

Multihop Progressive Decentralized Estimation in Wireless Sensor Networks

Yi Huang and Yingbo Hua, *Fellow, IEEE*

Abstract—A multihop progressive decentralized estimation scheme is presented for 1-D, 2-D, or 3-D sensor networks where routing trees are available. The bit allocation for each sensor is optimized in a multihop setting. This new scheme is shown to be much more efficient in energy consumption than the nonprogressive decentralized estimation scheme available in the literature. The transmission energy model is based on spectrum-limited applications.

Index Terms—Decentralized estimation, distributed estimation, multihop progressive scheme, sensor networks.

I. INTRODUCTION

RECENTLY, a class of decentralized estimation algorithms have been introduced in [1]–[4]. A central idea in those algorithms is as follows. Consider a sensor network where the k th sensor measures a discrete-time observation $x_k(i)$ modelled as

$$x_k(i) = \theta(i) + n_k(i) \quad (1)$$

where $\theta(i)$ is the desired signal and assumed to be deterministic spatially, and $n_k(i)$ is the measurement noise with zero mean and k -dependent variance $\sigma_{n_k}^2$. A primary purpose of the decentralized estimation algorithms is to find the number of bits B_k to quantize $x_k(i)$ into $m_k(i)$ at sensor k and time i such that the quantized data, $m_k(i)$ for all k , will consume the minimum amount of transmission energy for them to be transmitted to a fusion center and then used at the fusion center to estimate $\theta(i)$ within a given error margin. Since this process is repeated for any given i , we will drop i for notational convenience. In all of the above-referenced works, the choice for B_k is based on a primary assumption that each sensor can transmit B_k bits to a fusion center either directly or indirectly. For this reason, we will refer to those algorithms as nonprogressive decentralized estimation algorithms or simply nonprogressive scheme.

In this letter, we present a multihop progressive decentralized estimation scheme or simply called progressive scheme. In the progressive scheme, each sensor only forwards B_k bits to a neighboring sensor along a path toward a fusion center. Hence, each sensor must first perform the best estimation based on the data received from all of its immediate upstream sensors and the data collected by itself and then forward B_k bits of the best estimation to its immediate downstream sensor. For the pro-

Manuscript received March 12, 2007; revised June 7, 2007. This work was supported in part by the U.S. National Science Foundation under Grant TF-0514736. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Andreas Jakobsson.

The authors are with the Department of Electrical Engineering, University of California, Riverside, CA 92521 USA (e-mail: yhuang@ee.ucr.edu; yhua@ee.ucr.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/LSP.2007.904701

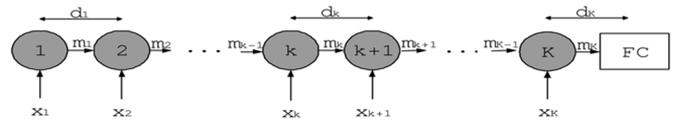


Fig. 1. Illustration of the multihop progressive decentralized estimation scheme for a 1-D network. The square on the far right is the fusion center.

gressive scheme, we take advantage of a multihop routing tree from sensors to a fusion center. Such a routing tree is feasible for networks of zero or low mobility. We develop an algorithm for computing B_k for each sensor such that the total network transmission energy is minimized subject to an error tolerance at the fusion center. The progressive scheme is shown to be much more energy efficient than the nonprogressive scheme in [1].

In Section II, the progressive scheme is presented in detail. This scheme exploits an optimization of bit allocation and will also be referred to as proposed progressive (PP) scheme. In Section III, the performance of the PP scheme is compared to that of the nonprogressive (NP) scheme in [1] as well as a uniform progressive (UP) scheme. The UP scheme uses a uniform bit allocation for each sensor. Unless specified otherwise, by progressive scheme, we imply the PP scheme.

