Empirical Evaluation of the Limits on Localization Using Signal Strength

Gayathri Chandrasekaran*, Mesut Ali Ergin*, Jie Yang[†], Song Liu*,
Yingying Chen[†], Marco Gruteser*, Richard P. Martin*

*WINLAB, Rutgers University
North Brunswick, NJ 08902
{chandrga,ergin,song,gruteser,rmartin}@winlab.rutgers.edu

[†]Dept. of ECE, Stevens Institute of Technology
Castle Point on Hudson, Hoboken, NJ 07030
{jyang,yingying.chen}@stevens.edu

Abstract—**This work investigates the lower bounds of wireless localization accuracy using signal strength on commodity hardware. Our work relies on trace-driven analysis using an extensive indoor experimental infrastructure. First, we report the best experimental accuracy, twice the best prior reported accuracy for any localization system. We experimentally show that adding more and more resources (e.g., training points or landmarks) beyond a certain limit, can degrade the localization performance for lateration-based algorithms, and that it could only be improved further by "cleaning" the data. However, matching algorithms are more robust to poor quality RSS measurements. We next compare with a theoretical lower bound using standard Cramér Rao Bound (CRB) analysis for unbiased estimators, which is frequently used to provide bounds on localization precision. Because many localization algorithms are based on different mathematical foundations, we apply a diverse set of existing algorithms to our packet traces and found that the variance of the localization errors from these algorithms are smaller than the variance bound established by the CRB. Finally, we found that there exists a wide discrepancy from what free-space models predict in the signal to distance function even in an environment with limited shadowing and multipath, thereby imposing a fundamental limit on the achievable localization accuracy indoors.**

I. INTRODUCTION

Location is essential for many emerging applications from a diverse set of areas including asset tracking, workflow management, geographic routing, and physical security. Wireless networks offer an unprecedented potential for realizing many of these applications. Given that wireless devices are carried by many people and attached to many objects and all modern radio chipsets include the hardware necessary to measure and report the received signal strength (RSS) of transmitted packets, there is a tremendous cost and deployment advantage to re-using the existing RSS infrastructure of the communication network for signal strength-based localization purposes.

Over the past years, algorithm advances have yielded accuracy improvements from RADAR's [1] median 3 m error to less than 1 m median error [2]. A significant further improvement to about 40 cm median error has been obtained using a larger number of landmarks (base stations), 21 instead of the 3–5 used in previous experiments [3]. Since radio environments are becoming increasingly dense, this points to possible further accuracy improvements through using additional measurement nodes. Particularly, cooperative localization techniques [4], [5] where clients also contribute RSS measurements could provide readings from tens to hundreds of nodes. The limits of localization performance in such settings remain an open question.

The primary contribution of this paper is thus an empirical quantification of the accuracy limits of RSS localization on commodity wireless hardware. We conduct experiments to capture RSS data in a controlled extremely dense laboratory environment with a single transmitter and up to 369 landmarks, which represents an ideal scenario for localization algorithms. Traces were collected using the ORBIT testbed, which is a 400 node indoor wireless experimental apparatus placed in a 3600 sq ft. area. Using the ORBIT platform allowed us to capture long, high quality packet traces in a dense environment free of major shadowing and with limited multipath effects.

We use a combination of theoretical as well as trace-driven analysis on this dataset. Our theoretical work uses a traditional Cramér-Rao Bound (CRB) analysis, which has previously been used to establish bounds on location estimation variance [4]. We then use a trace-driven emulation to characterize the performance of different algorithms. In order to show the generality of our results as well as compare localization strategies, we used algorithms with widely divergent mathematical foundations. They range from classification approaches such as RADAR [1] over probability density exploration methods such as H1 [6] to multi-lateration such as Non-Linear Least Squares (NLS) [7].

Specifically, we found that:

- RSS based localization can achieve median errors as low as 0.24 m, with a maximum error of 1.5 m. Interestingly, while NLS performed the best under perfect synthetic conditions, it has the worst performance for real RSS observations, with a mean error of 1.6 m and a maximum error of 5.4 m.
- classification and probability density exploration algorithms had fundamentally worse performance using perfect synthetic input, because of a combination of their

inherent discretization effects as well as the leave-one-out technique we employ to create our testing dataset.

- for lateration-based approaches, which assume a signal-to-distance function, quality of the RSS measurements is more important than the quantity of measurements. A subset of 179 landmarks whose data yield a good signal-to-distance fit provided best localization performance. Simply increasing the number of landmarks over this actually increased the median error from 24 cm to 58cm.
- classification algorithms are qualitatively less sensitive to variances and noise in the input set than lateration-based algorithms. Given RSS measurements that deviate substantially from standard models, these algorithms maintained good average and worst-case performance.
- accuracy improvements leveled off with about 100 nodes, the lower-bound of localization performance appears still limited by discrepancies between the underlying algorithmic models and the actual signal-propagation effects of indoor environments. For example, individual node differences due to differences in connectors, thermal effects, and local noise floors, as well as multi-path effects caused by diffraction, reflection and shadowing are not explicitly accounted for in any of the algorithms and account for the remaining errors.

Three important implications of our results are the following. First, the CRB for unbiased estimators [5], [8], which is widely used for assessing localization performance limits, does not represent an actual lower bound on localization performance for all algorithms, likely because its assumptions of unbiased estimators or normally distributed measurement errors do not hold for these algorithms. Second, the strategy of minimizing the square root of the sum of the residuals leaves much room for improved accuracy, so localization systems based on classifiers or probability density exploration are preferable to least squares. We explore these effects more in Section IV. Third, significant accuracy improvements are still possible, likely by algorithms that incorporate more accurate models of measurement noise.

The remainder of the paper is organized as follows: In Section II, we provide a literature survey and briefly describe the algorithms that we use in this paper. Section III explains the experimental methodology. We present the results from our experimental testbed in Section IV. We discuss some open questions from our results in Section V. Finally, Section VI concludes the paper.

