BLIND DECENTRALIZED ESTIMATION 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

In this paper, we extend the decentralized estimation model to the case in which imperfect transmission channels are considered. The proposed estimators, which operate on additive channel noise corrupted versions of quantized noisy sensor observations, are approached from a maximum likelihood (ML) perspective. Complicating this approach is the fact that the noise distribution is rarely fully known to the fusion center. Here we assume the distribution is known but not the defining parameters, e.g., variance. The resulting incomplete data estimation problem is approached from a expectation–maximization (EM) perspective. The critical initialization and convergence aspects of the EM algorithm are investigated. Furthermore, the estimation of the source parameter is extended to the blind case where both the channel and sensor noise parameters are unknown. Finally, numerical experiments are provided to show the effectiveness of the proposed estimators.

Index Terms— decentralized estimation, sensor network.

1. INTRODUCTION

Many wireless sensor networks (WSNs) are constrained by the fact the that bandwidth is limited, imposing 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 [1–4]. 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) [1, 3, 4]. Alternatively, when the sensor noise pdf is unknown, pdf–unaware estimators based on quantized sensor data have also been introduced recently [4].

The distributed estimation techniques considered in the previously proposed methods are based on *quantized noisy sensor observations*. These methods subsequently assume that the transmissions of binary observations from sensors to fusion center are perfect. In this paper, we extend the distributed estimation model to admit transmission imperfections, i.e., we consider the case where the quantized noisy sensor observations are corrupted by additive noise during transmission from sensors to fusion center. The considered 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. This formulation is complicated by the fact that the noise statistics are rarely known entirely in practice. Here we consider the practical case in which the noise pdf is known (e.g., Gaussian) but some parameters of the distribution are unknown. For instance, a case frequently encountered in practice is when the noise pdf is known except for its parameter σ (or equivalently its variance). Moreover, the unlabeled nature of the fusion center observations makes the problem at hand a typical task with incomplete data. We focus on the so-called EM algorithm, which has attracted a great deal of interest over the past few years in a wide range of application involving tasks with incomplete data sets [5]. We integrate the EM algorithm to solve the estimation problem where the channel noise parameter is unknown. The critical initialization and convergence aspects of the EM algorithm are investigated. Furthermore, the estimation of the source parameter is extended to the blind case where both the channel and sensor noise parameters are unknown. Finally, numerical experiments are provided to show the effectiveness of the proposed estimators.

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, 2, 4]

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

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.

Due to the inherent bandwidth limitations in sensor networks, the $\{x(k) : k = 1, 2, \ldots, K\}$ observations need 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, 2, 4],

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

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. The channels between the source signal and the sensors, and the sensors and the fusion center, are modeled as additive white Gaussian noise (AWGN), the defining pdf of which is given by $f(u) = 1/(\sigma\sqrt{2\pi}) \exp\left(-u^2/(2\sigma^2)\right)$ where σ denotes the spread parameter.

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

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

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

$$f_y(y) = a_w(y)[1 - F_n(\tau - \theta)] + b_w(y) \tag{5}$$

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

An inspection of the pdf of the observed random variable reveals that y can be modeled as a two–component Gaussian mixture model: $f_y(u) = F_n(\tau - \theta) f_w(u) + [1 - F_n(\tau - \theta)] f_w(u - 1)$, where $F_n(\tau - \theta)$ and $[1 - F_n(\tau - \theta)]$ are the mixing probabilities.

A realistic approach to the estimation problem in WSNs is to assume that the noise pdf is known (e.g., Gaussian) but that some of its parameters are unknown [1]. A case frequently encountered in practice is when the noise pdf is known except for its parameter σ (or equivalently its variance). In the following, we consider the estimation problem when the channel noise parameter σ_w is unknown but the sensor noise parameter σ_n is known.

