PROGRESSIVE ESTIMATION AND DETECTION

Yingbo Hua and Yi Huang

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

ABSTRACT

For large-scale multi-hop wireless sensor networks, progressive estimation and detection algorithms are developed. For progressive estimation, the classical principles of best linear unbiased estimation (BLUE) and linear minimum mean square error (LMMSE) estimation are applied. The difference and similarity between the two are revealed. For progressive detection, a tradeoff between detection performance and limited decision propagation is studied. The progressive estimation and detection algorithms are distributed and performed at the application layer of wireless sensor networks, which are fully adaptive to the operations at the networking and other lower layers which can be designed for high spectral efficiency.

Index Terms— Progressive estimation and detection, distributed estimation and detection, wireless sensor networks, large-scale multi-hop networks

1. INTRODUCTION

The ultimate engineering purpose of wireless sensor networks is to estimate and detect phenomena of interest. A key feature of wireless sensor networks is its spatially distributed sensors which need to perform not only sensing but also signal processing and communications. The performance of wireless sensor networks is measured not only by the accuracy of estimation and detection but also by the cost associated with it. While the cost of many resources such as electronic computing devices will most likely continue to decline as the technology advances, the radio spectrum available to wireless sensor networks is constrained by the law of physics, which tends to remain constant. On the other hand, the fundamental theory of estimation and detection points out that a good performance of estimation and detection using wireless sensor networks often needs to be achieved by utilizing the data gathered from many spatially distributed sensors. Therefore, a large number of sensors can be involved in sharing information with each other, and the spectral efficiency of the communications between these sensors should be kept high due to limited spectrum. It has been established that for large networks, the best spectral efficiency of communications between all nodes is achieved by data transmissions only between neighboring nodes [1], [2], [3]. This means that for "flat" wireless sensor networks (where there is no mobile data collector), all large volumes of communications should be done via multi-hop routing. Also note that a high spectral efficiency can also translate into a high efficiency in transmission power.

Much of the past research on distributed estimation and detection does not take into account the multi-hop nature of wireless sensor networks, e.g., see [4], [5], [6]. In this paper, we present our effort in developing the theory of progressive estimation and detection for wireless sensor networks. The concept of progressive estimation and detection shown in this paper differs from [7], extends from [8], and has the following features: 1) The routing paths established by the networking layer is exploited at the application layer where estimation and detection are performed; 2) All local information at each sensor is exploited locally and optimally. 3) The result of local estimation and detection is relayed and progressively improved from sensor to sensor. 4) Each sensor can terminate the process of estimation and detection when a pre-specified performance criterion is satisfied.

In Section 2, we present the progressive estimation algorithms based on BLUE and LMMSE, where an interesting similarity will be observed. In Section 3, we present the progressive detection algorithms using either estimation propagation or decision propagation, the latter of which further reduces the communication burden. In Section 4, we illustrate the performance of progressive estimation and detection. Concluding remarks are given in Section 5.

2. PROGRESSIVE ESTIMATION

Consider the following model for the data observed at sensor k:

$$\mathbf{x}_k = \mathbf{G}_k\mathbf{u} + \mathbf{n}_k \tag{1}$$

where $\mathbf{x}_k$ is the observation data vector, $\mathbf{n}_k$ is the observation noise vector of zero mean, $\mathbf{G}_k$ is the observation matrix of full column rank, and $\mathbf{u}$ is an unknown vector parameter invariant to the space (or sensor) index k. The observation matrix $\mathbf{G}_k$ depends on the characteristics of sensor k and its location with respect to the unknown vector $\mathbf{u}$. Since $\mathbf{G}_k$ is arbitrary, a simple normalization of (1) can ensure that the covariance matrix of each noise vector is the identify matrix $\mathbf{I}$. We assume that only sensor k knows $\mathbf{G}_k$.

