Optimization and Sequence Search based Localization in Wireless Sensor Networks

Sen Wang, Huosheng Hu, Klaus McDonald-Maier

School of Computer Science and Electronic Engineering
University of Essex, Colchester CO4 3SQ, United Kingdom
E-mail: {swangi, hhu, kdm}@essex.ac.uk

Abstract—**Localization is critical for various applications of Wireless Sensor Networks. This paper presents a 3D localization algorithm for high accuracy localization of a wireless sensor network, which consists of three parts: pre-localization, refinement and sequence search. It is clear that the localization accuracy partly depends on the localization sequence of the unknown nodes, which has not been investigated so far. The proposed novel algorithm aims to address this problem by searching a localization sequence corresponding to a high localization accuracy and a robust algorithm. The simulation results show that the proposed algorithm can get rid of the flip ambiguity and is more robust than several existing algorithms in terms of the localization accuracy.**

Keywords-**Wireless Sensor Networks; localization; optimization; localization sequence;**

I. Introduction

Wireless Sensor Network (WSN), which can benefit various systems in terms of the communication, perception and cooperation, has received considerable attentions in recent research and application [1]. In the WSN, the data is usually tagged with the corresponding locations for its usability, and the location knowledge is also useful for improving the network performance, e.g. the location based routing protocol. Moreover, the WSN is a good alternative for the localization and tracking in the Global Positioning System (GPS) denied environments. Therefore, determining the locations of the sensor nodes is essential for WSN.

The WSN based localization can be generally divided into range-free and range-based localization. Our attention here focuses on the range-based one due to its higher localization accuracy. In [2], Shang *et al.* use the multidimensional scaling (MDS) technique to design the MDS-MAP algorithm, which has the good localization accuracy even there are few anchors. As a very popular and widely employed method, maximum likelihood estimation (MLE) formulates the localization as nonconvex optimization problem [3], [4], which is difficult to be solved. The conjugate gradient [3], the Newton-Raphson iterative [4], etc. have been applied to find the optimal result. However, they all suffer from a problem that the result may be the local minima rather than the global ones.

Biswas *et al.* [5] propose the semi-definite programming (SDP) relaxation technique to mitigate this local problem. However, the estimation error was high because of the max-rank property of the SDP solution. In [6], both the regularization term and optimization methods are designed to reduce the rank of SDP. Unfortunately, there is no efficient way to determine an important parameter used in the regularization term.

To the best of our knowledge, the serious influence of the localization sequence was not addressed in the literature so far, which is in common for the algorithms that employ the optimization and the recursive positioning method [7] together. This paper aims to investigate the influence of localization sequence. A 3D localization algorithm, namely LMS, is proposed, which combines the least squares (LS), MLE and the sequence search (SS) with the recursive positioning method to improve the localization accuracy and address the localization sequence problem.

The rest of this paper is organized as follows. The problem formulation is stated in Section II. Section III describes the details of the proposed localization algorithm. The simulation results and analysis are presented in Section IV. Finally, a brief conclusion and future work are given in Section V.

II. Problem Formulation

The notation used in this paper is introduced here to formulate the problem. Consider a 3D network in $\mathbb{R}^3$ with m anchors and n unknown nodes. Let $\boldsymbol{\theta} = [\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_m, \ldots, \boldsymbol{\theta}_{m+n}]^T$ denote the true locations, where $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_m$ and $\boldsymbol{\theta}_{m+1}, \ldots, \boldsymbol{\theta}_{m+n}$ are the locations of the anchors and the unknown nodes respectively, and $\boldsymbol{\theta}_i = [x_i, y_i, z_i]^T$. The locations of the anchor nodes are assumed to be exact without no error. The estimated locations are denoted by $\hat{\boldsymbol{\theta}}$ where $\hat{\boldsymbol{\theta}}_i$ equals $\boldsymbol{\theta}_i, i \leq m$.

