

Efficient Distributed Algorithms for Data Fusion and Node Localization in Mobile Ad-hoc Networks

Andrew P. Brown, Ronald A. Iltis, and Ryan Kastner
University of California, Santa Barbara
Electrical and Computer Engineering Department
andrew.p.brown@gmail.com, {iltis, kastner}@ece.ucsb.edu

Abstract—Efficient distributed algorithms are an important enabling technology for large-scale ad-hoc wireless sensor and communications networks. In this paper, optimal Bayesian data fusion under the assumption of linear Gaussian state and measurement models is presented. Within this framework, an efficient algorithm for distributed state estimation in ad-hoc networks is developed. Approximate algorithms are then developed for further improvements in network resource efficiency. These include a parameterizable tradeoff of improved communications efficiency for increased latency in the rate at which information propagates through the network. It is also shown that the algorithms are well-suited for use with non-linear measurements. Finally, for distributed node position estimation in a mobile ad-hoc network, simulation results show that accurate, efficient node localization is achieved.

I. INTRODUCTION

Distributed estimation algorithms which provide savings in communication, computation, and memory resources are required for deployment of ad-hoc sensor and communications networks with large numbers of nodes. The applications of such networks are nearly limitless. Networks of low-cost wireless sensors can autonomously monitor phenomena of commercial, domestic, scientific, or military interest. For example, wireless sensor networks are useful in industrial process control, inventory management, applications enabling human safety/efficiency/convenience, scientific observation, and surveillance/target tracking.

In wireless sensor networks (WSNs) [1], information may be communicated multi-hop to a sink or central processor, with some degree of data reduction or fusion *en route*. The central processor would then compute the central (global) estimate of the phenomena of interest. However, in large networks, nodes which are near the sink are heavily burdened by relaying packets from more distant nodes. As a result, the sink data rate may be insufficient, and relay nodes may quickly become energy-depleted. It is therefore necessary to distribute computation throughout the network, with each node performing data fusion and reduction. In ad-hoc networks operating in the absence of any infrastructure/network backbones, distributed computation is in fact the only viable option.

In the distributed estimation approach presented here, raw measurements collected at each sensor are processed locally to generate local estimates of the states of interest. The sufficient statistics of these estimates provide a lossless compression of all measurement data used to generate the local estimates.

Information packets based on these sufficient statistics are then transmitted to other nearby nodes and fused. The communication and computation burden is shared equally by the network nodes, leading to improved network performance and longevity. Furthermore, the overall algorithm is parameterizable in the tradeoff of improved communications efficiency for increased latency in the rate at which information propagates through the network.

Work on distributed estimation dates back at least to the 1970s [5] and continues vigorously [19][24][34][27][28]. However, for the case of time-varying parameters (states), the development of practical algorithms which can effectively handle realistic communication delays and throughput constraints is an open problem. This case appears very often in real-world applications, especially in large ad-hoc networks of mobile nodes, where communications delays (e.g., due to medium access contention) are non-negligible, and inequitable load sharing between nodes would quickly lead to node and link failures due to energy depletion and traffic congestion.

In many previous approaches for distributed/decentralized estimation, solutions were limited to time-invariant states, or failed to account for correlations between local estimates (which leads to re-use of mutual information, resulting in inaccurate estimation and instability [36]). In [24][23], correlations between local estimates of *static* states were considered, with limited applicability for non-linear state observations. In [11], optimal distributed estimation was investigated via information flow graphs, and a procedure for identifying mutual information contained in two or three local estimates was provided for the time-invariant case. In [19], efficient distributed estimation for non-linear observations of *static* states in networks with incomplete connectivity was investigated in detail, building upon the work in [11]. In this paper, the difficulties associated with estimating time-varying states in the presence of finite communication delays are explained, and a Bayesian framework for efficient estimate fusion is presented. The resulting distributed estimation algorithm is designed for communications resource efficiency, and approximations to this optimal algorithm yield further increases in communications, computation, and memory efficiency.

As an example application, distributed localization in mobile ad-hoc networks is discussed, and simulations show that accurate, efficient localization can be achieved using inter-node range measurements under practical conditions. The

application to distributed node position estimation is important because, for example, node position information is valuable for enabling efficient routing in networks [33][22], and is required for position-dependent services. Node localization is also required for providing context to collected data in sensor networks, for example in habitat monitoring and in surveillance and tracking applications [35][26][1]. GPS cannot always be relied upon due to node design constraints (e.g., cost, power consumption, and/or form factor) and reliability (e.g., position estimate unavailability due to shadowing, multipath, and/or jamming). Thus, GPS-less distributed localization is an active research area in wireless sensor networks for both civilian [26][31][12][29] and military [13][25][36][19] applications.

II. LOCAL STATE ESTIMATION

Consider an ad-hoc network of sensor nodes which obtain state measurements and which communicate sufficient statistics for these measurements to some subset of neighboring nodes with which they are *connected* (e.g., via a shared wireless channel). Information packets may also be communicated multi-hop to other more distant nodes. In this and Sections III through V, the standard assumption of linear Gaussian state and measurement equations is made. The extension of these results to nonlinear state and/or measurement models is discussed in Section VI, and the particular application of position estimation based on (nonlinear) range estimates is discussed in Sections VII through VIII.

The linear state dynamics model is

$$\mathbf{x}(n) = \mathbf{F}(n, n-1)\mathbf{x}(n-1) + \mathbf{v}(n-1) \in \mathbb{R}^{N_x}, \quad (1)$$

(see, for example, [4]), where the initial state $\mathbf{x}(0)$ is assumed Gaussian distributed, $\mathbf{F}(n, n-1)$ is the state transition matrix, and $\mathbf{v}(n-1) \in \mathbb{R}^{N_w}$ is white Gaussian process noise with covariance $\mathbf{Q}(n-1)$. In the sequel, it will be convenient to write the state transition from discrete time index m to n (with $m \leq n$) as

$$\mathbf{x}(n) = \mathbf{F}(n, m)\mathbf{x}(m) + \mathbf{v}(n, m), \quad (2)$$

where the white Gaussian process noise $\mathbf{v}(n, m)$ has covariance $\mathbf{Q}(n, m)$. The matrices $\mathbf{F}(n, m)$ and $\mathbf{Q}(n, m)$ can be obtained recursively using

$$\mathbf{F}(m', m) = \mathbf{F}(m', m'-1)\mathbf{F}(m'-1, m), \quad (3)$$

$$\mathbf{Q}(m', m) = \mathbf{F}(m', m'-1)\mathbf{Q}(m'-1, m)\mathbf{F}^T(m', m'-1) + \mathbf{Q}(m'-1), \quad (4)$$

where $m' = m+2, \dots, n$. However, for computational efficiency, $\mathbf{F}(n, m)$ and $\mathbf{Q}(n, m)$ should be computed directly, e.g., as demonstrated for a particular choice of linear state dynamics model in Section VII. The linear measurement model is

$$\mathbf{z}(n) = \mathbf{H}(n)\mathbf{x}(n) + \mathbf{w}(n) \in \mathbb{R}^{N_z}, \quad (5)$$

where $\mathbf{H}(n)$ is the observation matrix, and $\mathbf{w}(n)$ is white Gaussian measurement noise with covariance $\mathbf{R}(n)$. The index

n references the time t_n when some event (relevant to the distributed estimation) occurs asynchronously in the network; the delay $t_{n+1} - t_n$ between consecutive events is arbitrary. Note that for time-varying states, network-wide clock synchronization is required, but if unavailable, the clock biases can be estimated jointly with $\mathbf{x}(n)$.