We next show two algorithms for progressive estimation: one is based on the best (minimum variance) linear unbiased estimation (BLUE) and the other is based on the linear minimum mean square error (LMMSE) estimation.

2.1. BLUE

We denote an unbiased estimator of $\mathbf{u}$ at sensor k by $\hat{\mathbf{u}}_k$ and its covariance matrix by $\mathbf{C}_k$. We let sensor $k-1$ be the upper stream

This paper was prepared for the First SenSIP Workshop, May 11-14, 2008, Sedona, Arizona, USA. This work was supported in part by NSF under Grant No. TF-0514736, and ARL CTA Program.

sensor to sensor k as illustrated in Figure 1. We assume that sensor $k-1$ transmits $\hat{\mathbf{u}}_{k-1}$ and $\mathbf{C}_{k-1}$ to sensor k before sensor k computes $\hat{\mathbf{u}}_k$ and $\mathbf{C}_k$.

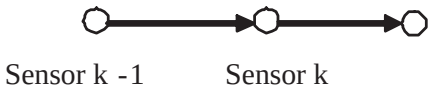

Fig. 1. Portion of a routing path where each sensor has only one upper stream sensor.

Assume that sensor 1 does not have any upper stream sensor, then the BLUE estimate of $\mathbf{u}$ at sensor 1 is given by

$$\hat{\mathbf{u}}_1 = \left(\mathbf{G}_1^H \mathbf{G}_1\right)^{-1} \mathbf{G}_1^H \mathbf{x}_1 \tag{2}$$

and its covariance matrix is

$$\mathbf{C}_1 = \left(\mathbf{G}_1^H \mathbf{G}_1\right)^{-1} \tag{3}$$

Provided that $\mathbf{x}_k$ and $\hat{\mathbf{u}}_{k-1}$ are uncorrelated (which is guaranteed if the noises $\mathbf{n}_k$ and $\mathbf{n}_l$ are uncorrelated for $k \neq l$), the BLUE estimate of $\mathbf{u}$ at sensor k based on $\mathbf{x}_k$ and $\hat{\mathbf{u}}_{k-1}$ is given by

$$\hat{\mathbf{u}}_k = \left(\mathbf{C}_{k-1}^{-1} + \mathbf{G}_k^H \mathbf{G}_k\right)^{-1} \left(\mathbf{C}_{k-1}^{-1} \hat{\mathbf{u}}_{k-1} + \mathbf{G}_k^H \mathbf{x}_k\right) \tag{4}$$

and the inverse of its covariance matrix is

$$\mathbf{C}_k^{-1} = \mathbf{C}_{k-1}^{-1} + \mathbf{G}_k^H \mathbf{G}_k \tag{5}$$

We see that $\mathbf{C}_k^{-1}$ is monotonically increasing with k and hence $\mathbf{C}_k$ is monotonically decreasing to zero as k increases. We also see that the update of $\mathbf{C}_k^{-1}$ is much simpler than the update of $\mathbf{C}_k$, which suggests that sensor $k-1$ should transmit $\mathbf{C}_{k-1}^{-1}$ (instead of $\mathbf{C}_{k-1}$) to sensor k. More importantly, the covariance matrix of the progressive estimation error at each sensor k is the same as that if sensor k performs the centralized estimation using all observations from its upper stream sensors. For sensor k to collect all observations $\{x_k\}$ from sensors 1 to $k-1$, there need to be $\sum_{i=1}^{k-1} = \frac{k(k-1)}{2} = O(k^2)$ transmissions in the network, which is in contrast to $O(k)$ transmissions required by the progressive estimation. Furthermore, the centralized estimator also needs to know all the associated observation matrices $\{\mathbf{G}_k\}$. Clearly, if a large number of sensors is needed to meet a required estimation accuracy, the progressive estimation is more desirable than the centralized estimation. It should be noted that at each hop of the progressive estimation, either $\mathbf{C}_{k-1}$ or $\mathbf{C}_{k-1}^{-1}$ needs to be transmitted along with a local estimate $\hat{\mathbf{u}}_{k-1}$ of the unknown vector $\mathbf{u}$. If the dimension of the unknown vector $\mathbf{u} \in C^{N\times 1}$ is so large that N^2 is comparable to K^2 where K is the "last" sensor, the advantage of the progressive estimation may not be realized.

