THE IMPACT OF FADING ON DECENTRALIZED DETECTION IN POWER CONSTRAINED WIRELESS SENSOR NETWORKS

Jean-François Chamberland and Venugopal V. Veeravalli

Department of Electrical and Computer Engineering
Coordinated Science Laboratory, University of Illinois
1308 W Main St, Urbana, IL 61801
chmbrlnd,vvv@uiuc.edu

ABSTRACT

We study a binary decentralized detection problem in which a set of sensor nodes provides partial information about the state of nature to a fusion center. Sensor nodes have access to conditionally independent observations, given the state of nature, and they transmit their data over separate wireless channels. The communication link between each node and the fusion center is subject to fading, with certain nodes possibly having much better connections than others. Upon reception of the information, the fusion center attempts to accurately reconstruct the state of nature. Large deviation theory is employed to obtain design guidelines for wireless sensor networks with a large number of nodes. The normalized Chernoff information is shown to be an appropriate performance metric to compare prospective sensor nodes. For the specific example of binary sensor nodes sending data over Rayleigh fading channels, the performance loss due to fading is found to be small.

1. INTRODUCTION

In recent years, signal processing systems with distributed sensor nodes have been gaining in popularity as means to collect, analyze, and transmit environmental data. This increasing interest in distributed sensing springs partly from the projected affordability of miniature sensing technology along with the wide availability of the computing resources necessary to handle complex environmental data. A typical wireless sensor network is one where a set of geographically dispersed sensor nodes gathers information about the properties or the likely occurrence of an event of interest. Each node sends a compressed version of its observation to a central entity, termed fusion center, for purpose of detection and control. Distributed sensing provides redundant and hence reliable information about the observed environment.

This research was supported by the National Science Foundation under the CAREER/PECASE award CCR-0049089.

Wireless sensor networks vary widely in structure and function. They will presumably have significant impact in many fields; including civil surveillance, national security, health care, and manufacturing. Detection often serves as the initial goal of a sensor network. For instance, the presence of an object has to be ascertained before the network can estimate its attributes, like position and velocity. For systems observing infrequent events, detection may very well be the prevalent function of the system. Furthermore, in some applications such as surveillance, the detection of an intruder is the sole purpose of the sensing system.

The topic of decentralized detection has received much attention in the past. The reader is referred to Tsitsiklis [1] and to the references contained therein for a survey of the early work in the field. In the canonical decentralized detection problem, a set of dispersed sensor nodes receives information about the state of nature H. Sensor node S_ℓ then selects one of D_ℓ possible messages and sends it to the fusion center via a noiseless channel. Based on the received data, the fusion center produces an estimate of the state of nature.

Most of the research on decentralized detection has been centered around finding optimal decision rules for sensor nodes and their fusion center. A celebrated accomplishment in this field has been the demonstration that, for the canonical decentralized detection problem, likelihood ratio tests at the sensor nodes are optimal when the observations are conditionally independent, conditioned on H (see [1]). This assumption of conditional independence is reasonable if the limited accuracy of the sensor nodes is responsible for noisy observations. However, if the observed signal is stochastic in nature or if the sensors are subject to external noise, this assumption may fail. Whether conditional independence holds or not, finding optimal quantization rules for the sensor nodes remains, in most cases, a hard problem.

The canonical decentralized detection problem formulation has limited application to wireless sensor networks, as the model disregards important features of wireless sensing technology and overlooks intrinsic aspects of wireless com-

0-7803-8484-9/04/$20.00 ©2004 IEEE

munication. For example, it assumes that messages from the sensor nodes are conveyed without error to the fusion center. This is hardly the case when information is transmitted over noisy channels [2]. Evidently, the performance of a sensor network may depend on the nature of the wireless environment available to the sensor nodes. In our previous work [3], we proposed a more adequate model for decentralized detection over wireless sensor networks. That model incorporates physical characteristics of the wireless channel employed by the sensor nodes. Specifically, we extended existing asymptotic results about large sensor networks to the case where the nodes are subject to a total power constraint and the communication channel from each node to the fusion center is corrupted by additive noise. Analysis derived from this limiting regime, where the number of sensors in the network becomes large, is especially relevant in the context of sensor networks since such networks are envisioned to contain in excess of thousands of nodes.

