Often, sensor localization bounds are derived assuming that received signal strength (RSS) measurements are performed with perfectly known sensor transmit powers. In this paper the Bayesian Cramér-Rao bound (CRB) is derived assuming transmit powers are random, with known prior distribution. Further, we use both directional measurements on each link, from $i$ to $j$ and from $j$ to $i$, and explicitly consider their correlation. Results show that random transmit powers have a small (typically 5-13%) impact on coordinate estimation bounds. However, using only the average of the directional measurements can significantly increase localization lower bounds.

1. INTRODUCTION

As part of the self-organizing nature of wireless sensor networks, sensors will perform cooperative localization to estimate their own coordinates. In cooperative localization, a small fraction of nodes will have $a$ priori known coordinates, either because they are fixed and have been given their coordinate by a human moderator, or because they are capable of estimating their coordinate using an independent method, such as GPS. This paper addresses the case in which sensors measure the received signal strength (RSS) of messages sent by nearby neighbors in the network. The RSS measurements and prior coordinate information are used to estimate the coordinates of all sensors.

A system designer bears the burden of ensuring that the accuracy of the coordinate estimates will meet the requirements of the application. Applications such as geographic-based routing, inventory management and logistics, environmental monitoring, and source localization, will all require sensor localization, at certain degrees of accuracy. The system designer should design sensors and algorithms such that, over the ensemble of sensor network deployments, the localization variance of the sensors meets the needs of the system.

The Cramér-Rao bound (CRB) \cite{1, 2, 3, 4, 5} and the Bayesian CRB \cite{6} have been used to provide a tool to aid in the design of cooperative localization systems. This paper complicates the analysis of cooperative localization from RSS measurements in two aspects:

1. Both directional measurements (from $i$ to $j$, and from $j$ to $i$) and their correlation are considered explicitly,

2. The transmit power is not assumed to be perfectly known.

Both represent realistic representations of RSS measurements in wireless sensor networks. Item (1.) helps quantify the benefit in measuring RSS in both directions of a link between two sensors. Item (2.) helps to quantify the loss in accuracy when the transmit powers of each sensor is not known exactly. In realistic sensor networks, due to the cost of device calibration, sensors do not know their precise transmit power level. While they may report that they are transmitting at a nominal power level, their actual transmit power may vary by a few dB, because of device manufacturing variations, because of battery level, or because of the position of and objects near the antenna.

2. PROBLEM STATEMENT AND MODELS

The cooperative localization problem is the estimation of sensor coordinates $z_i = [x_i, y_i]^T$ for sensors $i = 1 \ldots n$ (where $^T$ indicates transpose), given the $a$ priori known coordinates $z_i$ for $i = n + 1 \ldots N$, where $N = n + m$. This paper assumes 2D coordinates, although the analysis could be readily extended to the 3D case.

It is assumed that measurements of received signal strength are made between many pairs of sensors. In particular, $P_{i,j}$, the received signal strength measured at $i$ of the signal transmitted by $j$, is measured for each pair $(i, j)$ in a measurement set $H$. This set $H$ is a subset of $\{1, \ldots, N\}^2$, since not all pairs may be in range, or limited bandwidth may
are constant during short periods of RSS measurements. Also, \((i, i)\) is not in \(H\) since self-RSS measurement is not done.

2.1. Pair Measurement Correlation

Two measurements can be made between each pair of sensors \((i, j)\): one with \(i\) as the transmitter and \(j\) as the receiver, and the other with them switched. For the purposes of analysis, it would be simpler to assume that either these two measurements are either perfectly correlated, or that they are not correlated at all. In reality, the measurements are highly correlated, due to the identical multipath shadowing channel which affects both directional links identically. As a result, past research has assumed, for simplicity, that the two measurements were perfectly correlated [2].

In this paper, we complicate this analysis by considering that the two directional measurements are not perfectly correlated. These differences come from: additive noise, movement in the channel in the time between the two measurements, and differences in the transmitter and receiver characteristics of the two devices. We will assume that \(P_{i,j}\) and \(P_{j,i}\) are correlated with correlation coefficient \(0 \leq \rho \leq 1\). This model is specified in Section 2.3. We also consider in Section 4.4 the effect of estimating coordinates using only the average of \(P_{i,j}\) and \(P_{j,i}\), for all pairs \((i, j)\), instead of the full measurements.