The cumulative measurement set \mathbf{Z}_i^m is defined as the set of all measurements recorded at node i , along with the set of all measurements for which sufficient statistics are received via communication with other nodes, up to and including time (index) m . That is, $p(\mathbf{x}(n)|\mathbf{Z}_i^m)$ denotes the *a posteriori* probability distribution on $\mathbf{x}(n)$, given the cumulative information available at node i at time m . Then the state estimate, the state estimate error, and the covariance of this error (equivalently, it can be shown, of the *state*), are given, respectively, by

$$\hat{\mathbf{x}}(n|\mathbf{Z}_i^m) = E\{\mathbf{x}(n)|\mathbf{Z}_i^m\}, \quad (6)$$

$$\tilde{\mathbf{x}}(n|\mathbf{Z}_i^m) = \mathbf{x}(n) - \hat{\mathbf{x}}(n|\mathbf{Z}_i^m), \quad (7)$$

$$\begin{aligned} \hat{\mathbf{P}}(n|\mathbf{Z}_i^m) &= Cov\{\tilde{\mathbf{x}}(n|\mathbf{Z}_i^m)\} \\ &= Cov\{\mathbf{x}(n)|\mathbf{Z}_i^m\} \\ &= \mathbf{P}(n|\mathbf{Z}_i^m). \end{aligned} \quad (8)$$

For $n \geq m$, $\hat{\mathbf{x}}(n|\mathbf{Z}_i^m)$ and $\mathbf{P}(n|\mathbf{Z}_i^m)$ can be efficiently calculated using the well-known Kalman filter (KF) algorithm [18][2][3] running locally at node i . Together, this conditional mean and covariance constitute sufficient statistics for the *a posteriori* density $p(\mathbf{x}(n)|\mathbf{Z}_i^m)$, under the assumption of linear Gaussian state and measurement models.

III. STATE ESTIMATE FUSION

Consider a data fusion event at node i and time $n \geq m$: node i receives a transmission from node j containing sufficient statistics $\hat{\mathbf{x}}(m|\mathcal{Z}_j^m)$, $\mathbf{P}(m|\mathcal{Z}_j^m)$, with \mathcal{Z}_j^m denoting some set of measurements (though not necessarily a *cumulative* measurement set) up to time m . Using the KF prediction equations, node i can propagate $\hat{\mathbf{x}}(m|\mathcal{Z}_j^m)$, $\mathbf{P}(m|\mathcal{Z}_j^m)$ to the current time n :

$$\hat{\mathbf{x}}(n|\mathcal{Z}_j^m) = \mathbf{F}(n, m)\hat{\mathbf{x}}(m|\mathcal{Z}_j^m), \quad (9)$$

$$\mathbf{P}(n|\mathcal{Z}_j^m) = \mathbf{Q}(n, m) + \mathbf{F}(n, m)\mathbf{P}(m|\mathcal{Z}_j^m)\mathbf{F}^T(n, m). \quad (10)$$

Node i 's task is then to fuse $\hat{\mathbf{x}}(n|\mathcal{Z}_j^m)$, $\mathbf{P}(n|\mathcal{Z}_j^m)$ from node j with its previous local information $\hat{\mathbf{x}}(n|\mathbf{Z}_i^{n-1})$, $\mathbf{P}(n|\mathbf{Z}_i^{n-1})$ to form $\hat{\mathbf{x}}(n|\mathbf{Z}_i^n)$, $\mathbf{P}(n|\mathbf{Z}_i^n)$ (sufficient statistics for $p(\mathbf{x}(n)|\mathbf{Z}_i^n)$).

To fuse two Gaussian posteriori densities $p(\mathbf{x}(n)|\mathcal{Z}_1)$ and $p(\mathbf{x}(n)|\mathcal{Z}_2)$ conditioned on measurement sets \mathcal{Z}_1 and \mathcal{Z}_2 , the fusion formula

$$p(\mathbf{x}(n)|\mathcal{Z}_1 \cup \mathcal{Z}_2) = \frac{1}{c_f} \frac{p(\mathbf{x}(n)|\mathcal{Z}_1)p(\mathbf{x}(n)|\mathcal{Z}_2)}{p(\mathbf{x}(n)|\mathcal{Z}_1 \cap \mathcal{Z}_2)} \quad (11)$$

from [11] can be used if

$$\begin{aligned} p(\mathcal{Z}_1 \cup \mathcal{Z}_2|\mathbf{x}(n)) &= p(\mathcal{Z}_1 - \mathcal{Z}_2|\mathbf{x}(n)) \\ &\quad \cdot p(\mathcal{Z}_2 - \mathcal{Z}_1|\mathbf{x}(n)) \cdot p(\mathcal{Z}_1 \cap \mathcal{Z}_2|\mathbf{x}(n)), \end{aligned} \quad (12)$$

where ‘ \cup ’, ‘ \cap ’, and ‘ $-$ ’ denote the set union, intersection, and reduction operators. That is, the fusion formula (11) holds if the measurement sets $\mathcal{Z}_1 - \mathcal{Z}_2$, $\mathcal{Z}_2 - \mathcal{Z}_1$, and $\mathcal{Z}_1 \cap \mathcal{Z}_2$

are independent, given $\mathbf{x}(n)$. This is not generally the case for time-varying $\mathbf{x}(n)$. For example, if $\mathcal{Z}_1 - \mathcal{Z}_2$ contains the measurement $\mathbf{z}^{k_1}(m_{k_1})$ obtained at node k_1 at time m_{k_1} , and $\mathcal{Z}_2 - \mathcal{Z}_1$ contains the measurement $\mathbf{z}^{k_2}(m_{k_2})$ obtained at node k_2 at time m_{k_2} , where $n > m_{k_1} > m_{k_2}$, then $\mathbf{z}^{k_1}(m_{k_1})$ and $\mathbf{z}^{k_2}(m_{k_2})$ are dependent, since $\mathbf{x}(m_{k_2})$ and $\mathbf{x}(m_{k_1})$ are dependent. Condition (12) *does* hold for estimation of the joint posterior density $\mathbf{x}^n = \{\mathbf{x}(n), \mathbf{x}(n-1), \dots, \mathbf{x}(0)\}$ of the process at all measurement times:

$$p(\mathcal{Z}_1 \cup \mathcal{Z}_2 | \mathbf{x}^n) = p(\mathcal{Z}_1 - \mathcal{Z}_2 | \mathbf{x}^n) \cdot p(\mathcal{Z}_2 - \mathcal{Z}_1 | \mathbf{x}^n) \cdot p(\mathcal{Z}_1 \cap \mathcal{Z}_2 | \mathbf{x}^n). \quad (13)$$