Let us define $\psi \triangleq \psi(\theta) = 1 - F_n(\tau - \theta)$. Note that the ψ is the probability that the binary sensor observation $b(k)$ is unity, i.e., $\psi(\theta) = \Pr\{b(k) = 1\}$, and is restricted to the open interval $(0, 1)$. To simplify the problem, we first derive the estimate for ψ and utilize the invariance of the ML estimate to estimate θ using (4). The ML estimate of $\psi \in (0, 1)$ thus reduces to $\hat{\psi}_{\mathrm{ML}} = \arg\max_\psi \prod_{k=1}^{K} [1 - \psi] f_w(y(k)) + \psi f_w(y(k) - 1)$. Taking the natural $\log(\cdot)$ yields the log–likelihood function, denoted as $\Lambda_{\mathrm{L}}(\psi)$, and the ML estimate of ψ is then given by

$$\hat{\psi}_{\mathrm{ML}} = \arg\max_{\psi} \sum_{k=1}^{K} \log([1 - \psi] f_w(y(k)) + \psi f_w(y(k) - 1)). \tag{6}$$

The unknown parameter set for the above estimation is $\mathbf{p} = \{\psi, \sigma_w\}$. Due to the lack of information concerning the labels of $\{y(k) : k = 1, 2, \ldots, K\}$, the summation formulation and the unknown channel parameter, the typical ML estimation encounters difficulty. The missing label information makes the problem at hand a typical task with an incomplete data set. We focus on addressing this problem utilizing the so-called EM algorithm, which has attracted a great deal of interest over the past few years in a wide range of application involving tasks with incomplete data sets [5].

The followings are the M– and E– steps for the unknown parameter set estimation of finite Gaussian mixture models in the considered WSN application.

E–Step – Let the parameters estimated at the j–th iteration be marked by a superscript (j). Compute the posterior probabilities

$$q(k) = \frac{\hat{\psi}_{\mathrm{ML}}^{(j)} f_w(y(k) - 1|\mathbf{p}^{(j)})}{\hat{\psi}_{\mathrm{ML}}^{(j)} f_w(y(k) - 1|\mathbf{p}^{(j)}) + (1 - \hat{\psi}_{\mathrm{ML}}^{(j)}) f_w(y(k)|\mathbf{p}^{(j)})}. \tag{7}$$

M–step – The ML estimates, $\hat{\mathbf{p}}^{(j+1)} = \{\hat{\psi}_{\mathrm{ML}}^{(j+1)}, \hat{\sigma}_{w,\mathrm{ML}}^{(j+1)}\}$ are given by

$$\hat{\psi}_{\mathrm{ML}}^{(j+1)} = \frac{\Delta(\mathbf{q})}{K} \tag{8}$$

where $\Delta(\mathbf{q}) \triangleq \sum_{k=1}^{K} q(k)$ and $\mathbf{q} = \{q(1), q(2), \ldots, q(K)\}$, and

$$\hat{\sigma}_{w,\mathrm{ML}}^{(j+1)} = \left[\frac{\Delta(\mathbf{q} \odot (\mathbf{y} - 1)^2) + \Delta((1 - \mathbf{q}) \odot \mathbf{y})}{K}\right]^{1/2} \tag{9}$$

where $\mathbf{y} = \{y(k) : k = 1, 2, \ldots, K\}$, and $\odot$ and $(\cdot)^2$ denote the element–wise multiplication and squaring operations, respectively.

If $|\hat{\psi}_{\mathrm{ML}}^{(j+1)} - \hat{\psi}_{\mathrm{ML}}^{(j)}| > \epsilon_1$ and $|\hat{\sigma}_{w,\mathrm{ML}}^{(j+1)} - \hat{\sigma}_{w,\mathrm{ML}}^{(j)}| > \epsilon_2$, where $|\cdot|$ and ϵ_i for $i = 1, 2$, denote the absolute value operator and positive small numbers, respectively, the algorithm returns to the E–step. Otherwise, the iteration is ended.

Given the estimate of ψ, and utilizing the facts that $F_n(\cdot)$ is a bijection and symmetric, and the ML estimates are invariant, the ML estimate of θ is given by

$$\hat{\theta}_{\mathrm{ML}} = F_n^{-1}(\hat{\psi}_{\mathrm{ML}}) + \tau \tag{10}$$

where $\hat{\psi}_{\mathrm{ML}}$ is the solution obtained through the EM algorithm.

