

SENSOR NODE LOCALIZATION VIA SPATIAL DOMAIN QUASI-MAXIMUM LIKELIHOOD ESTIMATION

Seshan Srirangarajan and Ahmed H. Tewfik

Department of Electrical and Computer Engineering, University of Minnesota
200 Union Street SE, 4-174 EE/CSci Bldg., 55455, Minneapolis, USA
{seshan, tewfik}@umn.edu

ABSTRACT

A Sensor node localization algorithm for indoor quasi-static sensor environments using spatial domain quasi-maximum likelihood (QML) estimation is presented. A time of arrival (TOA) based algorithm is used to arrive at the “pseudo” range estimates from the base stations to the sensor nodes. The localization algorithm uses spatial domain quasi-maximum likelihood estimation to determine the actual sensor location. The algorithm is preceded by a calibration phase during which statistical characterization of the line-of-sight (LOS) and non-line-of-sight (NLOS) returns are derived. Using a synthesized bandwidth of 2GHz, a 4-bit analog-to-digital converter (ADC) and 5-10dB signal-to-noise ratio (SNR), localization with high accuracy is achieved.

1. INTRODUCTION

Wireless sensor networks are becoming increasingly popular as a result of recent advances in low-power circuit design, availability of simple yet reasonably efficient wireless communication equipment and reduced manufacturing costs [1]. These networks typically combine wireless communication components, minimal computation capabilities and some sensing of the physical environment into a network. All these components together in a single device form a sensor node.

Sensor networks are used in a number of surveillance type tasks, such as asset tracking, finding people in emergency situations etc. Sensor nodes also measure some physical quantity(s) at a given position. Thus having knowledge of the sensor locations is very essential to these applications. This feature is also emphasized in the IEEE 802.15.3a and IEEE 802.15.4a wireless personal area networks.

Traditional localization techniques use one or more of the following measures: received signal strength indicator (RSSI) [2][13], time of arrival (TOA) [3][4], time difference of arrival (TDOA) [5], or angle of arrival (AOA) [6]. Some approaches have explored the use of RSSI fingerprinting technique for localization [13], but this requires extensive offline calibration which is not practical in most applications. We concentrate on the time-based approaches which are most suitable for applications that need high ranging accuracy.

The presence of multipath components in the received signal due to non-line-of-sight (NLOS) propagation makes the problem of ranging especially challenging. Most of the

work in the literature is based on assumptions that the line-of-sight (LOS) signal is always present, for e.g. [3], [4]. In [4] it is assumed that the LOS signal is the earliest arrival and the results showed that the estimation error increases rapidly with the transmitter to receiver range. *These approaches, despite using bandwidths in excess of 1GHz, do not achieve the high ranging accuracy that some sensor network applications require.* Many approaches have also been proposed in the literature that find some distinct properties of the NLOS range measurements to distinguish them from LOS measurements, for e.g. [7] and [8]. Using pure statistical characteristics to distinguish NLOS measurements from LOS measurements is a difficult problem.

Most proposed approaches dealing with the NLOS challenge can be decomposed into two steps. In the first step, the TOA (or TDOA) measurements associated with the visible base stations (BS) are obtained. In the second step, a location algorithm is implemented, fusing the measurements obtained in the first step into a position estimate. It is seen that the geometric relationship between the TOAs is logically exploited in the second step to obtain the position estimate but is not used to assist in the TOA estimation of the first step [9].

We present here, a novel approach that fuses the range (or TOA) estimation and localization phases in such a way that the geometric relationship of the TOAs also aids in the range estimation. This lowers the probability of reporting NLOS range estimates. This additional information is available while estimating the range and improves the localization accuracy dramatically compared to the traditional approaches. At the same time our algorithm is computationally efficient.

In our approach, the signal (LOS/NLOS) and noise peaks are characterized during an initial calibration phase. The ranging algorithm gives multiple *pseudo* range estimates. One of these estimates will be due to the LOS component while others will be due to NLOS components or noise. The range estimates from all visible BSs are combined in a spatial-domain quasi-maximum likelihood (QML) estimation technique to arrive at the final location estimate. This method will be shown to give more accurate position estimates than previously reported approaches.

It will also be shown that the calibration phase is sufficient to gather all the necessary channel information and no other prior information about the channel is needed. Suitable

modifications to this algorithm enable it to be used in low bandwidth (BW) and multi-band systems.