However, since the number of states grows linearly with n , this approach is impractical due to computation and memory requirements.

There is an important case for which the fusion formula (11) can be used efficiently: condition (12) holds when either \mathcal{Z}_1 or \mathcal{Z}_2 is simply $\mathbf{z}^i(n)$, the vector of measurements obtained at the most recent measurement time n . Given $\mathbf{x}(n)$, these measurements are independent, under the assumption of white measurement noise in (5). The implications of this observation will be demonstrated in the following sections.

IV. OPTIMAL DISTRIBUTED STATE ESTIMATION

When node i obtains measurements $\mathbf{z}^i(n)$ at time n , it uses the KF to compute (sufficient statistics for) $p(\mathbf{x}(n) | \mathbf{Z}_i^n)$, given $p(\mathbf{x}(n-1) | \mathbf{Z}_i^{n-1})$. Node i can then use (11) to calculate sufficient statistics for the posterior density based only on the new measurements $\mathbf{z}^i(n)$. Specifically,

$$p(\mathbf{x}(n) | \mathbf{z}^i(n)) = \frac{1}{c'_f} \frac{p(\mathbf{x}(n)) p(\mathbf{x}(n) | \mathbf{Z}_i^n)}{p(\mathbf{x}(n) | \mathbf{Z}_i^{n-1})}, \quad (14)$$

where $p(\mathbf{x}(n))$ is the *a priori* probability distribution (efficient formulas for computing the multiplication and division of Gaussian densities are provided in the Appendix). The sufficient statistics $\hat{\mathbf{x}}(n | \mathbf{z}^i(n)), \mathbf{P}(n | \mathbf{z}^i(n))$ for $p(\mathbf{x}(n) | \mathbf{z}^i(n))$ can then be communicated via a sufficient statistics packet (SSP) to other nearby nodes, which can use the received information to update their local state estimates¹. Implicitly, it is assumed that SSPs are time-stamped with the most recent asynchronous measurement time, t_n in this case.

When node $j \neq i$ receives an SSP containing $\hat{\mathbf{x}}(n | \mathbf{z}^i(n)), \mathbf{P}(n | \mathbf{z}^i(n))$, it fuses this new information with its most recently-computed sufficient statistics $\hat{\mathbf{x}}(m | \mathbf{Z}_j^m), \mathbf{P}(m | \mathbf{Z}_j^m)$ for $p(\mathbf{x}(m) | \mathbf{Z}_j^m)$. If node j receives redundant copies of $\hat{\mathbf{x}}(n | \mathbf{z}^i(n)), \mathbf{P}(n | \mathbf{z}^i(n))$, it can identify (based on the index pair (i, n)) and discard them. For fusing new information, there are two cases to consider, $n \geq m$ and $n < m$. In the first case, it is simple to obtain optimal fusion:

IF $n \geq m$:

$$\begin{aligned} \hat{\mathbf{x}}(n | \mathbf{Z}_j^m) &= \mathbf{F}(n, m) \hat{\mathbf{x}}(m | \mathbf{Z}_j^m) \\ \mathbf{P}(n | \mathbf{Z}_j^m) &= \mathbf{Q}(n, m) + \mathbf{F}(n, m) \mathbf{P}(m | \mathbf{Z}_j^m) \mathbf{F}^T(n, m) \end{aligned}$$

¹For improved communications efficiency, SSP messages may be included in packets transmitted for other purposes.

$$p(\mathbf{x}(n) | \mathbf{Z}_j^n) = \frac{1}{c'_f} \frac{p(\mathbf{x}(n) | \mathbf{Z}_j^m) p(\mathbf{x}(n) | \mathbf{z}^i(n))}{p(\mathbf{x}(n))} \quad (15)$$

END IF

In the case $n < m$, optimal fusion is more difficult, and unfortunately, this case is commonly encountered when communications delays are non-negligible. In formulating the optimal fusion algorithm, the following definitions are adopted: $p(\mathbf{x}(n-1) | \mathcal{Z}_j^{n-1})$ denotes the posterior density given the set of measurements \mathcal{Z}_j^{n-1} up to time $n-1$ ² for which SSPs are known at node j , and $\{p(\mathbf{x}(m_k) | \mathbf{z}^k(m_k))\}$ denotes a set of posterior densities, with $k = 1, 2, \dots, K$ indexing the SSPs previously known at node j , for measurement times $m_k \geq n$ ($m_K = m$). Note that k references the node of origin for the SSP formed from $\mathbf{z}^k(m_k)$. Given $\hat{\mathbf{x}}(n-1 | \mathcal{Z}_j^{n-1}), \mathbf{P}(n-1 | \mathcal{Z}_j^{n-1})$ and $\{\hat{\mathbf{x}}(m_k | \mathbf{z}^k(m_k)), \mathbf{P}(m_k | \mathbf{z}^k(m_k))\}$, optimal fusion of $\hat{\mathbf{x}}(n | \mathbf{z}^i(n)), \mathbf{P}(n | \mathbf{z}^i(n))$ is performed as follows:

IF $n < m$:

$$\hat{\mathbf{x}}^{\text{temp}} \leftarrow \hat{\mathbf{x}}(n-1 | \mathcal{Z}_j^{n-1}),$$

$$\mathbf{P}^{\text{temp}} = \mathbf{P}(n-1 | \mathcal{Z}_j^{n-1})$$

$$m_{-1} \leftarrow n-1, m_0 \leftarrow n$$

$$\hat{\mathbf{x}}(m_0 | \mathbf{z}^0(m_0)) \leftarrow \hat{\mathbf{x}}(n | \mathbf{z}^i(n))$$

$$\mathbf{P}(m_0 | \mathbf{z}^0(m_0)) \leftarrow \mathbf{P}(n | \mathbf{z}^i(n))$$