More generally, as illustrated in Figure 2, if sensor k receives from its upper stream sensors more than one uncorrelated estimates of $\mathbf{u}$, denoted by $\hat{\mathbf{u}}_{k-1,i}$ with the corresponding covariance matrices $\mathbf{C}_{k-1,i}$ where $i = 1, ..., I_k$, then the BLUE estimate of $\mathbf{u}$ at sensor k is given by

$$\begin{aligned} \hat{\mathbf{u}}_k &= \left(\sum_{i=1}^{I_k} \mathbf{C}_{k-1,i}^{-1} + \mathbf{G}_k^H \mathbf{G}_k\right)^{-1} \\ &\quad \cdot \left(\sum_{i=1}^{I_k} \mathbf{C}_{k-1,i}^{-1} \hat{\mathbf{u}}_{k-1,i} + \mathbf{G}_k^H \mathbf{x}_k\right) \end{aligned} \tag{6}$$

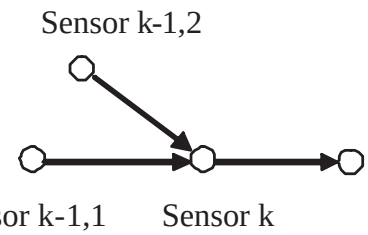

Fig. 2. Portion of a routing path where each sensor may have more than one upper stream sensors.

and the corresponding covariance matrix is

$$\mathbf{C}_k = \left(\sum_{i=1}^{I_k} \mathbf{C}_{k-1,i}^{-1} + \mathbf{G}_k^H \mathbf{G}_k\right)^{-1} \tag{7}$$

The above progressive estimation algorithm is distributed. Each sensor uses its locally available information. When the estimation accuracy at a sensor meets a pre-specified requirement, e.g., $tr(\mathbf{C}_k)$ becomes small enough, this sensor can stop the process of the progressive estimation and possibly send a message to trigger a network-wide operation.

2.2. LMMSE

The previous BLUE estimation treats $\mathbf{u}$ as an unknown deterministic vector. For the LMMSE estimation, we will treat $\mathbf{u}$ as an unknown random vector with zero mean and a known covariance matrix. Since $\mathbf{G}_k$ is arbitrary, a simple normalization of (1) can ensure that the covariance matrix of $\mathbf{u}$ and the covariance matrix of the noise $\mathbf{n}_k$ are both equal to the identity matrix $\mathbf{I}$.

At sensor 1, the LMMSE estimate of $\mathbf{u}$ is given by

$$\tilde{\mathbf{u}}_1 = \mathbf{G}_1^H \left(\mathbf{G}_1 \mathbf{G}_1^H + \mathbf{I}\right)^{-1} \mathbf{x}_1 \tag{8}$$

and its mean square error (MSE) matrix is

$$\mathbf{M}_1 = \mathbf{I} - \mathbf{G}_1^H \left(\mathbf{G}_1 \mathbf{G}_1^H + \mathbf{I}\right)^{-1} \mathbf{G}_1 \tag{9}$$

