An Algorithm for Distributed Beacon Selection

Dominik Lieckfeldt, Jiaxi You, Dirk Timmermann
Institute of Applied Microelectronics and Computer Engineering
University of Rostock, 18119 Rostock, Germany
Email: {dominik.lieckfeldt, jiaxi.you, dirk.timmermann}@uni-rostock.de

Abstract

This paper investigates wireless sensor networks where a small percentage of nodes are assumed to know their location a priori. These reference nodes enable absolute localization of other nodes in direct neighborhood. Having estimated their location, these nodes in turn provide their location to other nodes within transmission range. Therefore, location information spreads throughout the network. Consequently, in later state of the network, unknowns desiring to determine their location, or to improve it, will be able to choose from a large pool of nodes with known or estimated locations, which we refer to as beacons. We investigate a method to select a subset of beacons to minimize the error of localization. Regarding Cramer-Rao-Lower-Bound on localization error, the method proposed constitutes a significant improvement in comparison with the often used nearest-neighbors approach.

1. Introduction

Wireless sensor networks (WSNs) consist of small electronic devices, referred to as sensor nodes, which are capable of sensing environmental parameters and equipped with processors and transceivers to perform computations and to communicate wireless, respectively. Typical applications of WSNs include the detection of forest fires, surveillance of dikes and precision farming. For instance, fire trucks can be sent to the exact site where the forest fire is detected. Positions of nodes are essential to provide such location information, and are used for geographic routing and clustering.

Consequently, localization in WSNs has attracted great attention during the last decade which led to various approaches to determine locations of sensor nodes. These approaches can be classified into two groups: *fine-grained* methods try to achieve an exact and *coarse-grained* methods an approximate solution of the localization problem. In detail, fine-grained approaches are based on the solution of a system of equations that includes distances or angles between nodes, which makes this approach relatively resource intensive. Typically, fine-grained methods would provide exact locations if true distances/angles are given. However, coarse-grained approaches simply aim at approximate estimation of location which makes them less resource intensive and would, in most cases, lead to approximate and therefore defective results.

A common approach for localization in WSNs is to assume that the absolute locations of some nodes are known. These nodes, which are denominated *beacons*, enable absolute localization of nodes with unknown locations, which are called *unknowns*. It has often been stated in the literature, that the relative location of beacons and unknowns has strong impact on the accuracy of localization for both fine- and coarse-grained approaches. Therefore, best placement and optimal selection of beacons in terms of accuracy of localization has been spotted as an interesting field for investigations.

As complexity of localization depends on the number of beacons used, it is desirable to select those beacons first which contribute most to a high accuracy. In contrast, selecting a subset of beacons in order to optimize localization accuracy has hardly been studied in the literature. In [8, 7, 3], range measurements are weighted according to their variance and distance or beacons are selected based on the difference between distances and estimated locations [2]. Others apply tests to detect outliers in order to exclude them from calculations or just choose the nearest beacons for estimation of location [9, 11]. For coarse-grained localization, it has been reported in [6] that choosing the nearest three beacons increases localization accuracy when estimating distances with the DV-Hop method. Furthermore, in [4], localization errors are simulated at beacons to decide where additional beacons have to be placed to decrease errors effectively. However, this leads to large computational overhead on beacons. In [12], geometry of the situation is considered. Here, the set of all beacons is divided into groups of three beacons. The beacons of one group form a triangle whose angles must meet a certain requirement for this group to be selected for localization. Drawbacks of this approach are high computational complexity as all possible groups of beacons are considered and the need for global knowledge to be available at beacons.

Existing approaches either do not avoid that insignificant beacons communicate during localization process or they require global knowledge which also assumes additional communication. Furthermore, while the impact of geometry has often been stated in the literature, we are not aware of any work that considers geometric information to select beacons and uses local knowledge only.

This paper contributes a formalized view on the selection of beacons and the investigation of a promising algorithm for this task which is based on the Cramer-Rao-Lower-Bound (CRLB) on the variance of fine-grained localization. As the CRLB poses a fundamental bound on the variance of any unbiased estimator, it can be effectively used to quantify and compare the contribution of each beacon to the accuracy of localization. The evaluation considers the cases where global and where local knowledge is available and develops an algorithm based solely on local knowledge. The proposed algorithm constitutes a significant improvement compared with the often used nearest-neighbors.