II. BACKGROUND

In this section, we briefly survey several localization algorithms and provide background on the Cramér-Rao bound (CRB). A full treatment of localization algorithms is beyond the scope of this paper. Additionally, we also provide a brief background on the representative subset of the algorithms we use in this paper. Because our purpose is to explore the limits of algorithmic performance, our descriptions focus on each algorithm's broad strategy. The reader is encouraged to pursue the references for additional details.

A. Lateration Based Algorithms

Lateration-based algorithms [7], [12], [13] explicitly model the signal-to-distance effect on RSS. They estimate the position of the transmitter by measuring the distance to multiple receivers. In [11], the authors use a Bayesian graphical model based on lateration to find a location estimate. We select our representative subset from lateration-based algorithms as *Non-Linear Least Square (NLS)* [7], and *Bayesian Networks (M1)* [11], and explain them briefly next.

1) Non-Linear Least Square (NLS): In NLS, estimating the true location of the transmitter (x, y) can be viewed as an optimization problem where the actual locations of the reference points (x_i, y_i) are known apriori and the distance estimates d_i are obtained from the signal-to-distance relationship. The problem then becomes solving for the optimal $(\hat{x}, \hat{y})$ that minimizes the sum of residuals:

$$(\hat{x}, \hat{y}) = \arg \min_{x,y} \sum_{i=1}^{N} \left[\sqrt{(x_i - x)^2 + (y_i - y)^2} - d_i \right]^2 \quad (1)$$

2) Bayesian Networks (M1): The M1 algorithm uses Bayes Nets which encode dependencies and relationships among a set of random variables. The vertices of a Bayes Net graph correspond to the variables and the edges represent dependencies [14]. The networks used for localization encode the relationship between the RSS and the (x, y) location using a simple log-distance propagation model.

The M1 strategy describes the joint probability density of (x, y) as a function of the observed RSS. However, in general, there is no closed form solution for the returned joint distribution. Therefore, we use Markov Chain Monte Carlo (MCMC) sampling to draw samples from the joint density [15]. The resulting samples allow us to approximate the true PDF of the (x, y). M1 selects the averages of the drawn samples as the (x, y) location estimate.

B. Classification Based Algorithms

Classification algorithms, a.k.a matching algorithms, do not rely on a model of signal strength and distance relationship. Rather, they match RSS observations against an existing signal map. The term classification, as used in the machine learning sense, implies that the goal of the classifier is to map a potentially large input space into a much smaller space of *labels*. In the case of localization, the labels are a set of discrete (x, y) locations.

Previous approaches [1], [9] and [2] are examples of classical fingerprint matching algorithms. The location of a node is estimated by matching its fingerprints to the closest one in the signal map. Matching algorithms in [3], and [10] employ probabilistic inference to estimate the location of a node.

Technique	Type	Area (m^2)	Num APs	Num Training	Median Error(m)	75th % Error(m)	Max Error(m)
RADAR [1]	Classification(Scene Matching)	22.50 × 43.49	3	70	2.93	4.69	24.99
AURA [9]							
CMU-PM	Classification(Scene Matching)	Not Specified	5	17	0.98	3.29	>9.75
CMU-TMI	Classification(Scene Matching with interpolation)	Not Specified	5	17	1.94	3.29	8.50
LEASE [2]	Classification(Scene Matching with Interpolated Grid)	68.58 × 43.90	5	100	2.29	n/a	n/a
		76.20 × 53.34	4	100	0.61	n/a	n/a
Ref [10]	Classification(Probabilistic)	68.28 × 35.94	4	110	1.07	1.22	7.32
HORUS [3]	Classification(Probabilistic)	68.28 × 35.94	21	172	0.39	0.55	4.99
		11.80 × 33.13	6	110	0.51	0.90	4.99
M1 [11]	Lateration(Bayesian inference)	60.96 × 24.38	4	115	5.49	6.71	27.43
		64.00 × 42.67	5	215	5.49	6.1	27.43
Ref [7]	Lateration with LLS	60.96 × 24.38	4	286	6.1	9.14	42.67
	Lateration with NLS	60.96 × 24.38	4	286	3.35	6.1	33.53

TABLE I

SUMMARY OF THE REPORTED EXPERIMENTAL ACCURACIES FOR VARIOUS LOCALIZATION ALGORITHMS

Also, [16] uses Bayesian inversion to return the location that maximizes the probability of the RSS vector. The authors of [17] apply the same technique to the robotics domain and experimentally show that 83% of the time, the location error is within 5 ft.

Deriving closed form solutions for the lower-bound of most of these algorithms is not trivial, or the bound may not be existent at all. Indeed, framing localization as a map matching problem naturally lends itself to machine-learning approaches, for which many algorithms have provably no closed form solution. We select our representative subset from classification-based algorithms as *RADAR*, *Gridded-RADAR (GR)* and *Highest Probability (H1)* and continue with a brief overview of these algorithms.

1) RADAR: [1] is a classical scene matching localization algorithm where the signal map, a set of fingerprints with known (x, y) locations, is provided as an input to the offline phase of the algorithm. In the online phase, when presented with a fingerprint of a node with unknown location, RADAR returns the location of the 'closest fingerprint from the signal map.

2) Gridded-RADAR (GR): is an improvisation over RADAR where measurement area is sub-divided into a regular grid and the signal map provided in the offline phase is interpolated over the entire grid. The online phase is similar to RADAR with the exception that the "closest" fingerprint in signal space is chosen from the interpolated signal map. This approach has an advantage of obtaining a much finer-grained resolution as the regions which are not covered by the signal map can also be returned as location estimates.