FOR $k = 0, 1, \dots, K$:

$$\hat{\mathbf{x}}^{\text{temp}} \leftarrow \mathbf{F}(m_k, m_{k-1}) \hat{\mathbf{x}}^{\text{temp}}$$

$$\mathbf{P}^{\text{temp}} \leftarrow \mathbf{Q}(m_k, m_{k-1})$$

$$+ \mathbf{F}(m_k, m_{k-1}) \mathbf{P}^{\text{temp}} \mathbf{F}^T(m_k, m_{k-1})$$

$$p^{\text{temp}} = \frac{1}{c'_f} \frac{p^{\text{temp}} p(\mathbf{x}(m_k) | \mathbf{z}^k(m_k))}{p(\mathbf{x}(m_k))} \quad (16)$$

(p^{temp} is represented by suff. stats. $\hat{\mathbf{x}}^{\text{temp}}, \mathbf{P}^{\text{temp}}$)

END FOR

$$p(\mathbf{x}(m) | \mathbf{Z}_j^m) \leftarrow p^{\text{temp}}$$

END IF

Due to SSP reprocessing required in this optimal fusion algorithm, the computation and memory requirement is $\mathcal{O}(K+1)$ times greater in the case $n < m$, compared with the case $n \geq m$.

V. APPROXIMATE DISTRIBUTED STATE ESTIMATION

Though the optimal distributed estimation algorithm of Section IV is a practical choice for many applications, there will be other applications for which greater efficiency is required. In Section V-A, an approximate estimate fusion algorithm is introduced for improved computation and memory resource efficiency, and in Section V-B, a method for improved communications efficiency is presented.

²Specifically, $n-1 = \max\{n_{k'} : \mathbf{z}^{k'}(n_{k'}) \in \mathbf{Z}_j^m, n_{k'} \leq n\}$

A. Improvements in Computation and Memory Resource Efficiency

In the case $n < m$, the computational complexity of optimal estimate fusion is $\mathcal{O}((K+1)N_x^3)$, which can be reduced to $\mathcal{O}(N_x^3)$, with a similar reduction in memory requirements, using the following approximate fusion algorithm:

IF $n < m$:

$$\hat{\mathbf{x}}(m|\mathbf{z}^i(n)) = \mathbf{F}(m, n)\hat{\mathbf{x}}(n|\mathbf{z}^i(n))$$

$$\mathbf{P}(m|\mathbf{z}^i(n)) = \mathbf{Q}(m, n) + \mathbf{F}(m, n)\mathbf{P}(n|\mathbf{z}^i(n))\mathbf{F}^T(m, n)$$

$$p(\mathbf{x}(m)|\mathbf{Z}_j^m) \leftarrow \frac{1}{c_f^{m'}} \frac{p(\mathbf{x}(m)|\mathbf{Z}_j^m) p(\mathbf{x}(m)|\mathbf{z}^i(n))}{p(\mathbf{x}(m))} \quad (17)$$

END IF

From Section III, (17) results from the approximation

$$p(\mathbf{Z}_j^m \cup \mathbf{z}^i(n)|\mathbf{x}(m)) \approx p(\mathbf{Z}_j^m|\mathbf{x}(m))p(\mathbf{z}^i(n)|\mathbf{x}(m)), \quad (18)$$

which holds exactly if the states are time-invariant. Notice that for $n = m$, (17) is the same as (15); i.e., the fusion is optimal if the SSP communication delay is 0.

B. Improvements in Communications Resource Efficiency

The algorithms of Sections IV and V-A can be modified to obtain an approximate distributed state estimation algorithm which is parameterizable in the tradeoff of improved communications efficiency for increased latency in the rate at which information propagates through the network. The approach is essentially the same, except that transmitted SSPs contain information for *blocks* of N_b consecutive measurements at a single node, rather than for a single measurement vector obtained at one node and one time³.

SSP formation for improved communications efficiency is accomplished as follows. As each new measurement vector $\mathbf{z}^i(n)$ is obtained, the new information is extracted via (14) using the predicted and updated KF estimates. The extracted information $\hat{\mathbf{x}}(n|\mathbf{z}^i(n)), \mathbf{P}(n|\mathbf{z}^i(n))$ is then fused into the current SSP sufficient statistics $\hat{\mathbf{x}}(n|\mathcal{Z}_{\text{SSP}}^n), \mathbf{P}(n|\mathcal{Z}_{\text{SSP}}^n)$ (predicted to the current time via the KF prediction equations), where $\mathcal{Z}_{\text{SSP}}^n$ is the set of measurements whose information is included in the current SSP. The fusion of new information into an SSP is accomplished via another application of (11):

$$p(\mathbf{x}(n)|\mathcal{Z}_{\text{SSP}}^n) \leftarrow \frac{1}{c_f^{m''}} \frac{p(\mathbf{x}(n)|\mathbf{z}^i(n)) p(\mathbf{x}(n)|\mathcal{Z}_{\text{SSP}}^n)}{p(\mathbf{x}(n))}. \quad (19)$$

When the full block of N_b measurements has been processed, the SSP is transmitted to nearby nodes. When node $j \neq i$ receives the SSP, it uses the approximate fusion algorithm of Section V-A to update its local estimate $\hat{\mathbf{x}}(m|\mathbf{Z}_j^m), \mathbf{P}(m|\mathbf{Z}_j^m)$ using $\hat{\mathbf{x}}(n|\mathcal{Z}_{\text{SSP}}^n), \hat{\mathbf{P}}(n|\mathcal{Z}_{\text{SSP}}^n)$. In Section VIII, simulation results are presented for various choices of N_b , and it is shown

³The resulting improvement in communications efficiency is especially large if SSPs are transmitted wirelessly as individual packets in a network using a carrier sense multiple access collision avoidance protocol (such as in 802.11).

that significant improvement in communications efficiency may be possible with minimal degradation of estimation accuracy.

VI. NONLINEAR STATE ESTIMATION

The distributed algorithms of Sections IV and V are well-suited for estimation with non-linear state dynamics and/or measurement equations. The extension to this case is now discussed, along with considerations affecting estimation accuracy. Consider the state dynamics model

$$\mathbf{x}(n+1) = \mathbf{f}^n(\mathbf{x}(n)) + \mathbf{g}^n(\mathbf{x}(n))\mathbf{w}(n), \quad (20)$$

where $\mathbf{f}^n(\cdot)$ is a linear or non-linear function of $\mathbf{x}(n)$, and $\mathbf{g}^n(\cdot)$ is an $N_x \times N_w$ matrix whose components depend on $\mathbf{x}(n)$. For the measurement model, let

$$\mathbf{z}(n) = \mathbf{h}^n(\mathbf{x}(n)) + \mathbf{v}(n), \quad (21)$$

where $\mathbf{h}^n(\cdot)$ is a linear or non-linear function of $\mathbf{x}(n)$.