The remainder of the paper is organized as follows: Section 2 gives an outline of the scenario and a short review on CRLB. Section 3 states the problem and considers algorithms to select subsets of beacons when global or only local knowledge is available. Simulation results and conclusions are given in sections 4 and 5, respectively.

2. Accuracy of Fine-Grained Localization

2.1. Scenario Set-Up and Variables

We consider random deployments of M sensor nodes where two different types of nodes exist: reference nodes with a priori known locations, and m_u unknowns to be localized. In the following, we use n_i to refer to a specific node. By considering a WSN in a later state of its lifetime, it is justified to assume that all unknowns in range of at least 3 reference nodes have obtained an estimate of their location. The class of beacons is formed by all nodes with known locations and, therefore, contains all reference nodes and unknowns with estimated locations. Consequently, in later states of the WSN, one node desiring to estimate its location or to improve the accuracy of its location estimate will have a relatively large number of beacons to choose from. This situation motivates the need for a resource-aware selection of the best subset of beacons in terms of localization accuracy.

First, we define the set of unknowns $U := \{n_i \mid i \in \{1, 2, \ldots, m_\mathrm{u}\}\}$, the set of beacons $B := \{n_i \mid i \in \{m_\mathrm{u}+1, m_\mathrm{u}+2, \ldots, m_\mathrm{u}+m_\mathrm{b}\}\}$ and the set of all nodes $N := B \cup U$ with $u_i \in U$, $b_i \in B$ referencing a specific unknown and beacon, respectively. Nodes are capable of wireless communication and thereby can estimate distances between communicating nodes by measuring received signal strength (RSS) of radio signals. We assume that nodes are synchronized and use time division multiple access (TDMA). Assuming a 2D cartesian coordinate system, the true locations of nodes are $\mathbf{z}_i = (x_i \; y_i)^T$ with distances $d_{i,j} = ||\mathbf{z}_i - \mathbf{z}_j||$ between nodes n_i and n_j. Estimates of parameters are indicated by a hat, e.g. estimates of distance are $\hat{d}_{i,j}$. Since wireless communication has limited range r_tx, we further define the set of all beacons being within transmission range of the unknown u_i: $B_i := \{b_j \mid d_{i,j} \leq r_\mathrm{tx}\}$.

2.2. Error of Localization

In real scenarios, measurements of RSS are disturbed due to multipath and noise. Therefore, measurements usually do not reflect the true situation, i.e. the true distances. Consequently, estimates of location that rely on these measurements will be erroneous and vary over time.

Naturally, we seek to find estimates of location with smallest error $|\mathbf{e}_i|$, which is the distance between the true location and the estimate:

$$\mathbf{e}_i = \begin{pmatrix} x_i - \hat{x}_i \\ y_i - \hat{y}_i \end{pmatrix} \tag{1}$$

The average variance of unbiased estimates over x- and y-directions is:

$$\tilde{\sigma}_i^2 = E\left\{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2\right\} \tag{2}$$

The CRLB poses a lower bound on this variance and therefore on the smallest mean square error of any unbiased estimator. As we use the CRLB to determine the best possible performance of localization, we briefly review the CRLB and the underlying statistical model here. A more detailed discussion along with experimental validation of the bound can be found in [10, 5]. In the following, we limit the considerations to $m_\mathrm{u} = 1$ unknown without loss of generality as this is the atomic situation which any other configuration can be decomposed into.

Assuming a log-normal shadowing model of the wireless channel and distances estimated using received signal strength, the variance $\tilde{\sigma}_1^2$ of any unbiased estimator of location is lower bounded by:

$$\sigma_1^2 = \frac{1}{a} \frac{\sum_{i=2}^{M} d_{1,i}^{-2}}{\sum_{i=2}^{M-1} \sum_{j=i+1}^{M} \left(\frac{d_{1 \perp i,j}\, d_{i,j}}{d_{1,i}^2 d_{1,j}^2}\right)^2} \tag{3}$$

