ORDERING FOR ENERGY EFFICIENT ESTIMATION AND OPTIMIZATION IN SENSOR NETWORKS

Rick S. Blum

ECE Department, Lehigh University
19 Memorial Drive West, Bethlehem, PA 18015
rblum@ece.lehigh.edu

ABSTRACT

A discretized version of a continuous optimization problem is considered for the case where data is obtained from a set of dispersed sensor nodes and the overall metric is a sum of individual metrics computed at each sensor. An example of such a problem is maximum likelihood estimation based on statistically independent sensor observations. By ordering transmissions from the sensor nodes, a method for achieving a saving in the average number of sensor transmissions is described. While the average number of sensor transmissions is reduced, the approach always yields the same solution as the optimum approach where all sensor transmissions occur. Further, for cases with $N$ sufficiently well designed sensors with sufficiently large signal-to-interference-plus-noise ratios, the average percentage of transmissions saved approaches 100 percent as the number of discrete grid points in the optimization problem $Q$ becomes significantly large. In these same cases, the average percentage of transmissions saved approaches $(Q-1)/Q \times 100$ percent as the number of sensors $N$ in the network becomes significantly large.

Index Terms— Sensor networks, optimization, estimation, energy efficient, ordering, maximum likelihood estimation.

1. INTRODUCTION

Networks of sensors are often used to solve estimation and optimization problems and these topics have been studied for many years, see for example [1, 2]. More recently, there has been great interest in energy efficient estimation in sensor networks, since these sensors often carry their own limited power sources. The issues have been nicely discussed in much of the recent literature, see for example [3]. Much of this pervious work [4]-[8] has focused on either optimizing estimation performance for a constraint on energy used or on minimizing energy used for a constraint on estimation performance. Some special emphasis has been paid to quantization, see for example [3], to multihop networks, see for example [9], and to compressive sensing-type approaches, see for example [10]. In the previous work which focused on energy efficiency, energy is saved at the expense of performance. Here we present a new method to save transmissions, and thus energy, without any performance loss.

Consider a sensor system with $N$ sensors, where the $k^{th}$ sensor calculates a metric $L_k(\theta)$, for example log-likelihood, and the overall optimization, or estimation of the vector $\theta$, is

$$\hat{\theta} = \arg \max_\theta \sum_{k=1}^N L_k(\theta)$$

(1)

where we assume a unique optimum value exists. An example of such a problem is maximum likelihood (ML) estimation based on statistically independent sensor observations. Now discretize the estimation problem so that we want to select $\theta$ from a discrete set consisting of \{\theta_i, i = 1, \ldots, Q\}, Thus selecting $\theta_i$ leads to a sensor metric of $L_k(\theta_i)$. To compute the best choice of $i$, thus the best $\theta_i$, one generally needs to have at a central location, the fusion center, all $Q$ values of $L_k(\theta_i), i = 1, \ldots, Q$ from all sensors $k = 1, \ldots, N$. This would, in general, require about $NQ$ transmissions. In Section 2, we use ordering [11] to reduce the number of transmissions. We show that the average number of transmissions can be reduced while always determining the same optimum solution as if all transmissions were made. For sufficiently well designed systems, with noise free sensor metrics (log-likelihood) which have a peak at the optimum value which is measurably (more than a factor of 3 for positive peak log-likelihood) larger than other peaks, we show in Section 3 that the savings will be very large in cases where the signal whose parameters are estimated is sufficiently large compared to the noise-plus-interference. Under these conditions, we show that the average percentage of transmissions saved approaches 100 percent as the number of grid points $Q$ in the optimization problem becomes significantly large and approaches $(Q-1)/Q \times 100$ percent as the number of sensors $N$ in the network becomes significantly large. Similar gains, which are sometimes even larger, are demonstrated for cases
with some sensors which are of poor quality.

2. ORDERING FOR ESTIMATION AND OPTIMIZATION

Let the $k^{th}$ sensor, $k = 1, \ldots, N$, transmit $L_k(\theta_i)$ after a time inversely proportional to $|L_k(\theta_i)|$. Thus the first transmission corresponds to the largest $|L_k(\theta_i)|$, over all $k$ and $\theta_i$. The next transmission corresponds to the next largest and so on. At any given time, let $|L_k(\theta_i)|$ correspond to the magnitude of the metric corresponding to the last transmission. Let $N_i$ correspond to the set of sensor indices, from $k = 1, \ldots, N$, corresponding to sensors who have transmitted a metric corresponding to $\theta_i$ so far. Thus these sensors have transmitted $L_k(\theta_i)$ for some $k$. In fact $N_i$ is the set of all such $k$. Further, let $n_i$ be the size of the set $N_i$. The fusion center accumulates $\sum_{k \in N_i} L_k(\theta_i)$ for each $i = 1, \ldots, Q$.

