DECENTRALIZED DETECTION IN WIRELESS SENSOR NETWORKS WITH CHANNEL FADING STATISTICS

Bin Liu

Syracuse University
biliu@ecs.syr.edu

Biao Chen

Syracuse University
bichen@ecs.syr.edu

ABSTRACT

Existing channel aware signal processing design for decentralized detection in wireless sensor networks typically assumes the clairvoyant case, i.e., global information regarding the transmission channels is known at the design stage. In this paper, we consider the distributed detection problem where only the channel fading statistics, instead of the instant channel state information (CSI), is available to the designer. We investigate the design of local decision rules for the following two cases: 1. Fusion center has the instant CSI; 2. Fusion center does not have the instant CSI. We show that, for both cases, the optimal local decision rules that minimize the error probability at the fusion center amount to a likelihood-ratio test (LRT), as in the previous work with known CSI. The proposed approach enables distributed design for a decentralized detection problem.

1. INTRODUCTION

While study of decentralized decision making can be traced back to the early 1960s in the context of team decision problems (see, e.g., [1]), the effort significantly intensified since the publication of [2]. In [2], Tenney and Sandell formulated the distributed detection problem using a Bayesian setting and showed that, for a two-sensor case and under the conditional independence assumption, the optimal local sensor decisions are likelihood-ratio tests (LRTs). This work was later generalized to multiple sensors by Reibman and Nolte [3] and by Hoballah and Varshney [4]. Similarly, under the Neyman-Pearson (NP) criterion, the optimality of the local LRT has been established in [5].

All of the above work assumed error-free transmission between the local sensors and the fusion center. This is overly idealistic in systems with stringent resource and delay constraints, such as the wireless sensor network (WSN) with geographically dispersed lower power low cost sensor nodes. Accounting for non-ideal transmission channels, channel aware signal processing for distributed detection problem has been developed in [6–8]. The optimal local decision rule was still shown to be a monotone likelihood ratio partition of its observation space, provided the observations were conditionally independent across the sensors.

The work in [6–8] assumed a clairvoyant case, i.e., global information regarding the transmission channels between the local sensors and the fusion center is available at the design stage. This is theoretically significant as it provides the best achievable detection performance to which any suboptimal approach needs to be compared. On the other hand, it lacks practical significance due to the requirement of exact knowledge of global CSI. This is further exacerbated by the potential mobility of sensors that leads to fast fading channels: decision rules for all sensors need to be synchronously updated for different channel realizations.

In this paper, we consider the distributed detection problem where the designer only has the channel fading statistics instead of the instant channel state information (CSI). In this case, a sensible performance measure is to use the average error probability at the fusion center where the averaging is performed with respect to the channel state. We restrict ourselves to binary local sensor outputs and derive the necessary conditions for optimal local decision rules that minimize the average error probability at the fusion center for the following two cases: 1) CSIF: the fusion center has the instant CSI. 2) NOCSIF: the fusion center does *not* know the instant CSI. We note here that the CSIF design itself does not require CSI even though it assumes the fusion center has knowledge of CSI. Its computation, however, is very involved and has to resort to exhaustive search. On the other hand, the NOCSIF case can be reduced to the channel aware design where one averages the channel transition probability with respect to the fading channel using the known fading statistics.

We show that the local decision rules amount to local LRTs for both cases. Compared with the channel aware design based on CSI, the proposed approaches have an important practical advantage: the sensor decision rules remain the same for different CSI, as long as the fading statistics remain unchanged. This enables distributed design as no global CSI is used in determining the local decision rules. We also demonstrate through numerical examples that the proposed schemes suffer small performance loss compared with the CSI based approach, as long as the CSI is available at the fusion center.