$$a = \left(\frac{10 n_\mathrm{p}}{\sigma_\mathrm{dB} \ln 10}\right)^2$$

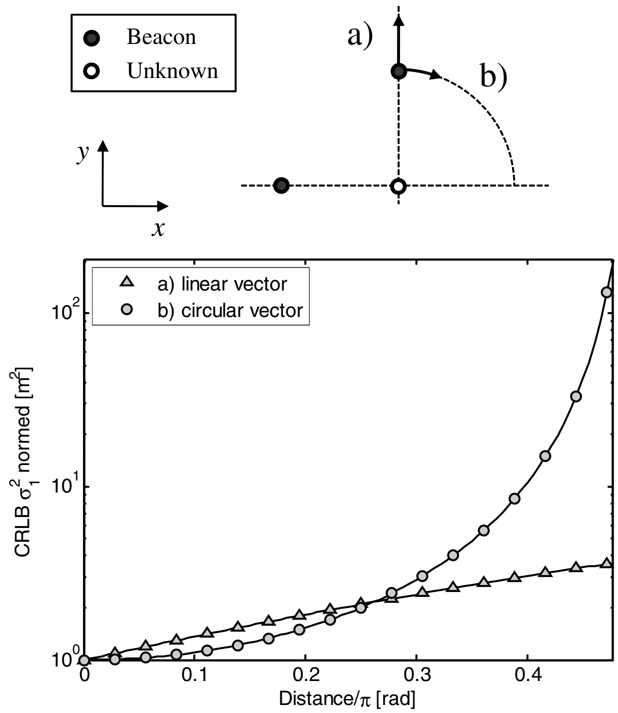

Figure 1. Top: Movement of beacons for example. Bottom: Impact of geometry on CRLB.

The well-known channel parameters n_{p} and σ_{dB} denote the path loss exponent and the standard deviation of the received signal strength. In [10] these parameters have been determined based on indoor experiments and we use these results ($n_{\mathrm{p}} = 2.3$; $\sigma_{\mathrm{dB}} = 3.92\,\mathrm{dBm}$) for our investigations. It is noted that in order to calculate the CRLB, for instance, the three distances between two beacons and the unknown are needed.

The distance $d_{1\perp i,j}$ is the length of the altitude through the unknown with base given by the interconnection of two beacons b_i, b_j and characterizes the "condition" of the triangle with vertices at the locations of unknown and beacons b_i, b_j. We will illustrate the impact of the "condition" of this triangle in the next section.

2.3. Impact of Geometry

While often stated in the literature, the impact of geometry on accuracy of localization has hardly been considered for optimal selection of beacons. The following example shown in figure 1 illustrates the strong dependency of accuracy on geometry in terms of CRLB: Two beacons (solid dots) and one unknown (hollow dot) are depicted. Arrows indicate the vector which beacons are moved along, i.e. in case a) along a linear vector and along a circular vector in case b). Distances are given in radians based on the movement of case b). For both a) and b), beacons move with the same step width and after each step, the CRLB is calculated individually.

The lower diagram in figure 1 depicts the behavior of CRLB for the two cases mentioned before. It is shown that after having passed a certain distance, movement along the circular vector of b) results in an increasingly larger CRLB compared with a). Thus, geometry is sometimes more important regarding accuracy of localization than distances between beacons and unknown.

3. Selecting Subsets of Beacons

3.1. Problem Statement

We aim at selecting a specific subset of beacons that minimizes the average mean square error of location estimates over all possible subsets. Let $\mathbb{S}(c)$ be the set of all subsets with cardinality c: $\mathbb{S}(c) := \{S \mid S \subseteq B;\ \mathrm{card}(S) = c\}$ and determine the subset $S(c)$ that minimizes $E\{|\mathbf{e}_1|^2\}$:

$$S(c) = \arg \min_{\{S \in \mathbb{S}(c)\}} E\left\{|\mathbf{e}_1\,(S)|^2\right\} \tag{4}$$

This funtion basically returns the subset S, which minimizes the expected error $E\{|\mathbf{e}_1\,(S)|^2\}$. Thereby, it is assured, that the beacons contained in $S(c)$ contribute most to accurate localization. For unbiased estimators, $E\{|\mathbf{e}_1(S)|^2\}$ can be substituted by the CRLB $\sigma_1^2(S)$ as indicated by (1), (2):

$$S(c) = \arg \min_{\{S \in \mathbb{S}(c)\}} \sigma_1^2\,(S) \tag{5}$$

For this optimization, full knowledge of locations of beacons and unknowns is needed. In practice, this information is not available and, therefore, we apply different algorithms to select a subset $\hat{S}(c)$ that nearly achieves the same lower bound on localization error as the optimal subset S(c).

3.2. Algorithms

In a later state of a WSN, the number of unknowns which have not been able to estimate their location or which desire to improve it, is relatively small. Therefore, it is reasonable to assume that the unknowns initiate localization by broadcasting a "request for localization" message. Beacons receiving this message, will respond in an order determined by a specific algorithm. Several algorithms for this task are investigated and explained in the following.