Let’s say the largest $\sum_{k \in N_i} L_k(\theta_i)$ for $i = 1, \ldots, Q$ occurs for $i = i^*$. Let $i = \hat{i}^{**}$ denote the next largest. $n_i, n_{\hat{i}}$ be the smallest value of $n_i$ for all possible $i$, thus $i = 1, \ldots, Q$.

Next we give a Theorem that allows us to save energy by avoiding transmissions without any loss in performance, thus the same $\hat{\theta}_i$ is chosen as if all $NQ$ transmissions took place.

**Theorem 1.** If the stopping condition $\sum_{k \in N_i} L_k(\theta_i) > \sum_{k \in N_\hat{i}} L_k(\theta_{\hat{i}}) + (N-n_{\hat{i},min})|L_k(\theta_{\hat{i}})|$ is reached then we can stop all sensor transmissions and choose the estimate $\hat{\theta}_i$, while being sure that we have the optimum discounted value of the estimate. In particular we can be sure that

$$\hat{\theta}_i = \arg \max_{\theta_i \in \{\theta_1, \ldots, \theta_Q\}} \sum_{k=1}^{N} L_k(\theta_i)$$

while employing a smaller number of sensor transmissions than $NQ$. If the stopping condition is never reached we keep transmitting until all $NQ$ transmissions have occurred. This approach will employ a smaller average number of sensor transmissions than $NQ$.

**Proof.** Since we have ordered the transmissions, we know that all the metrics (log-likelihoods) that have not yet been transmitted have a magnitude that is strictly smaller than the last one transmitted: $|L_k(\theta_{\hat{i}})|$. Thus we can bound the sum of the metrics for any $\theta_i$ using

$$\sum_{k \in N_i} L_k(\theta_i) - (N-n_i)|L_k(\theta_{\hat{i}})| < \sum_{k=1}^{N} L_k(\theta_i)$$

$$< \sum_{k \in N_i} L_k(\theta_i) + (N-n_i)|L_k(\theta_{\hat{i}})|.$$  

Thus for $i = i^*$ we can write

$$\sum_{k=1}^{N} L_k(\theta_{\hat{i}}) > \sum_{k \in N_i} L_k(\theta_i) + (N-n_{i,\min})|L_k(\theta_i)|.$$  

Further, for $i = i^{**}$ we can write

$$\sum_{k=1}^{N} L_k(\theta_{\hat{i}}) < \sum_{k \in N_i} L_k(\theta_i) + (N-n_{i,\min})|L_k(\theta_i)|.$$  

Thus, if the right-hand-side of (5) is larger than the right-hand-side of (6) then we know (from (5) and (6)) that

$$\sum_{k=1}^{N} L_k(\theta_{\hat{i}}) > \sum_{k=1}^{N} L_k(\theta_{i^{**}}).$$

By definition $\sum_{k \in N_i} L_k(\theta_i) > \sum_{k \in N_{\hat{i}}} L_k(\theta_{\hat{i}})$ for any $i$ other than $i = i^*$ or $i = i^{**}$. Thus, for any $i$ other than $i = i^*$ or $i = i^{**}$ we can write, using (4),

$$\sum_{k=1}^{N} L_k(\theta_{i}) < \sum_{k \in N_i} L_k(\theta_i) + (N-n_{i,\min})|L_k(\theta_i)| < \sum_{k \in N_i} L_k(\theta_{\hat{i}}) + (N-n_{i,\min})|L_k(\theta_{\hat{i}})|.$$  

Thus from (6), (7), and (8), when the right-hand-side of (5) is larger than the right-hand-side of (6), then for any $i$ other than $i = i^*$ we know that

$$\sum_{k=1}^{N} L_k(\theta_{\hat{i}}) > \sum_{k=1}^{N} L_k(\theta_{i}).$$

Thus, if the right-hand-side of (5) is larger than the right-hand-side of (6), then we know $\hat{\theta}_i$ solves the problem posed in (2). Note that the right-hand-side of (5) being larger than the right-hand-side of (6) is equivalent to $\sum_{k \in N_i} L_k(\theta_{i}) > \sum_{k \in N_{\hat{i}}} L_k(\theta_{\hat{i}}) + 2(N-n_{i,\min})|L_k(\theta_{\hat{i}})|$ by direct computation. It is easy to see that the event that we save transmissions occurs with nonzero probability.