Assume that sensor k receives the LMMSE estimate $\tilde{\mathbf{u}}_{k-1}$ and the MSE matrix $\mathbf{M}_{k-1}$ from sensor $k-1$. Then, the LMMSE estimate of $\mathbf{u}$ at sensor k based on $\tilde{\mathbf{u}}_{k-1}$ and $\mathbf{x}_k$ is given by

$$\tilde{\mathbf{u}}_k = \mathbf{C}_{u,y_k} \mathbf{C}_{y_k,y_k}^{-1} \mathbf{y}_k \tag{10}$$

and its MSE matrix is

$$\mathbf{M}_k = \mathbf{I} - \mathbf{C}_{u,y_k} \mathbf{C}_{y_k,y_k}^{-1} \mathbf{C}_{u,y_k}^H \tag{11}$$

where

$$\mathbf{y}_k = \begin{bmatrix} \tilde{\mathbf{u}}_{k-1}^T & \mathbf{x}_k^T \end{bmatrix}^T \tag{12}$$

$$\mathbf{C}_{u,y_k} = \begin{bmatrix} \mathbf{I} - \mathbf{M}_{k-1} & \mathbf{G}_k^H \end{bmatrix} \tag{13}$$

$$\mathbf{C}_{y_k,y_k} = \begin{bmatrix} \mathbf{I} - \mathbf{M}_{k-1} & (\mathbf{I} - \mathbf{M}_{k-1})\mathbf{G}_k^H \\ \mathbf{G}_k(\mathbf{I} - \mathbf{M}_{k-1}) & \mathbf{G}_k \mathbf{G}_k^H + \mathbf{I} \end{bmatrix} \tag{14}$$

We see that like the BLUE estimation, the LMMSE estimation can also be formulated into a progressive form. Furthermore, one can verify by using (11) that

$$\mathbf{M}_k^{-1} = \mathbf{I} + \mathbf{C}_{u,y_k} \left(\mathbf{C}_{y_k,y_k} - \mathbf{C}_{u,y_k}^H \mathbf{C}_{u,y_k} \right)^{-1} \mathbf{C}_{u,y_k}^H \quad (15)$$

and by using (9), (13) and (14) that

$$\mathbf{M}_1^{-1} = \mathbf{I} + \mathbf{G}_k^H \mathbf{G}_k \quad (16)$$

$$\mathbf{M}_k^{-1} = \mathbf{M}_{k-1}^{-1} + \mathbf{G}_k^H \mathbf{G}_k \quad (17)$$

It is useful to notice the similarity between (17) and (5).

If the covariance matrix $\mathbf{C}_{u,u}$ of $\mathbf{u}$ is arbitrary but nonsingular, we can show that (17) still holds but (16) needs to be replaced by

$$\mathbf{M}_1^{-1} = \mathbf{C}_{u,u}^{-1} + \mathbf{G}_k^H \mathbf{G}_k \quad (18)$$

In general, we have that $\mathbf{M}_k < \mathbf{C}_k$ for all finite k, and $\mathbf{M}_k \to \mathbf{C}_k$ as k becomes large. But if $\mathbf{C}_{u,u}$ is so large that $\mathbf{C}_{u,u}^{-1}$ becomes negligible from (18), then $\mathbf{M}_k = \mathbf{C}_k$ for all k, i.e., the performance of BLUE becomes identical to that of LMMSE. In other words, an unknown deterministic vector is equivalent to a random vector of a covariance matrix with all eigenvalues infinitely large.

However, if sensor k receives multiple LMMSE estimates $\tilde{\mathbf{u}}_{k-1,i}$ and multiple MSE matrices $\mathbf{M}_{k-1,i}$ where $i = 1, ..., I_k$ from its upper stream sensors (see Figure 2), it does not appear possible in general to obtain the LMMSE estimate of $\mathbf{u}$ from $\mathbf{x}_k$, $\mathbf{G}_k$, $\tilde{\mathbf{u}}_{k-1,i}$ and $\mathbf{M}_{k-1,i}$, $i = 1, ..., I_k$. The problem lies with the fact that sensor k does not have sufficient information to compute such a cross-correlation matrix $\mathbf{C}_{\tilde{u}_{k-1,i},\tilde{u}_{k-1,j}}$ for $i \neq j$. This is a disadvantage of the LMMSE estimation.