The measured Euclidean distance between a pair of sensors i and j is denoted as d_{ij}. Because of the inevitable disturbance on the range measurement, d_{ij} suffers from the following additive noise:

$$d_{ij} = r_{ij} + n_{ij} = \|\boldsymbol{\theta}_i - \boldsymbol{\theta}_j\| + n_{ij} \qquad (1)$$

where r_{ij} is the true distance between them, n_{ij} is the additive Gaussian noise of the distance estimation d_{ij} with mean μ_n and variance σ_n^2, and $\|\cdot\|$ denotes the 2-norm.

Since the communication range $\mathcal{R}$ is limited, the unknown node can only collect the information from the nearby anchors. The set of anchors whose distances to sensor i are smaller than $\mathcal{R}$ is denoted as $\mathcal{N}_i^a$. Because the recursive positioning method adds the newly estimated sensors into the anchor set, $\mathcal{N}_i^a$ also includes the localized unknown node, which is also called transformed node.

The WSN localization problem corresponds to the estimation of coordinates of the unknown nodes $\boldsymbol{\theta}_i$ with the knowledge of anchor positions, measured distances etc. Our goal is to derive an algorithm which can successfully address this problem with high and robust accuracy.

III. LOCALIZATION ALGORITHM

In this section, the LMS localization algorithm is presented. It comprises the LS based pre-localization, MLE based refinement and SS based localization sequence selection. The localization sequence problem is also introduced and further addressed by the sequence search method.

A. Pre-Localization

Because the optimization problem of WSN localization is nonconvex with many local minima, the initial point highly determines whether the global minimum can be converged to. In other words, a pure search method starting from a stochastic location may not result in the best localization result. Therefore, the LS method [8] is employed as the pre-localization to provide a good initial point for the optimization, although the point is not very accurate.

The recursive positioning method which extends the WSN coverage by transforming the localized sensors into the anchors is adopted. Suppose some unknown nodes have been successfully estimated and their locations can be used when the sensor i is being estimated. Therefore, the pre-location $\boldsymbol{\theta}_i^P$ of the unknown node i is

$$\|\boldsymbol{\theta}_i^P - \boldsymbol{\theta}_j\| = d_{ij}, 1 \leq j < i \,\&\, j \in \mathcal{N}_i^a. \quad (2)$$

For 3D localization, the number of the neighbor beacons for an unknown node should be no less than four. Assuming that the sensor i has N neighbor beacons and they are $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N$ (they can be easily changed to this if they are not). According to LS, $\boldsymbol{\theta}_i^P = [x_i^P, y_i^P, z_i^P]^T$ can be estimated by

$$\begin{cases} \sqrt{(x_i^P - x_1)^2 + (y_i^P - y_1)^2 + (z_i^P - z_1)^2} = d_{i1} \\ \sqrt{(x_i^P - x_2)^2 + (y_i^P - y_2)^2 + (z_i^P - z_2)^2} = d_{i2} \\ \quad\quad\quad \vdots \\ \sqrt{(x_i^P - x_N)^2 + (y_i^P - y_N)^2 + (z_i^P - z_N)^2} = d_{iN} \end{cases} \quad (3)$$

Hence, we have

$$\mathbf{A}_i \boldsymbol{\theta}_i^P = \mathbf{b}_i \quad (4)$$

where

$$\mathbf{A}_i = 2 \begin{bmatrix} x_1 - x_N & y_1 - y_N & z_1 - z_N \\ \vdots & \vdots & \vdots \\ x_{N-1} - x_N & y_{N-1} - y_N & z_{N-1} - z_N \end{bmatrix} \quad (5)$$

and

$$\mathbf{b}_i = \begin{bmatrix} x_1^2 - x_N^2 + y_1^2 - y_N^2 + d_{iN}^2 - d_{i1}^2 \\ \vdots \\ x_{N-1}^2 - x_N^2 + y_{N-1}^2 - y_N^2 + d_{iN}^2 - d_{iN-1}^2 \end{bmatrix} \quad (6)$$

are the coefficient matrix and constant terms of the overdetermined equation (3), respectively. Then, $\boldsymbol{\theta}_i^P$ can be estimated by

$$\boldsymbol{\theta}_i^P = (\mathbf{A}_i^T \mathbf{A}_i)^{-1} \mathbf{A}_i^T \mathbf{b}_i. \quad (7)$$