II. PROGRESSIVE DECENTRALIZED ESTIMATION

A. The 1-D Network

Let us start with a 1-D sensor network as shown in Fig. 1, where there are total K sensors. The sensor nearest to the fusion center is indexed by $k = K$, and the sensor farthest from the fusion center is indexed by $k = 1$. The input to the k th sensor consists of x_k and m_{k-1} , where x_k is the measurement at the k th sensor as described in (1), and m_{k-1} is the data received from sensor $k-1$. The output from the k th sensor is denoted by m_k , which is quantized from the best linear unbiased estimator (BLUE) of θ at sensor k . The BLUE is traditionally known as minimum variance unbiased linear estimation, which is discussed in many textbooks on estimation theory. Let the variances of x_k and m_{k-1} be denoted by $\sigma_{x_k}^2$ and $\sigma_{m_{k-1}}^2$, respectively. Obviously, $\sigma_{x_k}^2 = \sigma_{n_k}^2$. Then, assuming $B_{k-1} > 0$, the BLUE of θ based on x_k and m_{k-1} at sensor k is

$$\hat{\theta}_k = \left(\frac{1}{\sigma_{m_{k-1}}^2} + \frac{1}{\sigma_{x_k}^2} \right)^{-1} \left(\frac{m_{k-1}}{\sigma_{m_{k-1}}^2} + \frac{x_k}{\sigma_{x_k}^2} \right) \quad (2)$$

and the variance of $\hat{\theta}_k$ is

$$\sigma_{\theta_k}^2 = \left(\frac{1}{\sigma_{m_{k-1}}^2} + \frac{1}{\sigma_{x_k}^2} \right)^{-1}. \quad (3)$$

With B_k bits for quantization of $\hat{\theta}_k$, we have m_k as the output of sensor k . Assume that $\hat{\theta}_k$ is bounded within $[-W, W]$, and

then, the variance of quantization error at sensor k is $\sigma_{q_k}^2 = (c_k W^2)/(2^{2B_k}) \leq (W^2)/(2^{2B_k})$, where $c_k \leq 1$. If the quantization error is uniformly distributed, we have $c_k = 1/3$. Since it is difficult to know the exact distribution of the quantization error of $\hat{\theta}_k$, which generally also depends on k , we will use the upper bound of $\sigma_{q_k}^2$ corresponding to $c_k = 1$. The variance of m_k is $\sigma_{m_k}^2 = \sigma_{q_k}^2 + \sigma_{\theta_k}^2$, i.e.,

$$\sigma_{m_k}^2 \leq \frac{W^2}{2^{2B_k}} + \left(\frac{1}{\sigma_{m_{k-1}}^2} + \frac{1}{\sigma_{x_k}^2} \right)^{-1} \quad (4)$$

which holds for all k , even if we set $c_k = 1$ for all k .

Note that when $\sigma_{m_k}^2$ is computed by using (4), we obtain the upper bounds of $\sigma_{m_k}^2$ for all k . Using the upper bound of $\sigma_{m_{k-1}}^2$ in (2), the corresponding value of $\hat{\theta}_k$ is an approximate (not exact) BLUE of θ . For convenience, by $\sigma_{m_k}^2$, we will refer to its upper bound computed by (4).

As shown in (2), $\sigma_{m_{k-1}}^2$ is needed for BLUE at sensor k . This means that $\sigma_{m_k}^2$ for all k should be computed in advance using (4) and made available at the corresponding sensors before the distributed estimation of θ is conducted in the network. Note that $\sigma_{m_k}^2$ used in BLUE at each sensor is obtained by (4), not by (5) shown later. The inequality (4) is tighter than the inequality (5).

To determine B_k for all k , we use an idea also used in [1], i.e., we will formulate a criterion that minimizes a measure of the network transmission energy subject to a mean-square-error (MSE) constraint. However, the details are somewhat different, as shown next.

Since (4) is a nonlinear recursion for $\sigma_{m_k}^2$, it is hard to find the exact form of $\sigma_{m,K}^2$ that is the MSE of the final estimate m_K transmitted to the fusion center. However, we can use the following inequality:

$$\sigma_{m_k}^2 \leq \frac{W^2}{2^{2B_k}} + \frac{\sigma_{m_{k-1}}^2}{4} + \frac{\sigma_{x_k}^2}{4} \quad (5)$$

which follows from (4) and $2\sigma_{m_{k-1}}\sigma_{x_k} \leq \sigma_{m_{k-1}}^2 + \sigma_{x_k}^2$. The above inequality may or may not be tight, depending on whether or not $\sigma_{m_{k-1}}^2$ is close to $\sigma_{x_k}^2$. However, this does not matter for our application. The above inequality recursion leads to

$$\begin{aligned} \text{MSE} &\doteq \sigma_{m,K}^2 \\ &\leq \sum_{k=1}^K \left(\frac{1}{4} \right)^{K-k} \frac{W^2}{2^{2B_k}} \\ &\quad + \sum_{k=2}^K \left(\frac{1}{4} \right)^{K-k+1} \sigma_{x_k}^2 \\ &\quad + \left(\frac{1}{4} \right)^{K-1} \sigma_{x_1}^2 \doteq \text{MSE}_{\text{bound}}. \end{aligned} \quad (6)$$

