Design and Evaluation of Localization Protocols and Algorithms in Wireless Sensor Networks Using UWB

Di Wu[†‡], Lichun Bao[†], Min Du[‡], Renfa Li[‡]
[†]Donald Bren School of ICS, University of California, Irvine, USA
[‡]School of Computer and Communication, Hunan University, Changsha, China

***Abstract*—Localization has many important applications in wireless sensor networks (WSNs). A variety of technologies, such as acoustic, infrared, and UWB (ultra-wide band) media have been utilized for localization purposes. In this paper, we propose a holistic, bottom-up design of a UWB-based communication architecture and related protocols for localization in WSNs. A new UWB coding method, called U-BOTH (UWB based on Orthogonal Variable Spreading Factor and Time Hopping), is utilized for minimum interference communication, and an ALOHA-type channel access method and a message exchange protocol are used to collect distance information in WSNs. We derive the corresponding UWB path loss model in order to apply the maximum likelihood estimation (MLE) method to compute the distances between neighbor nodes using the RSSI information. Then, we propose NMDS-MLE (Non-metric Multidimensional Scaling and Maximum Likelihood Estimation) localization algorithms based on the two types of distance information: estimated distance and Euclidean distance. The performance of the system is validated using theoretic analysis and simulations.**

I. Introduction

Large-scale economic wireless sensor networks (WSNs) are widely deployed for environmental monitoring and control operations. Object tracking and localization are two important capabilities in many WSN applications [3]. The basic approach to a WSN localization is to infer distances to anchor locations, then to derive the location of a node by trilateration or other estimation algorithms. The first step is called "ranging", and the second step is called "localization".

So far, various ranging solutions have been proposed based on two major ranging techniques: 1) time of arrival (ToA) [16], time difference of arrival (TDOA) and angle of arrival (AOA) based ranging techniques, as used by GPS systems, 2) the path loss model based on radio RSSI signal strength [4] or acoustic signal strength [22] attenuation models. Other range-free techniques were also proposed for localization purposes, such as hop count or centroid methods [9]. We adopt the path loss model to derive range information because it is an efficient method in low-cost WSNs, in contrast to expensive synchronization requirements in the former approach [8].

Ranging algorithms based on path loss model depend on the wireless medium and signal transmission methods. In order to provide precision ranging, we utilize the UWB (ultra-wide band) transmission and coding technologies in both indoor and outdoor environments. Beside providing high data bandwidth, UWB exhibits excellent resistance to co-channel interference. IEEE 802.15.4a has appeared as the *de facto* standard to provide low power long distant low data-rate service for real-time communication and precise ranging and localization applications [1], [14].

Of the different UWB transmission techniques, Impulse Radio Ultra-wideband (IR-UWB) is most attractive for localization purposes in WSNs [23]. However, existing coding algorithms for IR-UWB communication systems, such as DS-UWB (Direct Sequence UWB) and TH-UWB (Time Hopping UWB) [13] have failed to guarantee high quality localization due to multipath and multi-user interference. In this paper, we apply the Orthogonal Variable Spread Factor (OVSF) coding algorithm in IR-UWB networks to solve the multi-user interference problem in data transmissions.

Once the approximate distances between a node and a subset of anchor points are derived in the WSNs, the coordinates of the node can be derive by localization algorithms. Savarese *et al.* presented a trilateration algorithm based on least squares (LS) method in large-scale WSNs [19]. Caplun *et al.* [7] proposed a GPS-free positioning system for mobile ad hoc networks, by first establishing the local coordinates of two-hop neighbors with each node as the origin, then tuning these local coordinates to the global coordinates of the entire system. The DV-coordinate algorithm [15] used similar idea.

Different from trilateration algorithms, the MDS (Multidimensional Scaling) method uses two types of maps — the relative map and the absolute map to derive locations using statistical techniques [5]. The relative map reflects partial and relative inter-nodal relationships in lower dimension space, whereas the absolute map is generated relative to the anchor nodes using the relative map. MDS requires less information and configuration overhead than other localization algorithms in WSNs, and provides strong resilience to measurement errors.

Several variants of MDS were proposed so far. MDS-MAP uses connectivity information (whether or not two devices are in range) for localization [21]. MDS-MAP(P) improved the basic MDS-MAP on anisotropic topologies [20] by building a local relative map of a small sub-network for each node using MDS, then merging them to form a global relative map. However, most of MDS algorithms were based on the assumption that proximity data between objects should be proportional to Euclidean distances by underlying quantitative transformation function, which is not flexible or robust.

In this paper, we present the NMDS-MLE (Non-metric MDS and Maximum Likelihood Estimation) localization algorithm, based on IR-UWB model and RSSI (Received Signal Strength Indication) information. Non-metric MDS is different from previous MDS variants in that the proximity data are only assumed to be related to Euclidean distances according to same ordinal level by some monotone transformation.

Overall, the contribution of this work is the following:

1) A new UWB coding method, called U-BOTH (UWB based on Orthogonal Variable Spreading Factor and Time Hopping), is proposed for minimum interference communication.
2) An ALOHA-type channel access protocol and a message exchange protocol are used to collect distance information in WSNs.
3) The UWB path loss model in U-BOTH is derived and applied in the maximum likelihood estimation (MLE) method to compute the distances between neighbor nodes using the RSSI information.
4) The NMDS-MLE (Non-metric Multidimensional Scaling and Maximum Likelihood Estimation) localization algorithm is proposed using two types of distance information: estimated distance and Euclidean distance.

This work was sponsored in parts by the National Natural Science Foundation of China under Grant No. 60673061 and the Raytheon Company under Grant No. RC-42621.

The rest of the paper is organized as follows. Section II describes the basic assumptions of the localization system, and the notation used in this paper. Section III presents a new IR-UWB coding method, called U-BOTH (UWB based on Orthogonal Variable Spreading Factor and Time Hopping), and provides the signal processing model. Section IV specified a WSN communication protocol to localization using U-BOTH. According to the path loss model and the RSSI information gathered by the target nodes, Section V and Section VI present the ranging and localization algorithms using the MLE and NMDS methods, respectively. Section VII evaluates the system using simulations. Section VIII concludes the paper.

II. Assumptions and Notation

Our ranging and localization in mainly based on the RSSI information. In order to collect the distance information as quickly as possible and avoid interference of multi-users, we assume that each node in the WSN is able to communicate through U-BOTH, proposed in this paper.

For convenience, the notation in Table I is used in this paper.