In this paper, we study the impact of fading on the performance of distributed sensing systems. Wireless channels are prone to attenuation and fading. If sensor nodes are to be scattered around somewhat randomly, it is conceivable that their respective communication channels will experience fading, with certain nodes possibly having much better connections than others. It is then of interest to quantify the performance loss due to fading in sensor networks.

2. STATEMENT OF THE PROBLEM

Let $\{Y_\ell\}$ be a sequence of random observations. Sensor node S_ℓ has access to observation Y_ℓ and is required to send a summary of its own observation to the fusion center. The information provided by sensor node S_ℓ is transmitted over a wireless communication channel. The fusion center then receives degraded information U_ℓ of the form

$$U_\ell = \Theta_\ell \gamma_\ell(Y_\ell) + W_\ell, \tag{1}$$

where $\gamma_\ell(Y_\ell)$ is the sent signal, Θ_ℓ is the channel gain, and W_ℓ is additive noise. We consider the problem where the observations are independent and identically distributed, conditioned on the state of nature, and where the channel gains and the communication noise are also independent and identically distributed across sensor nodes. Furthermore, we assume that the channel gains $\{\Theta_\ell\}$ are known at the fusion center but not at the sensor nodes. This setting is illustrated in Figure 1. We note that the total number of active sensor nodes L is not fixed a priori in this framework. An alternative problem formulation in which the channel state information is not available at the fusion center can be handled in a similar manner; this is exemplified in Section 4. The binary hypothesis testing problem consists of deciding, based on the data received at the fusion center, whether the law generating $\{Y_\ell\}$ is $\mathcal{P}_0$ corresponding to hypothesis H_0, or $\mathcal{P}_1$ corresponding to H_1.

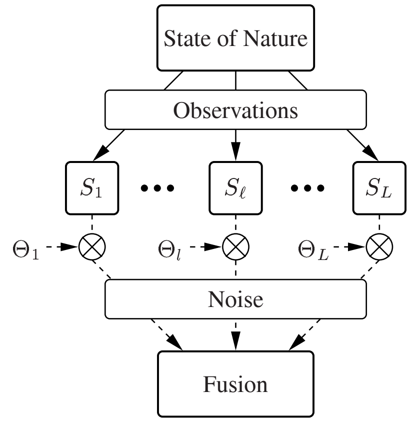

Fig. 1. Block diagram of a wireless sensor network where sensor nodes communicate over a wireless channel.

Let γ represent a function that maps an observation to a transmit signal, and denote the set of all such functions by Γ. A transmission strategy $\mathcal{G}$ is a vector function of the form $(\gamma_1, \ldots, \gamma_L)$ with the interpretation that, upon observing $Y_\ell = y$, sensor node S_ℓ transmits summary $\gamma_\ell(y)$ to the fusion center. We assume that the probability measures $\mathcal{P}_0$ and $\mathcal{P}_1$ are known beforehand, that they are mutually absolutely continuous, and that they are distinguishable. We also assume that the probability distributions of the channel gain and the additive noise are known beforehand.

We consider the problem where the channel gain of every sensor node is known at the fusion center. The fusion center then makes a decision based on the pairs $\{(U_\ell, \Theta_\ell)\}$, one pair per node. Under hypothesis H_i, the function γ induces a probability law $\mathcal{Q}_{i,\gamma}$ on the reception space $(\mathcal{U} \times \mathbf{\Theta})$. A decision test $\mathcal{S}$ is a map $\mathcal{S} : (\mathcal{U} \times \mathbf{\Theta})^L \mapsto \{0, 1\}$ such that hypothesis H_1 is accepted with probability $p \in [0, 1]$ when $\mathcal{S}((u_1, \theta_1), \ldots, (u_L, \theta_L)) = p$; and hypothesis H_0 is accepted otherwise. The performance of decision test $\mathcal{S}$ is characterized by the error probabilities