The transmissions can be halted by the sensor nodes themselves if they hear all transmissions and perform the needed computations. Alternatively, the fusion center can tell the sensor nodes to stop transmitting. The approach given in Theorem 1 is pretty simple to implement. We can save even more transmissions by employing more complicated approaches. We give one such approach in the next Theorem.
**Theorem 2.** If the stopping condition \( \sum_{k \in N_{i^*}} L_k(\theta_{i^*}) > \max_{i \neq i^*} \left( \sum_{k \in N_i} L_k(\theta_i) + (N - n_i + N - n_i)|L_k(\theta_i)| \right) \) is reached then we can stop all sensor transmissions and choose the estimate \( \theta_{i^*} \) while being sure that we have the optimum discretized value of the estimate. If this condition is reached, we can be sure that

\[
\theta_{i^*} = \arg \max_{\theta \in \{\theta_1, \ldots, \theta_P\}} \sum_{k=1}^{N} L_k(\theta)
\]  

while employing a smaller number of sensor transmissions than \( NQ \). If the stopping condition is never reached we keep transmitting until all \( NQ \) transmissions have occurred. This approach will employ a smaller average number of sensor transmissions than \( NQ \). Further this approach will generally save more transmissions, on average, than the approach in Theorem 1 and will in all cases save at least as many transmissions on average.

**Proof.** The proof follows closely the proof given in Theorem 1. Thus (3) applies for all \( i \), including \( i = i^* \).

Thus for \( i = i^* \) we can write

\[
\sum_{k=1}^{N} L_k(\theta_{i^*}) > \sum_{\{k \in N_{i^*}\}} L_k(\theta_{i^*}) - (N - n_{i^*})|L_k(\theta_{i^*})|. \tag{11}
\]

Further, for any \( i \neq i^* \) we can write

\[
\sum_{k=1}^{N} L_k(\theta_i) < \sum_{\{k \in N_i\}} L_k(\theta_i) + (N - n_i)|L_k(\theta_i)|. \tag{12}
\]

Thus, if the right-hand-side of (11) is larger than the right-hand-side of (12) for all \( i \) then we know that

\[
\sum_{k=1}^{N} L_k(\theta_{i^*}) > \sum_{k=1}^{N} L_k(\theta_i).
\tag{13}
\]

for any \( i \neq i^* \).

\[\square\]

### 3. CONDITIONS FOR LARGE SAVINGS

Consider a general optimization problem of the form of (1) where we seek to optimize some metric, as given in (1), where data is obtained from a set of dispersed sensor nodes and the overall metric is a sum of individual metrics computed at each sensor. Let us assume we have \( NQ > 1 \) good quality sensors, called sufficiently well designed sensors, in the system and \( N \) and \( NQ \) poor quality sensors in the system where \( NQ \) is an integer. This includes the case where all the sensors are good, \( P = 1 \) but includes other cases also. Poor quality sensors can be in poor positions, sensors of inferior quality, or faulty. Interestingly, if some of the sensors are poor quality, giving small magnitude returns for all \( \theta \), then the savings can be even larger in some cases. These ideas are formalized in the following theorem, after some more careful definitions.

Let \( \Delta = \min(\delta_1, \ldots, \delta_{NQ}) > 0 \). We formally define a sufficiently well designed sensor \( k = 1, \ldots, N \) as one where

\[
\lim_{\text{SINR} \to \infty} L_k(\theta) = \begin{cases} \delta_k, \text{if } \theta = \theta_0 \text{ (typically the optimum)} \\ \epsilon(\theta) < \mathcal{P}\left(\frac{\Delta}{3}\right), \text{ otherwise} \end{cases} \tag{14}
\]

The previous definition and discussion leads to the following Theorem.

**Theorem 3.** Let \( NQ \) denote a positive integer. Adopt the definition of a well designed sensor from (14) with \( \Delta = \min(\delta_1, \ldots, \delta_{NQ}) > 0 \). Consider a system, with \( NQ \) sufficiently well designed sensors numbered \( k = 1, \ldots, N \) and \( N - NQ \) poorly designed sensors such that for \( k = N + 1, \ldots, N \),

\[
\lim_{\text{SINR} \to \infty} L_k(\theta) < \mathcal{P}\left(\frac{\Delta}{3}\right) \quad \forall \theta,
\]