The initialization of the ψ and σ_w is an important step of the EM algorithm. Consider first the initialization of $\hat{\psi}_{\mathrm{ML}}$. The $\psi_{\mathrm{ML}}^{(0)}$ is set to:

$$\psi_{\mathrm{ML}}^{(0)} = \frac{\Delta(\mathbf{y})}{K}. \tag{11}$$

The reason for this setting can be seen as follows. By the strong law of large numbers: $\psi_{\mathrm{ML}}^{(0)} = (\Delta(\mathbf{y}))/K \rightarrow E\{y\}$ almost surely, which implies that $\psi_{\mathrm{ML}}^{(0)} \rightarrow E\{b\} + E\{w\} = E\{b\} = \psi$ following from the fact that the channel noise is zero–mean. Due to the finite number of fusion center observations, the initial estimate $\psi_{\mathrm{ML}}^{(0)} \approx \psi$.

Consider next the initialization of the estimate of σ_w. Note that $\phi(y) = [1-F_n(\tau-\theta)]F_n(\tau-\theta)+\sigma_w^2$ since $\phi(b) = [1-F_n(\tau-\theta)]F_n(\tau-\theta)$, where $\phi(\cdot)$ denotes the variance of its argument. Solving for σ_w yields $\sigma_w = \sqrt{\phi(y) - \phi(b)}$. Utilizing the unbiased ML estimator of the variance of y and the loose upper bound $[[1 - F_n(\tau-\theta)]F_n(\tau-\theta) \leq 1/4]$ gives

$$\sigma_w \geq \sqrt{\frac{1}{K-1}\sum_{k=1}^{K}\left(y(k) - \frac{\Delta(\mathbf{y})}{K}\right)^2 - \frac{1}{4}} \triangleq \mathcal{L}(\sigma_w). \tag{12}$$

Clearly, the upper bound of σ_w is estimated by

$$\sigma_w \leq \sqrt{\frac{1}{K-1}\sum_{k=1}^{K}\left(y(k) - \frac{\Delta(\mathbf{y})}{K}\right)^2} \triangleq \mathcal{U}(\sigma_w). \tag{13}$$

Noting that the $\mathcal{L}(\sigma_w)$ can take on imaginary numbers if the term inside the square root is negative, the $\hat{\sigma}_{w,\mathrm{ML}}$ is hence initialized as

$$\hat{\sigma}_{w,\mathrm{ML}}^{(0)} = \begin{cases} (\mathcal{L}(\sigma_w) + \mathcal{U}(\sigma_w))/2, & \mathrm{Im}(\mathcal{L}(\sigma_w)) = 0 \\ (\mathcal{U}(\sigma_w))/2, & \mathrm{Im}(\mathcal{L}(\sigma_w)) \neq 0 \end{cases} \tag{14}$$

where $\mathrm{Im}(\cdot)$ denotes the imaginary part of its argument.

It is known that it is possible for the EM algorithm to converge to local extrema or saddle points in unusual cases. In the following, we prove that the log–likelihood function is concave in ψ guaranteeing the convergence to global maximum. The log–likelihood function, $\Lambda_{\mathrm{L}}(\psi)$, is rewritten as $\Lambda_{\mathrm{L}}(\psi) = \sum_{k=1}^{K}\log(a_w(y(k))\psi + b_w(y(k)))$. Note that $a_w(y(k))\psi + b_w(y(k))$ is a concave function in ψ, and that $a_w(y(k))\psi + b_w(y(k))$ is monotonically decreasing (increasing) if $a_w(y(k)) < 0 (> 0)$. Also recalling that $\log(\cdot)$ function is strictly concave indicating that $\log(a_w(y(k))\psi + b_w(y(k)))$ is concave in ψ. Finally, noting that the summation preserves concavity concludes the proof.

