DECENTRALIZED ESTIMATION OVER NOISY CHANNELS FOR BANDWIDTH–CONSTRAINED SENSOR NETWORKS

Tuncer C. Aysal and Kenneth E. Barner

Signal Processing and Communications Group
Electrical and Computer Engineering, University of Delaware

ABSTRACT

Recently proposed decentralized estimation methods do no consider errors occurring during the transmission of binary observations from the sensors to fusion center. In this paper, we extend the decentralized estimation model to the case where imperfect transmission channels are considered. The proposed estimator, which operates on additive channel noise corrupted versions of quantized noisy sensor observations, is approached from maximum likelihood (ML) perspective. The ML estimate unfortunately has no closed–form solution. We analyze the log–likelihood function showing that it is logconcave, thereby indicating that numerical methods, such as Newtons algorithm, can be utilized to obtain the optimal solution. A suboptimal mean estimator that requires minimal information about the channel and sensing environment is also proposed. Simulation results evaluating the variances of the proposed optimal and suboptimal solutions are provided.

***Index Terms*—** distributed estimation, sensor networks.

1. INTRODUCTION

A constraint in many wireless sensor networks (WSNs) is that bandwidth is limited, necessitating the use and transmission of quantized binary versions of the original noisy observations. Many recent efforts address the estimation of a deterministic source signal from quantized noisy observations [1–6]. When the probability density function (pdf) of the sensor noise is known, transmitting a single bit per sensor leads to minimal loss in estimator variance compared with a clairvoyant estimator (estimator based on unquantized measurements) [4, 5, 7]. Alternatively, when the sensor noise pdf is unknown, pdf–unaware estimators based on quantized sensor data have also been introduced recently [3, 5, 6].

The distributed estimation techniques considered in the previously proposed methods are based on *quantized noisy sensor observations*. These methods thus subsequently assume that the transmission of binary observations from sensors to fusion center is perfect. In this paper, we extend the distributed estimation model to admit transmission imperfectness, i.e., we consider the case where the quantized noisy sensor observations are corrupted by additive noise during transmission from sensor to fusion center. Our estimator is hence based on *noisy quantized versions of noisy sensor observations*. Utilizing this extended WSN model, we derive the maximum likelihood (ML) estimate of a deterministic source signal. The ML estimate unfortunately has no closed–form solution. We analyze the log–likelihood function showing that it is logconcave, thereby indicating that numerical methods, such as Newtons algorithm, can be utilized to obtain the optimal solution. To further address the complexity and implementation issues of the optimal ML estimator, we propose fast, simple and practical suboptimal solution: mean estimator. The mean estimator simply averages the noisy observations received at the fusion center.

The remainder of this paper is organized as follows. The problem formulation and the extended WSN model admitting transmission noise is introduced in Section 2. The estimator of a deterministic source signal utilizing the corrupted quantized noisy sensor observations is derived in Section 3 along with the suboptimal mean estimator. Section 4 details the experiments evaluating the performance of the ML and mean estimator. Finally, conclusions are drawn in Section 5.

2. PROBLEM FORMULATION

Consider a set of K distributed sensors, each making observations of a deterministic source signal θ. The observations are corrupted by additive noise and are described by [1–6]

$$x(k) = \theta + n(k), \quad k = 1, 2, \ldots, K. \tag{1}$$

Noise samples $\{n(k) : k = 1, 2, \ldots, K\}$ are assumed zero–mean, spatially uncorrelated and independent. Furthermore, the density function of the sensor noise is denoted by $n(k) \sim f_n(u; \sigma_n)$, where σ_n denotes the scale parameter of f_n.

Suppose a fusion center is to estimate θ based on the noisy sensor observations $\{x(k) : k = 1, 2, \ldots, K\}$. If the fusion center has knowledge of the sensor noise density function and sensors are capable of sending the observations $\{x(k) : k = 1, 2, \ldots, K\}$ to the fusion center without distortion, then the fusion center can simply perform the ML estimate of θ, $\hat{\theta} = \arg\min_\beta \left[\sum_{k=1}^{N} \rho(x(k) - \theta)\right]$ where $\rho(u) = -\log f_n(u; \sigma_n)$.

