HIERARCHICAL CENSORING FOR DISTRIBUTED DETECTION IN WIRELESS SENSOR NETWORKS

Neal Patwari and Alfred O. Hero III

University of Michigan, Dept. of EECS
Ann Arbor MI, USA
E-mail: [npatwari, hero]@eecs.umich.edu.

ABSTRACT

In energy-limited wireless sensor networks, detection using 'censoring sensors' reduces the probability that a sensor must transmit, thereby saving energy. In this paper, we introduce a hierarchical distributed detection scheme designed specifically for multi-hop networks. If a sensor's local likelihood ratio (LLR) crosses a threshold, it is sent to the next higher level sensor. A simple feedback scheme is also considered. We study the performance of a Gaussian change-of-mean detection system using this hierarchical censoring scheme, with and without feedback. We show that good detection performance can be achieved while significantly reducing sensor transmissions compared to the optimal detection system.

1. INTRODUCTION

Large-scale wireless sensor networks are envisioned to monitor wide environments without network management for long lifetimes. Many example applications, such as temperature and VOC monitoring in buildings, moisture and fertilizer level sensing in agricultural fields, and detection of intruders across borders, involve detection of events. Distributed detection using multiple sensors has been studied extensively in the literature [1][2]. This paper applies a distributed detection framework to wireless sensor networks, which have the peculiarities of being energy-constrained and multi-hop. We argue that these characteristics lead naturally to a hierarchical topology and a "censoring sensors" [3] strategy. We present analysis of censoring in a hierarchical topology and suggest a simple feedback scheme. Finally, we show numerical results for a Gaussian change-of-mean detection system with and without the feedback scheme.

1.1. Energy Constraint

Energy is of primary concern in wireless sensor networks [4][5][6]. In most applications, the bit rate is very low, often less than one bit per second [5], and the bandwidth is wide. IC costs will fall with Moore's law, however, battery costs will remain relatively constant, thus energy will become an increasing fraction of the sensor cost. Economical deployment of thousands of sensors will require aggressive energy limitation. Equivalently, given a battery size, minimizing energy consumption maximizes system lifetime. If energy consumption can be sufficiently reduced, solar power or energy-mining can be used to power each sensor [5].

This material is based in part upon work supported under a NSF Graduate Research Fellowship.

Wireless sensor networks use low duty cycles, ie. low percentage of device 'on' time, on the order of 0.01% to 1%, allowing circuits to remain in a sleep state the vast majority of the time. When necessary, a device wakes up its sensor, processor, transmitter or receiver in order to sense, process, receive or transmit a message. Each wake-up consumes significant energy. Specifically, for the transmitter circuitry, it has been reported that there is a tradeoff between the time required for (and thus energy expended during) wake-up and the energy used during sleep mode [5]. Due to the large percentage of time spent in sleep mode, sleep mode energy is minimized, and as a result, wake-up energy is high. It has also been reported that wake-up energy can be significantly higher than the energy used during transmission [7].

Much distributed detection research has focused on capacity-constrained networks. Research has addressed quantization of sensor data [8] and exploiting source correlation [9] to reduce sensor bit rate. In particular cases, it has been shown that for a R sensor network with a capacity constraint of R bits per unit time, having each sensor send one bit is optimal [10].

However, from the perspective of energy, the cost of transmitting one bit involves wake-up energy and packet overhead such as synchronization and id bits. Considering all energy costs in an energy budget as in [6] shows that sending one bit of data consumes only marginally less energy than sending several bits. In fact, it can be argued that the appropriate constraint to bound energy consumption for many wireless sensor networks is the probability of transmission from each sensor, as used in [11]. In this paper, a sensor's decision regarding whether or not to transmit its data is a local decision based on its LLR [3].

1.2. Hierarchical Networks

In wireless sensor networks, due to devices' limited range, communication to a fusion center must be routed through intermediate devices in the network. Multi-hop is used for energy-efficiency and reduced device cost. If required transmit power $\propto 1/r^2$, where r is the path length, then total transmit energy is decreased by using multiple short hops. Network-wide power savings in multi-hop systems can be significant, especially for large-scale networks when reception energy costs are small compared to transmission costs. Technology scaling should reduce receiver energy consumption, while transmission costs will remain constant [5]. This projection underscores the importance of both multi-hop and minimizing the probability of transmission.