3) Highest Probability (H1): Given an area divided into a discrete set of points called tiles, the strategy used by H1 is to return the most likely (x, y) by finding the highest probable tile using Bayes' rule over the set of RSS values. In order to find the likelihood of the RSS-matching for each tile in isolation, H1 assumes that the distribution of the RSS for each receiver follows a Gaussian distribution. This assumption significantly simplifies the computations with little performance loss. Using Bayes' rule, H1 computes the probability of being at each tile on the floor, L_i, given the fingerprint of the localized object $\bar{S}_l$ as

$$P\left(L_i|\bar{S}_l\right) = \frac{P\left(\bar{S}_l|L_i\right) \times P\left(L_i\right)}{P\left(\bar{S}_l\right)}. \tag{2}$$

However fingerprint $\bar{S}_l = \left(\overline{s_{lj}}\right)$ is some constant c and with no prior information about the exact object's location, H1 assumes that the object to be localized is equally likely to be at any location on the floor, i.e., $P\left(L_i\right) = P\left(L_j\right), \forall i, j$. Thus, Equation 2 can be rewritten as

$$P\left(L_i|\bar{S}_l\right) = c \times P\left(\bar{S}_l|L_i\right). \tag{3}$$

Without having to know the value c, H1 can just return the tile L_{max}, where $L_{max} = argmax(P\left(\bar{S}_l|L_i\right))$, by computing $P\left(\bar{S}_l|L_i\right)$ for every tile i on the floor. Up to this step H1 is very similar to the traditional Bayesian approaches [10], [17], with the exception of the Gaussian and variance assumptions.

Finally, [4] and [18] studied establishing theoretical lower bounds for the achievable localization performance using estimation techniques that employ unbiased estimators. Table II summarizes the median, 75th percentile and maximum errors for various localization algorithms that were studied experimentally. Overall, we can see that none of the approaches have experimented with more than 21 access points and the state of the art approach [3] has its median, 75th percentile and the max errors as 0.39 m, 0.55 m, and 4.99 m, respectively.

In addition to studying RSS-based localization in a high density setting with hundreds of landmarks, and reporting localization errors an order of magnitude lower than the state-of-the-art, in this work, we also focus on understanding the factors that limit the RSS-based localization performance from achieving near-zero errors.

C. Cramér-Rao Bound

Localization can be defined as an estimation problem where measurements like wireless signal strength, angle or time of arrival are provided to an estimator (i.e. the localization algorithm) to obtain the most likely position in the assumed co-ordinate system. In estimation theory, the Cramér-Rao bound (CRB) has been derived as a lower-bound on the variance of an estimator [19]. Although CRB has been applied to certain classes of biased estimators [20], [21], it is commonly used to bound the variance of unbiased estimators [8]. The CRB has frequently been used by researchers to assess localization techniques [4], [5], [18].

The CRB for an unbiased estimator is obtained from the inverse of the Fisher Information Matrix (FIM) [8], given as

$$\mathcal{I}(\theta) = E\left\{\left[\frac{\partial \ln f(p|\theta)}{\partial \theta}\right]\left[\frac{\partial \ln f(p|\theta)}{\partial \theta}\right]^T\right\}, \quad (4)$$

where θ is the unknown parameter to be estimated from measurements p, which follows a probability density function $f(p|\theta)$. Intuitively, the FIM is an estimate of the curvature of the log likelihood function $\ln f(p|\theta)$. If the curvature is sharp, the parameter estimation becomes more accurate making the lower bound on the variance of the estimator(CRB) very small. The received power at a landmark location (x_n, y_n) from a transmitter at (x, y) can be modeled as [22]

$$P(x_n, y_n) = P_0 - 10\gamma \log_{10}(d_n/d_0) + S_n \quad \text{(dB)}, \quad (5)$$

where $d_n = \sqrt{(x_n - x)^2 + (y_n - y)^2}$ with $n = 1...N$ landmarks; P_0 is the received power at the reference distance d_0 from the source; γ is the path loss exponent; and S_n is the random variation of the signal measurements and assumed to come from an i.i.d. Gaussian distribution $\mathcal{N}(0, \sigma_{RSS}^2)$. For the case of a single unknown transmitter location and N landmarks, the CRB for the variance σ^2 of an unbiased location estimator is given as

$$\sigma^2 \geq \frac{\mathcal{I}(\theta)_{xx} + \mathcal{I}(\theta)_{yy}}{\mathcal{I}(\theta)_{xx}\mathcal{I}(\theta)_{yy} - \mathcal{I}(\theta)_{xy}^2} \quad (6)$$

where $\mathcal{I}(\theta)_{xx}$ and $\mathcal{I}(\theta)_{yy}$ are the diagonal blocks and $\mathcal{I}(\theta)_{xy}$ and $\mathcal{I}(\theta)_{xy}^T$ are the off-diagonal blocks of the FIM in Eq.4. Details of the derivation can be found in [5].

Note that the above CRB, for any localization technique using RSS information, critically depends on:

- the number and topology of the landmarks and the transmitters to be localized,
- the ratio of the RSS standard deviation to the propagation constant (σ_{RSS}/γ), characterizing the signal and the propagation environment
- the assumption of RSS fluctuations due to an i.i.d. Gaussian distribution with a common variance σ_{RSS}^2.

III. TESTBED EXPERIMENTS

In this section, we begin by describing our experimental objectives and then explain how we perform experiments in order to fulfill these objectives.

A. Objectives

The objectives of our experiments are:

- To quantify the limits of different localization algorithms—M1, H1, GR, and NLS—under a dense, indoor landmark deployment with limited shadowing and multipath fading
- To understand how these limits compare to the CRB
- To understand how variations in RSSI observations limit the localization algorithms from achieving perfect results in this laboratory environment

B. Experimental Methodology

We performed our experiments on ORBIT [23], a large scale indoor wireless testbed. The ORBIT testbed consists of 400 small form-factor PCs, with two IEEE 802.11a/b/g wireless interfaces per node. The nodes are suspended from the ceiling as shown in Figure 1(a) and are placed in a 20×20 regular grid with an inter-node separation of 91.44cm (3ft) spanning a total area of 3600 sq ft. We collected packet data traces from a subset of 369 nodes using one of the identical Atheros 5212 based 802.11a/b/g NIC in every node. The remaining 31 nodes were down for maintenance.