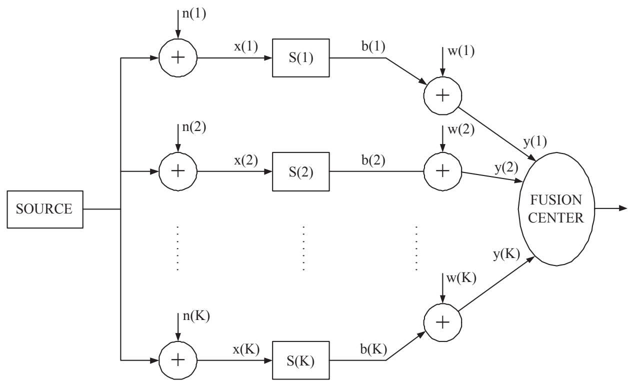

Fig. 1. A decentralized, noisy channels WSN scheme with a fusion center.

This scheme is only applicable in a centralized estimation situation where observations are either centrally located, or can be transmitted to a central location without distortion. Neither of these requirements is realistic in a WSN, where the sensor nodes are bandwidth constrained and the communication links between the fusion center and sensors are noisy. Due to bandwidth limitations, the $\{x(k) : k = 1, 2, \ldots, K\}$ observations have to be quantized. To this end, we consider the quantization operation as the construction of a set of indicator variables, which are binary observations [1–6]

$$b(k) = \mathbf{1}\{x(k) \in (\tau_k, +\infty)\}, \quad k = 1, 2, \ldots, K \tag{2}$$

where $\tau_k \in \mathbf{Z}$ is a threshold defining $b(k)$, $\mathbf{Z}$ denotes the set of real numbers, and $\mathbf{1}\{\cdot\}$ is the indicator function. In addition, due to imperfections of communication links between sensor nodes and the fusion center, we further extend the model to include channel noise,

$$y(k) = b(k) + w(k), \quad k = 1, 2, \ldots, K \tag{3}$$

where the $\{w(k) : k = 1, 2, \ldots, K\}$ are assumed to be zero–mean independent channel noise samples and $\{y(k) : k = 1, 2, \ldots, K\}$ are the noisy observations received at the fusion center. Moreover, the density function of the link noise is denoted by $w(k) \sim f_w(u; \sigma_w)$, where σ_w denotes the scale parameter of f_w. As a result, we consider the extended decentralized scheme shown in Fig. 1, where $\{S(k) : k = 1, 2, \ldots, K\}$ denote the sensors.

3. ESTIMATION BASED ON NOISY BINARY OBSERVATIONS

Consider the most demanding bandwidth constraint case, in which sensors are restricted to transmit one bit per $x(k)$ observation. Furthermore, let every sensor use the same threshold τ to form $\{b(k) : k = 1, 2, \ldots, K\}$, i.e., $b(k) = \mathbf{1}\{x(k) \in (\tau, +\infty)\}, \quad k = 1, 2, \ldots, K$. Instrumental to the WSN scheme presented in Section 2 is the fact that $b(k)$ is a Bernoulli random variable with parameter

$$1 - \alpha(\theta) \triangleq \Pr\{b(k) = 1\} = 1 - F_n(\tau - \theta) \tag{4}$$

where $F_n(\cdot)$ is the cumulative distribution function of $n(k)$. The probability density function of the noisy observations received at the fusion center, i.e., $y(k) = b(k) + w(k)$, for $k = 1, 2, \ldots, K$, is then given by

$$f_y(y) = a_w(y)F_n(\tau - \theta) + b_w(y) \tag{5}$$

where $a_w(y) \triangleq [f_w(y) - f_w(y-1)]$ and $b_w(y) \triangleq f_w(y-1)$.

Let us define

$$\alpha \triangleq \alpha(\theta) = F_n(\tau - \theta). \tag{6}$$

Note that the α is the probability that the binary sensor observation $b(k)$ is zero, i.e., $\alpha(\theta) = \Pr\{b(k) = 0\}$, and is restricted to the open interval $(0, 1)$. To simplify the problem, we first derive the estimate for α and utilize the invariance of the ML estimate to estimate θ using (6). The log–likelihood function of α based on the noisy observations is given by

$$\log\{f(\mathbf{y}|\alpha)\} = \sum_{k=1}^{K} \log\{a_w(y(k))\alpha + b_w(y(k))\} \tag{7}$$

where $\mathbf{y} = \{y(1), y(2), \ldots, y(K)\}$. Unfortunately, ML estimate of α has no closed–form solution. We utilize the Newton's iteration technique. Convergence of Newton's algorithm to the optimal solution is guaranteed in this case since in the following we prove that $\log\{f(\mathbf{y}|\alpha)\}$ is concave.