The extended Kalman filter (EKF) is the algorithm most commonly applied for non-linear estimation. In the EKF, equations (20) and (21) are simply *linearized* via a first-order Taylor series expansion about the most recent local state estimates, $\hat{\mathbf{x}}(n|\mathbf{Z}_i^n)$ and $\hat{\mathbf{x}}(n|\mathbf{Z}_i^{n-1})$. The EKF equations [2] are then obtained by applying the standard KF for the linearized model.

The distributed estimation algorithms of Sections IV and V can be applied using conditional means and covariances obtained from EKF algorithms running locally at each node. Due to the linearizations in the EKF, the estimates obtained are approximations to the optimal minimum variance estimates, and the estimation errors are only approximately Gaussian. For calculation via (14) of sufficient statistics $\hat{\mathbf{x}}(n|\mathbf{z}^i(n)), \mathbf{P}(n|\mathbf{z}^i(n))$ based on the current measurement vector $\mathbf{z}^i(n)$, good accuracy can be obtained if the EKF is closely tracking the states (i.e., if the estimation errors are small). In fact, the utility of (14) is most evident in the context of non-linear estimation. The reason for this can be seen as follows. The calculated density $p(\mathbf{x}(n)|\mathbf{z}^i(n))$ is the *a posteriori* distribution based on a single measurement vector $\mathbf{z}^i(n)$ and the *a priori* distribution $p(\mathbf{x}(n))$. For the sake of argument, ignore linearization errors. Then $p(\mathbf{x}(n)|\mathbf{z}^i(n))$ could be calculated by predicting the *a priori* density (which in practice may be based on some initial measurements shared among all nodes) to the current time using the KF/EKF prediction equations, and then calculating the updated density using $\mathbf{z}^i(n)$ in the KF/EKF update equations. However, bringing linearization error back into the picture, in the non-linear case the estimation of $p(\mathbf{x}(n)|\mathbf{z}^i(n))$ would be severely inaccurate because the *a priori* mean predicted to the current time would likely be very far from the true state, and the predicted *a priori* covariance would be very large [2]⁴. In contrast, (14) can provide an accurate method for estimating sufficient statistics for $p(\mathbf{x}(n)|\mathbf{z}^i(n))$. The utility of (14) is demonstrated in Section VIII for distributed EKF-based node

⁴In [2] it is shown that the linearization is highly accurate if the state estimate covariance is very small.

localization using inter-node range measurements (non-linear transformations of the node positions).

The method of Section V-B for improving communications efficiency at the expense of increased delays in information propagation should also be evaluated in the context of non-linear estimation. Increasing the number N_p of measurements represented by one SSP could cause the local estimation errors at each node to increase, since the most recent measurements obtained at other nodes are unavailable locally. This increase in instantaneous estimation errors could make the EKF linearization less accurate, leading to further increases in estimation errors. However, this effect may be negligible as long as the local estimation errors are kept small (e.g., by ensuring a sufficient local measurement rate). For the simulations described in Section VIII, N_p was varied between 1 and 10, with little effect on estimation performance.

VII. DISTRIBUTED POSITION ESTIMATION

In wireless ad-hoc networks in which GPS is unavailable to all or some fraction of the network nodes, localization can be based on any of a variety of measurements, including received signal strength [30], connectivity [7], angle of arrival (AOA), time of arrival (TOA), and time difference of arrival (TDOA) [8]. Received Signal Strength Indicator (RSSI) measurements are available from existing radios, but this position estimation method [15][14][30] is vulnerable to multipath fading [8] and requires extensive environment-specific calibration. Connectivity-based localization cannot yield high-resolution position estimates in many cases. When AOA measurements are used, a position can be estimated with only two reference nodes. However, performance may be severely degraded by scattering near antennas [8], and the use of antenna arrays corresponds to increased sensor node cost, power consumption, size, and complexity. In the simulations described in Section VIII, node positions are estimated using range measurements between pairs of nodes. Such range measurements can be obtained, for example, from round-trip travel time (RTT) measurements [17][25][16][21][20][19].

For a master node i located at $\mathbf{x}_i(n)$ obtaining range measurements to N_z ($N_z = 1$ for the simulations discussed) reference nodes $j = j_1, j_2, \dots, j_{N_z}$ located at $\mathbf{x}_j(n)$, the non-linear measurement equation is given by

$$\mathbf{z}^i(n) = \mathbf{h}^n(\mathbf{x}(n)) + \mathbf{v}(n) \in \mathbb{R}^{N_z \times 1}, \quad (22)$$

where

$$\mathbf{h}^n(\mathbf{x}(n)) = [\rho_{i,j_1}(n) \ \rho_{i,j_2}(n) \ \cdots \ \rho_{i,j_{N_z}}(n)]^T, \quad (23)$$

and $\rho_{i,j}(n) = \|\mathbf{x}_j(n) - \mathbf{x}_i(n)\|$. Node positions estimation is thus complicated by the non-linearity and state coupling inherent in the measurement equation. In this paper, time-varying node positions are estimated and tracked using the EKF. Other approaches to position estimation based on range measurements have included methods based on semidefinite programming [12] and local coordinate system construction based on geometry [9]. Combinations of the local coordinate and an EKF approach were developed in [31][32]. Indeed,

for time-varying states, a KF-based algorithm is best able to provide the required position tracking capabilities.

The linearization of (22), as required for the EKF algorithm, is specified in [6][19]. For the simulations presented in Section VIII, a common linear model, the discretized ‘‘continuous white noise acceleration model’’ [4], is assumed for the state dynamics. In this model, position $y(n)$ and velocity $v_y(n)$ in one coordinate are described by

$$\mathbf{y}_{1D}(n) = \mathbf{F}(n, m)\mathbf{y}_{1D}(m) + \mathbf{v}(n, m), \quad (24)$$

where

$$\mathbf{y}_{1D}(n) = \begin{bmatrix} y(n) \\ v_y(n) \end{bmatrix}^T, \quad (25)$$

$$\mathbf{F}(n, m) = \begin{bmatrix} 1 & t_n - t_m \\ 0 & 1 \end{bmatrix}, \quad (26)$$

and $\mathbf{v}(n, m)$ has covariance

$$\mathbf{Q}(n, m) = \begin{bmatrix} \frac{1}{3}(t_n - t_m)^3 & \frac{1}{2}(t_n - t_m)^2 \\ \frac{1}{2}(t_n - t_m)^2 & (t_n - t_m) \end{bmatrix} \sigma_{\Delta v}^2. \quad (27)$$

The overall state vector $\mathbf{x}(n)$ is of dimension $N_x = 4N_n$, for a network of N_n nodes distributed in two dimensions. Each node estimates the entire vector $\mathbf{x}(n)$ locally. This is necessary due to the coupled nature of the node pair positions in range measurements, if there is uncertainty in both node locations. The resulting cross-correlations between position and velocity coordinate estimates for all nodes must be accounted for in order to avoid information loops leading to instability [36]. Since the complexity of the overall distributed estimation is $\mathcal{O}(N_x^3)$, in large networks the formation of smaller subnetworks, though sub-optimal, is a practical necessity.

