On the Error Characteristics of Multihop Node Localization in Ad-Hoc Sensor Networks

Andreas Savvides[1], Wendy Garber[2], Sachin Adlakha[1], Randolph Moses[2], and Mani B. Srivastava[1]

[1] Networked and Embedded Systems Lab
Electrical Engineering Department
University of California, Los Angeles
{asavvide, sachin, mbs}@ee.ucla.edu
[2] Department of Electrical Engineering,
Ohio State University
{garberw, randy}@ee.eng.ohio-state.edu

Abstract. Ad-hoc localization in multihop setups is a vital component of numerous sensor network applications. Although considerable effort has been invested in the development of multihop localization protocols, to the best of our knowledge the sensitivity of localization to its different parameters (network density, measurement error and beacon density) that are usually known prior to deployment has not been systematically studied. In an effort to reveal the trends and to gain better understanding of the error behavior in various deployment patterns, in this paper we study the Cramer Rao Bound behavior in carefully controlled scenarios. This analysis has a dual purpose. First, to provide valuable design time suggestions by revealing the error trends associated with deployment and second to provide a benchmark for the performance evaluation of existing localization algorithms.

1 Introduction

Ad-hoc node localization is widely recognized to be an integral component for a diverse set of applications in wireless sensor networks and ubiquitous computing. Although several ad-hoc localization approaches have been recently proposed in the literature [2, 4–7], the trends in localization error behavior in multihop setups have not been studied in a systematic manner. The majority of previously proposed localization approaches evaluate the 'goodness' of their solution with randomly generated scenarios and comparison of the computed results to ground truth. While this is a good starting point, it does not provide an intimate understanding of the different error components that come into play in multi-hop localization systems.

Ideally node localization would result in error-free position estimates if sensor measurements were to be perfect, and the algorithms were not to make any approximations such as operating on partial information and ignoring finite-precision arithmetic effects. In reality however, sensor measurements are noisy

and produce in noisy location estimates. Zooming into the origins of these errors we observe that measurement errors consist two main components, *intrinsic* and *extrinsic*. The intrinsic component is caused by imperfections in the sensor hardware or software. The extrinsic component is more complex and it is attributed to the physical effects on the measurement channel such as obstructions or fading that vary significantly according to the deployment environment. Although, the first type of error can be easily characterized in a lab setup, it is important to note that it also induces additional error that affect both the network setup aspects as well the choice of localization algorithms used.

This paper investigates the different aspects of the error induced by the intrinsic measurement error component in multihop localization setups. To explore the different aspects of error trends we study the Cramér-Rao Lower Bound (CRLB) behavior of carefully controlled deployment scenarios under different configuration parameters. In particular, we study the effect of network density, pre-characterized measurement accuracy, beacon (or other landmark density) and network size. The analysis presented here serves a dual purpose. First, to provide algorithm-independent design time insight in to the error trends associated with the different network setup parameters. This can help optimize the multihop localization performance prior to deployment. Second, the CRLB results can be used as an evaluation benchmark for multihop localization algorithms.

The remainder of this paper is organized as follows. The next section motivates our work by providing an overview of the sources of error in multihop localization systems. Section 3 provides the formulation of CRLB for multihop localization and explains the scenario structures used in this evaluation. Section 4 presents our simulation results. Section 5 discusses the evaluation of existing localization algorithms and section 6 concludes the paper.

2 Sources of Error in Multihop Localization Systems

2.1 Multihop Localization Problem Statement

Assume we have a set of A sensors in a plane, each with unknown location $\{r_i = (x_i, y_i)\}_{i=1}^{A}$. In addition, a set of B beacon with known locations $r_i = (x_i, y_i)_{i=-B+1}^{0}$ are placed in the plane. Each beacon node advertises its location and this information is forwarded to the other nodes in the network. Furthermore, each sensor node and beacon node emits some known signals that allow neighboring nodes to estimate their distance from the emitting node.

The distance measurements contain measurement error. We denote the error as e_{ij}, where

$$\hat{d}_{ij} = d_{ij} + e_{ij} \tag{1}$$

$$d_{ij} = \|r_i - r_j\| = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{2}$$

and where d_{ij} is the true distance between nodes i and j.