3. PROGRESSIVE DETECTION

At each sensor, the estimate of $\mathbf{u}$ obtained by progressive estimation can be used for progressive detection if multiple hypotheses are associated with different regions of $\mathbf{u}$. The BLUE estimation does not need any prior knowledge about $\mathbf{u}$ while the LMMSE estimation needs to know the covariance matrix of $\mathbf{u}$.

To understand the performance of the progressive detection, we now consider a real-valued scalar data model for the observation at each sensor:

$$x_k = g_k u + n_k \quad (19)$$

which is simply a scalar version of (1) and the noise n_k is Gausssian with zero mean and unit variance. Furthermore, we treat a binary hypothesis problem where

$$u = \begin{cases} 0 & H_0 \\ a & H_1 \end{cases} \quad (20)$$

and $Pr(H_0) = Pr(H_1) = 0.5$.

3.1. Using estimation propagation

If the BLUE estimator $\hat{u}_k$ with the variance c_k is used at sensor k for detection, the detector with the minimum probability of error is: Decide H_1 if $\hat{u}_k > a/2$, or H_0 otherwise. The corresponding probability of error at sensor k is

$$p_{e,k} = Q\left(\frac{a}{2\sqrt{c_k}} \right) \quad (21)$$

where $Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx$.

If the LMMSE estimator $\tilde{u}_k$ with the MSE m_k is used at sensor k for detection, the decision rule is the same: Decide H_1 if $\tilde{u}_k > a/2$, or H_0 otherwise. The detection error probability is given by (21) with c_k replaced by m_k. Note that for u defined in (20), the variance of u is $\frac{a^2}{2}$.

3.2. Using decision propagation

For both BLUE and LMMSE shown above, sensor k needs to receive two real numbers from sensor $k-1$. To reduce the communication burden, let us consider "decision propagation" where sensor k receives the decision d_{k-1} (a binary number) and the corresponding probability of error $p_{e,k-1}$ (a real number) from sensor $k-1$. The detection at sensor k is solely based on d_{k-1}, $p_{e,k-1}$, x_k and g_k. Define

$$\lambda_k = \frac{P(d_{k-1}|H_1) f(x_k|H_1)}{P(d_{k-1}|H_0) f(x_k|H_0)} \quad (22)$$

where $P(\cdot|\cdot)$ denotes conditional probability and $f(\cdot|\cdot)$ denotes conditional probability density function. The optimal decision rule is: Decide H_1 if $\lambda_k > 1$, or H_0 otherwise.

The above decision rule can be simplified into:

- Decide H_1 if $d_{k-1} = 1$ and $x_k > \frac{a}{2} - \delta_k$ or if $d_{k-1} = 0$ and $x_k > \frac{a}{2} + \delta_k$;
- Decide H_0 if $d_{k-1} = 1$ and $x_k < \frac{a}{2} - \delta_k$ or if $d_{k-1} = 0$ and $x_k < \frac{a}{2} + \delta_k$.

Here, the offset δ_k is given by

$$\delta_k = \frac{1}{a} \ln\left(\frac{1 - p_{e,k-1}}{p_{e,k-1}} \right) \quad (23)$$

which is positive as long as $p_{e,k-1} < 0.5$. A quick observation from the above rule is that a higher hurdle is required if the decision at sensor k differs from that at sensor $k-1$, and a lower hurdle is required if the decision at sensor k coincides with that at sensor $k-1$. This extra hurdle is determined by δ_k.

The corresponding probability of error at sensor k is

$$p_{e,k} = Q\left(\frac{a}{2} - \delta_k\right) p_{e,k-1} + Q\left(\frac{a}{2} + \delta_k\right)(1 - p_{e,k-1}) \quad (24)$$

It is not yet proven that $p_{e,k}$ converges to zero as k increases. But numerical examples suggest that such a convergence property holds, which is illustrated later. Intuitively, we see that when $p_{e,k-1}$ becomes smaller and hence δ_k becomes larger, $p_{e,k}$ is dominated by the first term in (24) and the ratio of $p_{e,k}$ over $p_{e,k-1}$ is dominated by $Q\left(\frac{a}{2} - \delta_k\right)$ which becomes closer to one.