The rest of the paper is organized as follows. Section 2 presents the sensor localization setting which we call, the localization model. Section 3 describes the calibration phase for statistical modeling. Section 4 describes the ML location estimation algorithm. Section 5 presents simulation and preliminary experimental results. Section 6 gives some concluding remarks and future research direction.

2. LOCALIZATION MODEL

Consider a model in which a number of sensor nodes and base stations communicate with each other for localization of the nodes. Each node that needs to be localized transmits a ranging pulse. All base stations and previously localized nodes which receive the transmission, record the received pulse. Each of these BSs (or nodes) executes the first phase of the localization algorithm to identify the pseudo range estimates. The pseudo range estimates are then communicated to a central BS along with the position of the node that measured them. At the central BS this information is fused with similar information from other BSs or nodes to produce an estimate of the sensor location. The BSs and nodes need to be time-synchronized using techniques such as [14]. We also briefly discuss some of the issues related with time-synchronization in our experimental results section.

This model uses an asymmetric setup where most of the computation is done at a central BS while minimal processing is done at the sensor nodes. This is a very desirable feature for sensor networks since sensor nodes are very resource constrained and hence should be required to do only a minimal amount of computation.

2.1 Ranging Pulse Characteristics

We assume that the sensor nodes and BSs transmit a short duration Gaussian monopulse with a bandwidth of 528 MHz. In multi-band communication systems, the whole bandwidth is divided into several sub-bands. In each time interval, a signal is transmitted in one of the sub-bands. At the receiver, signals from 4 sub-bands are combined to give a virtual large bandwidth ($\approx 2\text{GHz}$) signal using the technique in [10]. Signal in each of the sub-bands is sampled at 1GHz and then upsampled by a factor of 4 to give an effective sampling rate of 4GHz. Higher sampling rates are achieved via processing in the digital domain. The transmit signal after going through the channel is input to a matched filter receiver. The matched filter output is subject to thresholding to detect local peaks. The threshold is chosen based on the desired error performance and the estimated signal-to-noise ratio (SNR). Figure 1 shows the matched filter output plotted as signal amplitude (y-axis) versus the sample number (x-axis), in the presence of Gaussian noise.

The waveform shown in Figure 1 has significant multipath components and the signal peak does not occur at the leading edge of the waveform. We record the received signal peaks at each BS and use these to estimate the pseudo (or candidate) range estimates. We call these *pseudo* range esti-

mates since only one of them is due to the LOS component; all others are due to NLOS components or noise.

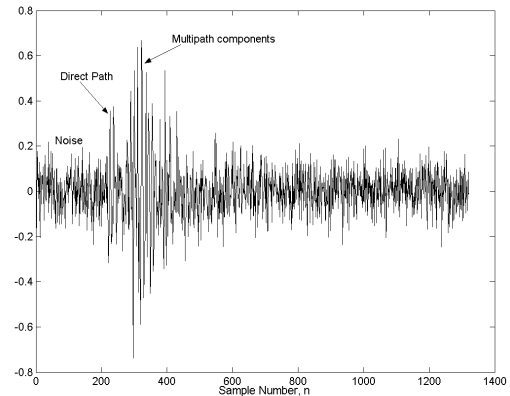


Figure 1- Matched filter output in the presence of Gaussian noise

The goal of the ranging technique would be to pick only the LOS estimates to be used in the next phase for location estimation. The calibration phase, explained next, aids in characterizing the signal components (LOS and NLOS) and noise, which will be used in the ML estimation algorithm.

3. CALIBRATION AND STATISTICAL MODELING

In the calibration phase, two sets of training runs are carried out. The first set is carried by averaging several measurements over a short time interval, to achieve a virtual high SNR environment. Alternatively, we can use a reliable channel model, such as the IEEE 802.15.3a. In a high SNR environment the local peaks detected will either be due to the LOS or the NLOS components. The measured peak strengths are normalized by the strength of the global peak in the output of the matched filter. *Global peak* is defined as the largest peak detected in a single received ranging pulse. Histograms are estimated for the strengths of the LOS and NLOS peaks using these measurements.

In this paper, we rely on the IEEE 802.15.3a channel model 3 (CM3). The second set of training runs are carried out in the presence of Gaussian noise but in the absence of a transmit signal, i.e., under noise-only conditions. The procedure outlined above is followed to estimate a histogram for the relative strength of the noise peaks. The histograms for the relative strengths are obtained using a channel model or results from the first set of training runs. Here we used the IEEE 802.15.3a channel model 3 (CM3) [11].