In this paper we assume the measurement errors are independent Gaussian random variables with zero mean and known variance σ^2; more general cases are considered in [9]. We denote the availability of a measurement using the indicator function I_{ij} where $I_{ij} = 1$ if node j receives a calibration signal from node i, and $I_{ij} = 0$ otherwise.

The general localization problem statement is as follows: Given noisy measurements of $\hat{d}_{ij}$ and known locations r_i for $i = -B + 1, \ldots, 0$, estimate the locations $\hat{r}_i$ for $i = 1, \ldots, A$.

2.2 A Classification of Error Components

As a first step in understanding the different sources of errors in multihop localization systems we categorize them in three broad classes *setup error*, *channel error* and *algorithmic error*.

Setup error is induced by intrinsic measurement error and it is reflected in the network configuration parameters such as network density, concentration of beacons (or other landmarks), network size and measurement error characteristics known prior to deployment and certainty of beacon locations. For the purposes of our discussion we assume that intrinsic measurement error can be characterized in a lab setup to provide an indication of the measurement accuracy of a particular ranging technology. Table 1 lists the measurement accuracies of four different ranging systems, an ultrasonic ranging system used in the AHLoS project [6], an ultra wide band (UWB) system [3], and RF Time-of-Flight system from Bluesoft [1] and a SICK laser range finder [10].

Table 1. Accuracy of different measurement technologies

Technology	System	Measurement Accuracy	Range
Ultrasound	AHLoS	2cm	3m
Ultra Wide Band	PAL UWB	1.5m	N/A
RF Time of Flight	Bluesoft	0.5 m	100m
Laser Time of Flight	Laser range finder	1cm	75m

Channel error is a result of the extrinsic measurement error and represents the physical channel effects on sensor measurements. Multipath and shadowing, multiple access interference, the presence of obstructions that results in unpredictable non-line of sight components, and fluctuations in the signal propagation speeds are just a few of these effects that can introduce error into the computation of locations. The magnitude of these effects on the distance measurement process is typically specific to the particular measurement technology and the environment in which they operate; hence different considerations should be applied for each technology.

Finally, the multihop nature of the problem and the different operational requirements introduce another level of complexity and subsequently more er-

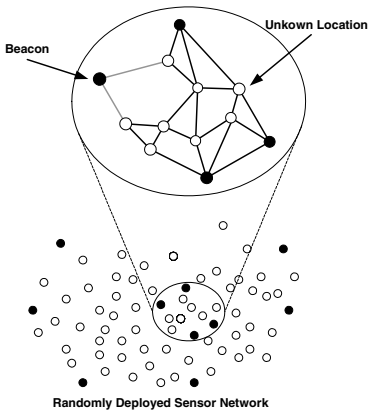

Fig. 1. Typical in network localization region

ror. Many settings that require the random deployment of low cost resource constrained sensor nodes, calls for fully distributed operation that has limited power consumption overhead. Such requirements may lead in approximate localization algorithms that have some additional error associated with them. The distributed computation model of *collaborative multilateration* described in [7] is an example of such an algorithm. In this case, the proposed fully distributed algorithm is an approximation of a centralized algorithm that conserves computation and communication energy. This design choice however introduces a small, yet tolerable error. We refer to this error as algorithmic error.

Although the goal of our research is to explore all aspects of error by building an operational ad-hoc localization system[3], in this paper we focus on setup error. The analysis presented here examines the setup error behavior inside specific segments within a sensor network. These segments are comprised of a small number of beacon nodes surrounding a large number of sensor nodes with unknown locations as shown in figure 1. These sensor nodes are expected to estimate their locations by combining their inter-node distance measurements and beacon locations. We evaluate the error trends in such setups using the Cramér-Rao Bound.

3 Localization Bounds in a Multihop Setup

3.1 The Cramér-Rao Bound

The accuracy for mean square location estimate can be evaluated using Cramér Rao (CR) bound [9]. The CR bound is a classical result from statistics that

[3] We refer the reader to our project website http://nesl.ee.ucla.edu/projects/ahlos for the details specific to our implementation including all released hardware and software

gives a lower bound on the error covariance matrix for an unbiased estimate of parameter θ (see, eg, [11]). The lower bound is given in terms of Fisher Information Matrix $J(\theta)$. Let $\hat{\theta}$ be any unbiased estimate of parameter θ based on observation vector X having a pdf of $f_X(x)$. The error covariance matrix is defined as

$$C = E\{(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T\} \tag{3}$$

This error covariance matrix is bounded below by the CR bound, which is given by

$$CRB = [J(\theta)]^{-1} \tag{4}$$

where the matrix $J(\theta)$ has elements given by

$$[J(\theta)]_{mn} = E\left\{\left[\frac{\partial \ln(f_X(X))}{\partial \theta_m}\right]\left[\frac{\partial \ln(f_X(X))}{\partial \theta_n}\right]\right\} \tag{5}$$

The matrix $J(\theta)$ is called the Fisher Information Matrix (FIM).