The use of multi-hop in wireless sensor networks can also be exploited for improved detection performance. Often, distributed detection literature assumes that all sensor data is sent to a fusion

0-7803-7663-3/03/$17.00 ©2003 IEEE

center. Rather than simply relaying messages to a fusion center, intermediate nodes can perform data aggregation and make local decisions, preventing a bottleneck at the fusion center. Furthermore, feedback has been suggested to enable sensors to make reliable decisions on certain events [12]. In this paper, we consider a hierarchical or 'spanning tree' topology (eg., Fig. 1) for the purposes of censoring and feedback.

We focus on the case where wireless sensor networks operate in weak signal environments, i.e., $P(H_1)$ is low. High $P(H_1)$ applications are typically incompatible with low-duty cycle operation. Thus we constrain the average probability of sensor transmission given H_0, and then optimize detection performance.

2. DERIVATION OF HIERARCHICAL CENSORING

In a hierarchical network of N sensors, we denote G_k as the set of sensors on level k of the hierarchy, $k = 1 \ldots M$ (eg., $G_2 = \{3, 6\}$ in Fig. 1). Each sensor (except for the fusion center) has a parent node. We denote the set of children of node i as K_i. At a given round of sensing, a sensor records data X_i. We assume X_i are i.i.d. conditional on the hypothesis $H_j, j = 0, 1$.

Censoring sensors was presented in [3] and [13]. The hierarchical version presented here expands censoring to multiple layers of sensors with feedback. Each sensor forms its LLR $L_{F,i}$ from both its own data X_i and the LLRs of its children. The F subscript denotes that $L_{F,i}$ is a fusion of data not only from sensor i but also its children. If $L_{F,i}$ falls in a send region, R_i, then sensor i sends its LLR to its parent. Since the number of data bits is not constrained, we can assume for analysis that the real-valued LLR is sent to the parent unquantized [3]. If the LLR falls in the no-send region, $\bar{R}_i$, sensor i doesn't transmit, and its silence is used as information by its parent. Define the constraint ρ as the mean probability of sensor transmission given H_0,

$$\frac{1}{N-1} \sum_{\substack{i=1 \\ i \neq FC}}^{N} P_0[L_{F,i} \in R_i] \leq \rho \leq 1. \qquad (1)$$

Here, FC is the index of the highest level sensor (7 in Fig. 1). Its false alarm rate is the global P_F that directly determines detection performance, and is constrained independently of the energy constraint. In this paper we use P_j to indicate the probability given H_j, for $j = \{0, 1\}$.

In [3], the optimal censoring region was shown to be a single interval, $\bar{R} = [\nu_i, \tau_i)$. Moreover, in cases where the prior probability of H_1 is sufficiently small and limited communication is allowed, it is optimal to set $\nu_1 = 0$. Sufficient conditions for the optimality of $\nu_i = 0$ are given in [13]. In this analysis, we assume $\nu_i = 0$ because of the assumptions on $P(H_1)$ and limited transmission probability detailed in Section 1. Thus the optimal fusion rule for sensor i's decision whether or not to transmit is a local likelihood ratio test (LLRT) with threshold τ_i,

$$L_{F,i} = L_{l,i} \prod_{j \in \kappa_i} L_{F,j} \times \prod_{j \in \bar{\kappa}_i} c_j \underset{H_0}{\overset{H_1}{\gtrless}} \tau_i \qquad (2)$$

where $L_{l,i}$ is the likelihood ratio based only on the data X_i, the sets $\kappa_i = \{j \in K_i : L_{F,j} \in R_j\}$ and $\bar{\kappa}_i = \{j \in K_i : L_{F,j} \in \bar{R}_j\}$ are the subsets of K_i which send, and do not send their LLR, respectively. Note that for $i \in G_1$, $L_{F,i} = L_{l,i}$ since sensors on level 1 have no children. The constant c_i is the effect on the LLR of a non-transmitting child node,

$$c_i = \frac{P_1(L_{F,i} \in \bar{R}_i)}{P_0(L_{F,i} \in \bar{R}_i)}. \qquad (3)$$

In terms of log-likelihoods $l_{F,i} = \log L_{F,i}$ and $l_{l,i} = \log L_{l,i}$,

$$l_{F,i} = l_{l,i} + \sum_{j \in \kappa_i} l_{F,j} + \sum_{j \in \bar{\kappa}_i} \log c_j \underset{H_0}{\overset{H_1}{\gtrless}} \log \tau_i \qquad (4)$$

Let the density of $l_{F,i}$ and $l_{l,i}$ be given by $f_{F,i}$ and $f_{l,i}$, respectively. Then $f_{F,i}$ can be seen as a mixture pdf. This is derived explicitly for a binary tree in Section 2.2.