The paper is organized as follows. Section 2 describes the system model and problem formulation. In Section 3, we establish, for both the CSIF and NOCSIF cases, the optimality

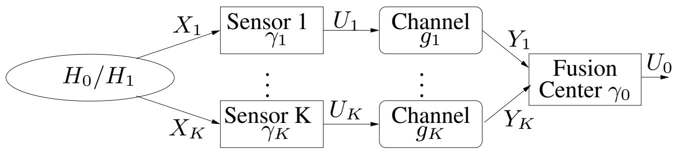

Fig. 1. A block diagram for a wireless sensor network tasked for binary hypothesis testing with channel fading statistics.

of LRTs at local sensors for minimum average error probability at the fusion center. Numerical examples are presented in Section 4 to evaluate the performance of these two cases. Finally, we conclude in Section 5.

2. STATEMENT OF THE PROBLEM

Consider the problem of testing two hypotheses, denoted by H_0 and H_1, with respective prior probabilities π_0 and π_1. A total number of K sensors are used to collect observations X_k, for $k = 1, \cdots, K$. We assume throughout this paper that the observations are conditionally independent, i.e.,

$$p(X_1, \cdots, X_K|H_i) = \prod_{k=1}^{K} p(X_k|H_i), \qquad i = 0, 1. \qquad (1)$$

Upon observing X_k, each local sensor makes a binary decision

$$U_k = \gamma_k(X_k) \qquad k = 1, \cdots, K.$$

The decisions U_k are sent to a fusion center through parallel transmission channels characterized by

$$p(Y_1, \cdots, Y_K|U_1, \cdots, U_K; g_1, \cdots, g_K) = \prod_{k=1}^{K} p(Y_k|U_k, g_k)$$

where $\mathbf{g} = \{g_1, \cdots, g_K\}$ represents the CSI. For the CSIF case, the fusion center takes both the channel output $\mathbf{y} = \{Y_1, \cdots, Y_K\}$ and the CSI $\mathbf{g}$ and makes a final decision using the optimal fusion rule to obtain $U_0 \in \{H_0, H_1\}$,

$$U_0 = \gamma_0(\mathbf{y}; \mathbf{g}).$$

For the case of NOCSIF, the fusion output depends on the channel output and the channel fading statistics:

$$U_0 = \gamma_0(\mathbf{y})$$

where the dependence of fading channel statistics is implicit in the above expression. An error happens if U_0 differs from the true hypothesis. Thus, the error probability at the fusion center, conditioned on a given $\mathbf{g}$, is

$$P_{e0}(\gamma_0, \cdots, \gamma_K|\mathbf{g}) \triangleq P_r(U_0 \neq H|\gamma_0, \cdots, \gamma_K, \mathbf{g}) \qquad (2)$$

where H is the true hypothesis. Our goal is, therefore, to design the optimal mapping $\gamma_k(\cdot)$ for each sensor and the fusion center that minimizes the average error probability, defined as:

$$\min_{\gamma_0(\cdot), \cdots, \gamma_K(\cdot)} \int_{\mathbf{g}} P_{e0}(\gamma_0, \cdots, \gamma_K|\mathbf{g}) p(\mathbf{g}) d\mathbf{g} \qquad (3)$$

where $p(\mathbf{g})$ is the distribution of CSI. A simple diagram illustrating the model is given in Fig. 1.

3. DESIGN OF OPTIMAL LOCAL DECISION RULES

As in [7], [8], we adopt in the following a person-by-person optimization (PBPO) approach, i.e., we optimize the local decision rule for the kth sensor given fixed decision rules at all other sensors and a fixed fusion rule. As such, the conditions obtained are necessary, but not sufficient, for optimality. The fusion center, as usual, is assumed to implement the maximum *a posteriori* probability (MAP) decision rule. Denote by

$$\begin{aligned} \mathbf{u} &= [U_1, U_2 \cdots, U_K], \\ \mathbf{x} &= [X_1, X_2, \cdots, X_K], \end{aligned}$$