TABLE I
NOTATION AND MEANING

Notation	Meaning
T_f	The frame time.
T_c	The chip time.
T_b	The bit time.
N_s	The number of pulses for every bit.
N_c	The number of chips for every frame.
d_j^n	The OVSF code of transmitter n.
SF	The spreading factor of OVSF codes.
N_s	The period of OVSF code.
E_{TX}^n	The transmission energy of transmitter n.
E_{RX}^n	The received energy of transmitter n.
$p_0(t)$	The energy normalized pulse waveform.
c_j^n	The time-hopping code with period N_s.
$a_{\lfloor j/N_s \rfloor}^n$	The indication of information bit b.
$r_u(t)$	The input useful signal of the receiver.
$r_{mui}(t)$	The input multiple users interference signal of the receiver.
$n(t)$	The input additive white Gaussian noise of the receiver.
$m(t)$	The correlation template of the receiver.
Z_u	The output useful signal of the receiver.
$Z_{mui}(t)$	The output multiple users interference of the receiver.
Z_n	The output additive white Gaussian noise of the receiver.
N_0	The noise spectral density.
τ	The delay of the other transmitter's interfering pulse.
μ_x	The mean value of variable x.
σ_x	The standard deviation of a random variable x.
$\mathrm{erfc}(x)$	The complementary error function of value x.
Pr_b	The bit error rate (BER).

III. Physical Layer Model

A. UWB Signal Spreading and Modulation

In order to achieve accurate localization, we need a reliable physical layer communication technique that reduces bit error rate (BER), while mitigating the multi-users-interference (MUI) and Gaussian noise interference. Our physical layer is a UWB system based on time-hopping (TH) signal transmission as well as OVSF (orthogonal variable spread factor) for spreading out the symbols.

OVSF (Orthogonal Variable Spread Factor) was extensively used in CDMA systems to provide variable spreading codes [2]. Shorter OVSF code lengths are usually optimized for short-distance and high-data-rate transmission in less crowed environments due to its smaller spreading factor. TH (time hopping) is one of many signal modulation methods used by UWB. We describe a system, called U-BOTH (UWB modulation Based on OVSF and Time Hopping), which applies the time-hopping pulse position modulation (TH-PPM) algorithm to encode UWB pulse streams, and OVSF direct sequence to spread the user data bit stream.

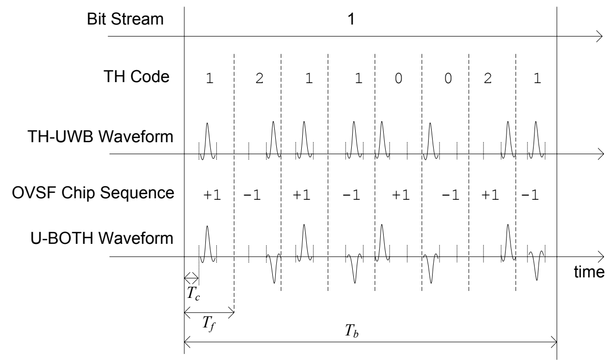


Fig. 1. U-BOTH: Interference Resistant UWB Modulation Using Time Hopping and OVSF.

Fig. 1 illustrates the utilization of time hopping (TH) pulse position modulation and OVSF spreading to encode a single bit in the user data stream. First, U-BOTH sends each bit in the bit time, denoted by T_b. Then it modulates the bit 1 using a TH code, 12110021, in which each digit denotes a chip slot position within a frame time, T_f, to send a broadband radio pulse. The number of pulses is denoted by N_s. Therefore, each bit duration is $T_b = T_f \times N_s$. Each chip slot lasts for T_c, sufficient to send a short UWB pulse signal.

After the initial pulse position modulation using UWB signals, the pulse sequence is again applied with OVSF code so that the phases are shifted by π to provide orthogonality between multiple users. The length of the OVSF code is called the spread factor SF, which is equal to N_S.

In U-BOTH, the TH code is a pseudo-random sequence generated from foreknown seeds, such as node IDs. While the OVSF codes are selected from a well-defined set of orthogonal spreading codes.

To formally analyze the system in this paper, we represent the transmitted signal by the nth transmitter in Eq. (1):

$$s^n(t) = \sum_{j=-\infty}^{+\infty} d_j^n a_{\lfloor j/N_s \rfloor}^n \sqrt{E_{TX}^n} p_0(t - jT_f - c_j^n T_c), \qquad (1)$$

in which, $d_j^n = \pm 1$ is the OVSF code with the period N_s, E_{TX}^n is the energy of the nth transmitter, $p_0(t)$ is the energy normalized pulse waveform, $c_j^n \in [0, N_c - 1]$ is the TH code with period N_s and $a_{\lfloor j/N_s \rfloor}^n$ indicates the data stream bit. If the data bit is 1, $a_{\lfloor j/N_s \rfloor}^n = +1$. Otherwise, $a_{\lfloor j/N_s \rfloor}^n = -1$.

At the receiver side, the received signal consists three source of information:

$$r(t) = r_u(t) + r_{mui}(t) + n(t),$$

in which, $r_u(t)$ is the desired user signal, $r_{mui}(t)$ is co-channel interference from multiple users, and $n(t)$ is the additive white Gaussian noise (AWGN).

Denote the pulse energy of the n-th transmitter as E_{RX}^n. Without loss of generality, we assume that the first user's transmission is the desired signal at the receiver for simplicity, then Eq. (2) provides the desired signal function at the receiver:

$$r_u(t) = \sum_{j=-\infty}^{+\infty} d_j^1 a_{\lfloor j/N_s \rfloor}^1 \sqrt{E_{RX}^1} p_0(t - jT_f - c_j^1 T_c). \qquad (2)$$

We define the correlation template of the receiver:

$$m(t) = \sum_{j=iN_s}^{(i+1)N_s-1} d_j^1 p_0(t - jT_f - c_j^1 T_c); i \in (-\infty, +\infty). \quad (3)$$

B. Single User System Analysis

As the first step, we assume that the channel is AWGN multipath-free channel, and that the transmitter and the receiver are synchronized. In a single user signal processing system, the input of the receiver has two parts: $r_u(t)$ and $n(t)$, and the output of the receiver in time interval $[0, T_b]$ is represented by:

$$Z = Z_u + Z_n = \int_0^{T_b} (r_u(t) + n(t))m(t)dt. \quad (4)$$

In Eq. (4), the useful output signal is:

$$Z_u = \sum_{j=0}^{N_s-1} \int_{jT_f+c_j^1T_c}^{jT_f+c_j^1T_c+T_c} d_j^1 d_j^1 a^1_{\lfloor j/N_s \rfloor} \sqrt{E^1_{RX}} \omega(t)dt,$$

where $\omega(t) = p_0(t - jT_f - c_j^1 T_c)p_0(t - jT_f - c_j^1 T_c)$.