B. Refinement

The pre-localization result can be refined by the means of the MLE and optimization method. As the further step of the pre-localization, the above assumptions can be still held. An assumption is also that the mean and variance of the additive noise have known. Given the unknown node's location $\boldsymbol{\theta}_i$ and its nearby anchor j, the conditional probability density of the measured distance d_{ij} is

$$p(d_{ij}|\boldsymbol{\theta}_i) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp\{-\frac{(d_{ij} - r_{ij} - \mu_n)^2}{2\sigma_n^2}\} \quad (8)$$

Further, consider all the neighbor anchors of sensor i. Let $D_i = \{d_{ij}, j \in \mathcal{N}_i^a\}$ denote the set of distance measurements. Due to the assumption that the anchors are mutually independent, (8) can be extended to the following joint conditional probability density function of D_i:

$$\begin{aligned} p(D_i|\boldsymbol{\theta}_i) &= \prod_{j \in \mathcal{N}_i^a} \frac{1}{\sqrt{2\pi}\sigma_n} \exp\{-\frac{(d_{ij} - r_{ij} - \mu_n)^2}{2\sigma_n^2}\} \\ &= \frac{1}{(2\pi\sigma_n^2)^{\frac{N}{2}}} \exp\{-\sum_{j \in \mathcal{N}_i^a} \frac{(d_{ij} - r_{ij} - \mu_n)^2}{2\sigma_n^2}\}. \end{aligned} \quad (9)$$

Because the MLE method aims to find the unknown position $\boldsymbol{\theta}_i$ which maximizes the $p(D_i|\boldsymbol{\theta}_i)$ and the mean μ_n and variance σ_n^2 are known constant, MLE of $\boldsymbol{\theta}_i$ is

$$\begin{aligned} \hat{\boldsymbol{\theta}}_i &= \arg\max_{\boldsymbol{\theta}_i} p(D_i|\boldsymbol{\theta}_i) \\ &= \arg\min_{\boldsymbol{\theta}_i} \sum_{j \in \mathcal{N}_i^a} (d_{ij} - r_{ij} - \mu_n)^2. \end{aligned} \quad (10)$$

The WSN localization is often formulated as (10), which is an optimization problem of the nonlinear least square cost function [9], [10]. Due to the disturbances, the measured distance d_{ij} cannot be so accurate that it equals the real one. Then, (10) cannot be zero in general. However, the equation can be solved by minimizing it with the aid of various optimization methods. Since this problem may have many local minima, the initial value is so important that it
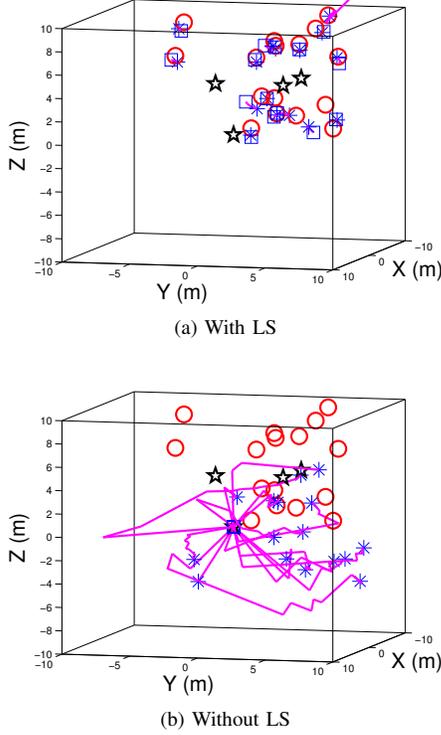
(a) With LS

(b) Without LS

Figure 1. Optimization with different initial points. a) Starting from the result of pre-localization. b) Starting from (0,0,0).

(a) 7-6-5

(b) 7-5-6

(c) 6-7-5

(d) 6-5-7

(e) 5-6-7

(f) 5-7-6

Figure 2. Results of different localization sequences. All the permutations of 5, 6 and 7 are given.

determines if the global optimal result can be reached. Figure 1 shows the two localization results produced by MLE with the same data except the initial values of the iteration.