The second derivative of the log–likelihood function, denoted as $\log\{f(\mathbf{y}|\alpha)\}''$ is given by

$$\log\{f(\mathbf{y}|\alpha)\}'' = -\sum_{k=1}^{K} \frac{a_w^2(y(k))}{[a_w(y(k))\alpha + b_w(y(k))]^2} \tag{8}$$

indicating that $\log\{f(\mathbf{y}|\alpha)\}'' < 0$ for all α. This subsequently indicates that $\log\{f(\mathbf{y}|\alpha)\}$ is concave.

Newton's algorithm is utilized to obtain optimal $\hat{\alpha}$. Newton's algorithm is based on the following iteration:

$$\hat{\alpha}(j+1) = \hat{\alpha}(j) - \frac{\log\{f(\mathbf{y}|\alpha)\}'}{\log\{f(\mathbf{y}|\alpha)\}''} \tag{9}$$

where j denotes the iteration number and

$$\log\{f(\mathbf{y}|\alpha)\}' = \sum_{k=1}^{K} \frac{a_w(y(k))}{a_w(y(k))\alpha + b_w(y(k))} \tag{10}$$

Newton's algorithm is guaranteed to converge to the optimal solution regardless of the initialization since $\log\{f(\mathbf{y}|\alpha)\}$ is concave.

Furthermore, the ML estimate for θ is now given by

$$\hat{\theta} = \tau - F_n^{-1}(\hat{\alpha}) \tag{11}$$

where we utilized the invariance of ML estimate. Note that the fact that $\hat{\theta}$ is the ML estimate (even though numerical methods are utilized to obtain) guarantees (at least asymptotically) unbiasedness, consistency and efficiency.

Although well established, Newton's algorithm requires the evaluation of $\Lambda'_L(\alpha(j), \mathbf{y})$ and $\Lambda''_L(\alpha(j), \mathbf{y})$ at each iteration. In addition, since Newton's algorithm is recursive, these function evaluations are performed M times, where M denotes the total number of iterations. Due to these issues, we develop practical, easy–to–implement and fast suboptimal solution in the following.

An estimate for θ that requires minimum information and complexity is defined as

$$\hat{\theta} = \tau - F_n^{-1}\left(1 + \nu_w - \frac{1}{K}\sum_{k=1}^{K} y(k)\right) \tag{12}$$

where $\nu_w = E\{w\}$. This estimator assumes that the channel noise has a finite mean, a constraint that, for instance, holds for the Gaussian density that enjoys the central limit theorem, but not for the Cauchy density that follows the generalized central limit theorem and a member of the alpha–Stable density family. To see the effectiveness of the mean estimator in finite mean valued channel noise cases, consider

$$\frac{1}{K}\sum_{k=1}^{K} y(k) \rightarrow E\{y\} = E\{b\} + E\{w\} \tag{13}$$

$$= 1 - F_n(\tau - \theta) + \nu_w \tag{14}$$

where the first and last lines follow from the weak law of large numbers and the fact that b is a bernoulli random variable with $\Pr(b(k) = 1) = 1 - F_n(\tau - \theta)$.

The mean estimator reduces, in the perfect transmission case, to the estimator of θ based strictly on quantized noisy sensor observations [1–5]:

$$\hat{\theta} = \tau - F_n^{-1}\left(1 - \frac{1}{K}\sum_{k=1}^{K} b(k)\right). \tag{15}$$

This is seen by noting that, for $\nu_w = 0$ and $\sigma_w \rightarrow 0$, $\{y(k) : k = 1, 2, \ldots, K\} \rightarrow \{b(k) : k = 1, 2, \ldots, K\}$.