2.1. Using Feedback

In this paper, we test a simple feedback scheme in which a device listens for transmissions from its 'siblings' (sensors with the same parent). If one sibling transmits its LLR to the parent, then all of the other siblings also transmit their LLRs to the parent. This scheme doesn't require the parent to transmit back to its children to request feedback. Although feedback from the parent could include more information than feedback from a sibling, analyzing this basic scheme helps to determine when feedback is valuable.

To achieve the same probability of transmission as without feedback, the thresholds τ_i of the LLRTs must be reduced. Eqs. (2)-(4) are still valid, but now, $\bar{\kappa}_i = K_i$ if $L_{F,j} \in \bar{R}_j, \forall j \in K_i$, or $\bar{\kappa}_i = \emptyset$ otherwise. Similarly, $\kappa_i = K_i \cap (\bar{\kappa}_i)^C$. If $\#\{K_i\}$ is the number of children of sensor i, the constraint becomes,

$$\frac{1}{N-1} \sum_{i=1}^{N} \#\{K_i\} P_0\Big[\bigcup_{j \in K_i} L_{F,j} \in R_j\Big] \leq \rho \leq 1, \qquad (5)$$

2.2. Binary Tree Simplification

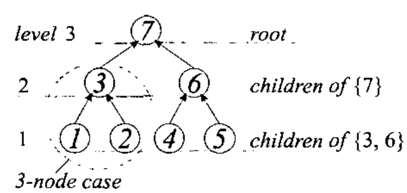

Fig. 1. Diagram of an example hierarchical network of sensors. Subset of sensors $\{1, 2, 3\}$ is also used as an example.

Assume that N sensors are arranged in a binary tree as the $N = 7$ case shown in Fig. 1. We simplify the parameter space by assuming that the thresholds for sensors on level k are identical. We let $P_{FT}(k)$ be the probability of false transmission from a sensor on level k. Without feedback, $P_{FT}(k) = P_0[L_{F,i} \in R_i]$, for any $i \in G_k$. In the feedback case, $P_{FT}(k) = 1 - (1 - P_0[L_{F,i} \in R_i])^2$, for any $i \in G_k$. The constraint from (1) is now,

$$\frac{1}{N-1} \sum_{k=1}^{M-1} 2^{M-k} P_{FT}(k) \leq \rho \leq 1. \qquad (6)$$

The density of sensor i's LLR given H_j can be shown to be,

$$f_{F,i|H_j} = \epsilon_0 g_{0|H_j} + \epsilon_1 g_{1|H_j} + \epsilon_2 g_{2|H_j}, \qquad (7)$$

where ϵ_k is the probability that k children of sensor i have LLRs in the send region, and $g_{k|H_j}$ is the density of the $l_{F,i}$ given H_j and given k children have LLRs in the send region. The binomial probability ϵ_k is,

$$\epsilon_k = \binom{2}{k} \left[P_j(l_{F,u} \in R_u)\right]^k \left[P_j(l_{F,u} \in \bar{R}_u)\right]^{2-k}, \quad (8)$$

where $u \in K_i$. The densities $g_{k|H_j}$ of (7) can be calculated via convolution. Defining $S_u = \{l_{F,u} \in R_u\}$ and $\bar{S}_u = \{l_{F,u} \in \bar{R}_u\}$, the densities of $f_{F,u}$ conditioned on H_j and S_u or $\bar{S}_u$ are,

$$f_{F,u|H_j,S_u}(t) = \begin{cases} \frac{f_{F,u|H_j}(t)}{P_j(l_{F,u} \in R_u)} & t \in R_u \\ 0 & t \in \bar{R}_u \end{cases} \quad (9)$$

$$f_{F,u|H_j,\bar{S}_u}(t) = \begin{cases} 0 & t \in R_u \\ \frac{f_{F,u|H_j}(t)}{P_j(l_{F,u} \in \bar{R}_u)} & t \in \bar{R}_u \end{cases} \quad (10)$$