In the picture, the (black) pentagrams correspond to the locations of the anchor nodes, and the (red) circles, the (blue) squares and the (blue) asterisks to the true, pre-localized and final estimated locations of the unknown nodes respectively. The (magenta) lines joining the squares and the asterisks represent the refinement trajectories. Figure 1(a) where the iteration starts with the pre-localization result describes that most of the estimated locations are accurate, while the trajectories stating with $(0, 0, 0)$ in Figure 1 (b) show few unknown nodes can converge to their locations. These verify that the pre-localization provides a good initial value for the optimization even though it is not quite accurate. Moreover, the refinement improves the accuracy along with the update trajectory, especially for the inaccurate nodes.

C. Localization Sequence of Unknown Nodes

1) Problem Description: It is found that the performance of the algorithm which only combines the pre-localization and MLE is unstable. The localization result sometimes is inaccurate even this situation is seldom encountered. By analyzing the reasons, it is considered as a general problem of the localization algorithms which introduce the optimization with the recursive positioning method. For simplicity yet
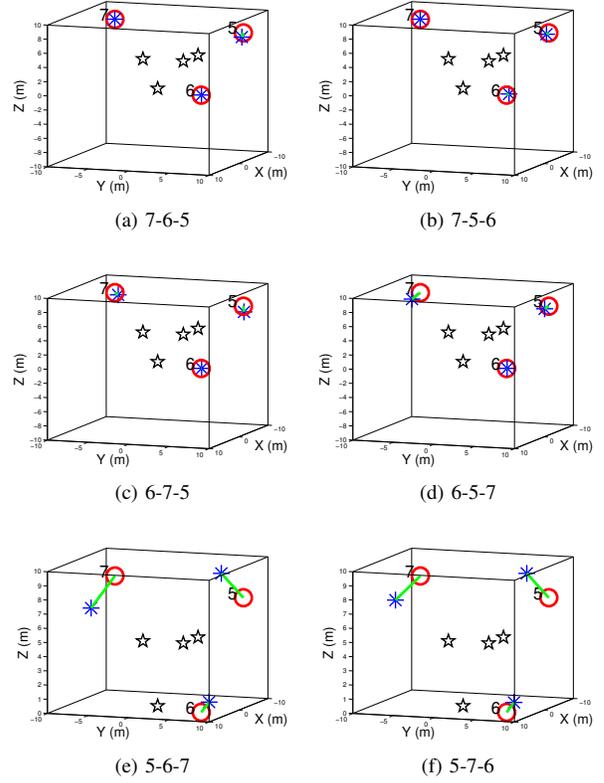
without loss of generality, a three nodes example is given in Figure 2. Despite only three nodes in this case, it still can clearly demonstrate that the localization accuracies change dramatically between the different localization sequences.

The corresponding localization errors are illustrated in Figure 3. The localization error of f is about six times bigger than the b's. Furthermore, its minimum error is much greater than the maximum of b. Note that these six results are obtained using the same raw data and the localization algorithm (the integration of the pre-localization and MLE) except the localization sequence. It suggests that a more accurate localization result may be achieved only by searching a proper localization sequence. Since this method does not require any additional hardware and amelioration of range precision, it is an easy yet effective method to highly improve the accuracy. Therefore, determining a localization sequence which can produce an excellent localization result and enhance the algorithm robustness is extremely significant.

2) Reasons: In order to devise an effective solution to the localization sequence problem, the reasons should be studied first. Recall the objective function in (10). Suppose μ_n is zero, and rewrite it by separating the anchors and the
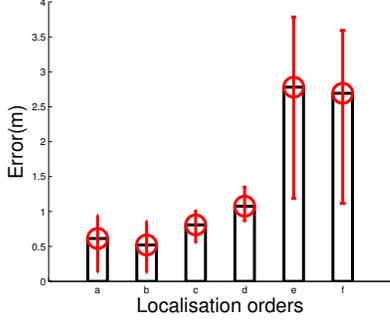
Figure 3. Localization errors of different localization sequences in Figure 2. a-f respectively correspond to the 6 situations.

transformed nodes:

$$f(\boldsymbol{\theta}_i) = \underbrace{\sum_{j \leq m} (d_{ij} - r_{ij})^2}_{\text{anchors}} + \underbrace{\sum_{m < j \leq M} (d_{ij} - r_{ij})^2}_{\text{transformed sensors}}, j \in \mathcal{N}_i^a \quad (11)$$

where $M = m + n$.