3.2 Obtaining CRLB in multihop setups

In the multihop problem, the parameter vector θ of interest is the $2A \times 1$ vector

$$\theta = [x_1, y_1, x_2, y_2, \ldots, x_A, y_A]^T \tag{6}$$

The measurement vector X is a vector formed by stacking the distance measurements $\hat{d}_{ij}$. Since it is assumed that the measurement is white Gaussian, the measurement pdf is the vector Gaussian pdf

$$f_X(x;\theta) = \mathcal{N}(\mu(\theta), \Sigma) = \frac{1}{(2\pi)^{2A}|\Sigma|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}[X - \mu(\theta)]^T \Sigma^{-1}[X - \mu(\theta)]\right\} \tag{7}$$

where the mean vector $\mu(\theta)$ is a vector of true distances whose elements are given by 2. The covariance matrix in equation 7 is given by

$$\Sigma = \sigma^2 I \tag{8}$$

where I is the $2A \times 2A$ identity matrix and where σ^2 is the variance of each measurement error e_{ij} in 1. Note that for this application the pdf depends on θ only through its mean value.

The vector X contains measurements of distances $\hat{d}_{ij}$ stacked in some order, and $\mu(\theta)$ is a vector of d_{ij} distances stacked in the same order. Let M denote the total number of $\hat{d}_{ij}$ measurements.

The CRB can be computed from the Fisher Information Matrix of θ from equation 5. The Fisher Information Matrix is given by

$$J_\theta = E\left\{[\nabla_\theta \ln f_X(X;\theta)]\,[\nabla_\theta \ln f_X(X;\theta)]^T\right\}$$

The partial derivatives are readily computed from equations (2), (6), and (7); we find that

$$J_\theta = \frac{1}{\sigma^2}[G'(\theta)]^T[G'(\theta)] \tag{9}$$

where $G'(\theta)$ is the $M \times 2A$ matrix whose mnth element is $\partial\mu_m(\theta)/\partial\theta_n$. Each element of $G'(\theta)$ is readily computed from equation 2. Let the mth element of $\mu(\theta)$ be d_{ij} for some corresponding values of i and j, and note that θ_n is either $x_{i'}$ or $y_{i'}$ for some corresponding i'. Then from equation 2,

$$G'(\theta)_{mn} = \begin{cases} 0 & \text{if } i' \neq i \text{ and } i' \neq j \\ \frac{x_i - x_j}{d_{ij}} & \text{if } \theta_n = x_i \\ \frac{x_j - x_i}{d_{ij}} & \text{if } \theta_n = x_j \\ \frac{y_i - y_j}{d_{ij}} & \text{if } \theta_n = y_i \\ \frac{y_j - y_i}{d_{ij}} & \text{if } \theta_n = y_j \end{cases} \tag{10}$$

The CR bound is then given by the inverse of the FIM as in 4.

3.3 Scenario Setup

To evaluate the effects of density variation and measurement error on the overall localization result, we generated a set of scenarios for which density and therefore node connectivity can be controlled. For the purposes of our experiments we define node density D to be the number of nodes per unit area. For N nodes deployed on a circular area A, $D = \frac{N}{A}$. Given this we can control the radius L of a circular field to be

$$L = \sqrt{\frac{A}{\pi}} = \sqrt{\frac{N}{D\pi}} \tag{11}$$

In a circular field, the probability of a node having d neighbors can be expressed as

$$P(d) = \binom{N-1}{d} P_R^d (1 - P_R)^{N-d-1} \tag{12}$$

where P_R is the probability that a node is within transmission range R from another node

$$P_R = \frac{\pi R^2}{A} = \frac{D\pi R^2}{N} \tag{13}$$

As N goes to infinity, the binomial distribution in equation 12 converges to a Poisson distribution (equation 14) with $\lambda = NP_R$