The data trace collection spanned two days and all 369 fixed ORBIT nodes were configured as IEEE 802.11a receivers operating at 5GHz, channel 44. The receivers used the Tshark packet sniffer utility to log the received signal strength indicator (RSSI) for every received packet. We used a movable ORBIT node attached to a portable antenna mast as our packet transmitter. Figure 1(b) shows the rubber-duck antenna mounted on the mast. The antenna rig was used to raise the movable node's antenna to the same plane receiver antennas were located as well as keeping transmitter antenna orientation the same throughout all transmitter locations. The transceiver diversity options were disabled in all our radios to eliminate

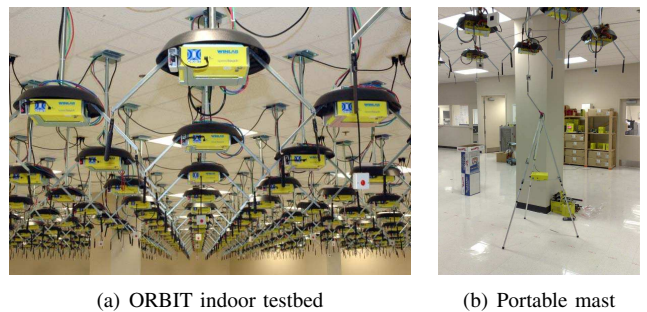(a) ORBIT indoor testbed   (b) Portable mast

Fig. 1. (a) The 400 node ORBIT experimental testbed (b) The movable node (Transmitter) attached to an antenna on a portable mast.
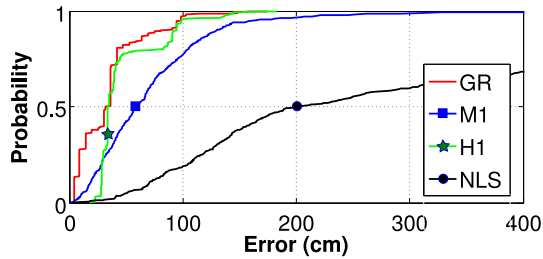
Fig. 2. Error CDF plots representing localization performance of the selected algorithms from ORBIT experiment using all available 369 landmarks

	Scaled Results			Sanitized Results		
	Median (m)	Max (m)	Std.Dev. (m)	Median (m)	Max (m)	Std.Dev. (m)
M1	.58	26.87	1.87	.24	1.60	0.25
NLS	2.01	13.44	2.41	1.62	5.37	0.79
GR	.31	1.74	0.30	.36	1.97	0.32
H1	.33	1.82	0.29	.39	1.70	0.33
CRB	—	—	0.62	—	—	0.75

TABLE II
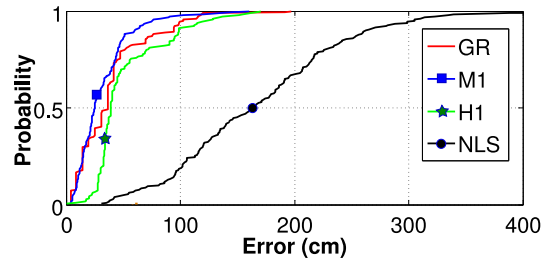SUMMARY OF LOCALIZATION ACCURACY AND PRECISION

Fig. 3. Error CDF plots representing localization performance from ORBIT experiment using the "sanitized" dataset with 179 Training and 179 Landmarks

unwanted RSSI oscillations. The transmitter was placed at 400 different locations, one below each one of the suspended ORBIT node for collecting the training dataset. The per-packet data trace collected at each of the receiver was then post processed to yield an average RSSI measurement over 1000 packets for the transmitter at every location. This resulted in a dataset with 400 training Points and 369 receivers for each training point. Note that our testing point locations overlapped with the landmark locations.

To evaluate the different algorithms, we use the well-known leave-one-out approach where the 400 point training set is split into 400 sets of 399 training points and one testing point. The difference between the known actual location (x, y) of the testing point and the localization algorithm's estimate $(\hat{x}, \hat{y})$ of the testing point derives the estimation error.

IV. RESULTS

In this section, we begin by highlighting our important contributions and provide detailed insights into our results. Table II summarizes our key findings. They are:

1) The M1 algorithm achieved the lowest median localization error of about 0.24 m, exceeding the best prior experimental result [3] by a factor of two
2) Adding a lot of landmarks can reduce the localization accuracy for lateration based algorithms. The best performance was achieved with a carefully selected cleaned subset of about half the landmarks
3) The general CRB for unbiased estimators is a poor benchmark of localization precision for the algorithms tested

The following subsection describe each of these results in detail.

A. High-density Localization Performance

Figure 2 plots the error CDF for the lateration algorithms (M1 and NLS) and the matching algorithms (GR and H1) with the training data collected from all 400 locations and 369 landmarks in the 20 by 20m space. We call this training set the *scaled dataset*. This dataset allows conclusions about how far localization error can be reduced with extreme measurement resources.

GR and H1 have the best performance with median errors of 0.31 m and 0.33 m, respectively exhibiting only an 16% improvement over the best prior reported median accuracy of 0.38 m with 21 landmarks.

Table II shows the median and maximum errors for all four algorithms. The lateration algorithms, in particular NLS, show much higher errors both in terms of median and maximum. M1's maximum error of about 27 m exceeds that of H1 and Gridded RADAR 15-fold. This motivates us to further explore the causes for such outliers.