the average error probability at the fusion center is

$$\begin{aligned} P_{e0} &= \int_{\mathbf{g}} \sum_{i=0}^{1} \pi_i P(U_0 = 1 - i|H_i, \mathbf{g}) p(\mathbf{g}) d\mathbf{g} \\ &= \int_{\mathbf{g}} \sum_{i=0}^{1} \pi_i \int_{\mathbf{y}} P(U_0 = 1 - i|\mathbf{y}, \mathbf{g}) \sum_{\mathbf{u}} p(\mathbf{y}|\mathbf{u}, \mathbf{g}) \\ &\quad \int_{\mathbf{x}} P(\mathbf{u}|\mathbf{x}) p(\mathbf{x}|H_i) p(\mathbf{g}) d\mathbf{x} d\mathbf{y} d\mathbf{g} \qquad (4) \end{aligned}$$

where, different from the CSI based channel aware design, the local decision rules do *not* depend on the instant CSI. Next, we will further expand the error probability with respect to the kth decision rule $\gamma_k(\cdot)$ for the two different cases.

3.1. The CSIF Case

We first consider the case where the fusion center knows the instant CSI. Define, for $k = 1, \cdots, K$ and $i = 0, 1$,

$$\begin{aligned} \mathbf{u}^k &= [U_1, \cdots, U_{k-1}, U_{k+1}, \cdots, U_K], \\ \mathbf{u}^{ki} &= [U_1, \cdots, U_{k-1}, U_k = i, U_{k+1}, \cdots, U_K], \end{aligned}$$

we can expand the average error probability in (4) with respect to the kth decision rule $\gamma_k(\cdot)$, and we get

$$\begin{aligned} P_{e0} &= \int_{X_k} P(U_k = 1|X_k)[\pi_0 p(X_k|H_0) A_k \\ &\quad -\pi_1 p(X_k|H_1) B_k] dX_k + C \qquad (5) \end{aligned}$$

where

$$C = \int_{X_k} \left[\sum_{i=0}^{1} \pi_i p(X_k|H_i) \int_{\mathbf{y}} \int_{\mathbf{g}} P(U_0 = 1 - i|\mathbf{y}, \mathbf{g}) \sum_{\mathbf{u}^k} p(\mathbf{y}|\mathbf{u}^{k0}, \mathbf{g}) p(\mathbf{g}) p(\mathbf{u}^k|H_i) d\mathbf{g} d\mathbf{y} \right] dX_k$$

is a constant with regard to U_k, and

$$A_k = \int_{\mathbf{y}}\int_{\mathbf{g}} P(U_0=1|\mathbf{y},\mathbf{g}) \left[\sum_{\mathbf{u}^k} \left(p(\mathbf{y}|\mathbf{u}^{k1},\mathbf{g}) - p(\mathbf{y}|\mathbf{u}^{k0},\mathbf{g})\right) p(\mathbf{g}) P(\mathbf{u}^k|H_0)\right] d\mathbf{g}d\mathbf{y}, \quad (6)$$

$$B_k = \int_{\mathbf{y}}\int_{\mathbf{g}} P(U_0=0|\mathbf{y},\mathbf{g}) \left[\sum_{\mathbf{u}^k} \left(p(\mathbf{y}|\mathbf{u}^{k0},\mathbf{g}) - p(\mathbf{y}|\mathbf{u}^{k1},\mathbf{g})\right) p(\mathbf{g}) P(\mathbf{u}^k|H_1)\right] d\mathbf{g}d\mathbf{y}. \quad (7)$$

To minimize P_{e0}, one can see from (5) that the optimal decision rule for the kth sensor is

$$P(U_k=1|X_k) = \begin{cases} 0, & \pi_0 p(X_k|H_0)A_k > \pi_1 p(X_k|H_1)B_k \\ 1, & \text{Otherwise.} \end{cases}$$

Let's further take a look at A_k in (6). We can rewrite it as

$$A_k = \int_{\mathbf{y}^k} [P(U_0=1|\mathbf{y}^k, U_k=1) - P(U_0=1|\mathbf{y}^k, U_k=0)] p(\mathbf{y}^k|H_0) d\mathbf{y}^k.$$

Then, following the similar derivation in [8], we can show that $A_k > 0$ as long as

$$L(U_k) \triangleq \frac{P(U_k|H_1)}{P(U_k|H_0)}$$

is a monotone increasing function of U_k (monotone LR index assignment), i.e.,

$$L(U_k=1) > L(U_k=0). \quad (8)$$

Similarly, $B_k > 0$ if condition (8) is satisfied. This immediately leads to the following result.