2.2. Random Transmit Power

This paper further complicates the assumption that the transmit power is known. In particular, we consider \(\Pi_{0i}\), the signal power (dBm) at a distance of \(\Delta_0 = 1\)m from the antenna of sensor \(i\). By considering the power slightly away from the transmitter antenna, we effectively consider both the device and battery variation and the variation in the antenna and the environment immediately around the antenna. In an idealized environment, \(\Pi_{0i}\) is given by the Friis transmission equation, and in a real channel, the near-field environment affects the antenna impedance and thus the antenna radiation efficiency [7].

We also assume that \(\Pi_{0i}\) is constant across many measurements. While battery levels and the near-field environment may change over time, we assume that these effects are constant during short periods of RSS measurements.

2.3. Combined Channel Model

In particular, we model the vector \(p_{ij} = [P_{i,j}, P_{j,i}]^T\) as bivariate Gaussian with mean \(\mu_{ij}\) and covariance \(C_{ij}\), where

\[
\mu_{ij} = \begin{bmatrix}
\Pi_{0j} - 10n_p \log_{10} \frac{\|s_i - z_i\|^2}{\Delta_0^2} \\
\Pi_{0i} - 10n_p \log_{10} \frac{\|s_j - z_j\|^2}{\Delta_0^2}
\end{bmatrix},
\]

\[
C_{ij} = \sigma_{dB}^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix},
\]

where \(n_p\) is the path loss exponent of the environment of interest, and \(\rho\) is the correlation coefficient, \(0 \leq \rho \leq 1\). To provide an intuitive feel for practical values of the parameters in (1), from the RSS measurements reported in [2], we can calculate that \(\sigma_{dB} = 4.25\) dB and \(\rho = 0.704\). We further assume that \(\{p_{ij}\}_{i,j}\) are statistically independent. The independence assumption is for simplification; future work should address the effects of correlations between different links in the network.

3. BOUNDS ANALYSIS

The unknown parameters defined in Section 2 are \(\theta\),

\[
\theta = [x^T, y^T, \Pi^T]^T
\]

\[
x = [x_1, \ldots, x_n]^T
\]

\[
y = [y_1, \ldots, y_n]^T
\]

\[
\Pi = [\Pi_{01}, \ldots, \Pi_{0N}]^T
\]

There are \(2n + N\) parameters since none of the \(N\) sensors have perfect knowledge of their transmit power, and \(n\) of the sensors have no prior coordinate knowledge. We assume perfect prior coordinate information for sensors \(n + 1\) through \(N\), and an i.i.d. Gaussian prior on \(\Pi_{0i}\) for all sensors \(i\), with variance \(\sigma^2_{\Pi}\) and mean \(\bar{\Pi}_{0i}\).

The Bayesian Cramér-Rao Bound [8, pp. 72-73] (also called the van Trees inequality or posterior CRB) states that any estimator, \(\hat{\theta}\), must have error correlation matrix \(R_e\) satisfying

\[
R_e \geq F^{-1} \triangleq [F_{\theta} + F_P]^{-1},
\]

where \(R_e \triangleq \mathbb{E}[(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T]\) (4)

where \(F_{\theta}\) is the Fisher information matrix, and \(F_P\) is the prior information matrix,

\[
F_{\theta} = -\mathbb{E} \left[ \nabla_{\theta} (\nabla_{\theta} \ln f(X|\theta))^T \right],
\]

\[
F_P = -\mathbb{E} \left[ \nabla_{\theta} (\nabla_{\theta} \ln f(\theta))^T \right],
\]

where \(\mathbb{E}[]\) indicates expected value, \(\nabla_{\theta}\) is the gradient operator w.r.t. \(\theta\).
3.1. Average and Difference of Measurement Pairs