A further reduction of the communication burden is possible if we use a finite number B of bits to quantize $p_{e,k-1}$ nonlinearly for each k. Specifically, we use p_e^* to denote a desired value of the probability of error, and ϵ for a small number such that $p_{e,1} < 0.5 - \epsilon$. We then choose Δ such that $\Delta 2^B = \log_{10}(0.5 - \epsilon) - \log_{10} p_e^*$. In other words, the number of bits required in terms of p_e^* is given by

$$B = \log_2\left(\frac{1}{\Delta} \log_{10} \frac{0.5 - \epsilon}{p_e^*} \right) \quad (25)$$

If the conventional linear quantization is used for $p_{e,k-1}$ for each k, the number of bits required would be $\log_2\left(\frac{1}{p_e^*}\right)$. At sensor $k-1$, an

integer l_{k-1} is chosen such that $\log_{10} p_{e,k-1} \leq \Delta l_{k-1} + \log_{10} p_e^* \doteq \log_{10} \hat{p}_{e,k-1}$. Sensor k only receives d_{k-1} and l_{k-1} from sensor $k-1$, and reconstructs $\hat{p}_{e,k-1}$ from the above formula.

In order for the above strategy to work, we need to show that with $\hat{p}_{e,k-1}$, sensor k is able to compute an upper bound $\hat{p}'_{e,k}$ on $p_{e,k}$ and hence l_k such that $\log_{10} p_{e,k} \leq \log_{10} \hat{p}'_{e,k} \leq \log_{10} \hat{p}_{e,k} = \Delta l_k + \log_{10} p_e^*$. Fortunately, such a $\hat{p}'_{e,k}$ is given by

$$\hat{p}'_{e,k} = Q\left(\frac{a}{2} - \hat{\delta}_k\right) \hat{p}_{e,k-1} + Q\left(\frac{a}{2} + \hat{\delta}_k\right) (1 - \hat{p}_{e,k-1}) \qquad (26)$$

where

$$\hat{\delta}_k = \frac{1}{a} \ln\left(\frac{1 - \hat{p}_{e,k-1}}{\hat{p}_{e,k-1}}\right) \qquad (27)$$

In fact, since the detection at sensor k is optimal given d_{k-1}, $p_{e,k-1}$, x_k and g_k, the probability error $p_{e,k}$ as given in (24) must be non-decreasing function of $p_{e,k-1}$ as long as $p_{e,k-1} < 0.5$. This fundamental property implies that $\hat{p}'_{e,k}$ given in (26) must be no less than $p_{e,k}$ given in (24) when $p_{e,k-1} \leq \hat{p}_{e,k-1} < 0.5$.

4. PERFORMANCE ILLUSTRATION

For performance illustration, we use (19) and (20) and further assume that $g_k = 1$ and $a = 1$. The probability of detection error $p_{e,k}$ versus the sensor index k is illustrated in Figure 3.

For the case of "estimation propagation", we used the BLUE estimation where the real-valued estimate $\hat{u}_{k-1}$ and its real-valued variance c_{k-1} are transported from sensor $k-1$ to sensor k. As expected, $p_{e,k}$ converges to zero rapidly.

For the case of "decision propagation - a", the binary decision variable d_{k-1} and the real-valued probability of detection error $p_{e,k-1}$ are transported from sensor $k-1$ to sensor k. It should be noted that $p_{e,k}$ apparently converges to zero when k is very large. (We omitted the portion for $k > 200$.) But a mathematical proof of this property is not yet available.

