1923.

[2] Y. Liu, Z. Yang, X. Yan, G. Liu, B. Ho, A novel multi-hop algorithm for wireless network with unevenly distributed nodes, Computers, Materials & Continua 58(2019) 79-100. DOI: 10.32604/cmc.2019.03626.

[3] D. Gao, S. Zhang, F. Zhang, T. He, J. Zhang, RowBee: a routing protocol based on cross-technology communication for energy-harvesting wireless sensor networks, IEEE Access 7(2019) 4066340673.

[4] Z. Yang, C. Wu, Y. Liu, Location-based Computing: Localization and Localizability of Wireless Networks, Tsinghua Univ. Press, Beijing, China, 2014.

[5] S. Wang, J. Cao, P. Yu, Deep learning for spatio-temporal data mining: a survey. <https://arxiv.org/abs/1906.04928>, 2019.

[6] S. Vural, E. Ekici, On multihop distances in wireless sensor networks with random node locations, IEEE Transactions on Mobile Computing 9(4)(2010) 540-552. DOI: 10.1109/TMC.2009.151.

[7] G. Schwarz, Estimating the dimension of a model, The Annals of Statistics 6(2)(1978) 461-464.

[8] D. Niculescu, B. Nath, DV based positioning in ad hoc networks, Telecommun. Syst. 22(1-4)(2003) 267-280.

[9] L. Gui, T. Val, A. Wei, R. Dalce, Improvement of range-free localization technology by a novel DV-hop protocol in wireless sensor networks, Ad Hoc Networks 24(2015) 55-73.

[10] B. Xiao, L. Chen, Q. Xiao, M. Li, Reliable anchor-based sensor localization in irregular areas, IEEE Trans. Mobile Comput. 9(1)(2010) 60-72.

[11] R. Nagpal, H. Shrobe, J. Bachrach, Organizing a global coordinate system from local information on an ad hoc sensor network, in: F. Zhao, L. Guibas (Eds.), Information Processing in Sensor Networks, Springer, 2003, pp. 333-348.

[12] X. Yan, L. Sun, Z. Sun, J. Zhou, A. Song, Improved hop-based localisation algorithm for irregular networks, IET Communications 13(5)(2019) 520-527. DOI: 10.1049/iet-com.2018.5583.

[13] X. Yan, J. Cao, L. Sun, J. Zhou, S. Wang, A. Song, Accurate analytical-based multi-hop localization with low energy consumption for irregular networks, IEEE Transactions on Vehicular Technology 69(2)(2020) 2021-2033.

[14] J. Lee, B. Choi, E. Kim, Novel range-free localization based on multidimensional support vector regression trained in the primal space, IEEE Transactions on Neural Networks and Learning Systems 24(7)(2013) 1099-1113.

[15] C. Schaffer, Overfitting avoidance as bias, Mach Learn 10(1993) 153-178.