Before calculating the Fisher information, we transform the two directional measurements \( p_{ij} \) for the purposes of discussion. It can be seen intuitively or from (11) that a full-rank transformation of the measurements does not change their Fisher information. We choose to transform each pair of measurements \( p_{ij} \) by orthogonal matrix \( A \),

\[
\hat{p}_{ij} = Ap_{ij}, \quad A = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.
\]

(6)

For notational purposes, we denote \( \hat{p}_{ij} = [\hat{p}_{ij}, p_{ij}^\Delta]^T \). The top element \( \hat{p}_{ij} \) is the average of the two measurements \( P_{ij,i} \) and \( P_{ij,j} \), and the bottom element, \( \hat{p}_{ij}^\Delta \) is half of their difference. These two random variables are still Gaussian, but they are independent, since

\[
\hat{R}_{ij} = A^T C_{ij} A = \frac{\sigma^2}{2} \begin{bmatrix} 1 + \rho & 0 \\ 0 & 1 - \rho \end{bmatrix}.
\]

(7)

where \( \rho \) is the correlation coefficient from (1). Because of (6), \( \hat{p}_{ij} \) and \( p_{ij}^\Delta \) have means \( \hat{\mu}_{ij} \) and \( \mu_{ij}^\Delta \), respectively, where

\[
\hat{\mu}_{ij} = \frac{\Pi_{0j} + \Pi_{0i} - 10n_p \log_{10} \| z_i - z_j \|^2}{2 \Delta^2_0},
\]

\[
\mu_{ij}^\Delta = \frac{1}{2}(\Pi_{0j} - \Pi_{0i}).
\]

(8)

We define \( \hat{\mu} \) and \( \mu^\Delta \) as the list of the means for all measured pairs \((i, j)\),

\[
\hat{\mu} = [\hat{\mu}_{i_1j_1}, \ldots, \hat{\mu}_{i_Sj_S}]^T,
\]

\[
\mu^\Delta = [\mu^\Delta_{i_1j_1}, \ldots, \mu^\Delta_{i_Sj_S}]^T
\]

(9)

where \( i_1j_1, \ldots, i_Sj_S \) is a listing of each unique pair which makes measurements, and \( S \) is the number of unique pairs measured (taken from set \( H \) defined in Section 2).

3.2. Derivation of Fisher Information

As a result of their independence, we can split the Fisher information into two matrices, \( F_\theta \) for the average measurements \( \{\hat{p}_{ij}\} \), and \( F^\Delta_\theta \) for the difference measurements \( \{p_{ij}^\Delta\} \), and

\[
F_\theta = F^\Delta_\theta + \bar{F}_\theta.
\]

(10)

We know that for a vector of multivariate Gaussian measurements \( P \) with mean \( \mu(\theta) \) and covariance \( C \) (for \( C \) not a function of \( \theta \)), that the Fisher information matrix is given by [9],

\[
F_\theta = [\nabla_\theta \mu(\theta)]^T C^{-1} [\nabla_\theta \mu(\theta)]
\]

(11)

Thus, for the two terms of (10),

\[
F^\Delta_\theta = [\nabla_\theta \mu^\Delta]^T (C^\Delta)^{-1} [\nabla_\theta \mu^\Delta]
\]

\[
F^\Delta_\theta = [\nabla_\theta \mu^\Delta]^T (C^\Delta)^{-1} [\nabla_\theta \mu^\Delta]
\]

(12)

where the covariance matrices are,

\[
C = \frac{(1 + \rho)\sigma^2}{2} I_{2n + N}, \quad C^\Delta = \frac{(1 - \rho)\sigma^2}{2} I_{2n + N},
\]

and \( I_{2n + N} \) is the \( 2n + N \times 2n + N \) identity matrix. The elements of \( \nabla_\theta \mu^\Delta \) are given as,

\[
\frac{\partial \mu_{ij}^\Delta}{\partial \Pi_{0k}} = \begin{cases} 1/2, & \text{if } k = i \\ 1/2, & \text{if } k = j \\ 0, & \text{otherwise} \end{cases}
\]

(13)

where \( \alpha = 10n_p/\log 10 \). Basically, there are six non-zero elements of \( \nabla_\theta \mu^\Delta \) for each pair \((i, j)\).