Assume that the communication channel between sensors has additive white Gaussian noise with power spectral density N_k , and the channel power attenuation factor is $a_k = d_k^r$, where d_k is the transmission distance from sensor k to sensor $k+1$ and r is the path loss exponent. Then, to transmit B_k bits reliably from sensor k to sensor $k+1$, the minimum required transmission energy E_k must satisfy the following, based on Shannon theory:

$$E_k = a_k N_k (2^{B_k} - 1) < a_k N_k 2^{B_k}. \quad (7)$$

The above transmission energy model is valid for spectrum-limited applications. If the spectrum is unlimited, B_k should be close to zero and the transmission energy becomes linearly proportional to B_k . In this letter, we are interested in spectrum-limited applications. One such example is an opportunistic sensor network that operates in a spectrum mostly occupied by other users. Under a practical coding and modulation scheme, the right side of (7) should be multiplied by a factor larger than one but independent of k . This factor, however, does not affect our theory on the choice of B_k .

It is therefore meaningful to set up the following criterion for determination of B_k :

$$\min_{\{B_k\}} \sum_{k=1}^K a_k^2 N_k^2 2^{2B_k} \quad (8)$$

subject to

$$\text{MSE}_{\text{bound}} \leq \text{MSE}_0 \quad (9)$$

where (8) is the L_2 -norm of an upper bound of the minimum required network transmission energy, and MSE_0 is a pre-specified tolerance of the upper bound on the MSE at the fusion center. An equivalent form of (9) is

$$\sum_{k=1}^K \left(\frac{1}{4} \right)^{K-k} \frac{1}{2^{2B_k}} \leq \eta \quad (10)$$

where

$$\begin{aligned} \eta &= \frac{1}{W^2} \\ &\times \left(\text{MSE}_0 - \sum_{k=2}^K \left(\frac{1}{4} \right)^{K-k+1} \sigma_{x_k}^2 - \left(\frac{1}{4} \right)^{K-1} \sigma_{x_1}^2 \right). \end{aligned} \quad (11)$$

Then, by applying the fact that for any real numbers x_i and y_i , $(\sum_i x_i^2)(\sum_i y_i^2) \geq (\sum_i x_i y_i)^2$ with equality when $x_i = \lambda y_i$, we have

$$\begin{aligned} \left(\sum_{k=1}^K a_k^2 N_k^2 2^{2B_k} \right) \left(\sum_{k=1}^K \left(\frac{1}{4} \right)^{K-k} \frac{1}{2^{2B_k}} \right) \\ \geq \left(\sum_{k=1}^K a_k N_k \left(\frac{1}{2} \right)^{K-k} \right)^2 \end{aligned} \quad (12)$$

with equality when

$$a_k N_k 2^{B_k} = \lambda \left(\frac{1}{2} \right)^{K-k} \frac{1}{2^{B_k}}. \quad (13)$$

Combining (10) and (12), we have

$$\left(\sum_{k=1}^K a_k^2 N_k^2 2^{2B_k} \right) \geq \frac{1}{\eta} \left(\sum_{k=1}^K a_k N_k \left(\frac{1}{2} \right)^{K-k} \right)^2. \quad (14)$$

Note that the right-hand side of (14) is independent of B_k and hence, if achieved, must be the minimum of (8). This minimum is indeed achieved if (13) holds, and (13) implies that

$$B_k = \frac{1}{2} (k - K + \log_2 \lambda - \log_2 (a_k N_k))^+ \quad (15)$$

where $(x)^+ = x$ if $x \geq 0$ and $(x)^+ = 0$ if $x < 0$. Note that we require B_k to be nonnegative.

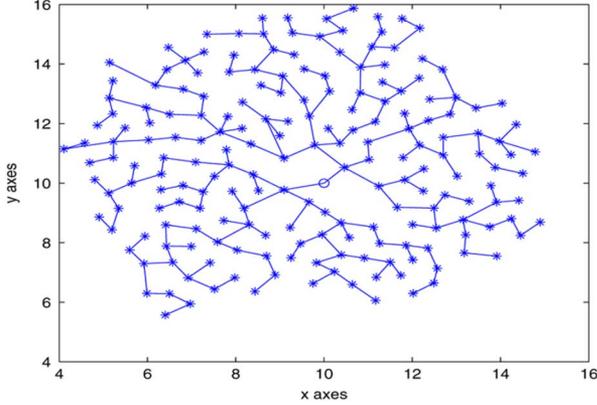


Fig. 2. Example of 2-D sensor network with a routing tree. The fusion center is denoted by the circle in the center. Here, there are 200 nodes, each denoted by *.