We divide the algorithms for selecting the subset $\hat{S}(c)$ into two classes. The first class is called *local knowledge*. The corresponding algorithms only use information which is available on a specific beacon. Algorithms of this class do not rely on additional communication to perform the selection of subsets. The second class is called *global knowledge* indicating that corresponding algorithms require additional communication among beacons, for example, to exchange and compare estimates $\hat{d}_{i,j}$ of the distance to the unknown. Naturally, the algorithms of class global knowledge are expected to select optimal or nearly optimal subsets in terms of (5) and perform better than those of class local knowledge because more information of the network topology is available.

An often used approach is to select the nearest k beacons which we refer to as *global-nearest-neighbor* algorithm which belongs to class global knowledge. This approach uses the distances between beacons and unknown to characterize their contribution to accurate localization. Naturally, beacons being nearer to the unknown are regarded as more valuable in terms of high accuracy of localization than beacons being farther away from it. This approach results in a suboptimal selection of beacons as geometry is not considered. The selected subsets are:

$$\hat{S}(c) = \arg \min_{\{S \in \mathbb{S}(c)\}} \sum_{b_i \in S} \hat{d}_{1,i} \tag{6}$$

We propose to use the CRLB on localization error (3) for the selection. Since the unknown broadcasts the request for localization, the beacons can estimate the distance to the unknown by means of received signal strength measurements, and use the CRLB to quantify their contribution to the accuracy of localization. Thereby, every beacon can determine how important it is for an accurate localization of the unknown in comparison to all other beacons if global knowledge is available. We refer to this approach as *global-crlb* algorithm. This algorithm selects optimal subsets in terms of (5). The subsets are:

$$\hat{S}(c) = S(c) = \arg \min_{\{S \in \mathbb{S}(c)\}} \sigma_1^2(S) \tag{7}$$

In the following, we adapt the algorithms of class global knowledge to be applicable if only local information is available. Starting with the initial request for localization, all beacons in transmission range of the unknown estimate the distance to the unknown.

We follow the approach to assign each beacon a probability $P^{(i)}$ $(i = m_\mathrm{u} + 1, \ldots, m_\mathrm{u} + m_\mathrm{b})$ of response which is key to avoid additional communication between beacons. Probabilities are based on distance and/or CRLB with the aim to ensure that beacons will respond in approximately the same order as optimal. Local information is obtained by a beacon from the unknown's initial broadcast and subsequently by overhearing the responses of other beacons. We emphasize, that in contrast to class global knowledge, local information is not explicitly exchanged but will be made

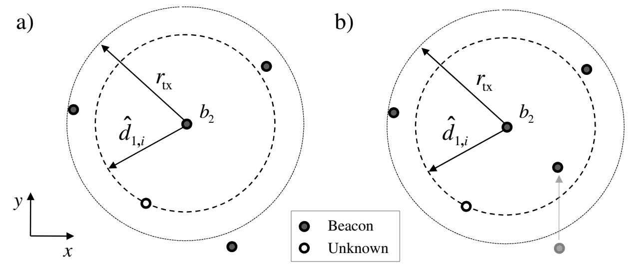

Figure 2. Circle rule applies for b_2 in a) but not in b).

available to other beacons during the process without additional communication effort.

The algorithm *local-nearest-neighbor* uses the ratio of transmission range r_tx of the unknown's request and distance to assign probabilities of response:

$$P^{(i)}_{\mathrm{nearest-k}} = 1 - \frac{\hat{d}_{1,i}}{r_\mathrm{tx}} \tag{8}$$

The beacons' responses include the originator's address and its location.

We propose a new hybrid algorithm, called *local-crlb*, which uses distances and the CRLB to assign beacons a probability of response. Responses are broadcast and include the originator's address and location and its estimate of the distance to the unknown. After the first beacon has answered, subsequent beacons can use the additional information provided by the former responses, namely distances of other beacons to the unknown. Consequently, the CRLB can be calculated and used as local selection criteria after the first beacon has answered.