VIII. POSITION ESTIMATION RESULTS

The distributed estimation algorithms of Sections IV-V and the EKF-based position estimation and tracking algorithm of Section VII were simulated for networks of $N_n = 20$ and $N_n = 40$ nodes. All nodes were randomly deployed in the network area (667-by-667 m for $N_n = 20$ and 1000-by-1000 m for $N_n = 40$) and moved according to the white noise acceleration model of Section VII with 10 m/s initial velocity standard deviation and $\sigma_{\Delta v} = 1$ m/s² acceleration standard deviation. The measurement and one-hop communication range of each node was limited to 275 m (a shorter range could be used in denser networks), and each node ranged (with measurement errors of 10 m standard deviation) to its nearest five neighbors (*if within range*) once every 0.5 to 1.5 seconds (one range measurement was obtained every 0.1 to 0.3 sec., selected randomly from a uniform distribution). Nodes communicated SSPs to neighbors located a maximum of $N_h = 1, 2$, or 3 hops away, where the number of hops was modeled according to the distance d (m) between source and destination nodes: $0 \leq d \leq 275 \rightarrow 1$ hop, $275 < d \leq 478 \rightarrow 2$ hops, and $478 < d \leq 662 \rightarrow 3$ hops. Note that in this model the number of hops required to guarantee that a message could be communicated between two nodes at far corners of the network would be 5 hops for the $N_n = 20$ network and

8 hops for $N_n = 40$ network. Source nodes communicated SSPs to one-hop neighbors with a delay (including processing delay) of 0.3 sec., or more, and the communications delay for any additional hops was 0.2 sec., or more, per hop. The standard deviation for initial node position and velocity estimates was 150 m and 5 m/s, respectively, for 70% of the nodes. The other 30% of nodes were “assisted” (e.g., via GPS) and obtained independent estimates of their own position and velocity once per second with error standard deviations of 10 m and 0.333 m/s, respectively. The existence of “assisted” nodes was assumed in order to enable geographic, as opposed to relative, localization (the assisted nodes serve a role similar to that of “anchor” or “beacon” nodes in the localization literature). Because estimation errors are coupled across nodes (as discussed in Section VII), assisted nodes communicated measurements of their own positions to other nodes up to N_h hops away, with the same communication delays as for SSPs.

The approximate algorithms of Sections IV-V were simulated over 100 (for $N_n = 20$) or 5 (for $N_n = 40$) ensemble runs, and their performance was compared with that of the centralized algorithm (optimal except for EKF linearization). In Figure 1, for $N_n = 20$, $N_h = 1$, and SSP block size $N_b = 1$, the mean absolute position estimation error and the rms position estimation error (equivalently, the standard deviation of the zero-mean errors) for the unassisted nodes is shown versus time, illustrating the speed with which the EKF-based distributed estimation algorithm converges and the efficiency with which it tracks the node positions. As expected, the centralized algorithm achieves lower average errors in these simulations, since it is given all measurements with no delays. Note that the centralized EKF rms error increases after $t = 4$ sec., perhaps due to decreasing network connectivity caused by the random node motion. In Figures 2-3 (20 nodes) and 4-5 (40 nodes), the performance of the approximate algorithms is further evaluated, in terms of the final mean absolute and rms error, for various choices of N_h and N_b . Significant savings in communications resources are achieved for smaller values of N_h and/or larger values of N_b , and there seems to be little degradation in estimation accuracy for SSP block sizes of up to at least $N_b = 5$. Finally, while the development of approximate fusion algorithms for time-varying states and non-negligible communications delays deserves further attention, the distributed estimation algorithms presented here have been shown to provide an attractive choice for EKF-based localization in ad-hoc networks of mobile nodes.

IX. CONCLUSIONS

A computationally-efficient algorithm for optimal Bayesian data fusion in mobile ad-hoc networks has been presented. It was shown that this algorithm leads to practical overall distributed estimation algorithms which are resource-efficient, and which are well-suited for use with non-linear measurements. Finally, for distributed node position estimation based on range measurements in a mobile ad-hoc network, simulation results show that accurate, efficient node localization is achieved.

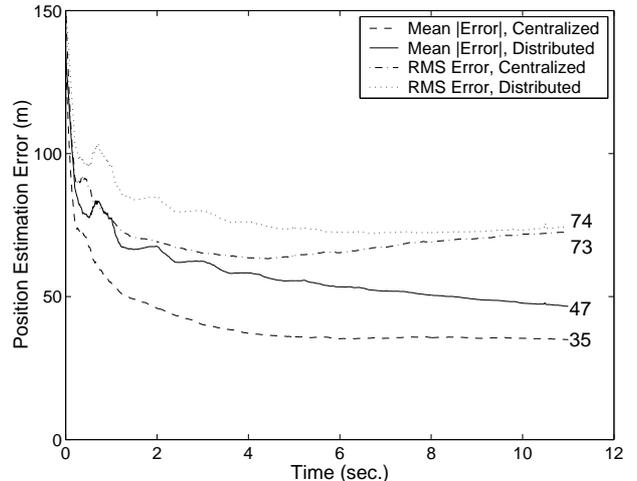


Fig. 1. Comparison of approximate distributed and optimal centralized algorithms: Mean absolute and rms position estimation errors versus time for $N_n = 20$ network nodes, with communication over a maximum of $N_h = 1$ hop.