$$P(d) = \frac{\lambda^d}{d!} e^{-\lambda} \tag{14}$$

Also from this the probability of a node having n or more neighbors is

$$P(d \geq n) = 1 - \sum_{i=0}^{n-1} P(i)$$

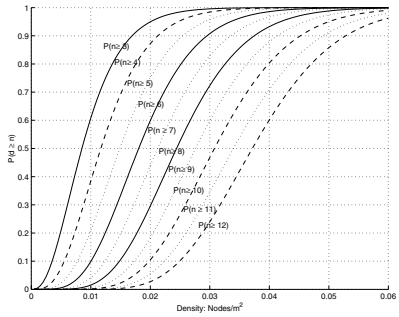

Fig. 2. Probability of n or more neighbors at different densities

In our simulation experiments sensor nodes have a 10-meter range. For this range the corresponding probabilities of having n or more neighbors and for different network densities are shown in figure 2. The scenarios used in this study are generated on a circular plane following the above analysis. To ensure even distribution of nodes per unit area, we divide the circle into rings of width $\frac{1}{\pi\sqrt{D}}$. In each ring we generate node positions in polar coordinates by generating a radius r and an angle θ for each node. The number of nodes in each ring is proportional to the area of the circle covered by the ring.

As it will be shown in section 4.1, this scenario pattern generation method was chosen to isolate error incurred from bad geometry setups. These effects arise when angles between beacons (or other anchor nodes) as seen by the node trying to determine their location are very small. This effect can be prevented when nodes are deployed using the circular pattern described above.

4 Simulation Results

Using the CRLB bounds derived in the previous section, we try to answer some fundamental questions related to setup error. This evaluation is performed by computing the Cramer-Rao bound on a comprehensive set of approximately 2,000 scenarios generated using the algorithm described in the previous section.

4.1 How does deployment geometry affect the solution?

Geometry setup alone can affect localization accuracy. This is a known effect frequently referred to as *geometric dilution of precision* (GDOP). The same effects come into play in a multihop setup where neighboring nodes with unknown positions help other nodes to estimate their locations by acting as *anchor points*. In both cases, the best estimates can be obtained when the nodes are placed within the convex hull of the beacons. These effects are demonstrated on a small set of simple scenarios. Figures 3 to 6 show the CLRB bound on a 10×10 grid. In

figure 3, four beacons are deployed on the vertices of a 8×8 square. The error is maximum outside the square, and behind the beacons where the angles between diagonally adjacent beacons is very small or zero. Figure 4 shows how the error behaves if the beacon square is shrunk to 2×2. By scaling the beacon square from 1×1 to 10×10 we note that the variance in the bounds at different points on the grid changes significantly. This effect can be seen by comparing figures 3 and 4. This also explains or choice of scenario generation algorithms. By keeping the beacons on the perimeter of the network we ensure that for the rest of our experiments, we operate in the places where the variance of the bounds is more uniform (i.e. similar to the flat region within the beacon square in figure 3.

Figures 5 and 6 show the error bounds when three beacons are used in a triangular configuration. In the first case the beacons are found at locations $B1 = 3, 3$, $B2 = 3, 8$, $B3 = 5.5, 5.5$. In the second case the beacons are placed at locations $B1 = 3, 3$, $B2 = 3, 4$, $B3 = 3.2, 5.5$. These two cases show the effect of geometry then the angles between each of the beacons as seen by a sensor node change. The largest error occurs when the angles to each beacon are very small.

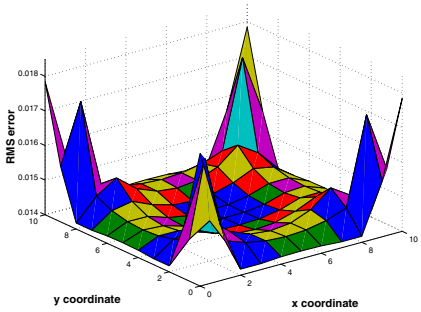

Fig. 3. Effects of geometry on $8 \times 8m$ square beacon pattern

4.2 How does network density affect localization accuracy?

Intuitively, one would expect that localization accuracy would improve as the network density increases. This is because increasing network density, and subsequently the number of neighbors for each node with unknown location adds more constraints to the optimization problem. After some critical density, the effect of density on location accuracy becomes less apparent. Our simulation results in figure 7 verify this expectation. The critical point occurs in the case where the majority of the nodes have at least 6 neighbors. For the particular range used in our experiments, this takes place at a density of 0.35 nodes/m^2. This result is consistent for a test suite of more than 1,000 scenarios at different ranging accuracies as shown in the figure.