Additionally, we apply the circle-rule (figure 2) to ensure that most important beacons respond first: By applying the circle rule, each beacon is assigned $P^{(i)}_{\mathrm{local-crlb}} = 1$ if there is no other beacon within distance $\hat{d}_{1,i}$ around it (Line 4 in table 2). Thus, each beacon assigns/updates its probability of response based on the number of overheard communications:

$$P^{(i)}_{\mathrm{local-crlb}} = \begin{cases} 1 - \frac{\hat{d}_{1,i}}{r_\mathrm{tx}} & \mathrm{card}\,(H_i) < 2 \\ \left| 1 - \frac{\sigma_1^2(H_i)}{\sigma_1^2(H_i \cup \{b_i\})} \right| & \mathrm{card}\,(H_i) \geq 2 \\ 1 & \text{circle-rule} \end{cases} \tag{9}$$

H_i contains those beacons b_j whose response has been overheard by beacon b_i, thus $H_i := \{b_j \mid \text{"}b_i \text{ has overheard response from } b_j\text{"}\}$ and $\mathrm{card}\{S\}$ refers to the cardinality of the set S Table 2 gives an outline of the algorithms of class local-knowledge in pseudo code. The code is executed by beacon $b_i \in B_1$ and

Table 1. Overview of algorithms.

algorithm	Knowledge	Metric used
global-crlb	global	CRLB
global-nearest-neighbor	global	Distances
local-nearest-neighbors	local	Distances
local-crlb	local	Distances, CRLB

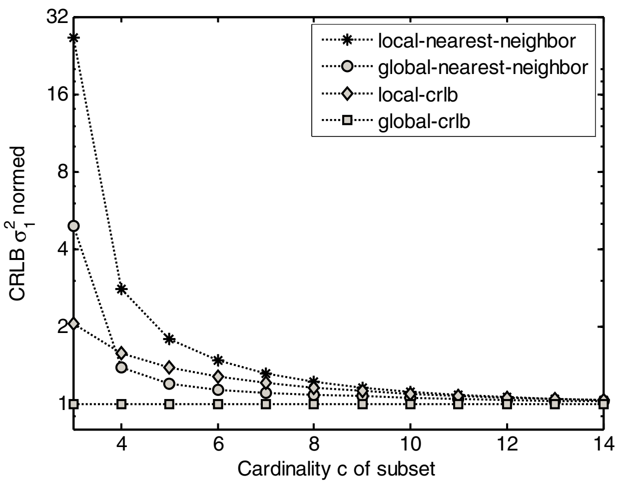

Figure 3. Simulation results.

it is assumed that all beacons in B_1 have received request for localization.

For our investigation, we assume that unknowns reduce transmission power to half the maximal transmission range while beacons respond with maximal transmission range. This effectively avoids the known problem of the "hidden terminal". Further, we assume ideal estimation of distances, i.e. $\hat{d}_{1,i} = d_{1,i}$ and that locations $\mathbf{z}_j$ of or distances $d_{i,j}$ to other beacons $b_j \in B_1$ are known at b_i. This assumption is justified because, for example due to communication during earlier localization processes or other communications, transmissions of beacons have been overheard by their direct neighbors.

4. Simulations and Results

We conducted simulations using Matlab [1] to compare the performance of the algorithms proposed in terms of CRLB with the following parameters: $M = 41$, $m_{\mathrm{u}} = 1$, $r_{\mathrm{tx}} = 100\,m$. Impact of borders is avoided by deploying all beacons in an area of size $3r_{\mathrm{tx}} \times 3r_{\mathrm{tx}}$ and placing randomly the unknown in the middle sector of size $r_{\mathrm{tx}} \times r_{\mathrm{tx}}$. The average number of beacons within transmission range of the unknown is $E\{\mathrm{card}(B_1)\} \approx 14$, thus $m_{\mathrm{b}} = 14$. Results are averaged over 10000 independent simulations and

Table 2. Algorithms of class local-knowledge. Code is executed at b_i.

Local-crlb

```
procedure ProcessLocalizationRequest(d̂_{1,i},
    {d̂_{i,j} | b_j ∈ B_1})
  if no other beacon within distance d̂_{1,i} to
      unknown then
    % assign probability of responding
    P_i ← 1
  else
    % calculate probability of responding
    P_i ← 1 − d̂_{1,i}/r_tx
  end if
  H_i ← ∅
  for each tdma cycle
    % update set of beacons which have
        responded
    H_i ← H_i ∪ {beacons b_j whose response has been
        overheard by b_i}
    if card(H_i) = 1 then
      if ²√(σ_1²(H_i ∪ {b_i})) > r_tx
        % quit current iteration without
            responding and try again next tdma
            cycle
        break
      end if
    elseif card(H_i) ≥ 2 then
      % update probability of responding
      P_i ← |1 − σ_1²(H_i)/σ_1²(H_i ∪ {b_i})|
    end if
    if randomNumber < P_i then
      % broadcast response
      respond(ownAddress, d̂_{1,i}, z_i)
      % exit procedure
      return
    end if
  end for
end procedure
```