First, the optimization minimizes (11) without distinguishing the anchor and the transformed node. However, the location of anchor is accurate while the transformed node's is with error. What is worse, this error is inevitably accumulated due to the employed recursive positioning method. A good example is the Figure 2(f) where the later localized sensors are seriously influenced by the inaccurate sensor 5 localized previously.

Therefore, the optimization method consumes the majority of energy to decrease the errors which come from these inaccurate transformed nodes rather than the reliable anchors, making the final result unacceptable. Secondly, the refinement is not omnipotent. The flip ambiguity, which is worse in the 3D space, can totally reflect the estimated position across the plane of the neighbor anchors without changing the value of the cost function. However, a transformed node supplied by an appropriate localization sequence can make the flipped node get rid of the coplanar anchors, mitigating the flip ambiguity.

3) Sequence Search Method: In general, it is impossible to traverse all the permutations of the localization sequence because the amount dramatically increases along with the node increment. Therefore, a sequence search method which can perceive the error accumulation is proposed to seek an acceptable localization sequence for the pre-localization and MLE.

If the errors have propagated via the recursive positioning method, it is difficult for some sensors to minimize the cost function (11), and their transformed nodes terms become very big. Therefore, by checking the following variance of the transformed nodes terms of all the unknown nodes the

error accumulation can be perceived:

$$Variance = \frac{1}{n} \sum_{i=m+1}^{m+n} (T_i - \overline{T}) \quad (12)$$

where

$$T_i = \sum_{m < j \leq M} (d_{ij} - r_{ij})^2, j \in \mathcal{N}_i^a \quad (13)$$

$$\overline{T} = \frac{1}{n} \sum_{i=m+1}^{m+n} T_i. \quad (14)$$

The big variance means that some unknown nodes are not localized accurately. Hence, another localization sequence should be utilized to localize the network again until the variance is eligible. The localization sequence can be generated randomly because it is almost impossible to return a used permutation. Meanwhile, this method can avoid the flip ambiguity. Although the cost function of the flipped sensor has no change, its inaccurate location will provoke serious influences on some of nodes which will be localized later using it as an anchor.

A network which comprises 4 anchors and 16 unknown nodes is used to test the proposed LMS algorithm and verify this sequence search method. Because the permutations of sixteen are too enormous to list, only some typical examples are given here. By searching the variance which is bigger than a threshold, several bad cases are detected by our method, see Figure 4(a)-(e). The error offsets between the true and finally estimated positions are denoted by the (green) lines, whose lengths represent the estimation errors. These cases all suffer from the low accurate localization or/and flip ambiguity.

Because the proposed method can successfully detect them, they can be easily avoided. Figure 4(f) illustrates a relatively accurate result, which may not be the best one but much better than the detected bad cases. The results demonstrate that the proposed sequence search method can choose a proper localization sequence to efficiently control error accumulation and prevent the flip ambiguity, guaranteeing the accurate localization and improving the robustness of the localization algorithm.

