ENERGY COST FOR ESTIMATION IN MULTIHOP WIRELESS SENSOR NETWORKS

Yi Huang and Yingbo Hua

Department of Electrical Engineering, University of California, Riverside, CA 92521.
yhuang@ee.ucr.edu, yhua@ee.ucr.edu

ABSTRACT

This paper addresses a transmission energy problem for estimation in multihop wireless sensor networks. We consider two very different schemes of estimation: progressive estimation and consensus estimation. Progressive estimation is designed for networks about which a central planner knows a routing tree and all channel state information. Consensus estimation is commonly known for more dynamic networks. In this paper, we develop energy planning algorithms to minimize the transmission energy cost for both schemes, and compare their energy costs. This study shows that the total energy cost for consensus estimation can be much higher than that for progressive estimation, but the peak energy for the former is less than that for the latter.

Index Terms— Multihop wireless sensor networks, distributed estimation, energy planning algorithms.

1. INTRODUCTION

Transmission energy cost is an important issue for wireless sensor networks. Reduction of transmission energy cost can prolong the lifetime of a wireless sensor network. While the computing energy cost can be reduced by efficient designs of microprocessors, the transmission energy cost can only be reduced primarily by efficient networking protocols. The networking protocols include medium access control, which governs how sensors exchange their information, and estimation schemes, which govern how sensors compress the information flowing between them. In this paper, we study two very different schemes of estimation for wireless sensor networks: progressive estimation and consensus estimation, and evaluate their transmission energy costs.

In a progressive estimation scheme, data packets are compressed as they hop from one sensor to another along a routing tree. A study of the progressive estimation scheme was presented in [1]. It is shown there that with progressive estimation and a fixed fusion center, multihop wireless sensor networks typically consume much less energy than single-hop wireless sensor networks. The work [1] is a multihop extension of decentralized estimation as developed in [2] and [3]. The importance of progressive estimation was also shown in [4]. In a more recent work [5], the progressive estimation scheme was further developed where analog communication and digital communication are compared. It is shown in [5] that as long as the communication spectrum is not too limited, digital communication is more energy efficient than analog communication. In this paper, we will present a generalized energy planning algorithm to further reduce the energy cost for progressive estimation.

In a consensus estimation scheme, each sensor repeatedly exchanges information with neighboring sensors and iteratively updates its own information. The consensus estimation scheme has been widely documented in the recent literature, e.g., see [6]-[7] and the references therein. Consensus estimation is particularly useful for peer-to-peer multihop networks with no central agent. However, the convergence behavior of any consensus algorithm always depends on the network topology. The transmission energy cost is directly affected by the number of iterations required. In this paper, we will present an in depth study of the transmission energy cost for consensus estimation. Although the energy planning algorithm developed for consensus estimation may not be practical for most situations where consensus estimation is intended, this study illustrates a lower bound on the energy cost for consensus estimation.

The rest of this paper is organized into sections of progressive estimation, consensus estimation, and performance evaluation. One conclusion of this study is that the sum energy required for progressive estimation under a good energy planning algorithm such as the one in this paper can be much smaller than that for consensus estimation, but the peak energy cost of individual sensors for progress estimation is generally higher than that for consensus estimation. Many details omitted due to space limitation are available in [8].

2. PROGRESSIVE ESTIMATION

2.1. System Model

For progressive estimation, we assume a multihop network with a routing tree and a fusion center (the destination node).
The data collected by each sensor can be of any dimension and any statistical properties. But we assume that each sensor can use its collected data to obtain an estimate of a desired unknown vector denoted by \( \theta(n) \in R^{M \times 1} \) at the time instant \( n \). This estimate at sensor \( k \) is denoted by \( \theta_k(n) \). We assume \( \theta_k(n) \) has mean equals to \( \theta(n) \) and its covariance matrix is denoted by \( C_{\theta,k} \) (but not required except for an upper bound on its trace). Furthermore, we assume that \( \mathbb{E}\{\theta_k(n)\theta_l(n)^T\} = 0 \) \( \forall \, k \neq l \), and the estimation of \( \theta(n) \) is done independently of \( \theta(l) \) for \( l \neq n \). All computations and communications required for estimating \( \theta(n) \) are performed within the time window \( n \). For this reason, we will drop the index \( n \) in \( \theta(n) \) hereafter.