$$\alpha = \int_{(\mathcal{U} \times \mathbf{\Theta})^L} \mathcal{S}((U_1, \Theta_1), \ldots (U_L, \Theta_L)) d\mathcal{Q}_{0,\mathcal{G}}$$
$$\beta = \int_{(\mathcal{U} \times \mathbf{\Theta})^L} [1 - \mathcal{S}((U_1, \Theta_1), \ldots (U_L, \Theta_L))] \, d\mathcal{Q}_{1,\mathcal{G}},$$

where $\mathcal{Q}_{i,\mathcal{G}}$ denotes the product measure on $(\mathcal{U} \times \mathbf{\Theta})^L$ induced by strategy $\mathcal{G}$ under hypothesis H_i along with the probability measure corresponding to the channel state. We emphasize that the expected values α and β are taken over both the aggregate observation space $\mathcal{U}^L$ and the possible realizations of the channel vector $(\Theta_1, \ldots, \Theta_L)$. Taking an expectation over the channel gains is justified when the realization of the channel vector is unknown to the designer of the system, but becomes available to the fusion center once the system is deployed. The likelihood ratio between $\mathcal{Q}_{1,\gamma}$ and $\mathcal{Q}_{0,\gamma}$ is given by

$$\mathcal{L}_{\mathcal{Q},\gamma}(u,\theta) = \frac{d\mathcal{Q}_{1,\gamma}}{d\mathcal{Q}_{0,\gamma}}(u,\theta). \tag{2}$$

Likelihood ratio tests at the fusion center are known to be optimal in minimizing the Bayes probability of error.

We focus on the detection problem where the wireless sensor network is subject to a total power constraint A. Let a priori probabilities $\mathrm{P}(H_0)$ and $\mathrm{P}(H_1)$ be given. An admissible transmission strategy $\mathcal{G}$ is a vector function $(\gamma_1, \ldots, \gamma_L)$ such that $F(\mathcal{G}) = \sum_{\ell=1}^{L} f(\gamma_\ell) \leq A$, where $f(\gamma_\ell) > 0$ denotes the expected power consumed by sensor node S_ℓ. Given a strategy $\mathcal{G}$, define

$$P_e(\mathcal{G}) = \inf_{\mathcal{S}} \left[\alpha \mathrm{P}(H_0) + \beta \mathrm{P}(H_1)\right],$$

where the infimum is over the set of all decision tests. We want to select an admissible transmission strategy $\mathcal{G}$ such that the Bayes probability of error at the fusion center $P_e(\mathcal{G})$ is minimized.

Being interested in large sensor networks, we consider the case where the power constraint tends to infinity. This corresponds to the asymptotic regime where the number of nodes and, possibly, the area covered by those nodes grow to infinity. The goal in considering this regime is to capture the dominant features of the studied sensor network, and to exploit these features to derive design guidelines for the original system. For any reasonable collection of transmission strategies, the Bayes probability of error at the fusion center goes to zero exponentially fast as $A \to \infty$. It is then natural to compare sets of strategies based on their rate of convergence to zero,

$$\liminf_{A \to \infty} \frac{\log P_e(\mathcal{G}_A)}{A}.$$

Throughout, we use $\mathcal{G}_A$ as a convenient notation for a transmission strategy with total power $F(\mathcal{G}_A) \leq A$.

3. BASIC CONCEPTS AND RESULTS

Let G be a finite subset of Γ, and let $G^{\mathbb{N}}$ be the set of transmission strategies of the form $\mathcal{G} = (\gamma_1, \ldots, \gamma_L)$, where $L \in \mathbb{N}$ and $\gamma_\ell \in G$ for all ℓ. That is, $G^{\mathbb{N}}$ is the set of all strategies with a finite number of sensor nodes where each node employs a transmission mapping γ_ℓ contained in G. Similarly, $\{\gamma\}^{\mathbb{N}}$ is the set of strategies for which all sensor nodes use transmission mapping γ.