For the case of "decision propagation - b", four curves are illustrated in this figure. The best (lowest) of the four corresponds to $B = 10$ bits used for quantization of each $p_{e,k-1}$ before the bits are transported from sensor $k-1$ to sensor k. The other three curves correspond to $B = 8$, $B = 6$ and $B = 4$ bits, respectively. We see that the performance of "decision propagation - b" has a nonzero floor when k is beyond a threshold.

5. CONCLUSION

In this brief paper, we have introduced the concept of progressive estimation and detection (PED) for wireless sensor networks. The classical estimation principles: BLUE and LMMSE, have been shown to be naturally embeddable into PED. The progressive detection based on estimation propagation appears promising in both theory and practice.

While several useful results have been found, PED is still in its infancy and ready to be further developed rapidly. There are many open issues for future research. One is the robustness of PED against unknown errors in the observation matrices. Another is the extension of PED from the linear model (1) to nonlinear models.

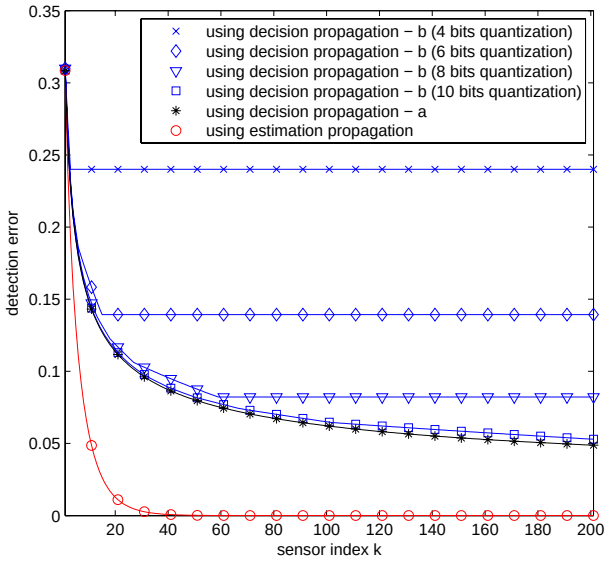

Fig. 3. The probability of the progressive detection error $p_{e,k}$ versus the sensor index k. The numbers of bits for quantization shown in this plot are used to quantize $p_{e,k-1}$ at sensor $k-1$ nonlinearly. The other parameters used for the quantization are $p_e^* = 0.04$ and $\epsilon = 0.19$.

6. REFERENCES

[1] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, Vol. 46, pp. 388-404, Mar 2000.

[2] Y. Hua, Y. Huang and J. J. Garcia-Luna-Aceves, "Maximizing the throughput of large ad hoc wireless networks," *IEEE Signal Processing Magazine*, Vol. 23, pp. 84-94, Sept 2006.

[3] Y. Yu, Y. Huang, B. Zhao and Y. Hua, "Throughput analysis of wireless mesh networks," *Proceedings of IEEE ICASSP'2008*, Las Vegas, NV, April 2008.

[4] J.-J. Xiao, S. Cui, Z.-Q. Luo, A. J. Goldsmith, "Power scheduling of universal decentralized estimation in sensor networks," *IEEE Trans. on Signal Processing*, vol. 54, pp. 413-422, Feb. 2006.

[5] P. Venkitasubramaniam, L. Tong, and A. Swami, "Quantization for maximin ARE in distributed estimation," *IEEE Trans. on Signal Processing*, vol. 55, no. 7, pp. 3596-3605, July 2007.

[6] A. Swami, Q. Zhao, Y.-W. Hong and L. Tong, editors, *Wireless Sensor Networks*, Wiley, 2007.

[7] M. G. Rabbat and R. D. Nowak, "Quantized incremental algorithms for distributed optimization," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 4, Apr. 2005.

[8] Y. Huang and Y. Hua, "Multi-hop progressive decentralized estimation in wireless sensor networks," *IEEE Signal Processing Letters*, Vol. 14, No. 12, pp. 1004-1007, Dec 2007.