Local-nearest-neighbor

```
procedure ProcessLocalizationRequest(d̂_{1,i})
  % calculate probability of responding
  P_i ← 1 − d̂_{1,i}/r_tx
  for each tdma cycle
    if randomNumber < P_i then
      % broadcast response
      respond(ownAddress, z_i)
      % exit procedure
      return
    end if
  end for
end procedure
```

depicted in figure 3.

Figure 3 depicts the average CRLB for the four algorithms discussed in section 3.2. The y-axis is equalized to the average optimal CRLB and the cardinality of subsets used to calculate (5) is depicted on the x-axis. It is shown that the algorithms of class global knowledge have a better overall performance despite for $c = 3$. Here local-crlb benefits from preventing some beacons from responding which would lead, otherwise, to a CRLB larger than the transmission range (see line 14 in table 2). Further, local-crlb achieves a significantly lower CRLB than local-nearest-neighbor.

5. Conclusions

This paper investigated algorithms to find the subset of beacons that minimizes the CRLB on localization error of the unknown over all possible subsets of a given cardinality. Since relatively small subsets of beacons often contribute disproportionately high to the accuracy of localization, these methods can be used to reduce complexity of localization. An algorithm has been developed, called local-crlb, which uses the CRLB on localization based on RSS for selection and which does not require additional communication between beacons. In order to avoid additional communication, the algorithm assigns each beacon a probability of response which can be calculate using only local information. Another approach which is often used in the literature selects the nearest beacons as subset. The performance of the algorithms considered has been compared based on simulations using Matlab. The simulations showed that when the algorithms are constrained to use local information only, the proposed algorithm allows for a more accurate localization of the unknown than the nearest neighbors approach.

6. Acknowledgement

This work was partially financed by the German Research Foundation (DFG) within the graduate school MuSAMA (GRK 1424).

References

[1] http://www.mathworks.com/products/matlab/.

[2] J. Albowicz, A. Chen, and L. Zhang. Recursive position estimation in sensor networks. In *Proceedings IEEE ICNP*, pages 35–41, Los Alamitos, CA, USA, 2001. IEEE Computer Society.

[3] J. Blumenthal, F. Reichenbach, and D. Timmermann. Minimal transmission power vs. signal strength as distance estimation for localization in wireless sensor networks. In *Proceedings of IEEE IWWAN*, pages 761–766, June 2006. New York, USA.

[4] N. Bulusu, J. Heidemann, and D. Estrin. Adaptive beacon placement. In *Proceedings of ICDCS*, pages 489–503, Washington, DC, USA, 2001. IEEE Computer Society.

[5] C. Chang and A. Sahai. Estimation bounds for localization. In *Proceedings of IEEE SECON*, pages 415–424, October 2004.

[6] K.-Y. Cheng, V. Tam, and K.-S. Lui. Improving aps with anchor selection in anisotropic sensor networks. In *Proceedings of ICAS/ICNS*, page 49, Los Alamitos, CA, USA, 2005. IEEE Computer Society.

[7] J. A. Costa, N. Patwari, and I. Alfred O. Hero. Distributed weighted-multidimensional scaling for node localization in sensor networks. *ACM Transactions on Sensor Networks*, 2(1):39–64, February 2006.

[8] J. Liu, Y. Zhang, and F. Zhao. Robust distributed node localization with error management. In *Proc. ACM MobiHoc '06*, pages 250–261. ACM Press, 2006.

[9] E. Olson, J. J. Leonard, and S. Teller. Robust range-only beacon localization. In *IEEE AUV*, 2004.

[10] N. Patwari, A. O. Hero III, M. Perkins, N. Correal, and R. O'Dea. Relative location estimation in wireless sensor networks. In *IEEE TSP*, volume 51, pages 2137–2148, August 2003.

[11] N. B. Priyantha, A. Chakraborty, and H. Balakrishnan. The cricket location-support system. In *ACM/IEEE MobiCom*, 2000.

[12] K. Sinha and A. D. Chowdhury. A beacon selection algorithm for bounded error location estimation in ad hoc networks. In *IEEE ICCTA*, pages 87–93. IEEE Computer Society, 2007.