B. Sensitivity to Data Quality

We define the *data quality of a landmark* based on how well the signal strength measurements for the landmarks match a distance to RSS propagation model fitted on the data. The lateration algorithms estimate the propagation parameters from the measurements based on this distance to RSS fit. Recall that in free space, the signal power decays linearly with log distance.
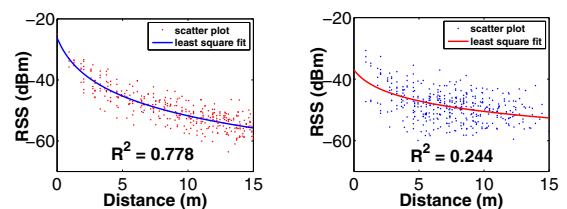
(a) High Quality Landmark   (b) Low Quality Landmark

Fig. 4. Plots showing the co-efficient of determination(R^2) for 2 different landmarks (a) Landmark with Good fit, $R^2 = 0.778$ (b) Landmark with poor fit, $R^2 = 0.2443$
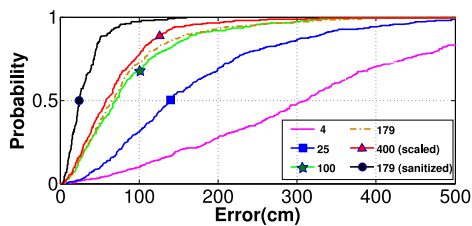
Fig. 5. Effect of Scaling the Number of APs

Parameter	Value
Path gain P_0 @ 2.4 GHz @ 1 meter	-42.934 dB
Path Loss Exponent (γ)	1.96
Transmit Power	10 dB
Antenna Gain	2 dB
Cable Losses	1 dB

TABLE III
SUMMARY OF THE PARAMETERS USED IN SYNTHETIC DATA GENERATION

Figure 4 plots the distance to RSS relationship together with the fitted free-space propagation model for two different landmarks using the same set of training points. We observe that the quality of the fit in terms of R^2 differs significantly.

To investigate the relatively poor performance of the lateration algorithms in the scaled dataset, we sanitize the data by removing low quality landmarks whose co-efficient of determination $R^2 < 0.5$. We empirically determined that this threshold significantly improves localization performance. After filtering, 179 landmarks and the corresponding 179 training points remain. We refer to this dataset as the *sanitized* (or cleaned) dataset.

Figure 3 plots the performance of different localization algorithms with the sanitized dataset. We can see that M1 outperforms all other algorithms, it achieves a median error of 0.24 m, about half the error reported by state of the art RSS-based localization algorithms. The median error for NLS has also improved from 2.01 m for the scaled dataset to 1.62 m for the sanitized dataset. Note also that the max errors for M1 dropped from 26.87 m to 1.60 m showing a 94% improvement.

Figure 5 plots the error CDF for M1 for varying number of landmarks (or access points) —4,25,100,179 and 400. In each scenario, the access points were deployed in a regular, equally spaced fashion. Additionally, we also plot M1's result for the sanitized dataset. While increasing the number of access points significantly reduces the error for M1, the results show diminishing returns. The reduction in error from 100 to 400 landmarks is minor compared to reduction from 4 to 25. Note also that, the sanitized dataset with 179 landmarks significantly outperforms all results with arbitrary selection of landmarks.

These results suggest that lateration algorithms generally are very sensitive to data from low-quality landmarks that cannot be fitted on a propagation model. For the lateration algorithms, increasing landmark density is less effective than selecting high-quality landmarks. The matching algorithms, however, remain very robust to these data quality issues.

C. Performance with Synthetic Data

We have shown that after data sanitization, a median localization accuracy of 0.24 m is achievable. To explore possibility of further improved localization, we study the localization performance starting with an ideal noise-less RSSI dataset, which we create synthetically to conform to the well-known

path-loss model given in Eq. 5. Then, by using empirical observations from our ORBIT experiments, we model two different classes of noise that affect RSSI data and perturb this perfect synthetic dataset according to the models. Finally we compare the performance of the localization algorithms using the perturbed data with the ORBIT experiments to validate our modeling.

The parameters used to create the synthetic dataset are given in Table III. These were obtained from a detailed channel measurement study in the ORBIT room [24], which determined the path gain at reference distance P_0, and the path loss exponent γ. To facilitate comparisons with the best case ORBIT experiments, our synthetic dataset consists of 179 landmarks with RSSI information from 179 different locations that fall 30cm away from each landmark. This replicates the transmitter locations used in the actual ORBIT experiments for the sanitized dataset.

Noise-less Performance: Figure 6 presents localization performance for five different algorithms using the noise-less dataset. Lateration algorithms M1 and NLS perform very well—both result in sub-centimeter accuracy for 99% of the time, and strictly below 1.5 cm all the time. Classification algorithms RADAR and GR are limited with the discrete number of fingerprints (classes) to which a given testing point can be associated. Due to the leave-one-out method of testing, RADAR can only match a testing point to the nearest possible landmark, which is 3-feet away in the ORBIT grid setting. Consequently we observe that RADAR has 99% of its error accumulated exactly at 3-feet (about 91 cms). The CDF for GR, which works with 2 inch (5.08 cm) grid-sizes, shows a step-like behavior and achieves a median error of 15 cm
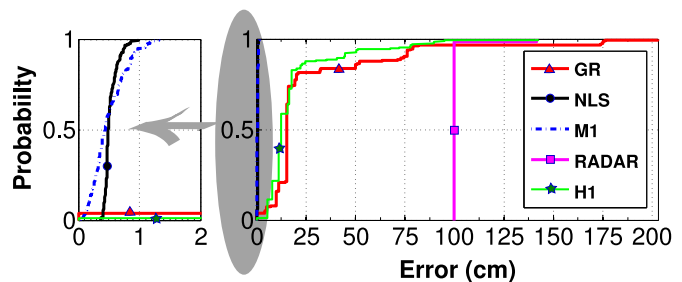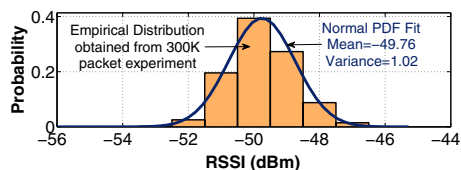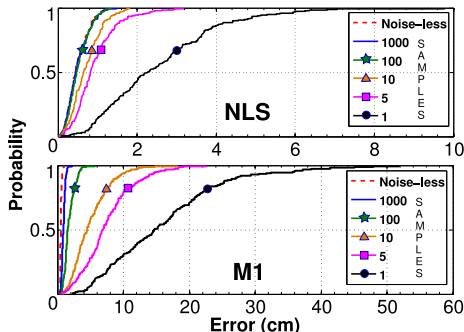