In the case without feedback, the densities $g_{k|H_j}(t)$ are,

$$\begin{aligned} g_{0|H_j}(t) &= f_{l,i|H_j}(t - 2\log c_u) \quad (11) \\ g_{1|H_j}(t) &= f_{l,i|H_j}(t - \log c_u) \star f_{F,u|H_j,S_u}(t) \\ g_{2|H_j}(t) &= f_{l,i|H_j}(t) \star f_{F,u|H_j,S_u}(t) \bullet f_{F,u|H_j,S_u}(t), \end{aligned}$$

where $\star$ indicates convolution. Note that $g_{2|H_j}(t)$ is the $f_{l,i|H_j}$ density convolved twice with the density in (9), since it represents the density of the sum of sensor i's own LLR and the LLRs of two children given both send their data.

In the feedback case, it can be shown that (11) holds except that $g_{1|H_j}(t)$ is replaced by,

$$g_{1|H_j}(t) = f_{l,i|H_j}(t) \star f_{F,u|H_j,S_u}(t) \star f_{F,u|H_j,\bar{S}_u}(t). \quad (12)$$

In this case, $g_{1|H_j}(t)$ is the $f_{l,i|H_j}$ density convolved with both densities in (9) and (10) since it represents the density of the sum of sensor i's own LLR and the LLRs of two children given exactly one child with LLR in the send region. Derivations of analytical results for high N are complicated by multiple convolutions with (9) and (10). Approximations exist for Gaussian data, but for brevity, we report numerical results.

3. NUMERICAL RESULTS

Consider the Gaussian change-of-mean detection system,

$$\begin{aligned} H_0 &: X_i \sim \mathcal{N}(0, \sigma^2) \\ H_1 &: X_i \sim \mathcal{N}(\mu, \sigma^2) \end{aligned} \quad (13)$$

In this example, the log-likelihood for sensor data X_i,

$$l_{l,i} = \frac{\mu}{\sigma^2} X_i - \frac{\mu}{2\sigma^2}, \quad (14)$$

is also Gaussian. In this section, $\mu = 1$ and $\sigma^2 = 1$ are used.

First, consider the 3-sensor hierarchy shown as a sub-tree in Fig. 1. We set the thresholds on level 1 to achieve a given probability of false transmission for sensors 1 and 2, $\rho = P_{FT}(1)$. We calculate via Matlab the densities $g_{k|H_j}(t)$ by sampling the required densities in (11) and (12) and using numerical convolution. The weights ϵ_k are calcalated from (8), and then (7) is used to find the densities of the LLR at sensor 3, $f_{F,3|H_0}$ and $f_{F,3|H_1}$.

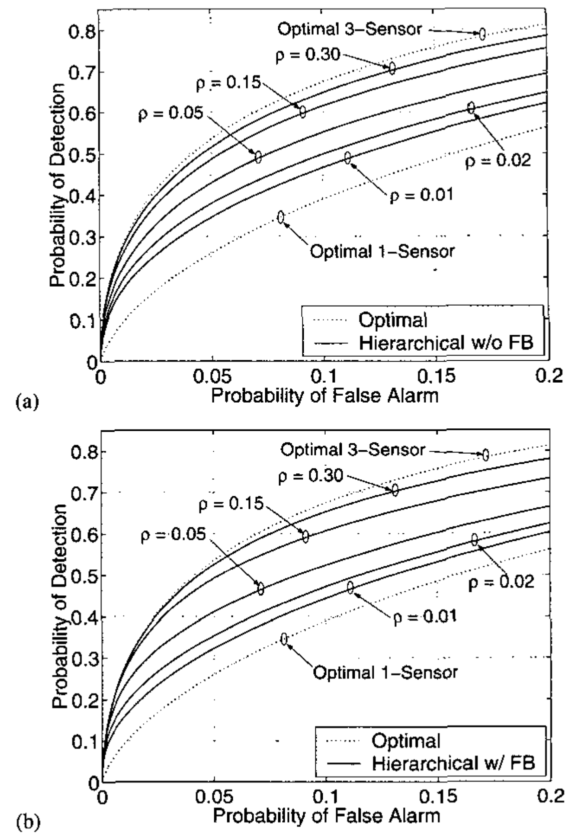

Fig. 2. ROCs for 3-sensor example (a) without feedback and (b) with feedback. Plots show the optimal 1 and 3-sensor ROC and the hierarchical ROC for $\rho = 0.30, 0.15, 0.05, 0.02$, and 0.01.