Because $d_j^1 d_j^1 = 1$, $p_0(t)$ is the energy normalized pulse waveform, we have

$$\begin{aligned} Z_u &= \textstyle\sum_{j=0}^{N_s-1} \int_0^{T_c} a^1_{\lfloor j/N_s \rfloor} \sqrt{E^1_{RX}} p_0(t)p_0(t)dt \\ &= N_s a^1_{\lfloor j/N_s \rfloor} \sqrt{E^1_{RX}} \textstyle\int_0^{T_c} p_0(t)p_0(t)dt \\ &= a^1_{\lfloor j/N_s \rfloor} N_s \sqrt{E^1_{RX}} \end{aligned}$$

In Eq. (4), the output noise signal is:

$$Z_n = \sum_{j=0}^{N_s-1} \int_0^{T_c} d_j^1 p_0(t)n(t)dt = \sum_{j=0}^{N_s-1} d_j^1 n_j,$$

where n_j is Gaussian random variable with mean 0 and variance $N_0/2$. Because d_j^1 is not a random variable, the variance of Z_n is:

$$D(Z_n) = D(\sum_{j=0}^{N_s-1} d_j^1 n_j) = N_s \frac{N_0}{2},$$

$$Z_n \sim N(0, N_0 N_s/2).$$

Suppose that the statistical probabilities of data bit $b = 0$ and $b = 1$ are equal, we obtain the BER (bit error rate) of the single user system in AWGN channel as follows:

$$Pr_b = \frac{1}{2} P(Z > 0|b = 0) + \frac{1}{2} P(Z < 0|b = 1) = P(Z > 0|b = 0).$$

Because $a^1_{\lfloor j/N_s \rfloor} = -1$ if $b = 0$, then the useful output is $Z_u = a^1_{\lfloor j/N_s \rfloor} N_s \sqrt{E^1_{RX}} = -N_s \sqrt{E^1_{RX}}$. Using Eq. (4), the BER become:

$$\begin{aligned} Pr_b &= P(Z > 0|b = 0) = P(-N_s \sqrt{E^1_{RX}} + Z_n > 0) \\ &= P(Z_n > N_s \sqrt{E^1_{RX}}) \end{aligned}$$

It can be rewritten by complementary error function $\mathrm{erfc}(x)$ as follow:

$$Pr_b = \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{N_s E^1_{RX}}{N_0}}\right).$$

Where $erfc(x) = \frac{2}{\sqrt{\Pi}} \int_x^\infty \exp(-t^2)dt$.

Because U-BOTH is a rate variable system using OVSF, we analyze the relation between BER and the bit rate. Suppose the system's OVSF code is a code tree of 6 layers [6], and the spreading factor is 2, 4, 8, 16, 32, 64, respectively. Further suppose the basic rate of our system is R_0, then the corresponding bit rate of U-BOTH is $R_b = iR_0$ (i= 32, 16, 8, 4, 2, 1, respectively).

Denote the bit rate as R_b, where $R_b = iR_0$, $i = 1, 2, \cdots, 32$, we can get the relation between BER and the bit rate:

$$\begin{aligned} Pr_b &= \tfrac{1}{2} \operatorname{erfc}\left(\sqrt{\tfrac{SF \cdot E^1_{RX}}{N_0}}\right) \\ &= \tfrac{1}{2} \operatorname{erfc}\left(\sqrt{\tfrac{64R_0 \cdot E^1_{RX}}{R_b N_0}}\right) \end{aligned} \quad (5)$$

Eq. (5) shows that the BER decrease when the spreading factor SF increases or when the bit rate decreases. Therefore, we can adjust SF to adapt different environments with various noise levels while maintaining the same bandwidth of the signal. This is the main reason we adjust OVSF codes in our system.

C. Multi-User Interference Analysis

In multi-user communication system, the received signal includes multi-user interference Z_{mui} and noises. The $Z_u + Z_n$ part is the same as Eq. (4), but the multi-user interference Z_{mui} is additional. Because the phase and delay τ of interfering pulses is random as shown in Fig. 2, we have to compute the interference's variance.

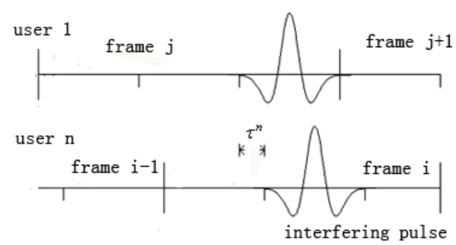


Fig. 2. The Interference to User 1 by The n-th User.

Suppose that τ^n is uniformly distributed over $[0, T_f)$, then the interference variance of the desired signal, *i.e.* the signal from the 1st user, caused by transmitter n is [13]:

$$\sigma^2_{bit} = \frac{N_s}{T_f} \int_0^{T_f} \left(\sqrt{E^n_{RX}} \int_0^{T_c} d_j^1 d_i^n p_0(t - \tau^n)p_0(t)dt) \right)^2 d\tau^n.$$

Therefore, the total interference variance σ^2_{mui} from all other transmitters is:

$$\sum_{n=2}^{N_u} \left(\frac{N_s E^n_{RX}}{T_f} \int_0^{T_f} \left(\int_0^{T_c} d_j^1 d_i^n p_0(t - \tau^n)p_0(t)dt \right)^2 d\tau^n \right).$$

Because the delay τ for all transmitters has the same distribution, we get the following formula:

$$\begin{aligned} &\sigma^2_{mui} = \\ &\tfrac{N_s}{T_f} \textstyle\sum_{n=2}^{N_u} E^n_{RX} \left(\int_0^{T_f} (\int_0^{T_c} d_j^1 d_i^n p_0(t - \tau^n)p_0(t)dt)^2 d\tau^n \right) \\ &= \sigma^2_M \tfrac{N_s}{T_f} \textstyle\sum_{n=2}^{N_u} E^n_{RX} \end{aligned}$$

in which,

$$\begin{aligned} \sigma^2_M &= \textstyle\int_0^{T_f} \left(\int_0^{T_c} d_j^1 d_i^n p_0(t - \tau^n)p_0(t)dt \right)^2 d\tau \\ &= \textstyle\int_0^{T_f} R^2(\tau)d\tau. \end{aligned}$$

According to [13], and noticing that $R_b = \frac{1}{N_s N_f}$ and $N_s = SF =$

$\frac{64R_0}{R_b}$, Eq. (6) gives the BER in multi-user interference environments.