For \( \nabla_\theta \mu^\Delta \), we have simply that \( \frac{\partial \mu^\Delta_{ij}}{\partial x_k} = \frac{\partial \mu^\Delta_{ij}}{\partial y_k} = 0 \), \( \forall i, j, k \), and that

\[
\frac{\partial \mu^\Delta_{ij}}{\partial \Pi_{0k}} = \begin{cases} -1/2, & \text{if } k = i \\ 1/2, & \text{if } k = j \\ 0, & \text{otherwise} \end{cases}
\]

(14)

\[
3.3. Prior Information
\]

As stated at the start of this section, we assume an i.i.d. Gaussian prior on \( \Pi_{0i} \) for all sensors \( i \), with variance \( \sigma^2_0 \) and mean \( \bar{\Pi}_0 \). We assume no prior exists for the coordinates of sensors 1 through \( n \), although this analysis could be readily modified to include such effects, as reported in [6]. Because of these assumptions, the prior information matrix \( F_P \) is,

\[
F_P = \text{diag} \left\{ \begin{bmatrix} 0_n & 0_n^T & \frac{1}{\sigma_\Pi^2} T_N \end{bmatrix} \right\}
\]

(15)

where \( 0_n \) is an \( n \)-length vector of zeros, and \( 1_N \) is an \( N \)-length vector of ones.

4. NUMERICAL RESULTS

Given the analysis of Section 2, in this section, we compute the Bayesian CRB \( F^{-1} \) from (3) for a number of examples. In particular, in the figures, we plot the lower bound on the 1-\( \sigma \) uncertainty ellipse for \( z_i \), the estimate of the \( i \)th sensor coordinate, for each sensor \( i = 1 \ldots n \). Also, we summarize each lower bound example by reporting the lower bound on RMS localization error \( \bar{\sigma}_z \) and RMS transmit power estimation error \( \bar{\sigma}_\Pi \),

\[
\bar{\sigma}_z = \frac{1}{n} \text{tr} F^{-1}_z, \quad \bar{\sigma}_\Pi = \frac{1}{N} \text{tr} F^{-1}_\Pi
\]

(16)

where \( F^{-1}_z \) and \( F^{-1}_\Pi \) are defined as the upper left \( 2n \times 2n \) and the lower right \( N \times N \) submatrices of \( F^{-1} \), respectively.
4.1. Grid with Measured Parameters

First, we test a grid of fully-connected sensors (all pairs \((i, j)\) for \(i \neq j\) are measured) using the parameters measured in [2], \(\sigma_{dB} = 4.25\) and \(\rho = 0.704\). Here, \(N = 16\), and the sensors are located in a 4 \(\times\) 4 grid contained within a 10m by 10m square area. If the transmit powers were known exactly, the lower bound on the 1-\(\sigma\) uncertainty ellipses would be as shown in Fig. 1(a) by the solid lines (——), and \(\bar{\sigma}_z = 1.29\)m. In the case in which transmit powers are random with prior standard deviation \(\sigma_\Pi = 4\) dB, the lower bound on the 1-\(\sigma\) uncertainty ellipses increases slightly, as shown in Fig. 1(a) by the dashed lines (-----), and \(\bar{\sigma}_z\) has increased by 4.6\% to 1.35m. Note that \(\sigma_\Pi = 0.93\) dB, thus, actual powers \(\{\Pi_{0i}\}\) can be accurately estimated.

4.2. Limited Range

For this example, for the identical 4 \(\times\) 4 grid, only pairs for which \(\|z_i - z_j\| < r = 5\)m are measured. The results are shown in Fig. 1(b). While the lower bounds on RMSE have increased dramatically compared to the fully-connected case (by about 32\%), the difference between the \(\bar{\sigma}_z\) in the known and random transmit power cases is still low at 4.8\%.