IV. RESULTS AND ANALYSIS

In order to evaluate the proposed LMS algorithm, the LS and MLE without sequence search method (LMnoS) and MDS-MAP method which outperforms other methods [11] are chosen as the comparisons. 4 anchors are deployed in a $5m \times 5m \times 5m$ region, while the 46 sensors, which are randomly distributed in a square area of $20m \times 20m \times 10m$, are out of the boundary of the anchors to test the nonconvex optimization. The maximum communication distance is assumed to be $10m$, and the range measurement suffers from the Gaussian noise whose mean and variance are given by $\mu_n = 0$ and $\sigma_n^2 = 0.09$.
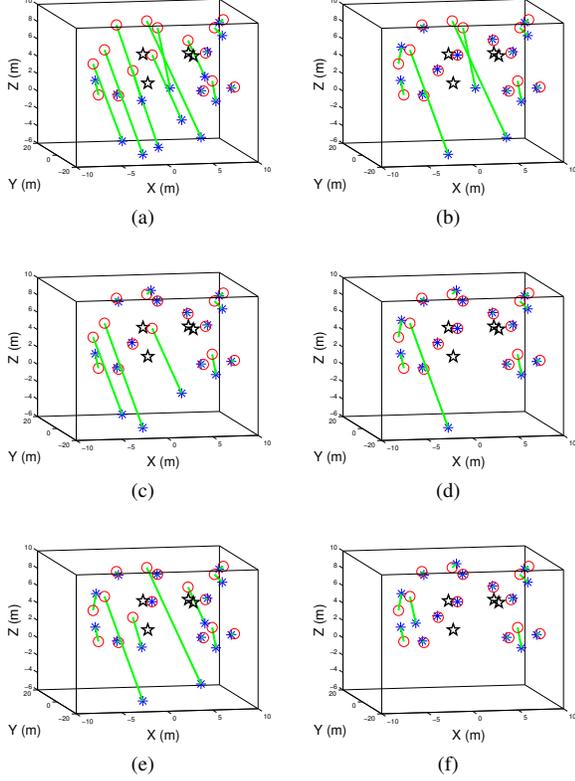
Figure 4. Six localization results produced by LMS with the different localization sequences searched by the sequence search method. a)-e) are the bad cases suffered from low accurate localization or/and flip ambiguity, while f) is the good one with accurate localization result.

Because there are a number of sensors in WSN, it is difficult and meaningless to analyze a single node. The performance of the algorithm is typically evaluated by the root mean square distance (RMSD):

$$\mathcal{E} = \frac{1}{\sqrt{n}} \left(\sum_{j=m+1}^{m+n} \| \boldsymbol{\theta}_j - \hat{\boldsymbol{\theta}}_j \|^2 \right)^{\frac{1}{2}}. \qquad (15)$$

Figure 5 is RMSD of two different anchor placements: one is the anchors are at $(0,0,0), (4,3,2), (2,2,5)$ and $(1,4,3)$, and the other is $(0,0,0), (5,5,5), (0,5,5)$ and $(5,0,5)$. For each anchor placement the algorithms are run for 10 independently generated sensor distributions. The threshold of the sequence search is set to be 100 times the quantity of σ_n^2. It can be seen that the proposed LMS method can get rid of the flip ambiguity and is more robust than the other two even though the good localization results of the three schemes are similar. Moreover, its precision is higher than LMnoS thanks to the sequence search.

Because the anchor deployment in Figure 5 (b) is more regular than that in Figure 5(a), it is easier for the algorithms to be influenced by the flip ambiguity in the first anchor placement. Therefore, the MDS-MAP seriously suffers from
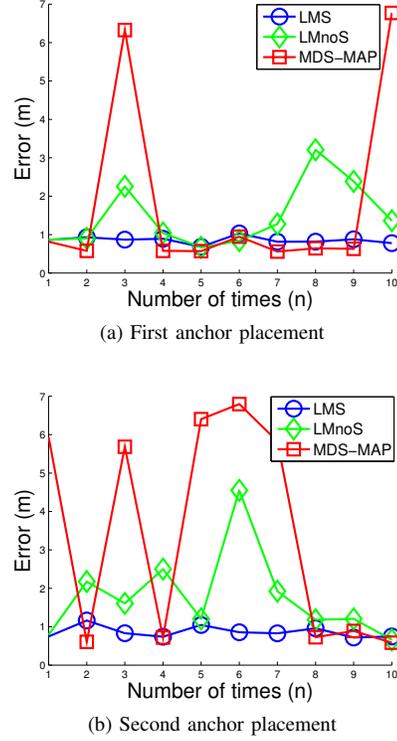

(a) First anchor placement

(b) Second anchor placement

Figure 5. Localization errors of LMS, LMnoS and MDS-MAP in two different anchor placements. a) Anchors at $(0,0,0), (4,3,2), (2,2,5)$ and $(1,4,3)$. b) Anchors at $(0,0,0), (5,5,5), (0,5,5)$ and $(5,0,5)$.

the flip ambiguity in 3D and it is more sensitive to the regular anchor deployment than LMS although some of its results are more accurate.