$$Pr_b = \frac{1}{2}\operatorname{erfc}\left(\sqrt{\frac{1}{2}\left(\left(\frac{2N_s E_{RX}^1}{N_0}\right)^{-1} + \left(\frac{N_s E_{RX}^1}{\sigma_M^2 \frac{1}{T_f}\sum_{n=2}^{N_u} E_{RX}^n}\right)^{-1}\right)^{-1}}\right) = \frac{1}{2}\operatorname{erfc}\left(\sqrt{\frac{1}{2}\left(\left(\frac{128 R_0 E_{RX}^1}{R_b N_0}\right)^{-1} + \left(\frac{E_{RX}^1}{\sigma_M^2 R_b \sum_{n=2}^{N_u} E_{RX}^n}\right)^{-1}\right)^{-1}}\right). \quad (6)$$

IV. Network Protocol Operations

Our localization algorithms depend on a two-step process — the first step is for the target node to acquire the signal strength information from neighbor nodes in the network using U-BOTH based communication protocols, and the second step is for the target node to calculate the distances to the neighbor nodes, and infer its own coordinate.

In ad hoc networks, code assignments are categorized into transmitter-oriented, receiver-oriented or a per-link-oriented code assignment schemes (also known as TOCA, ROCA and POCA, respectively) [10], [12]. Depending on the ways of assigning the OVSF-TH codes and encoding the MAC data frames for transmissions, we propose two different ways to implement multiple access protocols using U-BOTH.

a) ROCA-Based Protocol Operations: The first approach is based on the receiver-oriented code assignment (ROCA), in which case the data packet transmissions are encoded using the unique OVSF-TH code assigned to the receiver. Beside ROCA, there is a common OVSF-TH code for bootstrapping and coordination purposes.

In ROCA scheme, when a target node needs to find out its coordinate, it sends a location request message using the common OVSF-TH code to the neighbor nodes. The request message includes the request command, and the receiver's OVSF-TH code. Upon receiving the request message, neighbor nodes sends back a response message using the receiver's OVSF-TH code using a random backoff mechanism. The distance information could be derived from the response message.

b) TOCA-Based Protocol Operations: The second approach is based on transmitter-oriented code assignment (TOCA), in which case each packet transmission is encoded using two OVSF-TH codes — one is a common OVSF-TH code to encode the common physical layer frame header, and the other transmitter-specific code is to encode the physical layer frame payload. The frame head includes the transmitter-oriented OVSF-TH code for encoding the frame payload.

Because the physical layer headers are sent on a common OVSF-TH code, the physical layer header transmissions resemble those of ALOHA networks with regard to packet collision. Because the headers are usually short, the collision probability is low.

On the other hand, because the data frame payload is transmitted on unique OVSF-TH codes, the interference between the payload and other frame headers and payloads is dramatically reduced.

In both ROCA- and TOCA-based systems, packets from the neighbor nodes can be lost. However, this does not affect the overall performance of our localization algorithms because they tolerate such losses.

After getting the respective signal strength information from neighbor nodes, a target node calculates its coordinate in two steps — ranging and localization.

V. Ranging Algorithm

As mentioned before, ranging is to estimate the approximate distance between the target node and neighbor nodes. We use the MLE (maximum likelihood estimation) method for such calculations. First of all, we need to establish the path loss model of the UWB channel in order to inversely derive the distance information from received signal qualities.

A. The Path Loss Model

It is well-known that the path loss model can be expressed by the log-distance path loss law in many indoor or outdoor environments, as shown by Eq. (7).

$$PL(d) = \left(PL_0 + 10\gamma \log_{10}(\frac{d}{d_0})\right) + S;\ d \geq d_0, \quad (7)$$

in which

- d_0 is the reference distance (e.g. 1 meter in UWB medium),
- PL_0 means the path loss in dB at d_0,
- d is the distance between the transmitter (Tx) and receiver (Rx),
- γ refers to the path loss exponent which depends on channel and environment,
- S is the log-normal shadow fading in dB. Usually, S is a Gaussian-distributed random variable with zero mean and standard deviation σ_S.

Eq. (7) could construct a statistical path loss model for UWB propagation in different environments. The path loss $PL(d)$ can be expressed as a Gaussian-distributed random variable with:

$$S \sim N(0, \sigma_S^2),$$

$$PL(d) \sim N(PL_0 + 10\gamma \log_{10} d, \sigma_S^2).$$

The probability density function (*pdf*) of path loss $PL(d)$ is:

$$p(PL) = \frac{e^{-\frac{[PL-(PL_0+10\gamma\log_{10} d)]^2}{2\sigma_S^2}}}{\sqrt{2\pi\sigma_S^2}}. \quad (8)$$

IEEE 802.15.4a Task Group provided Channel Model 1-9 by taking limited real measurements to determine the values of γ, σ_S and other variables in different situations. When deploying real UWB networks, people could approximately choose the corresponding channel model with the parameters specified in IEEE 802.15.4a.

B. Ranging Algorithm based on Maximum Likelihood Estimation

The distance between the transmitter Tx and the receiver Rx in Eq. (7) can be calculated by the general ranging method between two nodes using the RSSI information:

$$\hat{d} = 10^{\frac{PL(d)-PL_0-S}{10\gamma}}.$$

Receiver computes the distance between the transmitter Tx and the receiver Rx using random values S. However, in above single random ranging, the random variables S selected by the sensor nodes are not exactly those in the real time-variant channel. In order to avoid the ranging errors caused by the large deviation between the estimated S values and the real S values in each round of ranging estimation, we propose an iterative ranging based on MLE (maximum likelihood estimation) in UWB wireless sensor networks.

Suppose PL_i is the ith observation value, we get the joint conditional *pdf* $p(PL|d)$ using Eq. (9).

$$p(PL|d) = \prod_{i=1}^{N} \frac{e^{-\frac{[PL_i-(PL_0+10\gamma\log_{10} d)]^2}{2\sigma_S^2}}}{\sqrt{2\pi\sigma_S^2}}. \quad (9)$$

The necessary condition to compute the MLE of d is:

$$\frac{\partial \ln p(PL|d)}{\partial d} = \frac{10N\gamma}{\sigma_S^2 d \ln 10}\left(\frac{1}{N}\sum_{i=1}^{N} PL_i - PL_0 - 10\gamma \log_{10} d\right) = 0. \tag{10}$$

We solve Eq. (10) and have:

$$\widehat{\log_{10}} d = \frac{1}{10N\gamma}\sum_{i=1}^{N} PL_i - \frac{PL_0}{10\gamma}$$

Therefore, the MLE based RSSI UWB ranging is:

$$\hat{d} = 10^{\frac{1}{10N\gamma}\sum_{i=1}^{N} PL_i - \frac{PL_0}{10\gamma}}. \tag{11}$$

VI. Localization Algorithm

A. Multi-Dimensional Scaling (MDS)

MDS (Multidimensional Scaling) is a statistical technique for exploratory data analysis or information visualization. MDS collects the proximity data between each pair of spatial objects as reference. Then it visualizes objects as points in a low dimensional Euclidean space and represents these proximity data as distances between points. In order to derive accurate results, MDS has to find some solutions that relate distance information to proximity information as closely as possible.