Fig. 6. Localization performance with synthetically generated noise-less data for 179 landmarks. The Fig in the left plots the Zoomed-in error CDF to show M1 and NLS performance and the one on right plots the Error CDF for all five algorithms

(a) Single Link RSSI Distribution

(b) 179 Landmark Localization Performance with Type I Noise

Fig. 7. (a) Empirical RSSI distribution from single-link 300K packet experiment and the normal distribution fit for Type I Noise model (b) Effect of the number of RSSI samples used for averaging on Type I Noise perturbed synthetic data for Nonlinear LS and Bayesian Network Algorithms

(worth almost 3 grid points). Similarly, the tile-discretization effect inherent to H1 results in a median error of 12 cm with this ideal noise-less dataset.

The performance discrepancy between Figures 3 and 6 indicate that the input dataset in our ORBIT experiments contains significant noise on RSSI observations.

RSSI Noise Analysis: To investigate this discrepancy, we measured the distribution and stability of RSSI readings on a single transmitter-receiver pair. Without environmental mobility and a time-invariant channel, the RSSI observations from an ideal pair of transmitter-receiver should be constant over the time. Figure 7(a) shows the RSSI distribution of 300.000 packets over a given ORBIT link.[1] Clearly, variations of 1–2 dB exist and we categorize this as Type I noise. Detailed discussions on the potential causes of the noise observed on RSSI are deferred to Section V.

Type I Noise Case: To understand how this variance affects localization, we have perturbed our noise-less synthetic dataset according to the variance of the normal distribution fit on the observed noise of 1.02 db as shown in Figure 7(a). The RSS for the synthetic data follows $\mathcal{N}(\mu(\theta), 1.02)$ where $\mu(\theta)$ is the mean received power which is $P_0 - 10\gamma \log_{10}(d_n/d_0)$ as given by Eq 5. We tested both M1 and NLS localization algorithms with this perturbed dataset. We observed that using single RSSI measurements resulted in increased median localization

[1]Note that these experiments were conducted remotely 4AM in the morning with no human presence and no 802.11 interference. Our tests with other combinations of off-the-shelf cards (i.e., Atheros 5212 and Intel ProWireless 2945) exhibited the same behavior.
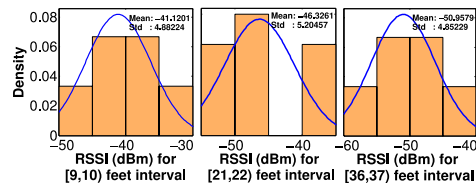
(a) Three example PDF fits for Landmark #5
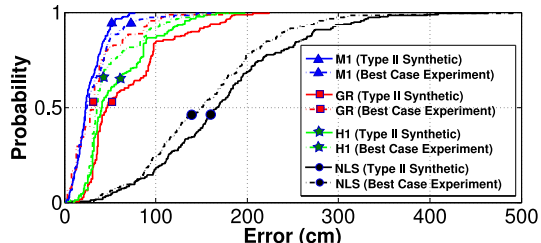
(b) Localization Performance with TypeII Noise

Fig. 8. (a) Example empirical RSSI distributions obtained per landmark at each foot distance separation for Type II Noise (b) Performance of localization under Type II Noise

error, from sub-centimeters up to of 14.6 cm and 2.2 cm for M1 and NLS respectively, as shown in Figure 7(b). RSSI averaging, however, reduced this effect—the mean of 1000 RSSI observations removed nearly all localization error due to this noise. Thus, the Type I noise do not account for the observed discrepancies, since experiments throughout this paper were conducted with the mean of 1000 packet RSS readings.

Type II Noise Case: Next we model the noise observed in Fig. 4, by dividing TX-RX distance into 1 ft buckets for each receiver and fitting a normal distribution on the data that fall in each bucket, as illustrated in Figure 8(a). We categorize this noise as Type II . Using these standard deviations($\hat{\sigma}$) obtained from the PDFs, we create another synthetic dataset whose RSS follows $\mathcal{N}(\mu(\theta), \hat{\sigma}^2)$ and again evaluate the localization performance. Results are depicted in Figure 8(b). Note that these results match the experimental results well, the discrepancy is less than 20 cm for all algorithms.

These results suggest that the noise that limits localization performance is not due to short term measurement noise on individual nodes, but rather due to variations across nodes and locations. We discuss more on RSSI noise in Section V.

D. Comparisons with Cramér Rao Lower Bound

Following the Gaussian distribution assumption in Section II-C, we have calculated the standard deviation (*stddev*) of the averaged RSS sample residuals (σ_{RSS}) from our experimental dataset of 400 nodes to be 8.880, as illustrated in Figure 9(a). Also, the path-loss exponent γ in the ORBIT room was previously measured to be 1.701, using precise measurement equipment [24] at 5.1GHz UNII band. Using this σ_{RSS}/γ ratio of 5.220 together with the 179 (i.e., sanitized) and 369 (i.e., scaled) landmark topologies as inputs, we

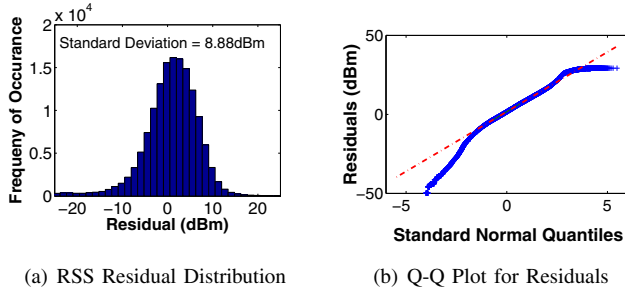(a) RSS Residual Distribution   (b) Q-Q Plot for Residuals

Fig. 9. Investigating Gaussian distribution assumption of RSS samples from 400 node dataset

calculated the CRB for each unknown transmitter position with the help of a Matlab script. The median value of the stddevs obtained from this CRB calculation is reported in Table II (in meters) together with stddevs of errors from the localization algorithms we have evaluated.