We assume digital communication between sensors. We denote \( B_{k,m} \) the number of bits allocated for quantization of the \( m \)th element of \( \theta_k \) for each \( n \). So, for each \( n \), sensor \( k \) needs to transmit \( \sum_{m=1}^{M} B_{k,m} \) bits to its down-stream sensor. We assume that there are \( L \) communication subchannels from any sensor to its down-stream sensor within each time window. We let \( B'_{k,l} \) be the number of bits transmitted from sensor \( k \) to its down-stream sensor in subchannel \( l \). It follows that

\[
\sum_{m=1}^{M} B_{k,m} = \sum_{l=1}^{L} B'_{k,l}.
\]

The energy cost for transmitting \( B'_{k,l} \) from sensor \( k \) to its down-stream sensor in subchannel \( l \) is, as discussed in [5], given by

\[
E_{k,l} = a_{k,l} \left( 2^{B'_{k,l}/\varphi} - 1 \right)
\]

where \( a_{k,l} = \frac{N_0}{\phi \mu|h_{k,l}|^2} \), \( N_0 \) is the noise spectral density of the RF communication channel, \( h_{k,l} \) is the gain of the \( l \)th sub-channel from sensor \( k \), the parameters \( 0 < \varphi < 1, \mu > 0 \), and \( 0 < \varphi < 1 \) are associated with particular digital coding scheme, analog waveform modulation and data packet overhead, respectively.

We assume that during each time window, the leaf sensors start to transmit their quantized estimates to their down-stream sensors. As soon as sensor \( k \) receives data from all its upper-stream sensors, it computes an updated estimate \( \hat{\theta}_k \) of \( \theta \) by

\[
\hat{\theta}_k = \frac{1}{|S_k| + 1} \left( \theta_k + \sum_{i \in S_k} m_i \right)
\]

where \( S_k \) is the set of indices of all upper-stream sensors of sensor \( k \), \( |S_k| \) the size of \( S_k \), and \( m_i \) the data received from sensor \( i \). Sensor \( k \) then quantizes \( \hat{\theta}_k \) into \( m_k \), which is then transmitted to its down-stream sensor. This process continues sequentially until the fusion center receives all data from its neighbors and performs the final estimation for the time window \( n \). We assume that in the process there is no packet collision and all transmissions are carried out correctly.

With uniform probabilistic quantization [2], [5], the quantization error of the \( m \)th element of \( \theta_k \) has zero mean and a variance upper bounded by \( \sigma^2_{\theta,k,m} \geq \frac{4W_m^2}{2^{2B_{k,m}}} \).

With this progressive estimation scheme, we can show [5] that the final estimate at the fusion center which we refer to as sensor \( K \) is an unbiased estimate of \( \theta \) with the MSE given by

\[
MSE = \text{Tr}\{C_K\} \leq \xi + \sum_{k=1}^{K-1} \sum_{m=1}^{M} \alpha_k \frac{4W_m^2}{2^{2B_{k,m}}} \quad (3)
\]

where \( \xi = \sum_{k=1}^{K} \frac{\alpha_k}{|S_k|+1} \text{Tr}\{C_{\theta,k}\} \), which is invariant to \( B_{k,m} \). \( \alpha_K = 1 \), and \( \alpha_l = \frac{\alpha_K}{|S_k|+1} \) for \( l \in S_k \). The actual values of \( \alpha_k \) depend on the routing tree. If \( MSE_0 \) is the desired MSE value at the destination node, it is meaningful to set up the following constraint

\[
\sum_{k=1}^{K-1} \sum_{m=1}^{M} \alpha_k \frac{4W_m^2}{2^{2B_{k,m}}} \leq \eta \quad (4)
\]

where \( \eta = \frac{1}{4}(MSE_0 - \xi) \).

Given the above preparations, we can now formulate the following optimization problem to determine \( B_{k,m} \) and \( B_{k,l} \) for all \( k, m \) and \( l \).

\[
\begin{align*}
\min_{B_{k,m},B'_{k,l}} & \quad J_p = \sum_{k=1}^{K-1} \sum_{l=1}^{L} a_{k,l}^{p} \left( 2^{B'_{k,l}/\varphi} - 1 \right)^p \\
\text{subject to} \quad & \quad (4) \\
& \quad \sum_{m=1}^{M} B_{k,m} = \sum_{l=1}^{L} B'_{k,l} \quad \text{for all} \quad k \\
& \quad B_{k,m} \geq 0, \quad B'_{k,l} \geq 0 \quad \text{for all} \quad k, m, \text{and} \ l
\end{align*}
\]

where \( J_p^{1/p} \) is the \( p \)th norm of all components of the energy cost in the network. This energy planning problem does not need the approximation and constraint used in [5] and hence is more general.