Suppose that n denotes the number of different objects, and the proximity for objects i and j is denoted by p_{ij}. Thus, we derive a proximity matrix $\mathbf{P}_{n\times n} = p_{ij}$. The coordinates of mapping points are represented by a matrix $\mathbf{X}_{n\times m}$, where m is the dimensions of the solution, e.g. 2D or 3D.

Now, let $d_{ij}(\mathbf{X})$ be the Euclidean distance between points i and j with coordinates in $\mathbf{X}_{n\times m}$, respectively. The objective of MDS is to find a matrix $\mathbf{X}$ so that $d_{ij}(\mathbf{X})$ proportionally matches p_{ij} as closely as possible, which is presented by $f(p_{ij}) \sim d_{ij}(\mathbf{X})$. The closeness is measured by metric STRESS as follows:

$$STRESS = \sum [f(p_{ij}) - d_{ij}(\mathbf{X})]^2.$$

MDS algorithms are taxonomized into several types, depending on whether the similarity data is quantitative or qualitative, and are called metric MDS and non-metric MDS, respectively.

Classical metric MDS formulates the relationship between proximity data of objects and distances in the Euclidean space by transformation functions. In order to find a perfect fitness between proximity data and Euclidean distance, the transformation formula $d_{ij}(\mathbf{X}) = f(p_{ij})$ is pursued, such as a linear model: $d_{ij}(\mathbf{X}) = a + bp_{ij}$. Because $d_{ij}(\mathbf{X})$ represents the Euclidean distance between points i and j in coordinate matrix $\mathbf{X}$, MDS rests on the fact that the coordinate matrix $\mathbf{X}$ can be derived by double centering and eigenvalue decomposition from the proximity matrix $\mathbf{P}$ with the least error.

The relationship between the proximity of objects and the Euclidean distances of points in Non-metric MDS is not as strict as metric MDS. Non-metric MDS only requires a monotonic relationship between them.

When Non-metric MDS takes proximity data of different objects to construct corresponding spatial coordinates, it only requires that the rank order of the proximity p_{ij} have to keep the same ordinal level as the distances d_{ij}. That is,

$$\forall i,j,k,l : p_{ij} < p_{kl} \Rightarrow d_{ij}(X) < d_{kl}(X).$$

Compared with metric MDS, the monotonic assumption that the data is measured at the ordinal level in Non-metric MDS makes it more flexible and applicable for localization in wireless sensor networks.

B. The NMDS-MLE Localization Algorithm

NMDS-MLE localization algorithm combines the ranging and localization processes. Ranging is based on the iterative RSSI information collected by above U-BOTH UWB system and refined by the MLE method. Localization is based on the NMDS algorithm. As a whole, NMDS-MLE localization consist of 5 steps:

1 Gather iterative RSSI from neighbors by U-BOTH system in the network, and form a sparse matrix $\mathbf{R}$, which is derived from the estimated distances denoted by r_{ij}. r_{ij} is estimated by iterative ranging based on RSSI information and MLE method. For the nodes that is out of the communication range, r_{ij} is zero.

2 Construct the proximity data matrix $\mathbf{P}$ based on sparse matrix $\mathbf{R}$. The estimated distance p_{ij} between every pair of nodes in the network is computed by the shortest path algorithm, such as Dijkstra's or Floyd's algorithm.

3 Construct the coordinate system to plot the objects in the Euclidean space and obtained the distance matrix D composed by the Euclidean distance d_{ij}.

4 Compare the ordinal level between aforementioned two types of distance information: estimated distance p_{ij} and Euclidean distance d_{ij}, and refine the relative coordinate $\mathbf{X}$ of nodes in Non-metric MDS.

5 Transform relative coordinate into global absolute coordinate by the anchor nodes in the network.

In Step 3 and Step 4, localization is executed by NMDS-MLE as Algorithm 1.

Algorithm 1: NMDS-MLE

Input: node set N, initial coordinate matrix $X^{(0)}$, proximity data matrix P, threshold ε, iteration number $k \longleftarrow 0$
Output: relative coordinate $X^{(n)}$
1 **for** *each* $i, j \in N$ **do**
2     $d_{ij}^k \longleftarrow \sqrt{(x_i^k - x_j^k)^2 + (y_i^k - y_j^k)^2}$
3   construct the Euclidean distance matrix $D^{(k)}$
4 **end**
5 **while** $STRESS \geq \varepsilon$ **do**
6   **for** *each* $i, j, u, v \in N$ **do**
7   **if** $p_{ij} < p_{uv}$ **and** $d_{ij} > d_{uv}$ **then**
8             $\hat{d}_{ij}^k \longleftarrow (d_{ij}^k + d_{uv}^k)/2$
9             $\hat{d}_{uv}^k \longleftarrow (d_{ij}^k + d_{uv}^k)/2$
10   **else if** $p_{ij} < p_{uv}$ **and** $d_{ij} \leq d_{uv}$ **then**
11             $\hat{d}_{ij}^k \longleftarrow d_{ij}^k$
12             $\hat{d}_{uv}^k \longleftarrow d_{uv}^k$
13   **end**
14   **end**
15     $k \longleftarrow k + 1$
16   update the coordinate matrix $X^{(k)}$
17   update the distance matrix $D^{(k)}$
18 **end**

In Algorithm 1, a monotonic transformation between proximity data and Euclidean distance is calculated in line 6 to 14, which yields an intermediate distance value $\hat{d}_{ij}$. By performing a monotone regression with the current distances d_{ij} as targets and proximity p_{ij} as inputs, NMDS-MLE generates $\hat{d}_{ij}$ to reflect the ordinal level of p_{ij} in each iteration, where $\hat{d}_{ij}$ should be subjected to:

$$\forall i,j,k,l : p_{ij} < p_{kl} \Rightarrow \hat{d_{ij}}(X) < \hat{d_{kl}}(X).$$

Because of above relation between p_{ij} and $\hat{d}_{ij}$, NMDS-MLE takes following STRESS applied in line 5 to evaluate the accuracy of the fitting:

$$STRESS = \sqrt{\sum_{ij,i\neq j} (\hat{d_{ij}} - d_{ij})^2 / \sum_{ij,i\neq j} d_{ij}^2}. \quad (12)$$

A small STRESS indicates a good fit, whereas a high value indicates a bad fit. Kruskal [11] provide some guide lines of stress value with respect to the goodness of fit of the solution, shown in Table II.