In both the scaled and sanitized cases, the CRB does not provide a lower bound on the variance of the localization error, except for the NLS algorithm (both cases) and M1 algorithm (only for scaled case). It is important to note that the CRB is only a bound on the variance of the localization error, but not a bound on the mean of the error. Therefore the CRB should be used as a benchmark for the precision of various unbiased localization algorithms but not the accuracy of them.

We identify two reasons why the CRB provided little value in comparing the precision of localization algorithms we tested. **RSS Distribution**: The CRB (Eq.6) assumes that the RSS sample residuals come from a zero mean Gaussian distribution with a stddev of σ_{RSS}. To verify if this assumption holds, we examined averaged RSS samples shown in Figure 9(a) with a quantile-quantile (Q-Q) plot given in Figure 9(b). We observe that RSS data fit the normality assumption only between -2 and +2 quantiles. Also, a chi-square test for the normality of the data indicates that the normality hypothesis can not be accepted with 95% confidence. It is likely that our RSS samples come from a more complex composite distribution, similar to the observation in [5]. Since the Gaussian distribution assumption does not strictly hold, the CRB can not be expected to provide a strict bound for the localization error variance. In theory [25], the RSS residual in the ORBIT testbed environment which has a strong LOS component with minimal shadowing should be characterized using the Ricean distribution instead of Gaussian.
Estimator Bias: The CRB in the form of Eq.6 is not applicable to biased estimators. It is quite possible that the majority of localization algorithms are biased. In fact, non-linear least squares is known to generally be a biased estimator. Also, RADAR intuitively appears biased due to the limited number of training points that positions are matched to. The median error for biased estimators can be lower than the bound for an unbiased estimator. Deriving a general form of the CRB for a biased estimator would require knowledge on the gradient of the estimator bias [26].

V. DISCUSSION

Our results leave unanswered questions with regards to the lower bounds of localization performance using RSS. In this section, we discuss how resources, node quality and algorithm choice impact the lower bounds of localization performance.

First, high accuracy requires a large number of observation points, which is in agreement with prior work [3]. However, motivation for many of the prior works has been building a localization system using only minimal additional infrastructure. We have also shown that the additional infrastructure can leverage the RSS measurements on existing communication waveforms by re-using the enormous investment in commodity chipsets.

Our results show that high quality RSS measurement is critical to localization performance, and that the measurement and reporting variances across devices limit the accuracy. Although the minute-scale averages of RSS observations are found to be stable (Figure 7), significant variance still exists when signal-to-distance fits are considered (Figure 8(a)). The exact breakdown for the causes of this distribution remains unknown. Table IV provides a non-exhaustive list of the sources of noise that might have potentially led to non-ideal RSSI observations for our experiments. Note that the list omits important items like mobility and external interference as they were not existent in our controlled experiment. *Calculation* of RSSI from a received IEEE 802.11 packet is only outlined by the standard [27] and implementation details for any given wireless card remain the manufacturers' intellectual property. Nonetheless in our experiments, we use 369 identical m-PCI Atheros 5212 IEEE 802.11a/b/g cards manufactured at close proximity in time, thus RSSI observations likely come from identical calculation algorithms. Also, in our experiments, *faults* are easily detectable as we have close proximity RSSI observations for every landmark from which outliers could be eliminated. Fading, as demonstrated by precise measurements in the ORBIT room [24], is time-invariant in the absence of environmental mobility, thus the multipath profile of the channel in the room is static. In our experiments, fading is visible not because it varies the received signal power for a given link over the time, but because the RSSI is observed from a static transmitter at 179 discrete locations in the room, each capturing a different fading profile. This makes several of the possible sources unlikely, pinpointing the exact reasons remains an open problem.

A last open issue is that we believe, there is still room for better algorithmic methods to extract localization performance from the traditional approach of finding the best fit that minimizes the residuals. The M1 algorithm is a first step in exploiting such prior information in a manner that goes beyond traditional classifiers, but our work raises the question if additional information could still be extracted by clever algorithms without resorting to classifiers.

EQUIPMENT	Calculation	How the RSSI is measured by the card (i.e., exact sampling from the packet training sequence, before or after AGC elements etc.). How interference and noise affect RSSI calculation. Details of the particular algorithm running in DSP.	
	Quantization	The way RSSI is quantized (i.e., 0-31, 0-63, 0-100). Averaging used. Non-linearity of rcvd power vs. reported RSSI	
	Faults	Leaking, or improperly terminated RF circuitry. Broken antenna cable or connectors.	
	Manufacturing/Design	Differences in the manufacturing process. Variance of the quality of the circuit components used. Cheap transceiver design with fluctuating TX power and receiver path gains.	
ENVIR.	Thermal Noise	Observed in the receiver electronics dependent on the ambient temperature. Also other forms of cyclo-stationary noises.	
	Shadow Fading	Caused by the blocking of direct, reflected, diffracted, and scattered signal copies from the transmitter.	
	Multipath Fading	Caused by multiple copies of the received signal through (a possible) line-of-sight component, and its reflections, diffractions, scattering, each delayed wrt power-delay profile of the physical environment around TX-RX pair.	

TABLE IV
SOURCES OF NOISE FOR RSSI RELEVANT TO OUR EXPERIMENTS

VI. CONCLUSION

In this work we investigated the lower bound obtainable using RSS based localization through a dense, high-precision wireless testbed. We found that high average accuracies, on the order of 0.2 m, are possible using commodity hardware in our configuration. We also showed that the maximum error can be reduced to about 1.6 m, which is also an encouraging result.

Our results show that the precision of a number of algorithms exceed the theoretical lower bounds commonly calculated for localization techniques using Cramér-Rao Bound (CRB) analysis for unbiased estimators. This raises questions about the validity of the assumptions underlying this analysis, particularly with regard to bias and normally distributed measurement errors.