The relative strength of the signal peaks follows an exponential distribution, whereas that of the noise peaks follows a lognormal distribution. By normalizing these histograms to unit area we obtain probability distributions for the relative strengths of the signal and noise peaks, which we denote by $f_{\text{signal}}(\rho)$ and $f_{\text{noise}}(\rho)$, respectively.

$$f_{\text{signal}}(\rho) = \frac{1}{\beta} \exp\left(-\frac{\rho}{\beta}\right)$$

$$f_{\text{noise}}\left(\ln\left(\frac{\rho}{\rho_0}\right)\right) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\ln^2\left(\frac{\rho}{\rho_0}\right)}{2\sigma^2}\right) \quad (1)$$

The histograms for the time difference between the location of the local peaks and the global peak in the matched filter output are also estimated. (The histograms are not shown here as they simply corroborate the channel model used in our simulations). This characterization of the signal and noise peaks is used in the location estimation algorithm described next.

4. LOCATION ESTIMATION ALGORITHM

4.1 Identifying Pseudo Range Estimates

The ranging pulse transmitted by a sensor node is received by all BSs within its radio range. At each BS, a threshold is chosen to detect peaks in the received signal. The detected peaks are recorded in terms of their signal strength (ρ_i , normalized with respect to the global peak in the received signal) and the time difference (δ_i) between the detected peak and the global peak. The time difference is used to estimate the pseudo range (r_i) due to each of the detected peaks (as if each of the peaks were due to the LOS component). Let t_{peak} denote the time stamp of the global peak, then:

$$r_i = c.(t_{\text{peak}} - \delta_i). \quad (2)$$

Each pseudo range estimate gives a circle centered on the corresponding BS with radius r_i , on which the sensor node could lie.

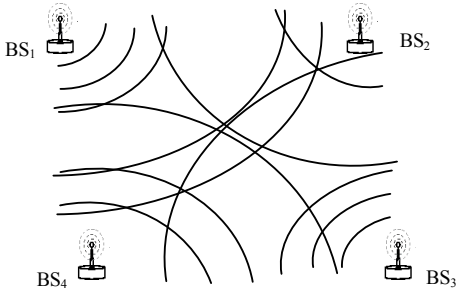


Figure 2 - Pseudo range estimates from 4 BSs to a sensor; the sensor node is located at the intersection of the LOS ranges

4.2 Spatial Domain Location Estimation Algorithm

The localization procedure is described below:

- Step 1. At each BS, detect peaks in the received signal by setting a threshold; record (ρ_i, δ_i) pair for each peak.
- Step 2. Estimate r_i for each peak detected in step 1.

Step 3. Each BS transmits the recorded information along with its location information to a central processing location.

Step 4. After information from all the BSs has been received at the central location, the *area of interest*, where the sensor node could possibly exist, is divided into a grid of cells. (Note: *Area of interest* is determined based on the known locations of the BSs and the largest pseudo range estimate from each BS. This is possible since pseudo range gives how far the sensor could be located from the BS). The cell size is determined by the smallest resolvable time interval during the pseudo range identification phase.

Step 5. Represent the grid of cells as a matrix. The matrix is populated such that the entry in each cell represents the likelihood of the cell containing a LOS range estimate. This computation is explained next.

4.2.1 Likelihood Matrix Computation

Let $y_i(t) = x(t) + n_i(t)$, be the received signal at the i^{th} BS, where $x(t)$ represents the transmitted signal and $n_i(t)$ the Gaussian noise. We assume that noise is independent for each BS and for each received pulse.

Let ρ_i be the strength of the peak, mapped to a cell, due to the signal received at the i^{th} BS. Then, at any BS the likelihood of the peak being due to signal or noise is given by $f(\text{signal}/\rho)$ or $f(\text{noise}/\rho)$, respectively.