We define $k_0 - 1 = \max\{1 \leq k \leq K \mid B_k = 0\}$. Since all sensors indexed with $k < k_0 - 1$ are upstream sensors of sensor $k_0 - 1$, we should set $B_k = 0$ for $k < k_0 - 1$ to save unnecessary energy consumption. Now using (15) in the equality of (10) yields

$$\lambda = \frac{1}{\eta} \sum_{k=k_0}^K a_k N_k 2^{k-K}. \quad (16)$$

The computation of (15) and (16) may require several iterations before convergence, and at each iteration, the value of k_0 may be altered. After convergence, we round up each of the final values of B_k .

B. The 2-D or 3-D Networks

For a 2-D or 3-D network, a routing tree must be first established as illustrated in Fig. 2, where each branch of the tree represents a path of data flow.

Let each sensor in the network have a unique label k , where $k = 1, 2, \dots, K$. The set containing the children of sensor k is denoted by \mathcal{C}_k , and the size of \mathcal{C}_k is denoted by e_k . Sensor k first conducts the BLUE $\hat{\theta}_k$ of the signal parameter θ using its local measurement x_k and the data $\{m_l, l \in \mathcal{C}_k\}$ received from \mathcal{C}_k . Then, sensor k uses B_k bits to quantize $\hat{\theta}_k$ to generate m_k , which is to be transmitted to its parent sensor. Following the same analysis as in the 1-D case, one can verify that the variance of m_k is

$$\sigma_{m_k}^2 = \sigma_{q_k}^2 + \sigma_{\hat{\theta}_k}^2 \quad (17)$$

with $\sigma_{q_k}^2 \leq (W^2)/(2^{2B_k})$ and

$$\begin{aligned} \sigma_{\hat{\theta}_k}^2 &= \left(\frac{1}{\sigma_{x_k}^2} + \sum_{l \in \mathcal{C}_k^+} \frac{1}{\sigma_{m_l}^2} \right)^{-1} \\ &\leq \frac{1}{(1+e_k^+)^2} \sum_{l \in \mathcal{C}_k^+} \sigma_{m_l}^2 + \frac{1}{(1+e_k^+)^2} \sigma_{x_k}^2 \end{aligned} \quad (18)$$

where $\mathcal{C}_k^+ = \{1 \leq l \leq K \mid l \in \mathcal{C}_k, B_l > 0\}$, and e_k^+ is the size of \mathcal{C}_k^+ . For convenience, we can let the fusion center be indexed

by $k = K + 1$ and set $\sigma_{x_{K+1}}^2 = 0$. Then, at the fusion center, the estimation variance (denoted by MSE) is bounded as follows:

$$\text{MSE} \leq W^2 \sum_{k \in K^+} \alpha_k \frac{1}{4^{B_k}} + \beta \quad (19)$$

where $K^+ = \{1 \leq k \leq K \mid B_k > 0\}$, α_k and β are independent of the number of bits assigned to sensors. (For the 1-D network shown earlier, we have $\text{MSE} \doteq \sigma_{m_{K+1}}^2 = \sigma_{m_K}^2$. However, for 2-D networks where the fusion center receives data from more than one sensor, we have $\text{MSE} \doteq \sigma_{m_{K+1}}^2 \neq \sigma_{m_K}^2$ in general.) The exact expressions of α_k and β depend on the tree topology. To compute α_k , we can start from the fusion center, where $\alpha_{K+1} = 1$. Then, for all $1 \leq k \leq K + 1$, compute recursively $\alpha_i = \alpha_k ((1)/(1+e_k^+))^2$, where $i \in \mathcal{C}_k^+$. Once we have α_k for all $1 \leq k \leq K$, we can compute $\beta = \sum_{k \in K^+} \alpha_k ((1)/(1+e_k^+))^2 \sigma_{x_k}^2$.