An estimator that requires the least amount of information is one that assumes that the sensor noise parameter σ_n is also unknown along with the channel noise parameter σ_w. This is the case considered in the following.

To estimate θ when σ_n and σ_w are unknown, while keeping the bandwidth constraint to one bit per sensor, we divide the sensors in two groups, with each group using a different region (i.e. threshold) to define the binary random observations [1]:

$$\mathcal{S}_i \triangleq (\tau_i, +\infty) \quad \text{for} \quad k = \mathcal{K}_i \tag{15}$$

for $i \in \{1, 2\}$, where $\mathcal{K}_1 = \{1, 2, \ldots, K/2\}$ and $\mathcal{K}_2 = K/2 + 1, K/2 + 2, \ldots, K$. That is, without loss of generality, the first $K/2$ sensors quantize their observations using the region $\mathcal{S}_1$, while the remaining $K/2$ sensors utilize the region $\mathcal{S}_2$. Furthermore, we assume, without loss of generality that $\tau_2 > \tau_1$.

The Bernoulli parameters of the resultant binary observations are expressed in terms of the cdf of the the *standard* Gaussian random variable,

$$\psi_i \triangleq 1 - F_s\left((\tau_i - \theta)/\sigma_n\right) \quad \text{for} \quad k = \mathcal{K}_i \tag{16}$$

for $i \in \{1, 2\}$. Given the noise independence across sensors, the ML estimate of ψ_i for $i \in \{1, 2\}$ is found as the solutions to EM algorithms operating on $\mathbf{y}_i = \{y(k) : k \in \mathcal{K}_i\}$. Mimicking the derivations in known sensor noise parameter case, we invert the cdf of *standard* $F_s(\cdot)$ and invoke the variance property of ML estimates to obtain the ML estimate of θ [1]:

$$\hat{\theta}_{\mathrm{ML}} = \frac{F_s^{-1}(1 - \hat{\psi}_{2,\mathrm{ML}})\tau_1 - F_s^{-1}(1 - \hat{\psi}_{1,\mathrm{ML}})\tau_2}{F_s^{-1}(1 - \hat{\psi}_{2,\mathrm{ML}}) - F_s^{-1}(1 - \hat{\psi}_{1,\mathrm{ML}})} \tag{17}$$

where $\hat{\psi}_{i,\mathrm{ML}}$ denotes the ML estimates obtained through the EM algorithm.

4. NUMERICAL EXPERIMENTS

This section presents numerical experiments, first analyzing the performance of EM algorithm in WSN problems and second evaluating the performances of EM–based ML estimator with unknown channel parameter (MLU) and the EM–based ML estimator with unknown channel and sensor noise, i.e., the blind ML estimator (MLB). Results are compared to the variance of the clairvoyant estimator (CE–operating directly on the analog observations, to the variance of binary estimator (BE–operating directly on the quantized noisy observations) [1, 3, 4], and to the variance of the ML estimator with known channel and sensor parameters (MLK–operating on the noisy quantized noisy observations) [6].

Consider a fusion center operating in a WSN with parameters, $\tau = 0.5$, $\sigma_n = 1$, $\sigma_w = 0.25$ and $K = 1000$, where

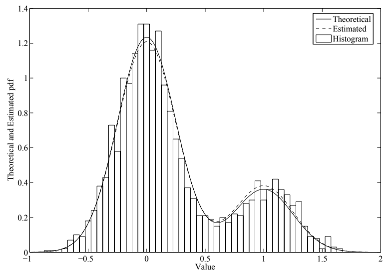

Fig. 1. Histogram of the fusion center observations, and theoretical and estimated pdf of the observed random variable.