attempting to solve an optimization problem with a unique solution of the form of (1), where one seeks to optimize some metric, as given in (1), where data is obtained from a set of dispersed sensor nodes and the overall metric is a sum of individual metrics computed at each sensor. Let \( Q > 1 \) denote the number of grid points in the discrete optimization and \( N > 1 \) denote the total number of sensors in the network. Assume the true optimum estimate falls on the grid. For sufficiently high SINR, the percentage of transmissions saved from using the approach in Theorem 1 compared to the case where all \( N \) sensors transmit all \( Q \) metric values, is strictly larger than \( \frac{Q - P}{QN} \times 100\% \) and approaches 100\% as \( Q \to \infty \) when \( N \) is held constant. For sufficiently high SINR, the percentage of transmissions saved approaches \( \frac{Q - P}{QN} \times 100\% \) as \( N \to \infty \) when \( Q \) is held constant. The savings from the approach in Theorem 2 must be as large or larger.

**Proof.** Since we assume \( N \) sufficiently well designed sensors, then for sufficiently high SINR, the first \( NQ \) transmissions in the ordering algorithm in Theorem 1 will involve transmitted log-likelihoods of \( \delta_1 \) through \( \delta_{NQ} \) from (14). Recall that \( \Delta > 0 \). Thus, the rest of the untransmitted log-likelihoods will have values smaller than \( \frac{\Delta}{3} \) from (14). Then after the \( NQ + 1 \)th transmission, the comparison to see if transmissions will be halted in Theorem 1 becomes successful since

\[
\sum_{k \in N_{i^*}} L_k(\theta_{i^*}) > NQ \Delta > \sum_{k \in N_{i^*}} L_k(\theta_{i^*}) + 2(N - n_{i^*})|L_k(\theta_{i^*})| \quad \text{since} \quad \sum_{k \in N_{i^*}} L_k(\theta_{i^*}) + 2(N - n_{i^*})|L_k(\theta_{i^*})| < \mathcal{P}\left(\frac{\Delta}{3}\right) + 2NQ \frac{\Delta}{3} < NQ \Delta \quad \text{(for all } N > 1 \text{)}
\]

so we can stop all sensor transmissions at this point. Thus we only need \( NQ + 1 \) transmissions, instead of \( QN \) transmissions.

---

2The notion of SINR can be generalized using the idea of distance measures, see [11].
Assuming $Q, N > 1$ (typically $Q, N >> 1$), the savings is $QN - (NP + 1) = N(Q - P) - 1$ and the percentage savings is

$$S = \left(\frac{(Q - P)N - 1}{QN}\right) \times 100\%$$

$$\rightarrow 100\% \text{ as } Q \rightarrow \infty$$

(16)

and

$$S = \left(\frac{(Q - P)N - 1}{QN}\right) \times 100\%$$

$$\rightarrow \frac{Q - P}{Q} \times 100\% \text{ as } N \rightarrow \infty.$$ 

(17)

The savings from Theorem 2 will be as large or larger as stated in Theorem 2.

The results in (16) and (17) imply large savings if $Q$ or $N$ is sufficiently large, as they might be made to obtain good resolution or accuracy in the estimations. Even for small $Q >> 1$ and $N >> 1$ the savings are pretty large. For $P = 1$ and $Q = N = 10$ we get $S = 89\%$. Larger $Q$ or $N$ implies larger savings since $S$ is monotonic decreasing in $Q$ and $N$ (for finite $Q, N$). Smaller $P < 1$ also implies larger savings, for finite $Q$, as $S$ is monotonic decreasing in $P$. It should be noted that Theorem 3 shows that large savings are achieved in some specific cases, but even larger savings are possible in some cases. For example, the factor $\frac{Q - P}{Q}$ in (14) can be replaced with another factor, that could depend on $N$ to achieve larger savings. It is also worth noting that we do not need to know $NP$ or $\delta_1, \ldots, \delta_N$ to achieve the savings, while they are useful in predicting the savings.

4. CONCLUSIONS

Here we study a discretized optimization problem where data is obtained from a set of dispersed sensor nodes and the overall metric is a sum of individual metrics computed at each sensor. By ordering transmissions from the sensor nodes, a method for reducing the number of sensor transmissions is described. While the average number of sensor transmissions is reduced, the approach always yields the same solution as the optimum approach where all sensor transmissions occur. Further, saving are significant for cases with some sufficiently well designed sensors with sufficiently large signal-to-interference-plus-noise ratio when the number of discrete grid points in the optimization problem $Q$ is not small.

5. ACKNOWLEDGEMENTS

The author would like to acknowledge the help of Qian He in providing excellent comments that helped guide the work.

6. REFERENCES