Following the same technique used for the 1-D network, one can verify the following solution for a 2-D or 3-D network:

$$B_k = \frac{1}{2} (\log_2 \sqrt{\alpha_k} + \log_2 \lambda - \log_2 (a_k N_k))^{+} \quad (20)$$

where $\lambda = (1/\eta) \sum_{k \in K^+} a_k N_k \sqrt{\alpha_k}$ and $\eta = (1/W^2)(\text{MSE}_0 - \beta)$. Since $\{\alpha_k\}$ in general depends on $\{B_k\}$, the above computation may need to iterate several times until convergence. The initial selection of $\{\alpha_k\}$ assumes that $B_k > 0$ for all k . After convergence, the final values of B_k for all k are rounded up.

III. PERFORMANCE EVALUATION

We now compare the performance of the PP scheme with that of the NP scheme in [1]. For comparison, we also include a UP scheme for which a constant number of bits for each sensor is assigned and the desired $\text{MSE} = \sigma_{m_{K+1}}^2$ at the fusion center is computed by following the recursion of (17) and the first equality in (18).

We will consider a 2-D network of 200 sensors as shown in Fig. 2. This network is constructed in such a way that the distance between a parent sensor and its child sensor is $D\delta$, where δ is uniformly distributed within the range $[0.5, 1.5]$, and D is unspecified. We further assume that $W = 1$, $\sigma_{x_k}^2 = 0.05$, $r = 4$, and $N_k = 1$.

Under the constraint $\text{MSE}_0 = 0.0043$ at the fusion center, Fig. 3 compares the bit allocations by the NP, PP, and UP schemes. The figure shows the number of bits for each sensor versus the Euclidean distance (divided by D) from the sensor to the fusion center. For the UP scheme, each sensor is allocated with the same number of bits. We see that the number of bits allocated by either the PP scheme or the UP scheme for a sensor at medium or high distance is much higher than that by the NP scheme. This is because of the short transmission range for each sensor under the progressive scheme. We also see that the PP scheme allocates a much smaller number of bits for each sensor at medium or high distance than the UP scheme. This is because of the optimization used in developing the PP scheme. In fact, under the UP scheme, much of the information from sensors at medium or high distance becomes lost as the information moves closer to the fusion center. Compared to the PP scheme, the NP scheme collects too little information from sensors at medium or high distance, and the UP scheme

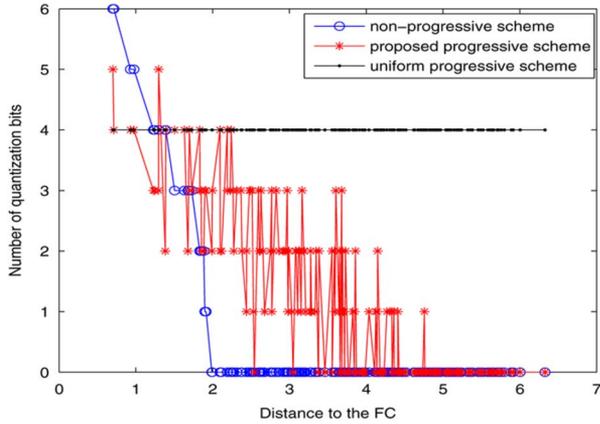


Fig. 3. Number of quantization bits allocated for each sensor versus the normalized Euclidean distance from the sensor to the fusion center. $MSE_0 = 0.0043$. The network used is Fig. 2.

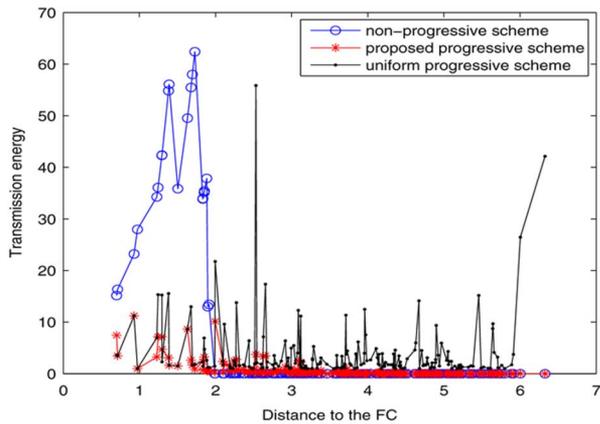


Fig. 4. Amount of normalized energy transmitted by each sensor versus the normalized Euclidean distance from the sensor to the fusion center. $MSE_0 = 0.0043$. The network used is Fig. 2.

collects too much information from sensors at medium or high distance.