TABLE II
STRESS AND GOODNESS OF FIT

Stress	Goodness of fit
>.20	poor
.10	fair
.05	good
.025	excellent
.00	perfect

Note in line 16 and line 17 in Algorithm 1, NMDS-MLE updates spatial coordinate matrix $\mathbf{X}^{k-1}$ to $\mathbf{X}^k$ according to d_{ij}^{k-1} and $\hat{d}_{ij}^{k-1}$, and then obtains a new Euclidean distance $d_{ij}(\mathbf{X})^k$. The spatial coordinate (x_i^k, y_i^k) is updated as follows:

$$\begin{aligned} x_i^k &= x_i^{k-1} + \tfrac{\alpha}{n-1}\textstyle\sum_{j\in M, j\neq i}(1 - \frac{\hat{d}_{ij}^{k-1}}{d_{ij}^{k-1}})(x_j^{k-1} - x_i^{k-1}), \\ y_i^k &= y_i^{k-1} + \tfrac{\alpha}{n-1}\textstyle\sum_{j\in M, j\neq i}(1 - \frac{\hat{d}_{ij}^{k-1}}{d_{ij}^{k-1}})(y_j^{k-1} - y_i^{k-1}). \end{aligned} \quad (13)$$

Where n is the number of target nodes, α is the iteration step length, which is set to be 0.2 in the paper.

In Step 4, the estimated location matrix $\mathbf{X}$ represents the relative coordinates of nodes, which have a different orientation and scaling than the original coordinates. And in Step 5, the transformation from relative coordinate $\mathbf{X}$ into absolute coordinates usually includes shift, rotation, scaling, and reflection of coordinates, which are implemented by some transformation to minimize the errors between the absolute coordinates of anchor nodes and their relative locations in the NMDS map. Suppose there are m anchor nodes whose relative locations are $\mathbf{X_R} = (\mathbf{X_{R_1}}, \mathbf{X_{R_2}}, \cdots, \mathbf{X_{R_m}})$, and real locations are $\mathbf{X_T} = (\mathbf{X_{T_1}}, \mathbf{X_{T_2}}, \cdots, \mathbf{X_{T_m}})$. We need firstly derive optimal transformation function $\mathbf{Q}$, and then transfer all the relative coordinates of nodes to the absolute coordinates by the optimal transformation function $\mathbf{Q}$.

VII. SIMULATION EVALUATIONS

In order to verify our localization algorithms based on U-BOTH system for WSNs, we simulated the following scenarios:

1) With regard to the BER (bit error rate), we evaluate U-BOTH system performance in single and multi-user scenarios.
2) Using NMDS-MLE localization algorithm, we evaluate our localization model both in random network and grid network.

A. U-BOTH System Performance

We assume the channel is AWGN multipath-free single user channel, the transmitter and the receiver are synchronized perfectly. Then we randomly generate 2000 bits, every bit uses 4 pulses to repeat coding ($N_s = 4$).

Fig. 3 illustrates the BER of the received signal using U-BOTH system, in contrast to DS-UWB that only uses direct sequence spreading, and TH-UWB that uses time-hopping pulse position modulation alone for UWB transmissions. We can see that the BER of U-BOTH and the DS-UWB system which use the π-phase shift keying modulation

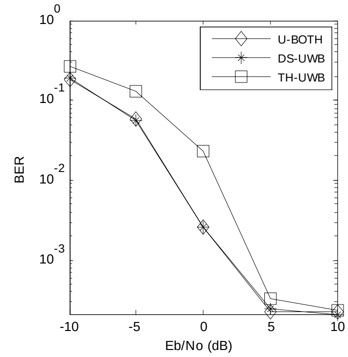


Fig. 3. Bit Error Rate in A Single User System with Additive White Gaussian Noise (AWGN).

are lower than TH-UWB. This is because the distance of two signals in binary phase shift keying (BPSK) modulation is $2\sqrt{E_{pulse}}$, but $\sqrt{2E_{pulse}}$ in TH-UWB [17].

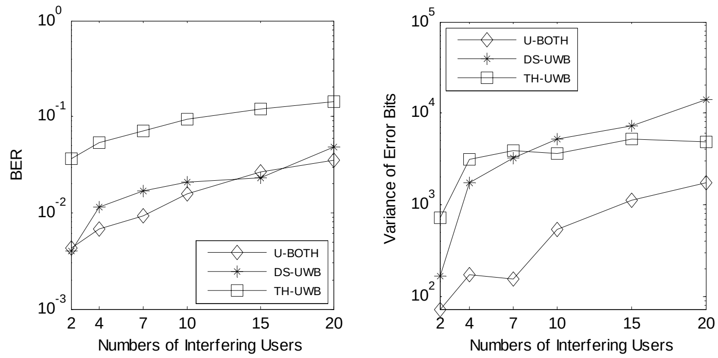


Fig. 4. Bit Error Rate and The Variance of The Number of Error Bits of 2000 Generated Bits.

Secondly, we let $E_b/N_0 = 0$ dB, $N_s = 4$ and generated 2000 bits randomly. Fig. 4 shows the relative performance of U-BOTH, TH-UWB and DS-UWB systems in multiple access scenarios. In this case, the received signal includes by noise and co-channel interference. In Fig. 4, although both the BER and the variance of error bits increase as the number of users increases, the performance of our U-BOTH system is still better than DS-UWB and TH-UWB, proving that the UWB coding based OVSF-TH effectively handle the burst errors.

B. Evaluation of the Localization Algorithms

We evaluate the performance of localization algorithms with mean estimation error, which is widely used in previous research works:

$$error = \frac{\sum_{i=m+1}^{n} \|X_{est}^i - X_{real}^i\|^2}{(n-m)\times R} \times 100\% \quad (14)$$

where n and m are the total number of sensors and the number of anchor nodes in the WNS, respectively, R represents communication range.

Based on the data in [18], we adopt values of UWB path loss model in outdoor NLOS environments for simulations as shown in Fig. 5.