These densities are used to calculate P_F and P_D, and the ROCs of the final decision are shown in Fig. 2 for cases with and without feedback. Values $\rho = P_{FT}(1) = 0.30, 0.15, 0.05, 0.02$, and 0.01 are shown. Optimal 3-sensor and 1-sensor cases are shown as bounds on the system performance. Note that with $\rho = 0.30$, there is a 70% reduction in transmission under H_0, but the detection performance remains close to optimal, especially at low P_F. Even with the two level 1 sensors transmitting only $1/100$ of the time, performance is significantly better than with only one sensor. Comparing Figs. 2(a) and 2(b), feedback allows performance closer to 3-sensor optimal at low P_F and high ρ, but noticeably degrades performance at high P_F and low ρ.

For the 7-sensor hierarchy in Fig. 1, two thresholds must be set. To meet the energy constraint in (6), $\rho = 4P_{FT}(1) + 2P_{FT}(2)$. Although τ_i for $i \in G_1$ can be set analytically to achieve a particular $P_{FT}(1)$, τ_i for $i \in G_2$ must be set using the numerically calculated density $f_{F,3|H_0}$. Several combinations of $\mathbf{P}_{FT} = [P_{FT}(1), P_{FT}(2)]$ can be tested in order to maximize P_D for a given P_F. In Fig. 3, three combinations that meet $\rho = 0.1$ are tested in the feedback case: $\mathbf{P}_{FT} = [0.05, 0.2]$, $[0.1, 0.1]$, and $[0.12, 0.06]$. At very low P_F, it is best to set $P_{FT}(1) > P_{FT}(2)$, while at high P_F, it is best to set $P_{FT}(1) < P_{FT}(2)$. In this case,

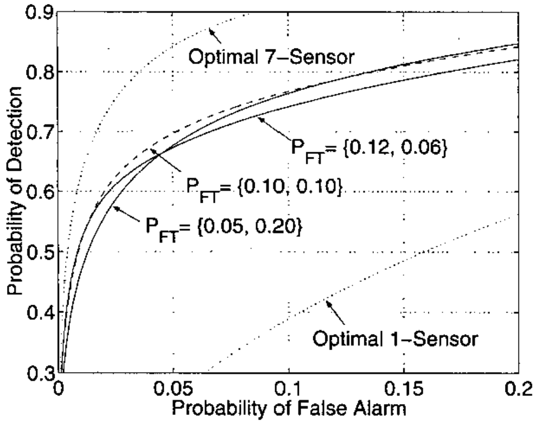

Fig. 3. 7-sensor system with feedback and with $\rho = 0.10$ for various combinations of $P_{FT}(1)$ and $P_{FT}(2)$.

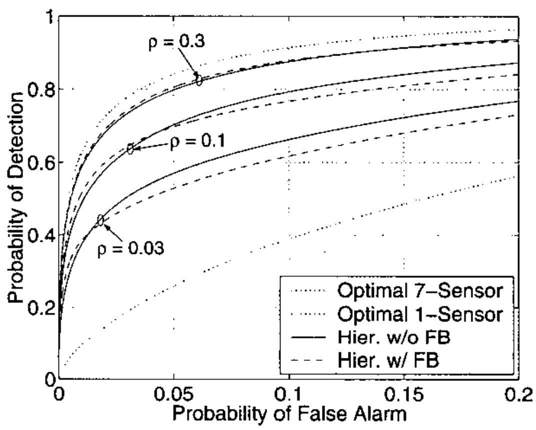

Fig. 4. 7-sensor system with and without feedback for $\rho = \{0.30, 0.10, 0.03\}$. $P_{FT}(1) = P_{FT}(2) = \rho$ for all three cases.

equal probability of false transmission on each level has very good overall performance. Letting $P_{FT}(k)$ be equal $\forall k$ also ensures an equal rate of energy consumption for all sensors in the network.