$$f(\text{signal} / \rho) = \frac{f(\rho / \text{signal}) \cdot f(\text{signal})}{f(\rho)} = \frac{f_{\text{signal}}(\rho) \cdot f(\text{signal})}{f(\rho)}$$

$$f(\text{noise} / \rho) = \frac{f(\rho / \text{noise}) \cdot f(\text{noise})}{f(\rho)} = \frac{f_{\text{noise}}(\rho) \cdot f(\text{noise})}{f(\rho)}$$

$$\dots (3)$$

To compute the location of the sensor, we need to combine the measurements from all BSs. There are two possibilities to consider. These correspond to two models for the path loss from the sensor node to each of the BSs, and the likelihood of any peak being due to signal or noise: (i) We may consider the path losses to be correlated since they are dominated by a deterministic loss that is a function of the BSs and sensor geometry. In this case, extensive calibration would be needed to model each of the individual path losses which is not practical. (ii) We may alternatively consider the path losses to be random and independent due to the arbitrary placement and dynamic structure of the obstructions between the sensor and each of the BSs. In this case, we can assume that the path losses are independent realizations drawn from the same distribution. We use the latter model as a reasonable approximation in the presence of multipath propagation.

At each cell, peaks from a number of BSs each with strengths ρ_i ($i = 1, \dots, M$), are reported. If each of these is an independent observation, the overall likelihood function would be a product of the individual likelihoods. Otherwise, the product of the individual likelihoods is not the overall likelihood. Nevertheless it provides a reasonable cost function that we will maximize below. We refer to this cost function as a quasi-likelihood function.

Let $A_{measured,i}$ represent the matrix of likelihood values for the i^{th} BS. Since each pseudo range estimate comes from a peak in the received signal, the likelihood of it being due to the LOS signal component is related, directly to the likelihood of the peak coming from the signal (LOS/NLOS) distribution, and inversely to the likelihood of the peak being from the noise distribution.

For each detected peak with relative strength, say ρ_l , the likelihood of it being due to noise, $f_{noise}(\rho_l)$, is obtained from the noise distribution. Locate ρ_l on the x-axis of the noise distribution and the corresponding value on the y-axis gives $f_{noise}(\rho_l)$. This computation is shown in Figure 3. The likelihood $f_{signal}(\rho_l)$ is calculated in a similar manner using the signal distribution. We denote $f_{signal}(\rho_l)$ and $f_{noise}(\rho_l)$ being assigned to cell (j, k) by $f_{signal}(j, k)$ and $f_{noise}(j, k)$, respectively.

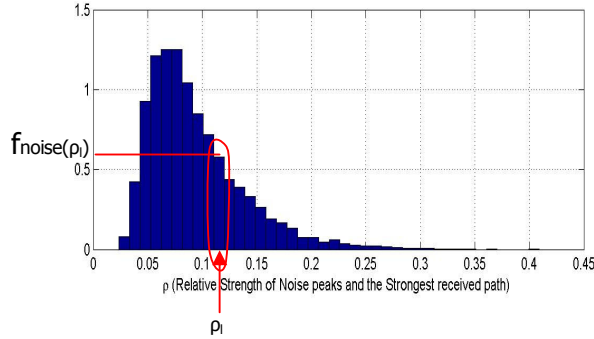


Figure 3 - Illustration of likelihood function evaluation

Using (3) we obtain:

$$A_{measured,i}(j,k) = \frac{f(signal / \rho_{j,k})}{f(noise / \rho_{j,k})} = \frac{f_{signal}(j,k)}{f_{noise}(j,k)} \quad (4)$$

In (4), i is the BS index and (j, k) are used to index the matrix or cell entries. Each of the matrices $A_{measured,i}$ indicates where the sensor is most likely to be present. Thus by overlaying each of these matrices, one over the other, on the cell grid would give the overall likelihood distribution.

$$A_{overall}(j,k) = \prod_i A_{measured,i}(j,k) \quad (5)$$

The cell with the maximum valued entry when mapped to the area of interest gives the location estimate (x_s, y_s) of the sensor node.

$$(x_s, y_s) \Leftrightarrow (j_s, k_s) = \max_{j,k} \left\{ \prod_i \left(\frac{f_{signal}(j,k)}{f_{noise}(j,k)} \right) \right\} \quad (6)$$

4.2.2 Intuitive Explanation

Each range estimates r_i can be represented by a circle around the BS on a map of the sensor network. The intersection of the circles from all the visible BSs would give a grid of possible locations for the target or the source. The position with the maximum likelihood (or quasi-likelihood) value would be chosen as the sensor node location estimate.

5. RESULTS

5.1 Simulation Results

Extensive simulations were carried out using 250 different channel realizations based on the IEEE 802.15.3a channel model. Noise is assumed to be independent for each signal return. A 4-bit analog-to-digital converter (ADC) is used in the receiver circuitry and the SNR reported includes the quantization noise effects due to the ADC.