1) Random Deployment: 100 nodes are deployed randomly in a $100m \times 100m$ square area as shown in Fig. 6(a), in which points represent nodes and edges represent the connections between neighbor nodes. The communication range is 12 m and the average connectivity is 4.6.

Notation	Meaning	Value	
		LOS	NLOS
d_0	The reference distance	1 *m*	1 *m*
PL_0	The path loss at reference distance	45.6 *dB*	73 *dB*
γ	The path loss exponent	1.76	2.5
σ_S	The standard deviation of shadow fading	0.83	2

Fig. 5. Portion of The Simulation Parameters.

Fig. 6(b) reflects the relative coordinate of every node generated by NMDS-MLE. It shows that the relative coordinates have a different orientation and scaling than the original network in Fig. 6(a). This is because that relative coordinate is derived only based on the distance relationship between every pair of nodes in the network. Fig. 6(c) derives the absolute coordinates of all the nodes. Their relative coordinates in Fig. 6(b) are transformed based on the location information provided by 4 random anchor nodes denoted by $\times$. The dots represent the real locations of the nodes, and the lines with arrows indicate the errors of the estimated locations from the real locations, the average localization error is about 5.3850%. The MDS-MAP algorithm is also applied in the case and the average localization error is about 18.4747%.

2) Grid Deployment: 100 nodes are deployed in a $45m \times 45m$ square area with grid deployment in Fig. 7(a). The communication range is 12 m and the average connectivity is 16.8. With the same symbol meaning in the figures, Fig. 7(b) represents the relative coordinate map using NMDS-MLE algorithm and Fig. 7(c) depicts the absolute coordinate map by transformation based on 4 random anchor nodes. The average localization error in the grid case is about 1.2876%. For MDS-MAP algorithm, it is about 6.0543%.

3) Performance analysis: The localization performance of NMDS-MLE in different scenarios under different degrees of connectivity is analyzed by Fig. 8, compared with the MDS-MAP by the same experimental settings. From the figure, we can see that localization error of NMDS-MLE algorithm is much lower and more stable than MDS-MAP in different scenarios. Furthermore, when NMDS-RSSI and MDS-MAP are applied in grid deployment with varies of connectivity. It shows that NMDS-RSSI obtain higher localization accuracy in the grid layout than in the random layout for the same connectivity level.

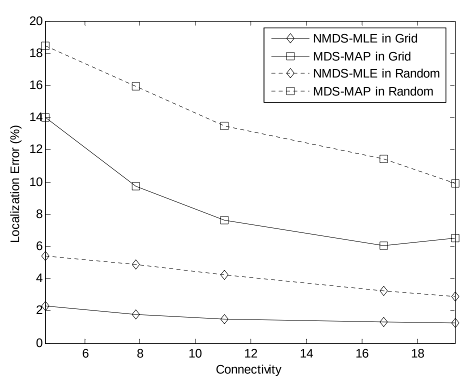


Fig. 8. Relation between The Connectivity and The Localization Error.

Fig. 9 presents the relation between localization error and the number of iteration N in NMDS-MLE algorithm. Because the accuracy of ranging is improved by MLE method based on the RSSI information provided by our U-BOTH system, it is obvious that the localization error decreases dramatically when the number of iterations in ranging increases in both random and grid deployment.

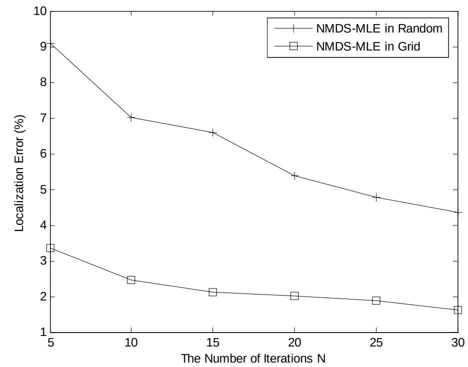


Fig. 9. Relation between The Number of Iteration N and The Localization Error.

VIII. Conclusion

In order to provide a localization algorithm using the NMDS-MLE methods, we have proposed the communication protocols based on a new UWB coding method, called U-BOTH (UWB based on Orthogonal Variable Spreading Factor and Time Hopping), and an ALOHA-type channel access method and a message exchange protocol to collect distance information in WSNs. Then we specified the NMDS-MLE algorithms using the UWB path loss model for ranging and localization purposes. The performance of NMDS-MLE algorithms in the U-BOTH based communication system are analyzed using communication theories and simulations. Results show that U-BOTH transmission technique can effectively reduce the bit error rate under the path loss model, and the corresponding ranging and localization algorithms can achieve comparable or better results than previous localization methods.

Acknowledgment

We would like to express our sincere appreciation of Fanzi Zeng, Ling Xiao, and Juan Luo for their insightful feedbacks during the preparation of this manuscript, and of the anonymous reviewers for their helpful comments. This work has been generously sponsored in parts by the National Natural Science Foundation of China under Grant No. 60673061 and the Raytheon Company under Grant No. RC-42621.

References

[1] IEEE Std 802.15.4a. Part 15.4a: Low Rate Alternative PHY Task Group (TG4a) for Wireless Personal Area Networks (WPANs). Technical report, IEEE, Jun. 2007.

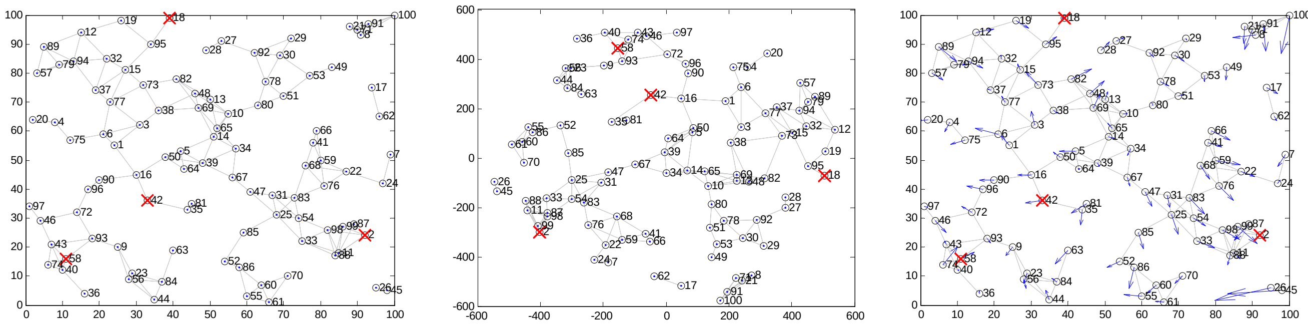
[2] F. Adachi, M. Sawahashi, and K. Okawa. Tree-structured generation of orthogonal spreading codes with dierent lengths for the forward link of DS-CDMA mobile radio. *IEE Electronics Letters*, 1(1):27–28, Jan. 1997.