### 2.2. Energy Planning for Progressive Estimation

We solve the energy planning problem by solving its Karush-Kuhn-Tucker (KKT) equations. Due to the limited space, we omit the derivation (available in [8]) but just give the solution below

\[
B_{k,m} = \left( \frac{1}{2} \log_2 \frac{\alpha_k W_m^2 |M_k^+|}{\eta_k - \sum_{l \notin M_k^+} \alpha_l W_m^2} \right)^{+} \quad (8)
\]

Here, \( (x)^{+} = \max(x, 0) \), \( |M_k^+| = \{ m | B_{k,m} > 0 \} \) and \( |M_k^-| \) is the size of the set \( M_k^- \).

\[
B'_{k,l} = [f_l(\xi_k)]^{+} \quad (9)
\]

where \( \xi_k \) is given by \( \sum_{l=1}^{L} [f_l(\xi_k)]^{+} = B_k \triangleq \sum_{m=1}^{M} B_{k,m} \), and \( f_l(\xi_k) \) is the value of \( B'_{k,l} \) from

\[
\alpha_k^{p} \ln \frac{2}{\varphi} p \left( 2^{B'_{k,l}/\varphi} - 1 \right)^{p-1} 2^{B'_{k,l}/\varphi} = -\xi_k
\]
with given $\xi_k$. The optimal $\eta_k$ is given by
\[
\eta_k = \eta' \frac{\xi_k |M^+_k|}{\sum_{j=1}^{K-1} \xi_j |M^+_j|} + \sum_{m \notin M^+_k} \alpha_k W^2_m \tag{10}
\]
where $\eta' = \eta - \sum_{j=1}^{K-1} \sum_{m \notin M^+_j} \alpha_j W^2_m$.

Based on (8), (9) and (10), an iterative algorithm can be shown as follows: 1) initialize all $\eta_i$; 2) update $B_{k,m}$ by computing (8) via finite iterations; 3) update $B_{k,l}$ by solving (9) via bisection search for each $k$; 4) determine a new set of $\eta_i$ by $\eta_i(s) = \beta \hat{\eta}_i(s) + (1 - \beta) \eta_i(s - 1)$ where $s$ is the outer loop iteration index, $1 > \beta > 0$ and $\hat{\eta}_i(s)$ is computed from (10); 5) go back to step 2) until convergence.

3. CONSENSUS ESTIMATION

3.1. System Model

The data model for $\theta_k(n)$ at each sensor is the same as in the previous section, and in particular $E\{\theta_k(n)\} = \theta(n)$. Consensus estimation does not require a routing tree but it requires iterative in-network computations within each time window $n$. At iteration time $t$, the estimate at the $k$th sensor of the unknown vector $\theta(n)$ is denoted by $\theta_k(n,t)$. For convenience, we will write $\theta_k(n,t)$ simply as $\theta_k(t)$. It is important to stress that the iteration time $t$ is different from the sampling time $n$. Typically, from the sampling time $n$ to the sampling time $n+1$, many iterations are required by a consensus algorithm.

The consensus algorithm that we consider is:
\[
\theta_k(t+1) = W_{k,k} \theta_k(t) + \sum_{j \in N_k} W_{k,j} m_j(t) \tag{11}
\]
where $N_k$ is the set of the indices of the neighboring transmitting sensors from which sensor $k$ receives a quantized vector, $m_j(t)$ is a quantized version of $\theta_j(t)$ from sensor $j$, and $W_{k,j}$ are weights.

We use $B_m$ bits for the $m$th element in quantization of $\theta_j(t)$. Therefore, for each iteration, each sensor first broadcasts a data packet containing $\sum_{m=1}^{M} B_m$ bits to its neighbors. For convenience, we assume the each sensor uses the same amount of energy $E_0$ for each broadcast, and there is no data packet collision (or otherwise more energy is needed). Assuming $L$ subchannels for each pair of transmitting and receiving sensors, sensor $j$ can successfully receive the data packets broadcasted from sensor $k$ if the following holds:
\[
\sum_{l=1}^{L} \psi \log_2 \left( 1 + \frac{E_0 |h_{j,k,l}|^2}{LN_0} \right) \geq \sum_{m=1}^{M} B_m \tag{12}
\]
where $h_{j,k,l}$ is the fading factor of the subchannel $l$ from sensor $k$ to sensor $j$. The other parameters are the same as those described for (1). We assume that $|h_{j,k,l}| = |h_{k,j,l}|$, i.e., all links are symmetric.

The total energy consumed by the network can be written as $E_{\text{total}} = TK E_0$ where $K$ is the total number of sensors in the network and $T$ is the number of iterations required until either convergence or a required performance is achieved.

One common choice of $W$ is $W = I - \gamma L$ where
\[
(L)_{j,k} = \begin{cases} -1, & j \in N_k \\ |N_k|, & k = j \\ 0, & \text{otherwise} \end{cases} \tag{13}
\]
and $0 < \gamma < \frac{2}{\lambda_1(L)}$, $\lambda_1(L)$ is the largest eigenvalue of $L$. With this $W$, we can show [8] that the mean of $\theta(t) = [\theta_1(t)^T, \theta_2(t)^T, \cdots, \theta_K(t)^T]^T$ is given by $E\{\theta(t)\} = 1 \otimes \theta$, and the MSE of $\theta(t)$ is given by
\[
\text{MSE}(t) = T \text{Tr} \{ D_c W^{2t} \} + \gamma^2 T \text{Tr} \left\{ \sum_{i=0}^{t-1} W^{2i} (L - L_D)^2 \right\} \text{Tr} \{ C_q \} \tag{14}
\]
where $D_c$ is a diagonal matrix whose $(k,k)$-th entry is $\text{Tr} \{ C_{\theta,k} \}$, and $L_D$ is the diagonal part of $L$.