We estimate the sensor position using range estimates from 5 BSs. The BS coordinates are generated randomly using a uniform distribution on a square grid and the SNR is fixed at 10dB. The simulation results from 50 trial runs are shown in Figure 4, where the BS locations are represented by different shapes $\{\triangle, \nabla, \square, \circ \text{ etc}\}$.

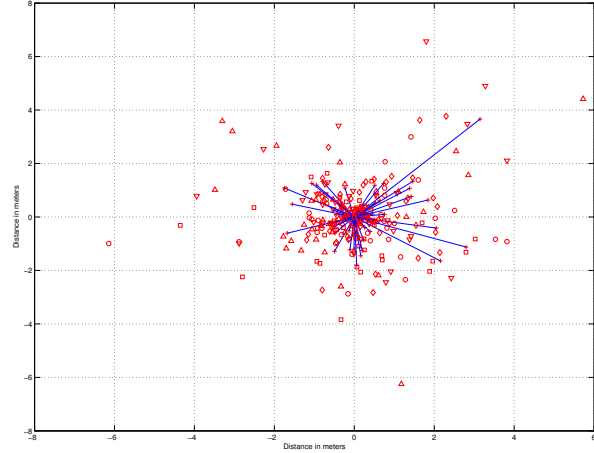


Figure 4 - Simulation results: Solid lines indicate the error between the estimated and true sensor positions.

The actual sensor location was $(0, 0)$. It can be seen from Table 1 that localization accuracy improves with SNR and there is no significant degradation in the accuracy even when the SNR is reduced to 5dB. The ranging accuracy achieved here is significantly better than the approaches reported in the literature.

TABLE 1. POSITION ESTIMATION ERROR (IN METERS)

SNR	Mean Error	RMS Error
5 dB	1.2080	0.9275
10 dB	1.1080	0.8860
20 dB	1.1055	0.7785

The approach in [4] looks for the LOS component in a small window of the received signal. Estimating the window becomes extremely difficult at long ranges due to the complex LOS blockage, resulting in large estimation errors. Our approach avoids this problem as we do not try to locate the LOS position in each received signal. Rather we combine the pseudo range estimates from a number of BSs and the final location estimate is the one with the maximum likelihood of having a signal component. This works well because it is highly unlikely that a majority of the BSs would report NLOS estimates which overlap at any given point (or cell).

5.2 Preliminary Experimental Results

We are conducting experiments using the MICAz mote module (2.4GHz, IEEE 802.15.4 compliant with an overall BW of 80MHz). In the absence of a way to tap into the raw RF transmit/receive signal at the sensor nodes, we are using the *start of frame delimiter* (SFD) signal to measure the time of flight (TOF) between nodes. We characterized the accuracy of the TOF measurements taken at a single node. Any improvements to the range estimation, as a result of calibrating the sensor nodes, would increase the localization accuracy as well.

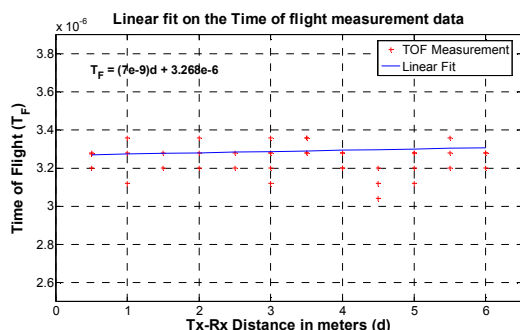


Figure 5 - Linear fit for the TOF measurements taken at a single node

The plot in Figure 5 shows the results of the TOF measurement between two nodes at distances between 0-6m with 0.5m increments. The linear fit shows a baseline measurement offset of 3.27 μ s. The offset could be due to a combination of packet processing delay, random back-off time due to the CSMA-CA protocol, clock offset between nodes and granularity due to the low accuracy clock. This error can be compensated for using a linear calibration based on these measurements. We characterize each sensor node in this manner and the averaged calibration coefficients are applied to every node. This would further reduce the range estimation error.