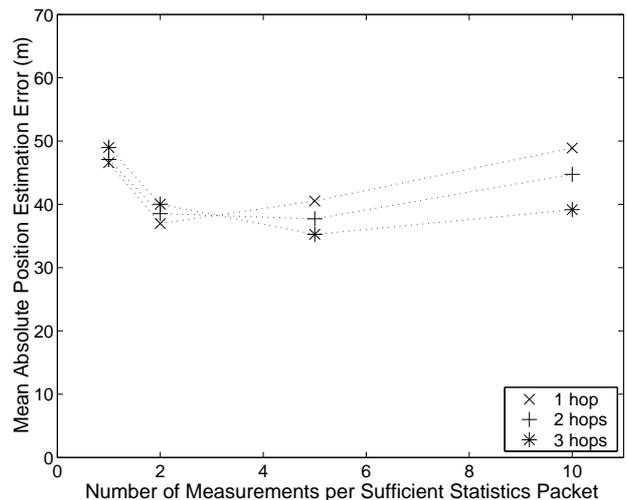


Fig. 2. Mean absolute position estimation errors versus sufficient statistics packet block size N_b , with communication over a maximum of $N_h = 1, 2$, or 3 hops, for $N_n = 20$ network nodes.

ACKNOWLEDGMENTS

This work was supported in part by NSF grant No. CNS-0411321.

APB thanks Dr. James V. Candy for many helpful discussions regarding nonlinear estimation.

APPENDIX: EFFICIENT MULTIPLICATION AND DIVISION OF GAUSSIAN DENSITIES

From [10][11][19], sufficient statistics for the Gaussian density

$$p(\mathbf{x}(n)|\mathcal{Z}_4) = \frac{1}{c_f} \frac{p(\mathbf{x}(n)|\mathcal{Z}_1) p(\mathbf{x}(n)|\mathcal{Z}_2)}{p(\mathbf{x}(n)|\mathcal{Z}_3)} \quad (28)$$

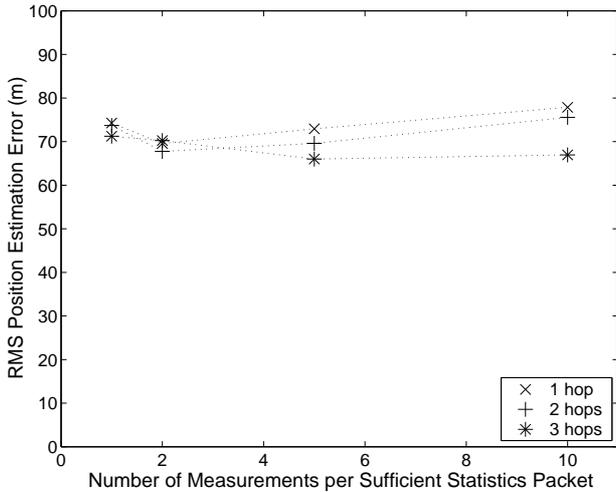


Fig. 3. *rms* position estimation errors versus sufficient statistics packet block size N_b , with communication over a maximum of $N_h = 1, 2$, or 3 hops, for $N_n = 20$ network nodes.

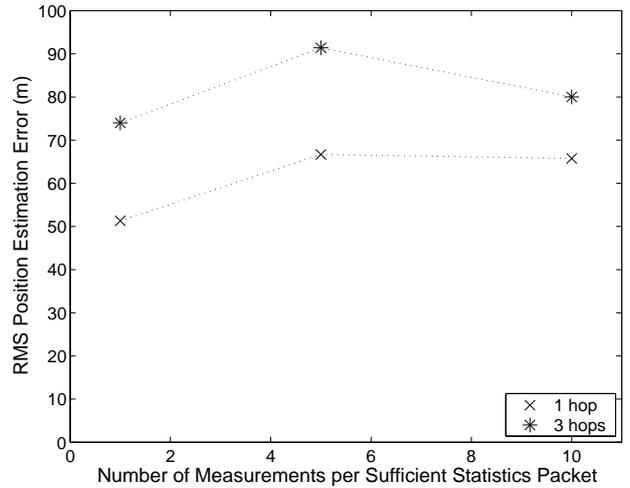


Fig. 5. *rms* position estimation errors versus sufficient statistics packet block size N_b , with communication over a maximum of $N_h = 1, 2$, or 3 hops, for $N_n = 40$ network nodes.

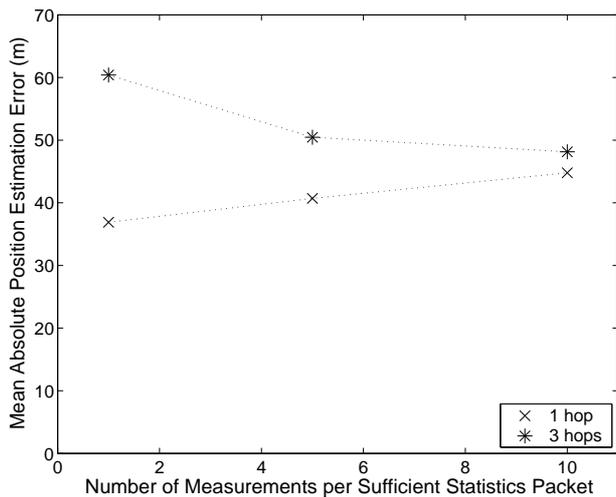


Fig. 4. *Mean absolute* position estimation errors versus sufficient statistics packet block size N_b , with communication over a maximum of $N_h = 1, 2$, or 3 hops, for $N_n = 40$ network nodes.

are given by

$$\mathbf{P}(n|\mathcal{Z}_4)^{-1} = \mathbf{P}(n|\mathcal{Z}_1)^{-1} + \mathbf{P}(n|\mathcal{Z}_2)^{-1} - \mathbf{P}(n|\mathcal{Z}_3)^{-1}, \quad (29)$$

$$\hat{\mathbf{x}}(n|\mathcal{Z}_4) = \mathbf{P}(n|\mathcal{Z}_4)[\mathbf{P}(n|\mathcal{Z}_1)^{-1}\hat{\mathbf{x}}(n|\mathcal{Z}_1) + \mathbf{P}(n|\mathcal{Z}_2)^{-1}\hat{\mathbf{x}}(n|\mathcal{Z}_2) - \mathbf{P}(n|\mathcal{Z}_3)^{-1}\hat{\mathbf{x}}(n|\mathcal{Z}_3)]. \quad (30)$$