Theorem 1 *For the distributed detection problem with unknown CSI only at local sensors, the optimal local decision rule for the kth sensor amounts to the following LRT assuming condition (8) is satisfied.*

$$P(U_k=1|X_k) = \begin{cases} 1, & \frac{p(X_k|H_1)}{p(X_k|H_0)} \geq \frac{\pi_0 A_k}{\pi_1 B_k} \\ 0, & \frac{p(X_k|H_1)}{p(X_k|H_0)} < \frac{\pi_0 A_k}{\pi_1 B_k} \end{cases} \quad (9)$$

where A_k and B_k are defined in (6) and (7) respectively.

Although the optimal local decision rule for each local sensor is explicitly formulated in (9), it is not amenable to numerical evaluation: In (6) and (7), the integrand involves the fusion rule that is a highly nonlinear function of the CSI $\mathbf{g}$, making the integration intractable. The only possible way of finding the optimal local decision rules appears to be an exhaustive search, whose complexity becomes prohibitive when K is large.

Instead of directly minimizing the average error probability as in (3), an alternative approach is to first average the channel transition probability with respect to the fading channel. That is, we compute $p(Y_k|U_k)$ by marginalizing out the channel g_k:

$$p(Y_k|U_k) = \int_{g_k} p(Y_k|U_k, g_k) p(g_k) dg_k. \quad (10)$$

With this marginalization, we can use the channel aware design approach [7] that tends to the 'averaged' transmission channel. This motivates the NOCSIF design. If the fusion rule does not depend on the instant CSI, i.e., $P(U_0=1-i|\mathbf{y},\mathbf{g}) = P(U_0=1-i,\mathbf{y})$, the average error probability in (4) can be rewritten as

$$P_{e0} = \sum_{i=0}^{1} \pi_i \int_{\mathbf{y}} \sum_{\mathbf{u}} P(U_0=1-i|\mathbf{y}) \left(\int_{\mathbf{g}} p(\mathbf{y}|\mathbf{u},\mathbf{g})p(\mathbf{g})d\mathbf{g}\right) \int_{\mathbf{x}} P(\mathbf{u}|\mathbf{x})p(\mathbf{x}|H_i)d\mathbf{x}d\mathbf{y}$$

where $\int_{\mathbf{g}} p(\mathbf{y}|\mathbf{u},\mathbf{g})p(\mathbf{g})d\mathbf{g}$ precisely describes the marginalized transmission channels (c.f. Eq. (10)). This directly leads to the following case.

3.2. The NOCSIF Case

Here we consider the case where the fusion center does not know the instant CSI. As such, the marginalization described in (10) will be implemented. This leads to the standard channel aware design where the transmission channels are characterized by the marginalized distribution $p(Y_k|U_k)$. From [7], we have a result resembling that of Theorem 1 except that A_k, B_k and C are replaced by A'_k, B'_k and C':

$$A'_k = \int_{\mathbf{y}} P(U_0=1|\mathbf{y}) \left[\sum_{\mathbf{u}^k} \left(p(\mathbf{y}|\mathbf{u}^{k1}) - p(\mathbf{y}|\mathbf{u}^{k0})\right) P(\mathbf{u}^k|H_0)\right] d\mathbf{y}, \quad (11)$$

$$B'_k = \int_{\mathbf{y}} P(U_0=0|\mathbf{y}) \left[\sum_{\mathbf{u}^k} \left(p(\mathbf{y}|\mathbf{u}^{k0}) - p(\mathbf{y}|\mathbf{u}^{k1})\right) P(\mathbf{u}^k|H_1)\right] d\mathbf{y}. \quad (12)$$

$$C' = \int_{X_k} \left[\sum_{i=0}^{1} \pi_i p(X_k|H_i) \int_{\mathbf{y}} P(U_0=1-i|\mathbf{y}) \sum_{\mathbf{u}^k} p(\mathbf{y}|\mathbf{u}^{k0}) p(\mathbf{u}^k|H_i) d\mathbf{y}\right] dX_k.$$

Contrary to the CSIF case, A'_k, B'_k for the NOCSIF case are much easier to evaluate.