6. CONCLUSION AND FUTURE WORK

A sensor localization algorithm, based on spatial domain quasi-maximum likelihood estimation, for multipath propagation conditions has been presented. This algorithm fuses the range estimation and localization phases such that the geometric relationship of the TOAs aids the range estimation phase as well. The algorithm is well suited for asymmetric implementation where most of the computation is done at a central BS. Simulation results have shown that this approach gives enhanced ranging accuracy. The accuracy achievable is limited only by the cell size chosen and the SNR. The cell size is related to the time resolution achievable with the ranging signal being used, which in turn depends on its bandwidth. Also, by not trying to directly estimate the position of the LOS component in each of the received signals, we overcome the difficult problem of NLOS identification. The measurement campaign with the MICAz motes is aimed at refining our location estimation algorithm for low bandwidth and multi-band systems.

REFERENCES

- [1] K. Holger, and A. Willig, "A Short Survey of Wireless Sensor networks", *TKN Technical Reports Series*, Technical University Berlin, Berlin, pp. 1-19, Oct 2003.
- [2] J. G. Castano, M. Svensson, and M. Ekstrom, "Local Positioning for Wireless Sensor Networks Based on Bluetooth," *IEEE Radio and Wireless Conf.*, pp. 195-198, Sep 2004.
- [3] W. C. Chung, and D. S. Ha, "An Accurate Ultra Wide-band Ranging for Precision Asset Location", *Int. Conf. UWB Systems and Technologies*, Reston, Virginia, Nov 2003, pp. 383-393.
- [4] J.-Y. Lee, and R.A. Scholtz, "Ranging in a Dense Multipath Environment Using an UWB Radio Link", *IEEE J. Selected Areas in Communications*, vol. 20, no. 9, pp. 1677-1683, Dec 2002.
- [5] K.C. Ho, and W. Xu, "An Accurate algebraic solution for moving source location using TDOA and FDOA measurements", *IEEE Trans. Signal Processing*, vol. 52, Issue 9, Sep 2004.
- [6] A. Pages-Zamora, J. Vidal, and D.H. Brooks, "Closed-form solution for positioning based on angle of arrival measurements", *The 13th IEEE Int. Symposium Personal, Indoor and Mobile Radio Communications*, vol. 4, Sep 2002.
- [7] J. Borras, P. Hatrack, and N.B. Mandayam, "Decision Theoretic Framework for NLOS Identification", *48th IEEE Vehicular Technology Conf.*, pp. 1583-1587, May 1998.
- [8] M.P. Wylie-Green, and S.S. Wang, "Robust Range Estimation in the Presence of Non-Line-of-Sight Error", *54th IEEE Vehicular Technology Conf.*, pp. 101-105, Oct 2001.
- [9] A. Urrela, E. Ström, and M. Rydström, "Novel Wireless Location Approach for W-CDMA Systems based on Multiple Sliding Correlators", *Int. Conf. on Acoustics, Speech, and Signal Processing*, pp. 945-948, Mar 2005.
- [10] E. Saberinia, and A.H. Tewfik, "Enhanced Localization in Wireless Personal Area Networks", *IEEE Globecom 2004*, pp. 2429-2434, Nov 2004.
- [11] S. Srirangarajan and A.H. Tewfik, "Localization in Wireless Sensor Networks under Non Line-of-Sight Propagation", *IEEE Globecom 2005*, pp. 3477-3481, Nov 2005.
- [12] D. Dardari and A. Conti, "A Sub-Optimal Hierarchical Likelihood Algorithm for Collaborative Localization in Ad-Hoc Networks", *1st IEEE Int. Conf. on Sensor and Ad Hoc Communications and Networks*, pp. 425-429, Oct 2004.
- [13] P. Bahl and V.N. Padmanabhan, "RADAR: An In-building RF-based User Location and Tracking System", *IEEE Infocom 2000*, pp. 775-784, Mar 2000.
- [14] K. Römer, "Time Synchronization in Ad Hoc Networks", *MobiHoc 2001*, Oct 2001.
- [15] A. M. Saleh, and R.A. Valenzuela, "A Statistical Model for Indoor Multipath Propagation", *IEEE J. Selected Areas in Communications*, vol. SAC-5, no.2, pp. 128-137, Feb 1987.
- [16] "Channel Modeling Sub-committee Final Report", *IEEE P802.15 Working Group for Wireless Personal Area Networks (WPANs)*, Feb 2003.
- [17] H.L. Van Trees, *Detection, Estimation, and Modulation Theory*, John Wiley and Sons, Inc., New York, 1968.