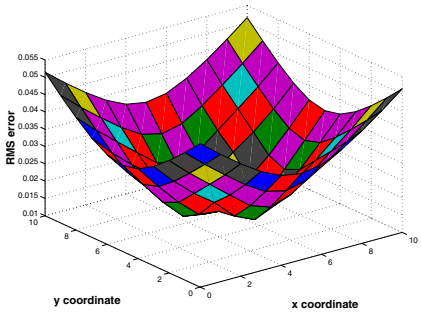

Fig. 4. Effects of geometry on $2 \times 2m$ square beacon pattern

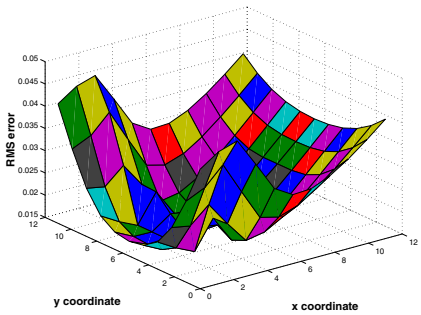

Fig. 5. Effects of geometry on an isosceles beacon triangle

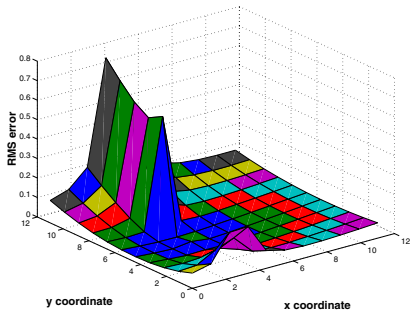

Fig. 6. Effects of geometry on a flat triangle

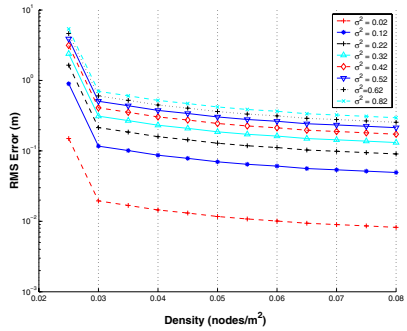

Fig. 7. Density trends at different values of measurement error variance σ^2

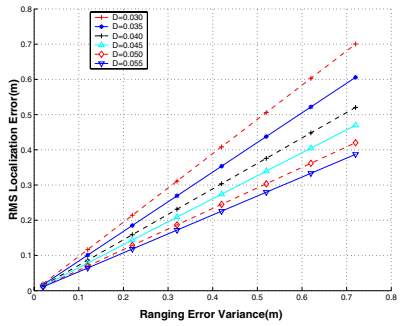

Fig. 8. Density and range error scaling

Furthermore, we note that the different CRLB plots for different values of ranging error σ^2 are scaled versions of one another. In figure 7 and in figures from subsequent sections, our plots show the bounds at different ranging errors to allow the user to associate these results to specific ranging technologies such as the ones listed in table 1.

4.3 What is the best solution one can achieve with a given measurement technology?

This question can be answered by observing the bounds on the same set of scenarios as the previous subsection. In general based on our simulations we not that if the network density is sufficient (6 or more neighbors per node), the bound predicts that the localization error will be close to (slightly lower) than the ranging error. The trend lines for different levels of ranging error are shown in figure 8.