The effect of varying measurement noises is described in Figure 6. For each specific noise factor, the algorithms are run 10 times independently, and the upper and lower RMSDs are represented by the error bars. The standard deviations of the recorded measurement noises are $0.1, 0.3, 0.5, 0.7$ and 0.9. From the picture, it can be observed that LMS is more insensitive to the noise factors, and it is always more accurate than LMnoS. For the low measurement noise, LMS outperforms MDS-MAP in terms of the average localization errors although the lowest of MDS-MAP is much better.

V. CONCLUSIONS AND FUTURE WORK

This paper presents a novel 3D localization algorithm LMS for WSN based localization based on the optimization and sequence search. It mainly investigates the localization sequence problem with good performance in terms of the localization accuracy and robustness. Simulation results are presented to the feasibility and good performance of LMS. It is clear that the proposed LMS is more robust to anchor placement than the popular MDS-MAP method and its average localization error is much smaller.

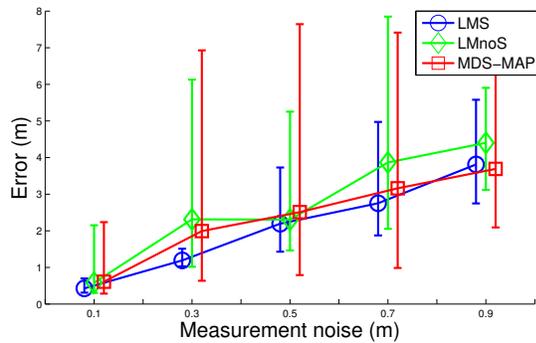However, the computation complexity may increase if a

Figure 6. Variation of localization errors under the different measurement noises.

appropriate localization sequence cannot be found, which is a drawback. Therefore, our future work will focus on how the localization sequence can be intelligently decided without human intervention.

ACKNOWLEDGMENTS

REFERENCES

[1] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Computer networks*, vol. 52, no. 12, pp. 2292–2330, 2008.

[2] Y. Shang, W. Rumi, Y. Zhang, and M. Fromherz, "Localization from connectivity in sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 15, no. 11, pp. 961–974, 2004.

[3] N. Patwari, A. Hero III, M. Perkins, N. Correal, and R. O'dea, "Relative location estimation in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, no. 8, pp. 2137–2148, 2003.

[4] X. Wang, M. Fu, and H. Zhang, "Target tracking in wireless sensor networks based on the combination of kf and mle using distance measurements," *IEEE Transactions on Mobile Computing*, no. 99, pp. 1–1, 2011.

[5] P. Biswas and Y. Ye, "Semidefinite programming for ad hoc wireless sensor network localization," in *Proceedings of the 3rd international symposium on Information processing in sensor networks*. ACM, 2004, pp. 46–54.

[6] P. Biswas, T. Liang, K. Toh, Y. Ye, and T. Wang, "Semidefinite programming approaches for sensor network localization with noisy distance measurements," *IEEE Transactions on Automation Science and Engineering*, vol. 3, no. 4, pp. 360–371, 2006.

[7] J. Albowicz, A. Chen, and L. Zhang, "Recursive position estimation in sensor networks," in *9th International Conference on Network Protocols*. IEEE, 2001, pp. 35–41.

[8] A. Savvides, C. Han, and M. Strivastava, "Dynamic fine-grained localization in ad-hoc networks of sensors," in *Proceedings of the 7th annual international conference on Mobile computing and networking*. ACM, 2001, pp. 166–179.

[9] S. Tennina and M. Di Renzo, "Esd: A novel optimization algorithm for positioning estimation in wireless sensor networks-analysis and experimental validation via a testbed platform," in *Proceedings of 17th International Conference on Computer Communications and Networks*. IEEE, 2008, pp. 1–7.

[10] Y. Kwon, K. Mechitov, S. Sundresh, W. Kim, and G. Agha, "Resilient localization for sensor networks in outdoor environments," *ACM Transactions on Sensor Networks*, vol. 7, no. 1, p. 3, 2010.

[11] P. Biswas, T. Lian, T. Wang, and Y. Ye, "Semidefinite programming based algorithms for sensor network localization," *ACM Transactions on Sensor Networks*, vol. 2, no. 2, pp. 188–220, 2006.