3.2. Energy Planning for Consensus Estimation

To simplify the problem, we will assume that $B_m = B$ for all $m$. For the fusion rule (11) subject to $W = I - \gamma L$, the total cost of energy can be minimized over the variables $E_0$, $B$ and $\gamma$. In other words, we need to solve the following problem:
\[
\min_{t, E_0, B, \gamma} : t K E_0 \tag{15}
\]
\[
s.t. \quad \text{MSE}(t)/K \leq \text{MSE}_0 \tag{16}
\]
where $\text{MSE}_0$ is the (averaged) target MSE per sensor, and the solution of $t$ here is the number of iterations required. We propose a two-loop search algorithm to this problem. In the inner loop, we fix $E_0$ and $B$, and look for the optimal $\gamma$ and the associated $t$. In the outer loop, we search for the optimal $E_0$ and $B$. For each loop, there are only two variables, and the brute force search is feasible. However, for the inner loop, a refined algorithm is outlined next.

With a given pair of $E_0$ and $B$, we can find $N_k$ for all $k$, and hence $L$. The problem of the inner loop search can be formulated as:
\[
\min_{\gamma, t} : t \tag{17}
\]
\[
s.t. \quad \text{MSE}(t)/K \leq \text{MSE}_0, \quad 0 < \gamma < \frac{2}{\lambda_1(L)} \tag{18}
\]
This problem can be solved by the KKT method (see [8] for details).

4. PERFORMANCE EVALUATION

We consider a network of $K = 400$ nodes. The routing tree is generated by following the minimal distance principle. The
distance between a sensor and its upper-stream sensor is \( D\delta \) where \( \delta \) is uniformly distributed within the range \([0.5, 1.5]\) and \( D \) is a normalizing factor.

For the simulation, we assume that sensor \( k \) observes the data vector \( x_k = G_k\theta + \omega_k \), where \( \theta \) is a deterministic unknown vector, \( G_k \) is the observation matrix, and \( \omega_k \) is white noise with identical covariance \( C_{\omega} = I \). With this observation model, an original local estimate of \( \theta \) at sensor \( k \) can be obtained by the best linear unbiased estimate: \( \theta_k = (G_k^H G_k)^{-1} G_k^H x_k \). We also have \( C_{\theta,k} = (G_k^H G_k)^{-1} \). Each \( G_k \) is a \( N \times M \) real matrix with elements randomly chosen from a Gaussian distribution with zero mean and standard deviation equals to 10.

The squared channel gain for each link in a routing tree is chosen to be \( |h_{k,l}|^2 = d_k^{-\alpha} \rho_{k,l} \) where \( d_k \) is the distance from sensor \( i \) to its down-stream sensor, \( \alpha = 4 \) and \( \rho_{k,l} \) is randomly chosen from an exponential distribution with mean equal to one. The same model is applied to \( |h_{j,k,l}|^2 = |h_{j,k,l}|^2 \) for the link between sensor \( k \) and sensor \( j \). We also choose \( N = 20 \), \( M = 10 \), \( L = 10 \), \( \mu = 1 \), \( \phi = 1 \), \( \varphi = 1 \), \( N_0 = 1 \), \( W_m = 1 \) for all \( m \).

For convenience of reference, we will refer to the algorithm proposed in Section 2.2 as the “generalized” algorithm, and the algorithm shown in [5] as the “previous” algorithm. We also have a “uniform” algorithm for which the same number of bits is assigned to each element of the estimate at each sensor, and the same number of bits is assigned to each sub-channel.

Fig. 1 illustrates the total energy cost \( J_1/D^\alpha \) versus the same target \( MSE_0 \) for both progressive and consensus estimation. We see that progressive estimation with the generalized algorithm is much more energy efficient than consensus estimation, especially when the target MSE is small. The curve for the previous algorithm is about two times higher than the curve for the generalized algorithm with \( p = 1 \). In this figure, several choices of \( p \) are shown for the generalized algorithm for progressive estimation. With a larger \( p \), the sum energy required by the generalized algorithm is slightly increased as expected. Under the uniform algorithm, progressive estimation can cost more sum energy than consensus estimation. However, as shown in Fig. 2, the peak energy required by progressive estimation (which always occurs near the fusion center) is generally larger than that by consensus estimation.

5. REFERENCES