REFERENCES

- [1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. A survey on sensor networks. *IEEE Commun. Magazine*, pages 102–114, Aug. 2002.
- [2] B. Anderson and J. Moore. *Optimal Filtering*. Prentice-Hall, Englewood Cliffs, NJ, 1979.
- [3] Y. Bar-Shalom and T. E. Fortmann. *Tracking and Data Association*. Academic Press, Orlando, FL, 1988.
- [4] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation*. Wiley-Interscience, New York, 2001.
- [5] Y. Bar-Shalom and E. Tse. Tracking in a cluttered environment with probabilistic data association. *Automatica*, 11(5):451–460, Sept. 1975.
- [6] A. P. Brown and R. A. Iltis. Distributed terrestrial radiolocation using the RLS algorithm. In *Proceedings of the International Telemetering Conference, ITC 2002*, San Diego, CA, Oct. 2002.
- [7] N. Bulusu, J. Heidemann, and D. Estrin. GPS-less low-cost outdoor localization for very small devices. *IEEE Personal Communications*, pages 28–34, Oct. 2000.
- [8] J. J. Caffery, Jr. *Wireless location in CDMA cellular radio systems*. Kluwer Academic Publishers, Norwell, MA, 2000.
- [9] S. Capkun, M. Hamdi, and J. Hubaux. GPS-free positioning in mobile ad-hoc networks. *Cluster Computing*, 5:157–167, 2002.
- [10] C. Y. Chong. Hierarchical estimation. In *Proceedings of the MIT/ONR Workshop on C3*, (Monterey, CA), 1979.
- [11] C. Y. Chong, E. Tse, and S. Mori. Distributed estimation in networks. In *Proceedings of the 1983 American Control Conference*, volume 1, pages 294–300, San Francisco, CA, Sept. 1983.
- [12] L. Doherty, K. S. J. Pister, and L. E. Ghaoui. Convex position estimation in wireless sensor networks. In *Proceedings IEEE INFOCOM 2001*, volume 3, pages 1655–1663, 2001.
- [13] M. Farley and S. Carlson. A new pseudolite battlefield navigation system. In *IEEE 1998 Position Location and Navigation Symposium*, pages 208–217, New York, NY, April 1998.
- [14] M. Hata and T. Nagatsu. Mobile location using signal strength measurements in a cellular system. *IEEE Trans. Veh. Technol.*, vt-29:245–252, May 1980.
- [15] M. Hellebrandt and R. Mathar. Location tracking of mobiles in cellular radio networks. *IEEE Trans. Veh. Technol.*, pages 1558–1562, Sep. 1999.
- [16] R. A. Iltis and S. Kim. Geometric derivation of expectation-maximization and generalized successive interference cancellation algorithms with CDMA channel estimation. *IEEE Trans. Signal Processing*, 51(5):1367–1377, May 2003.
- [17] M. Joa-Ng and I.-T. Lu. A novel spread spectrum-based synchronization and location determination method for wireless system. *IEEE Commun. Lett.*, 3(6):177–179, Jun. 1999.
- [18] R. E. Kalman. A new approach to linear filtering and prediction problems. *Trans. ASME J. Basic Eng.*, 82:32–45, Mar. 1960.
- [19] S. Kim, A. P. Brown, T. Pals, R. A. Iltis, and H. Lee. Geolocation in ad hoc networks using DS-CDMA and Generalized Successive Interference Cancellation. *IEEE J. on Select. Areas Commun.*, 23(5):984–998, May 2005.
- [20] S. Kim and R. A. Iltis. A matching pursuit/GSIC-based algorithm for DS-CDMA sparse channel estimation. *IEEE Signal Processing Letters*, 11(1):12–15, Jan. 2004.
- [21] S. Kim, T. Pals, R. A. Iltis, and H. Lee. CDMA multipath channel estimation using generalized successive interference cancellation algo-

- rithm for radiolocation. In *Proc. 37th Annual Conference on Information Sciences and Systems*, Princeton, NJ, Mar. 2002.
- [22] S. J. Lee, J. Hsu, R. Hayashida, M. Gerla, and R. Bagrodia. Selecting a routing strategy for your ad hoc network. *Computer Communications*, 26(7):723–733, May 2003.
- [23] X. R. Li and P. Zhang. Optimal linear estimation fusion—part III: Cross-correlation of local estimation errors. In *Proceedings of the 2001 International Conference on Information Fusion*, pages WeB1.11–WeB1.18, Montreal, QC, Canada, Aug. 2001.
- [24] X. R. Li, Y. Zhu, J. Wang, and C. Han. Optimal linear estimation fusion—I: Unified fusion rules. *IEEE Trans. Inform. Theory*, 49(9):2192–2208, Sept. 2003.
- [25] D. D. McCrady, L. Doyle, H. Forstrom, T. Dempsey, and M. Martorana. Mobile ranging using low-accuracy clocks. *IEEE Trans. Microwave Theory Tech.*, pages 961–957, Jun. 2000.
- [26] R. L. Moses, D. Krishnamurthy, and R. M. Patterson. A self-localization method for wireless sensor networks. *EURASIP Journal on Applied Signal Processing*, 2003:348–358, March 2003.
- [27] A. G. O. Mutambara and H. F. Durrant-Whyte. Estimation and control for a modular wheeled mobile robot. *IEEE Trans. on Control Systems Tech.*, 8(1):35–46, Jan. 2000.
- [28] E. M. Nebot, M. Bozorg, and H. F. Durrant-Whyte. Decentralized architecture for asynchronous sensors. *Autonomous Robots*, 6(2):147–164, April 1999.
- [29] N. Patwari, A. O. Hero, M. Perkins, N. S. Correal, and R. J. O’Dea. Relative location estimation in wireless sensor networks. *IEEE Trans. Signal Processing*, 51(8):2137–2003, Aug. 2003.
- [30] C. Savarese, J. M. Rabaey, and J. Beutel. Locationing in distributed ad-hoc wireless sensor networks. In *IEEE Intern’l Conf. on Acoustics, Speech and Sig. Proc.*, volume 4, pages 2037–2040, Salt Lake City, UT, 2001.
- [31] A. Savvides, C. C. Han, and M. B. Srivastava. Dynamic fine-grained localization in ad-hoc wireless sensor networks. In *Proceedings of the Fifth International Conference on Mobile Computing and Networking (MobiCom)*, pages 166–179, Rome, Italy, July 2001.
- [32] A. Savvides, H. Park, and M. B. Srivastava. The bits and flops of the N-hop multilateration primitive for node localization problems. In *First ACM International Workshop on Sensor Networks and Applications*, 2002.
- [33] I. Stojmenovic. Position-based routing in ad hoc networks. *IEEE Commun. Magazine*, pages 128–134, July 2002.
- [34] R. Vadigepalli and F. J. D. III. Structural analysis of large-scale systems for distributed state estimation and control applications. *Control Engineering Practice*, 11(8):895–905, Aug. 2003.
- [35] A. Wang and A. Chandrakasan. Energy-efficient DSPs for wireless sensor networks. *IEEE Signal Processing Magazine*, pages 68–78, July 2002.
- [36] W. S. Widnall and G. F. Gobbini. Stability of the decentralized estimation in the JTIDS relative navigation. *IEEE Trans. Aerosp. Electron. Syst.*, AES-19(2):240–249, Mar. 1983.