Theorem 1. *Fix finite set $G \subset \Gamma$. Using identical functions at all the sensor nodes is asymptotically optimal,*

$$\liminf_{A \to \infty} \min_{\mathcal{G}_A \in G^{\mathbb{N}}} \frac{\log P_e(\mathcal{G}_A)}{A} = \min_{\gamma \in G} \liminf_{A \to \infty} \min_{\mathcal{G}_A \in \{\gamma\}^{\mathbb{N}}} \frac{\log P_e(\mathcal{G}_A)}{A}.$$

Proof. This Theorem can be established rigorously using results from large deviation theory [4]. Although the proof is somewhat technical, the basic ideas contained in it are simple. First, we use Cramér's theorem to show that the Chernoff bound on the probability of error is asymptotically tight. Then, we prove that the class of strategies with identical maps is optimal in maximizing the error exponent. $\square$

In our previous work [3], we studied a simpler version of this theorem in which the wireless channels between the nodes and the fusion center are not subject to fading, i.e., $\Theta_\ell = 1$ for $1 \leq \ell \leq L$. This more general version of the theorem allows us to consider the effects of fading when channel state information is known at the fusion center.

In establishing Theorem 1, we get a useful corollary. For networks with large power budgets, prospective sensor nodes should be compare according to their normalized Chernoff information

$$-\frac{1}{f(\gamma)} \min_{\lambda \in [0,1]} \left\{ \log \mathrm{E}_{\mathcal{Q}_{0,\gamma}} \left[\left(\mathcal{L}_{\mathcal{Q},\gamma}(U,\Theta)\right)^{\lambda} \right] \right\}.$$

The normalized Chernoff information captures the tradeoff between energy consumption and information in large wireless sensor networks. Allocating more power per sensor node implies receiving more reliable information from each node. However, for a fixed power constraint A, a reduction in power consumption per node permits more active sensor nodes in the network. The normalized Chernoff information makes this tradeoff precise: doubling the Chernoff information provided by each sensor node results in the same gain in overall performance as reducing the power consumption per node by half and doubling the number of nodes.

4. NUMERICAL EXAMPLES

In this section, we use the normalized Chernoff information to compare the performance loss associated with fading. We study the simple scenario where sensor nodes have access to Gaussian observations,

$$p_{Y|H}(y|H_0) \sim \mathcal{N}\left(-m_y, \sigma_y^2\right)$$
$$p_{Y|H}(y|H_1) \sim \mathcal{N}\left(m_y, \sigma_y^2\right).$$

The fusion center receives a noisy version of the data transmitted by the sensor nodes, as described in equation (1). We assume that the communication noise W_ℓ is Gaussian, $\mathcal{N}\left(0, \sigma^2\right)$, and that the wireless connection Θ_ℓ is subject to Rayleigh fading: $f_\Theta(\theta) = 2\theta e^{-\theta^2}$, $\theta \geq 0$. Based on the received data, the fusion center admits one of the two hypotheses.

Finding a mapping γ that maximizes the normalized Chernoff information over all admissible functions is, in general, difficult. However, the normalized Chernoff information can be employed to assess the performance of a network where an optimal transmission mapping γ is to be selected from a reasonable collection of candidates G. For purpose of illustration, we consider the case where each node computes and sends a one-bit summary of the form

$$\gamma_a(y) = \begin{cases} a & : y \geq 0 \\ -a & : y < 0 \end{cases}, \quad (3)$$

where $a > 0$. When the channel gains are known at the fusion center, the measures on the reception space have probability density functions given by

$$\mathcal{Q}_{0,\gamma_a}(u,\theta) = \mathcal{Q}_{1,\gamma_a}(-u,\theta) = \frac{2\theta e^{-\theta^2}}{\sqrt{2\pi\sigma^2}} \left[Q\left(-\frac{m_y}{\sigma_y}\right) e^{-\frac{(u+a\theta)^2}{2\sigma^2}} + Q\left(\frac{m_y}{\sigma_y}\right) e^{-\frac{(u-a\theta)^2}{2\sigma^2}} \right]$$

where Q is the complementary Gaussian cumulative distribution function. The corresponding normalized Chernoff information is equal to

$$-\frac{1}{a^2} \log \left(\int_0^\infty \int_{-\infty}^\infty \sqrt{\mathcal{Q}_{0,\gamma_a}(u,\theta)\mathcal{Q}_{1,\gamma_a}(u,\theta)}\, du\, d\theta \right).$$