4.3. Random Deployment

Instead of the grid geometry, we next test an \(N = 16\) sensor network in which sensor coordinates were chosen from a uniform distribution on \([0m, 10m]^2\). The \(m = 4\) sen-
sensors closest to the corners of the square are chosen to have known-location. All other parameters, including the connectivity radius \( r = 5 \text{m} \), are kept the same as in the previous example, and the results are shown in Fig. 1(c). Here, for the cases of known and random transmit power, \( \bar{\sigma}_z = 1.22 \text{m} \) and 1.37m, respectively, a difference of 13%. Typically, for random deployments, we see that the locally higher densities of sensors helps to reduce the lower bound for most sensors. However, with limited range, other random deployments may run into situations in which the least-connected sensor might make no measurements, or the network as a whole will be disconnected, and as a result, there may not be a finite lower bound. The radius \( r \) for randomly deployed sensors may need to be higher to avoid such cases.

4.4. Using Only the Averages \( \{ \bar{p}_{ij} \} \)

It may seem natural to use only the link average RSS \( \bar{p}_{ij} \) and to discard \( p_{ij} \) in order to reduce by half the measured data in the sensor network. We analyze the effects in this section by replacing (10) with \( F_{\theta} = F_{\bar{\theta}} \), eliminating the information contained within the measurements \( \{ p_{ij} \} \).

In this example, we use a 5\( \times \)5 grid of devices. Other parameters and \( r = 5 \text{m} \) are kept as in previous examples. The results, as shown in Fig 1(d), show the lower bound for the case when only link averages are considered \( (\cdots) \), and we have that \( \bar{\sigma}_z = 1.63 \). Compared to the case when all information is used, \( \bar{\sigma}_z \) has increased 25\% – for the case when all measurements are used \( (-\cdots-) \), \( \bar{\sigma}_z = 1.30 \). In particular, the nodes at the edge of the network see their lower bounds expand dramatically along their major axis.

Note that \( \bar{\sigma}_{n1} = 3.40 \) in the average-only case, compared to \( \bar{\sigma}_{n1} = 1.24 \) in the all-information case. When difference measurements are ignored, it becomes difficult to estimate transmit power. As a result, there is less information about the size or extent of the network, and nodes at the edges are affected most strongly by this lack of knowledge.

4.5. Performance as a Function of \( \rho \) and \( \sigma_{n1} \)

Finally, we calculate the bound over a range of correlation coefficient \( \rho \) and standard deviation of transmit powers \( \sigma_{n1} \). We use the 5\( \times \)5 grid of sensors, with the \( m = 4 \) corner sensors having known-location, as in the previous example. The parameters \( \sigma_{dB} = 4.25 \), \( n_p = 2.30 \), and \( r = 5 \text{m} \) are constant through these calculations. We plot \( \bar{\sigma}_z \) in Fig 2(a) and \( \bar{\sigma}_{n1} \) in Fig 2(b).

Note that as \( \rho \to 1 \), the measurement \( p_{ij}^{\Delta} \) becomes exactly the difference \((\Pi_{ij} - \Pi_{0i})/2 \). Since the differences are known perfectly, the only unknown is the absolute, which represents only one degree of freedom. As a result, lower bounds (for localization or transmit power estimation) are not strongly impacted in this case.

Another limiting case is for \( \sigma_{n1} \to \infty \). Although it is not shown explicitly, the \( \sigma_{n1} = 10 \text{ dB} \) case in Fig 2 approximates the asymptotic bound very well. As the prior information on \( \{ \Pi_{0i} \} \) goes to zero, their estimation remains possible, because each sensor has redundancy in the form of many neighbors measuring the RSS of its transmissions.

5. CONCLUSION

The Bayesian CRB presents a means to bound the performance of sensor location estimation when two real-world complications are present in the measurements: correlation between the two measurements made on one link, and random sensor transmit powers. This paper presents this bound and compares it to the bound for the ideal case. These complications, for several numerical examples, are shown to increase the lower bound from 5-13\%. In particular, when transmit powers are random, the bound shows the importance of preserving RSS measurements for both directions.
on each link, instead of summarizing them by their average.

Future work may derive the lower bounds explicitly for the cases when $\rho \to 1$ or when $\sigma_{\Pi} \to \infty$. Also, it will be critical to develop algorithms robust to random sensor transmit powers, and to test them against these bounds.

6. REFERENCES