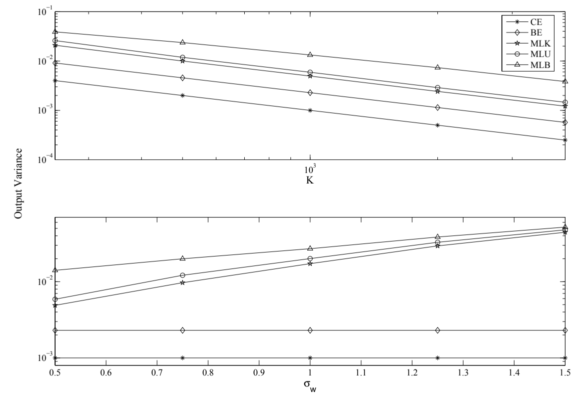

Fig. 2. Variances of MLB, MLU, MLK, BE and CE for (top:) varying K and (bottom:) varying σ_w.

the source parameter to be estimated is $\theta = 0.25$. The channel noise is unknown at the fusion center whereas the sensor noise is known. The (normalized) histogram of the fusion center observations are given in Fig. 1. The true pdf of the observed random variable y, along with the estimated pdf utilizing the EM procedures are given in Fig. 1. Note that the estimated pdf closely follows the true pdf.

Consider next a fusion center operating in a WSN with parameters $\tau = 0$, $\sigma_n = 1$, $\sigma_w = 0.5$, $\theta = 1$ and $K = \{250, 500, 1000, 2000, 4000\}$. The variances of the MLU and MLB are plotted in Fig. 2, along with the variances of the CE and MLK [6]. Note that the MLB performance loss compared to that of the MLU, and the MLU performance loss compared to that of the MLK are only marginal. Also, the estimators exhibits the expected performance order, i.e., $\phi(\hat{\theta}_{\mathrm{CE}}) < \phi(\hat{\theta}_{\mathrm{BE}}) < \phi(\hat{\theta}_{\mathrm{MLK}}) < \phi(\hat{\theta}_{\mathrm{MLU}}) < \phi(\hat{\theta}_{\mathrm{MLB}})$ where $\phi(\cdot)$ denotes the variance of the corresponding random variable. Also considered is the case where the estimators are operating in a WSN with $\sigma_w = \{0.5, 0.75, 1, 1.25, 1.5\}$. The MLU, MLK, BE and CE variances are plotted in Fig. 2 for $K = 1000$. Note that in this case, the CE and BE variances appear flat since they consider perfect transmission. Also, the performances of MLU and MLK approach the performance of BE as σ_w decreases since the extended WSN scheme, in this case, reduces to a WSN that disregards the transmission errors occurring between the sensors and the fusion center.

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, maximum likelihood estimators operating on corrupted quantized noisy sensor observations, for unknown channel noise and known sensor noise, and for unknown channel and sensor noise cases, are proposed. The missing label information makes the addressed problem a typical task with an incomplete data set, which we approach from a expectation–maximization (EM) perspective. The critical initialization and convergence aspects of the EM algorithm are investigated. Numerical examples showed that the near–optimal performance of the algorithms and determined that the variance of the estimators increase only marginally compared to the estimator with known channel parameter/known channel and sensor parameter.

6. REFERENCES

[1] A. Ribeiro and G. B. Giannakis, “Bandwidth-constrained distributed estimation for wireless sensor networks-part ii: Unknown probability density function,” *IEEE Trans. on Signal Proc.*, vol. 54, no. 7, pp. 2784–2796, July 2006.

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

[3] H. Papadopoulos, G. Wornell, and A. Oppenheim, “Sequential signal encoding from noisy measurements using quantizers with dynamic bias control,” *IEEE Trans. on Inf. The.*, vol. 47, no. 3, pp. 978–1002, Mar. 2001.

[4] Z.-Q. Luo, “Universal decentralized estimation in a bandwidth constrained sensor network,” *IEEE Trans. on Inf. The.*, vol. 51, no. 6, pp. 2210–2219, June 2005.

[5] A. Dempster, N. Laird, and D. Rubin, “Maximum likelihood from incomplete data via the em algorithm,” *Journal of the Royal Statistical Society, Series B*, vol. 39, no. 1, pp. 1–38, Nov. 1982.

[6] T. C. Aysal and K. E. Barner, “Constrained decentralized estimation over noisy channels for sensor networks,” *IEEE Trans. on Signal Proc.*, June 2006, submitted for publication.