These results also demonstrate that the choice of algorithms is important, in that we observed least squares approaches have the worst performance on real data sets. We also found that classification-based algorithms are more robust to poorer quality data than lateration approaches. Our results also point to the possibility of further improvements from increasing the quality of the RSS observations, raising the node density in real deployments, or adding algorithmic enhancements.

REFERENCES

[1] P. Bahl and V. N. Padmanabhan, "Radar: An in-building rf-based user location and tracking system," in *INFOCOM*, March 2000. [Online]. Available: citeseer.nj.nec.com/bahl00radar.html

[2] P. Krishnan, A. S. Krishnakumar, W.-H. Ju, C. Mallows, and S. Ganu, "A system for lease: Location estimation assisted by stationary emitters for indoor rf wireless networks," in *INFOCOM*, Oct. 2004.

[3] M. Youssef and A. Agrawala, "The horus location determination system," *Wirel. Netw.*, vol. 14, no. 3, pp. 357–374, 2008.

[4] N. Patwari, J. Ash, S. Kyperountas, I. Hero, A.O., R. Moses, and N. Correal, "Locating the nodes: cooperative localization in wireless sensor networks," *Signal Processing Magazine, IEEE*, vol. 22, no. 4, pp. 55 – 69, 2005.

[5] N. Patwari, A. O. H. III, M. Perkins, N. S. Correal, and R. J. OâĂŹdea, "Relative location estimation in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, no. 8, pp. 2137–2148, August 2003.

[6] E. Elnahrawy, X. Li, and R. P. Martin, "The limits of Localization Using Signal Strength: A Comparative Study," in *SECON*, 2004.

[7] Y. Chen, F. J.-Austen, W. Trappe, and R. P. Martin, "A Practical Appalsoroach to Landmark Deployment for Indoor Localization," in *SECON*, 2006.

[8] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Englewood Cliffs, NJ: PTR Prentice-Hall, 1993.

[9] A. Smailagic, D. P. Siewiorek, J. Anhalt, D. Kogan, and Y. Wang, "Location sensing and privacy in a context-aware computing environment," *IEEE Wireless Communications*, vol. 9, pp. 10–17, 2001.

[10] M. Youssef, A. Agrawal, and A. U. Shankar, "WLAN location determination via clustering and probability distributions," in *Proceedings of IEEE PerCom'03*, Fort Worth, TX, Mar. 2003.

[11] D. Madigan, E. Elnahrawy, R. P. Martin, W.-H. Ju, P. Krishnan, and A. Krishnakumar, "Bayesian Indoor Positioning Systems," in *Infocom*, 2005.

[12] D. Niculescu and B. Nath, "Ad hoc positioning system (APS)," in *GLOBECOM*, 2001.

[13] K. Langendoen and N. Reijers, "Distributed localization in wireless sensor networks: a quantitative comparison," *Comput. Netw.*, vol. 43, no. 4, pp. 499–518, 2003.

[14] D. Heckerman, "A tutorial on learning with bayesian networks," Microsoft Research, Tech. Rep. MSR-TR-95-06, March 1995.

[15] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, 2nd ed. Chapman and Hall, 2004.

[16] T. Roos, P. Myllymäki, and H. Tirri, "A statistical modeling approach to location estimation," *IEEE Transactions on Mobile Computing*, vol. 1, no. 1, pp. 59–69, 2002.

[17] A. M. Ladd, K. E. Bekris, A. Rudys, G. Marceau, L. E. Kavraki, and D. S. Wallach, "Robotics-based location sensing using wireless Ethernet," in *MOBICOM*, Atlanta, GA, sep 2002.

[18] S. Dulman, P. Havinga, A. Baggio, and K. Langendoen, "Revisiting the Cramer-Rao Bound for Localization Algorithms," in *4th IEEE/ACM DCOSS Work-in-progress paper*, June 2008.

[19] C. Rao, "Minimum variance and the estimation of several parameters," *Proc. of Cambridge Philos. Soc.*, pp. 280–283, 1946.

[20] H. L. V. Trees, *Detection, Estimation, and Modulation Theory. Part I: Detection, Estimation, and Linear Modulation Theory*. New York: John Wiley and Sons Inc, 1968.

[21] A. Hero, "A cramer-rao type lower bound for essentially unbiased parameter estimation," Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, MA, Tech. Rep. 890 DTIC AD-A246666, 1992.

[22] V. Erceg, L. J. Greenstein, S. Y. Tjandra, S. R. Parkoff, A. Gupta, B. Kulic, A. A. Julius, and R. Bianchi, "An empirically based path loss model for wireless channels in suburban environments," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 7, pp. 1205–1211, Jul. 1999.

[23] D. Raychaudhuri, I. Seskar, M. Ott, S. Ganu, K. Ramachandran, H. Kremo, R. Siracusa, H. Liu, and M. Singh, "Overview of the orbit radio grid testbed for evaluation of next-generation wireless network protocols," in *Proc. IEEE WCNC*, vol. 3, March 2005, pp. 1664–1669.

[24] H. Kremo, J. Lei, I. Seskar, L. Greenstein, and P. Spasojevic, "Characterization of the orbit indoor testbed radio environment," in *Proc. IEEE VTC-2007 Fall*, October 2007, pp. 946–950.

[25] T. S. Rappaport and T. Rappaport, *Wireless Communications: Principles and Practice (2nd Edition)*. Prentice Hall PTR, 2001.

[26] A. O. Hero, J. A. Fessler, and M. Usman, "Exploring estimator bias-variance tradeoffs using the uniform cr bound," *IEEE Tr. Sig. Proc*, vol. 44, pp. 2026–2041, 1996.

[27] IEEE, "IEEE Standard 802.11b - Wireless LAN Medium Access Control (MAC) and Physical layer (PHY) Specications: High Speed Physical Layer(PHY) in the 2.4 GHZz Band," 1999.