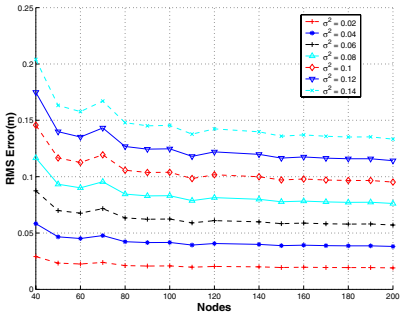

Fig. 9. Error propagation as network scales at 10% beacons, 6 neighbors per node

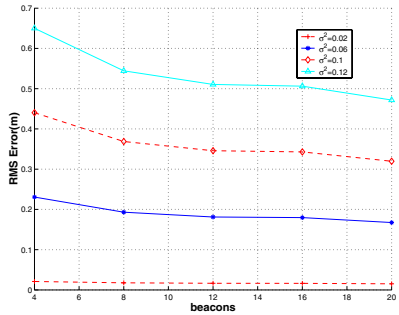

Fig. 10. The effect of beacon density on localization, 100 nodes, 4-20% beacons

4.4 How does error behave as the network scales?

If the density is kept fixed, the error bound degrades very showly as the network scales. A representative result from our experiments for different measurement accuracies is shown in figure 9. In this experiment, the network size is varied from 40 nodes to 200 nodes while the network density is kept constant at 6 neighbors per node, and 10% beacons.

4.5 What is the effect of beacon density on the computed solution?

To test the effect of beacon density on the localization bounds we used a set of scenarios with fixed density (0.45) and fixed number of nodes (100 nodes). The percentage of beacons was varied from 4% to 20%. The results are (shown in figure 10) indicate that increasing the number of beacon nodes does not dramatically reduce the localization bound. This is more profound when the ranging error is very small. As shown in the figure, at a range error variance of $0.02m$, using 4 beacon nodes on a 100 node network performs just as well as 20 beacons. Also for higher levels of ranging error, adding more beacons yields a modest improvement.

5 A Case Study on Algorithmic Error: Collaborative Multilateration

In this section we present a comparison of a specific multihop localization algorithm, *collaborative multilateration*, and corresponding the Cramér-Rao bound. The measurement error characteristics used for this evaluation are drawn from the lab characterization of the ultrasonic distance measurement system described in [6].

5.1 Collaborative Multilateration Overview

Collaborative multilateration is a method for performing node localization in multihop setups. The algorithm, which is described in detail in [7] relies on a small set of beacon nodes and inter-node distance measurements to estimate node locations in multihop setups while trying to prevent error accumulation inside the network. Collaborative multilateration supports two computation models, *centralized* and *distributed*. The centralized model estimates node locations at a central point in the network that has a global vantage point. All the inter-node distance measurements and beacons locations are used to set up a global non-linear optimization problem, which is solved using least squares.

Even though the centralized computation model can yield high quality estimates, it is not always suitable for sensor networks. First, it requires significant computation, which would require more processing and memory resources than what low cost sensor nodes can accommodate. Second, a centralized approach exposes a single point of failure in the network. Third, a centralized approach also requires some routing protocol support to propagate measurements and locations to the central computation point (and sometimes to also propagate the position estimates back to the nodes).

To address the issues of the centralized computation model, the distributed collaborative multilateration computation model was designed to operate in a fully distributed fashion. In this model, each node in the network is responsible for estimating its own location using distance measurements and location information from its one-hop neighbors. To compute an estimate of its location each node uses its neighbors as anchors. If these neighbors do not have a final estimate of their location, then an intermediate rough estimate of the node locations is used. All nodes compute an updated estimate of their locations and they pass it to their neighbors, which in turn use this information to update their own location estimate. The process continues until a certain tolerance is met. In this computation model, the estimate updates at each node happen in a consistent sequence that is repeated until the convergence criteria are met. This forms a gradient with respect to the global topology constraints that allows the nodes in the network to estimate their location with respect to the global constraint while computing their estimate locally.

The distributed computation mode of collaborative multilateration is an approximation of the centralized model that is designed to meet some of the operational requirements of a practical setup. This design decision however introduces

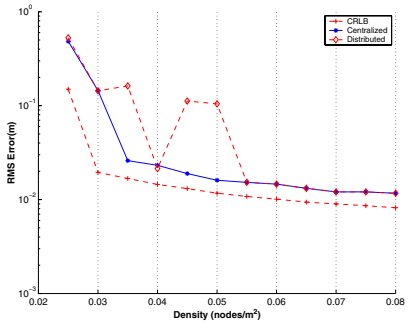

Fig. 11. Algorithmic error in centralized and distributed collaborative multilateration computation models

some algorithmic error in the location estimates. In the next subsection we compare the error from the two approaches to the bounds.