4. NUMERICAL EXPERIMENTS

The proposed optimal ML estimator (OMLE), and suboptimal mean estimator (ME) fusion centers are evaluated through illustrative numerical experiments. Considered are the output variances and processing times of the proposed estimators. The variances of the clairvoyant estimator (CE) [1, 8] and the estimator operating on quantized (binary) noisy sensor observations (BE) [1–3, 5] (no channel noise) are utilized as benchmarks.

To evaluate the fusion center performance for the various estimation techniques, consider an examples in which the sensor and channel noise are taken as Gaussian distributed random variables. The parameters of the experiment are: $\theta = 1.0$, $\tau = \{1.0, 1.2\}$, $K \in [50, 150]$, $\sigma_n = 0.5$, $\sigma_w \in [0.05, 1]$. The variance of the estimate $\hat{\theta}$ (ensemble average of 10000 experiments) is plotted as a function of the number of sensors, K, and as a function of the channel noise spread parameter, σ_w, in Fig 2 (a) and (b), respectively. In these two illustrations, the optimal $\theta = \tau$ case is considered. Fig 2 (c) and (d) present similar conditions, but in this case there is a discrepancy between the θ and τ, such that $\theta = 1$ and $\tau = 1.2$. Plotted are the variance of the CE, which, in this case, is given by $\mathrm{Var}(\hat{\theta}) = \sigma_n^2/K$ and the BE estimator in (15) operating on noise–free binary observations *in the ideal case*, i.e., when $\tau = \theta$, which is given by $\mathrm{Var}(\hat{\theta}) = 2\pi\sigma_n^2/4K$.

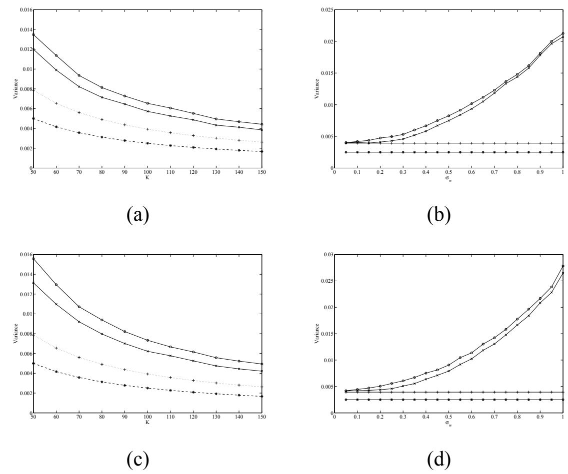

Fig. 2. Illustration of estimator variances for (a) optimal case ($\theta = \tau$) and $\sigma_w = 0.4$ with increasing number of sensors and (b) with increasing channel power where $K = 100$. (c–d) Similar illustration for non–ideal case ($\theta = 1$ and $\tau = 1.2$). Circles, crosses, pluses and asterisks represent the ME, OMLE, BE and CE estimators respectively.

As expected, the CE provides the smallest variances in all cases, followed by the BE. Also expected, amongst the proposed estimation techniques, the OMLE provides the best performance. On the other hand, ME estimator also provides close–to–optimal performance in environments characterized by Gaussian density.

Similar experiments are also performed for the case where the channel noise is modeled with the algebraic–tailed Cauchy distribution. Identical parameters are utilized and the scale parameter is set as $\gamma_w \in [0.05, 1]$. The results for ideal and non–ideal cases are given in Fig. 3 (a–b) and (c–d), respectively. The ME, as expected, provides results significantly worse than the OMLE in this heavy–tailed case due to its reliance on the averaging operation.

The processing time of the proposed OMLE and ME meth-

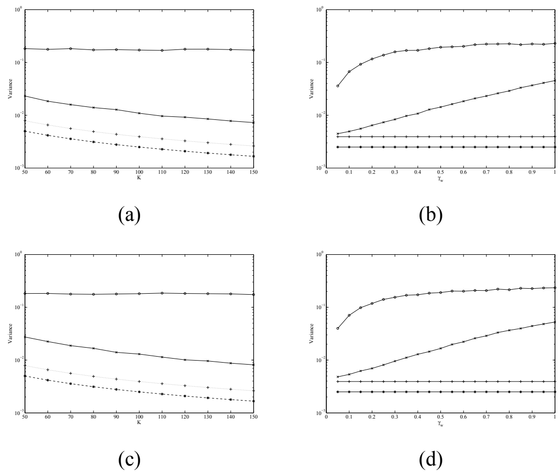

Fig. 3. Illustration of estimator variances for (a) optimal case ($\theta = \tau$) and $\gamma_w = 0.4$ with increasing number of sensors and (b) with increasing channel power where $K = 100$. (c–d) Similar illustration for non–ideal case ($\theta = 1$ and $\tau = 1.2$).