A similar derivation can be performed for the case where the channel state information is not known at the fusion center. In this latter setting, the effects of the channel gains are averaged out and the likelihood ratio of equation (2) becomes a function of u only. Figure 4 shows the normalized Chernoff information for the transmission map of equation (3). The plot contains three cases: Rayleigh fading with and without channel state information, and communication without fading. As seen in the figure, fading reduces overall performance. Yet, access to channel state information at the fusion center mitigates this performance loss. More importantly, we note that the quality of the observations has a much greater impact on system performance than fading.

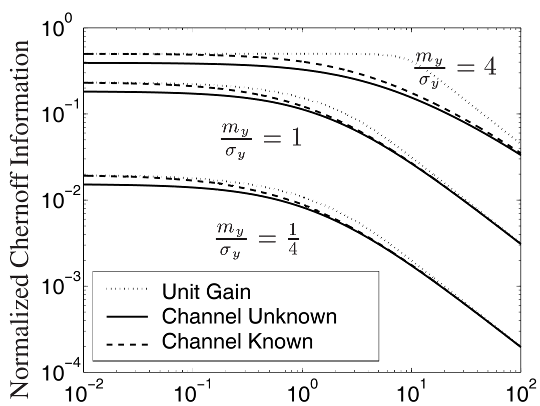

Fig. 2. Normalized Chernoff information corresponding to noise variance $\sigma^2 = 1$ and radiated power $f(\gamma_a) = a^2$.

5. DISCUSSION AND CONCLUSION

The binary decentralized detection problem in which nodes provide relevant information to the fusion center for purpose of detection was considered. We studied the problem where observations are conditionally independent, where sensor nodes send data over fading channels, and where the network is subject to a total power constraint. Using identical sensor nodes was found to be asymptotically optimal, as the power constraint and the number of nodes approach infinity. Moreover, the normalized Chernoff information was found to be an appropriate performance metric to compare and evaluate sensor node candidates. Numerical results suggest that fading only degrades performance slightly. The quality of the observations has a much greater impact on the probability of error at the fusion center than fading.

We focused on the setting where the channel gains are known at the fusion center. This assumption can easily be relaxed, as seen in Section 4. The case where the observations at the sensor nodes, or where the gains of the wireless channels, are not identically distributed across nodes is more difficult. Each node can then tailor its transmission function based on the distributions of its observation and channel. Symmetry is lost and, therefore, identical transmission functions are no longer necessarily optimal. This difficulty is due to the fact that the Chernoff information is non-additive. Hence, transmission functions must be selected jointly to maximize overall system performance. As a concluding remark, we mention briefly that similar mathematical tools can be used to solve the Neyman-Pearson variant of this detection problem. In the alternate formulation, the normalized Kullback-Leibler divergence becomes the appropriate performance metric for node selection.

6. REFERENCES

[1] J.N. Tsitsiklis, "Decentralized detection," *Advances in Statistical Signal Processing*, vol. 2, pp. 297–344, 1993.

[2] T.M. Duman and M. Salehi, "Decentralized detection over multiple-access channels," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 34, no. 2, pp. 469–476, April 1998.

[3] J.-F. Chamberland and V. V. Veeravalli, "Asymptotic results for detection in power constrained wireless sensor networks," in *The IEEE International Symposium on Information Theory Proceedings*, Yokohama, Japan, June 2003.

[4] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Jones and Barlett Publishers, Boston, MA, 1993.