Under the same constraint $MSE_0 = 0.0043$ at the fusion center, Fig. 4 shows the amount of normalized energy E_k/D^r transmitted by each sensor versus the Euclidean distance (divided by D) from the sensor to the fusion center. Here, we see that the energy consumed by the PP scheme is much less than that by the NP and UP schemes. The left equation in (7) is used to compute the transmission energy. For the NP scheme, d_k in (7) is replaced by the distance from sensor k to the fusion center. For the PP and UP schemes, d_k in (7) is the distance between sensor k and its immediate downstream sensor.

Fig. 5 shows the total normalized transmission energy consumed by the network, i.e., $\sum_{1 \leq k \leq K} E_k/D^r$, versus the target MSE_0 . The highest curve is for the NP scheme, the middle curve is for the UP scheme, and the lowest curve is for the PP scheme. The NP and UP curves cross each other at certain values of MSE_0 . We see that the PP scheme requires the least amount of energy throughout the whole MSE_0 region. At high MSE_0 , the PP scheme and the NP scheme consume approximately the same energy because almost all bits are allocated to the sensors right next to the fusion center.

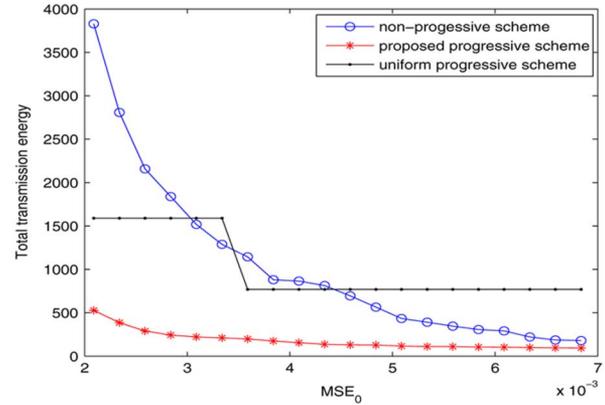


Fig. 5. Total amount of normalized transmission energy consumed by the network versus MSE_0 . The network used is Fig. 2.

IV. CONCLUSION

In this letter, we have presented a multihop progressive decentralized estimation scheme for 1-D, 2-D, or 3-D sensor networks. We have shown that this new scheme is much more energy efficient than the nonprogressive decentralized estimation scheme available in the literature. The proposed progressive scheme decentralizes the computational and communication burden to a larger degree than the nonprogressive scheme. Future research should consider more complex estimation problems than (1).

The work shown here is for both spectrum-limited and energy-limited applications. Although limited, the available spectrum has been assumed to be sufficient to avoid cochannel interference so that the energy model (7) is valid. If the available spectrum becomes tighter and the network is heavily loaded, one has to consider concurrent cochannel transmissions within the network [5]. Such scenarios may include the case where each sensor in the network needs to perform fusion of information from all other sensors. More research in this direction is needed.

A minimum distance routing tree is known as a Steiner tree. Techniques for finding a Steiner tree are available in [6]. However, the minimum distance tree may not be optimal for the purpose of estimation. Like the work in [7], the impact of the progressive scheme on the best choice of the routing tree should be another topic of interest.

REFERENCES

- [1] J. J. Xiao, S. Cui, Z. Q. Luo, and A. J. Goldsmith, "Power scheduling of universal decentralized estimation in sensor networks," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 413–422, Feb. 2007.
- [2] Z. Q. Luo, "Universal decentralized estimation in a bandwidth constrained sensor network," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2210–2219, Jun. 2005.
- [3] Z. Q. Luo, "An isotropic universal decentralized estimation scheme for a bandwidth constrained ad hoc sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 735–744, Apr. 2005.
- [4] J. J. Xiao and Z. Q. Luo, "Decentralized estimation in an inhomogeneous sensing environment," *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3564–3575, Oct. 2005.
- [5] Y. Hua, Y. Huang, and J. J. Garcia-Luna-Aceves, "Maximizing the throughput of large ad hoc wireless networks," *IEEE Signal Process. Mag.*, vol. 23, no. 5, pp. 84–94, Sep. 2006.
- [6] F. C. Harris Jr., "Steiner minimal trees: An introduction, parallel computation, and future work," in *Handbook of Combinatorial Optimization*. Norwell, MA: Kluwer, 1998.
- [7] Y. Sung, S. Mista, L. Tong, and A. Ephremides, "Signal processing for application-specific ad hoc networks," *IEEE Signal Process. Mag.*, vol. 23, no. 5, pp. 74–83, Sep. 2007.