In Fig. 4, we plot the ROC results for both with and without feedback cases when $P_{FT}(1) = P_{FT}(2) = \rho$ is set to 0.30, 0.10, or 0.03. Similar to the 3-sensor case, the feedback scheme results in increased P_D at low P_F, but has significantly lower P_D when ρ is small and P_F is high.

4. CONCLUSION

In this paper, we have applied censoring sensors to a hierarchical framework. This framework will be increasingly important in the design of energy-limited wireless sensor networks used for distributed detection. We have shown in a Gaussian change-of-mean detection example that close to optimal detection performance can be achieved when sensors may only transmit a fraction of their sensor data. We introduced a simple feedback scheme that can improve detection performance at low P_F. However, more general use of feedback must be studied in order to determine the best use of the mechanism. Furthermore, adapting the hierarchy of a network should be considered for possible energy savings and resilience to sensor failures.

5. REFERENCES

[1] Ramanarayanan Viswanathan and Pramod K. Varshney, "Distributed detection with multiple sensors: Part I–fundamentals," *Proc. IEEE*, vol. 85, no. 1, pp. 54–63, Jan. 1997.

[2] Rick S. Blum, Saleem A. Kassam, and H. Vincent Poor, "Distributed detection with multiple sensors: Part IIadvanced topics," *Proc. IEEE*, vol. 85, no. 1, pp. 64–79, Jan. 1997.

[3] Constantino Rago, Peter Willett, and Yaakov Bar-Shalom, "Censoring sensors: a low-communication-rate scheme for distributed detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 2, pp. 554–568, Apr. 1996.

[4] Jose A. Gutierrez, Marco Naeve, Ed Callaway, Monique Bourgeois, Vinay Mitter, and Bob Heile, "IEEE 802.15.4: A developing standard for low-power low-cost wireless personal area networks," *IEEE Network*, pp. 12–19, Sept.-Oct. 2001.

[5] Jan M. Rabaey, M. Josie Ammer, Julio L. da Silva Jr, Danny Patel, and Shad Roundy, "Picoradio supports ad hoc ultra-low power wireless networking," *IEEE Computer*, pp. 42–48, July 2000.

[6] Vijay Raghunathan, Curt Schurgers, Sung Park, and Mani B. Srivastava, "Energy-aware wireless microsensor networks," *IEEE Signal Processing Mag.*, pp. 40–50, Mar. 2002.

[7] Anantha Chandrakasan, Rajeevan Amirtharajah, SeongHwan Cho, James Goodman, Gangadhar Konduri, Joanna Kulik, Wendi Rabiner, and Alice Wang, "Design considerations for distributed microsensor systems," in *IEEE Custom Integrated Circuits Conf.*, May 1999, pp. 279–286.

[8] Maurizio Longo, Tom D. Lookabaugh, and Robert M. Gray, "Quantization for decentralized hypothesis testing under communication constraints," *IEEE Trans. on Inform. Theory*, vol. 36, no. 2, pp. 241–255, Mar. 1990.

[9] S. Sandeep Pradhan, Julius Kusuma, and Kannan Ramchandran, "Distributed compression in a dense microsensor network," *IEEE Signal Processing Mag.*, pp. 51–60, Mar. 2002.

[10] Jean-Francois Chamberland and Venugopal V. Veeravalli, "Decentralized detection in wireless sensor networks," in *2002 Conf. on Info. Sciences and Systems*, Princeton University, Mar. 2002.

[11] Teerasit Kasetkasem and Pramod K. Varshney, "Communication structure planning for multisensor detection systems," *IEE Proc. Radar, Sonar Navig.*, vol. 148, no. 1, pp. 2–8, Feb. 2001.

[12] Gregory J. Pottie, "Wireless sensor networks," in *IEEE Info. Theory Workshop 1998*, June 1998, pp. 42–48.

[13] Swaroop Appadwedula, Venugopal V. Veeravalli, and Douglas L. Jones, "Robust and locally-optimum decentralized detection with censoring sensors," in 5^{th} *Int. Conf. on Information Fusion*, July 2002, vol. 1, pp. 56–63.