[3] T.A. Alhmiedat and S.H. Yang. A Survey: Localization and Tracking Mobile Targets through Wireless Sensors Network. In *The Eighth Annual PostGraduate Symposium on the Convergence of Telecommunications, Networking and Broadcasting (PGNET)*, 2007.

[4] P. Bahl and V. Padmanabhan. RADAR: An In-Building RF-based User Location and Tracking System. In *INFOCOM*, 2000.

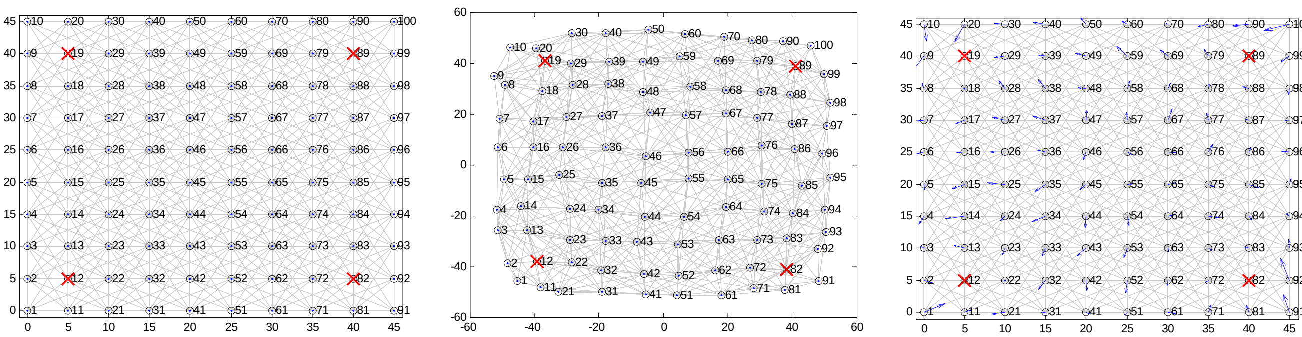
[5] I. Borg and P. Groener. *Modern Multidimensional Scaling: Theroy and Applications*. Springer-Verlag, New York, 1997.

[6] Hasan Cam. Nonblocking OVSF Codes and Enhancing Network Capacity for 3G Wireless and Beyond Systemsn. *Special Issue of Computer Communications on "3G Wireless and Beyond for Computer Communications"*, 26(17):1907–1917, 2003.



(a) Original Coordinates (b) Relative Coordinates (c) Absolute Coordinates

Fig. 6. Random Deployment.



(a) Original Coordinates (b) Relative Coordinates (c) Absolute Coordinates

Fig. 7. Grid Deployment.

[7] S. Caplun, M. Hamdi, and J.P. Hubaus. GPS-free Positioning in Mobile Ad Hoc Networks. In *Proc. of the 34th Annu. Hawaii Int. Conf. System Sciences*, pages 3481–3490, 2001.

[8] T. Gigl and G. J.M. Janssen. Analysis of a UWB Indoor Positioning System Based on Received Signal Strength. In *4th Workshop on Positioning, Navigation and Communication (WPNC)*, pages 97–101, 2007.

[9] T. He, J. A. Stankovic, C. Huang, T. Abdelzaher, and B. M. Blum. Range-Free Localization Schemes for Large Scale Sensor Networks. In *Proceedings of the Annual International Conference on Mobile Computing and Networking (MOBICOM)*, pages 81–95, 2003.

[10] Joa-Ng, M., and I.T. Lu. Spread spectrum medium access protocol with collision avoidance in mobile ad-hoc wireless network. In *Proc. of IEEE Conference on Computer Communications (INFOCOM)*, pages 776–83, New York, NY, USA, Mar. 21-25 1999.

[11] J. Kruskal. Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis. *Psychometrika*, 29:1–27, 1964.

[12] T. Makansi. Trasmitter-Oriented Code Assignment for Multihop Radio Net-works. *IEEE Transactions on Communications*, 35(12):1379–82, Dec. 1987.

[13] Maria-Gabriella, D. Benedetto, and G. Giancola. *Understanding Ultra Wide Band Radio Fundamentals*. Prentice Hall, New Jersey, 2004.

[14] A.F. Molisch, D. Cassioli, and C.-C. Chong. A Comprehensive Standardized Model for Ultrawideband Propagation Channels. *IEEE Transactions on Antennas and Propagation*, 54(11):3151–3166, 2006.

[15] D. Niculescu and B. Nath. DV based Positioning in Ad Hoc Networks. *Journal of Telecommunnication Systems*, 22(1/4):267–280, 2003.

[16] N. Patwari, A O Hero III, M. Perkins, N. S. Correal, and R. J. O'Dea. Relative location estimation in wireless sensor networks. *IEEE Transactions on Signal Processing*, 51:2137, 2003.

[17] John G. proakis. *Digital Communications: Fourth Edition*. McGraw-Hill, Columbus, 2001.

[18] Ghassemzadeh S. S., Jana R., Rice C. W., Turin W., and Tarokh V. Measurement and Modeling of an Ultra-Wide Bandwidth Indoor Channel. *IEEE Trans. on Comm.*, 52(10):1786–1796, 2004.

[19] C. Savarese, J.M. Rabaey, and J. Beutel. Locationing in Distributed Ad-Hoc Wireless Sensor Networks. In *Proc. IEEE International Conference on Acoustics, Speech, and Signal*, volume 4, pages 2037–2040, 2001.

[20] Y. Shang and W. Ruml. Improved MDS-Based Localization. In *Proc. of the IEEE Infocom*, pages 2640–2651, 2004.

[21] Y. Shang, W. Ruml, Y. Zhang, and M. Fromherz. Localization from mere connectivity. In *Proc. of the 4th ACM Int'l Symp. on Mobile Ad Hoc Networking Computing*, pages 201–212, 2003.

[22] K. Whitehouse and D. Culler. Macro-calibration in sensor/actuator networks. In *Mobile Networks and Applications (MONET)*, volume 8, pages 463–472, 2003.

[23] M.Z. Win and R.A. Scholtz. Ultra-Wide Bandwidth Time-Hopping Spread-Spectrum Impulse Radio for Wireless Multiple-Access Communication. *IEEE Transaction on Communication*, 48(4):679–691, 2000.