5.2 Comparison to the Bounds

To evaluate the quality of location estimates of the two computation models of collaborative multilateration, we compare the two computation models to the CRLB bounds, using the same scenarios as the ones used in section 4.1. For these scenarios the ranging error variance was set to $0.02m$ to match the characteristics of the ultrasonic ranging system described in [6]. The results from this comparison are shown in figure 11. The results shown here are averages from 10 different scenarios for each density.

From this comparison both computation models follow a similar trend to the CRLB bound but we also note some differences, which we classify as part of the *algorithmic error*. The critical density point has moved from 0.03 nodes/m^2 to 0.035 nodes/m^2. This corresponds to the point where the majority of nodes have 8 neighbors instead of 6 as predicted by the bound. We also note the discrepancy in the results of the distributed computation model. Although in most cases, the location estimates provided by distributed collaborative multilateration are almost identical to its centralized counterpart, for some cases the averages shown in figure 11 suggest that the position estimates are sometimes significantly different. A closer examination of the simulation data has shown that this discrepancy arises from very few isolated scenarios where the distributed process does not converge. Repeating the experiments has shown that this discrepancy can be prevented if some consistency checks are added to detect divergence. The distributed computation can converge it a different starting node is selected. This choice however would also incur increased algorithm complexity. At densities of 12 or more neighbors, the results of the two computation models are consistent. We attribute this to the fact that increased densities offer significantly more constraints that keep the process from diverging.

6 Conclusions

In this paper we explored some of the trends in localization error for multihop localization scenarios. Our simulation experiments have shown how intrinsic error from the sensor measurements incurs additional error with respect to different network parameters. This contributes some insight on what the deployment parameters should be for a multihop localization process to be successful. The beacon nodes should be deployed on the perimeter of the network to ensure that localization algorithms operate in the region where variance on the bounds is minimal. We also noted that there is a critical density after which localization improvement is much more gradual. By comparing this to collaborative multilateration we concluded that algorithmic error should also considered prior to deployment and deployment decisions should be more conservative than the ones predicted by the bounds. Furthermore, the study of the bounds has shown that multiple localization approaches are scalable and the position of a large number of nodes can be determined with a very small number of beacons that are found multiple hops away. As part of our future work, we plan to investigate the effects of the extrinsic measurement error on location estimates and how can this be handled at the network level by utilizing network redundancy and the trends exposed in this paper.

References

1. Bluesoft Inc *http://www.bluesoft-inc.com*
2. L. Doherty, L. El Ghaoui, K. S. J. Pister, *Convex Position Estimation in Wireless Sensor Networks*, Proceedings of Infocom 2001, Anchorage, AK, April 2001.
3. R. Fontana, S. Gunderson *Ultra Wideband Precision Asset Location System* Proceedings of IEEE Conference on Ultra Wideband Systems and Technologies, May 2002
4. D. Nicolescu and B. Nath *Ad-Hoc Positioning System* In proceedings of IEEE GlobeCom, November 2001
5. C. Savarese, J. Rabay and K. Langendoen *Robust Positioning Algorithms for Distributed Ad-Hoc Wireless Sensor Networks* USENIX Technical Annual Conference, Monterey, CA, June 2002
6. A. Savvides, C. C Han and M.B Srivastava *Dynamic Fine-Grained Localization in Ad-Hoc Networks of Sensors* Proceedings of fifth Annual International Conference on Mobile Computing and Networking, Mobicom pp. 166-179, Rome, Italy, July 2001
7. A. Savvides, H. Park and M. B. Srivastava, "The Bits and Flops of the n-Hop Multilateration Primitive for Node localization Problems", Proceedings of the first International Conference on Wireless Sensor Networks and Applications held with Mobicom, September 2002
8. R. L. Moses and R. M. Patterson, "Self-calibration of sensor networks," in *Unattended Ground Sensor Technologies and Applications IV (Proc. SPIE Vol. 4743)* (E. M. Carapezza, ed.), pp. 108–119, April1–4 2002.
9. R. L. Moses, D. Krishnamurthy, and R. Patterson, "A self-localization method for wireless sensor networks," *Eurasip Journal on Applied Signal Processing, Special Issue on Sensor Networks.* (submitted November 2001).

10. SICK http://www.sick.de
11. H. L. Van Trees, *Detection, Estimation, and Modulation Theory: Part I.* New York: Wiley, 1968.