Table 1. CPU Times in milisecs for Optimal and Suboptimal Algorithms Implemented in MATLAB

	Number of Sensors		
Gaussian	50	100	150
OMLE	$1.900\texttt{T}$	$1.975\texttt{T}$	$2.037\texttt{T}$
ME	$1.006\texttt{T}$	$1.025\texttt{T}$	$1.037\texttt{T}$
Cauchy	50	100	150
OMLE	$1.887\texttt{T}$	$1.950\texttt{T}$	$2.018\texttt{T}$
ME	$\texttt{T}$	$1.018\texttt{T}$	$1.037\texttt{T}$

ods are also evaluated through examples. The utilized experiment parameters are: $\hat{\theta} = 1.0, \tau = 1.2, \sigma_n = 0.5, \{\sigma_w, \gamma_w\} = 0.5$ and $K \in \{50, 75, 100, 125, 150\}$. The simulations are run in MATLAB on a 3.20 GHz, 2.00 GB RAM PC. Note that for each K, the processing time given is the ensemble average of 10000 trials. The results for both Gaussian and Cauchy channel noise cases are tabulated in Table 1 where $\texttt{T} = 0.160$ milisecs is a normalizing unit of time and corresponds to the simplest ME Cauchy $K = 50$ case.

The processing time of the estimators, as expected, increases with the number of sensors. The results indicate that the mean estimator is the most computationally efficient with the least processing time, which is approximatively two times faster than the OMLE.

5. CONCLUDING REMARKS

The decentralized WSN estimation scheme is extended to admit imperfections occurring during the data transmission from sensors to fusion center. Based on the extended decentralized estimation scheme, a maximum likelihood estimator operating on corrupted quantized noisy sensor observation is proposed. Noting that there is no–closed form maximum likelihood solution, we show that the log–likelihood function is concave and that numerical methods such as Newton's algorithm can be utilized to obtain the optimal solution. Due to complexity and implementation issues associated with numerical solutions, we derive the mean estimator suboptimal solution. The mean estimator, which requires the minimum amount of information and computational load, relies on simple averaging. Numerical examples evaluating and comparing the the proposed techniques in varying environments (characterized by Gaussian and Cauchy densities) indicate the effectiveness of the optimal ML estimator and mean estimators.

6. REFERENCES

[1] A. Ribeiro and G. B. Giannakis, "Bandwidth–constrained distributed estimation for wireless sensor networks–part i: Gaussian case," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 1131–1143, Mar. 2006.

[2] J.-J. Xiao, S. Cui, Z.-Q. Lao, and A. J. Goldsmith, "Power scheduling of universal decentralized estimation in sensor networks," *IEEE Transactions on Signal Processing*, vol. 54, no. 2, pp. 413–422, Feb. 2006.

[3] J.-J. Xiao and Z.-Q. Lao, "Decentralized estimation in an inhomogeneous sensing environment," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3564–3575, Oct. 2005.

[4] H. Papadopoulos, G. Wornell, and A. Oppenheim, "Sequential signal encoding from noisy measurements using quantizers with dynamic bias control," *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 978–1002, Mar. 2001.

[5] Z.-Q. Lao, "Universal decentralized estimation in a bandwidth constrained sensor network," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 2210–2219, June 2005.

[6] ——, "An isotropic universal decentralized estimation scheme for a bandwidth constrained ad hoc sensor network," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 4, pp. 735–744, Apr. 2005.

[7] M. Abdallah and H. Papadopoulos, "Sequential signal encoding and estimation for distributed sensor networks," in *Proceedings of the 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Salt Lake City, UT, 2001, pp. 2577–2580.

[8] Huber, *Robust Statistics*. John Wiley & Sons, Inc., 1981.