4. PERFORMANCE EVALUATION

In this section, we use a two-sensor example to evaluate the performance of both the CSIF and NOCSIF cases and compare them with the clairvoyant case where the global channel information is assumed known to the designer. For convenience, we call the clairvoyant case as the CSI case. Consider the detection of a known signal S in zero-mean complex Gaussian noises that are independent and identically distributed (i.i.d.) for the two sensors, i.e., for $k = 1, 2$

$$\begin{aligned} H_0: &\quad X_k = N_k, \\ H_1: &\quad X_k = S + N_k \end{aligned}$$

with N_1 and N_2 being i.i.d. $\mathcal{CN}(0, \sigma_1^2)$. Without loss of generality, we assume $S = 1$ and $\sigma_1^2 = 2$.

Each sensor makes a binary decision based on its observation X_k,

$$U_k = \gamma_k(X_k),$$

and then transmits it through a Rayleigh fading channel to the fusion center. The channel output is

$$Y_k = g_k X_k + W_k$$

where g_1, g_2 are i.i.d. zero-mean complex Gaussian distributed $\mathcal{CN}(0, \sigma_g^2)$ and W_1, W_2 are i.i.d. zero-mean complex Gaussian noises with distribution $\mathcal{CN}(0, \sigma_2^2)$. Without loss of generality, we assume $\sigma_g^2 = 1$.

In Fig. 2, for the equal prior probability case, the average error probabilities as a function of the average signal-to-noise ratio (SNR) of the received signal at the fusion center are plotted for the CSIF and NOCSIF cases, along with the CSI case. We also plot a curve, legended with 'CSIF1' in Fig. 2, where the local sensors use thresholds obtained via the NOCSIF approach but the fusion center implements a fusion rule that utilizes the CSI $\mathbf{g}$. The motivation is two-fold. First, estimating $\mathbf{g}$ at the fusion center is typically feasible. Second and more importantly, the threshold design for NOCSIF is much simpler compared with CSIF, as explained in Section 3.1.

As expected, the CSI case has the best performance and NOCSIF case has the worst performance since the designer has the most information in the clairvoyant case and has the least information in NOCSIF case. The CSIF case is only slightly worse than the CSI case but is much better than the NOCSIF case. The difference of CSIF1 and CSIF is almost indistinguishable. The explanation is that the performance is much more sensitive to the fusion rule than to the local sensor thresholds. This phenomenon has been observed before: the error probability versus threshold plot is rather flat near the optimum point, hence is robust to small changes in thresholds.

5. CONCLUSIONS

We investigated the distributed detection problem with channel fading statistics. Restricted binary local sensor decisions, we derive the necessary conditions for optimal local sensor decision rules that minimize the average error probability at the fusion center for both the CSIF and NOCSIF cases. We also establish the optimality of the LRT for local sensor decisions rules for both cases. Numerical results indicate that a mixed approach where the sensors use the decision rules from the NOCSIF approach while the fusion center implements a fusion rule using the CSI achieves almost identical performance to that of the CSIF case.

6. REFERENCES

[1] R. Radner, "Team decision problems," *Annals of Mathematical Statistics*, vol. 33, pp. 857–881, 1962.

[2] R.R. Tenney and N.R. Sandell Jr., "Detection with distributed sensors," *IEEE Trans. Aerospace Elec. Syst.*, vol. 17, pp. 501–510, July 1981.

[3] A.R. Reibman and L.W. Nolte, "Optimal detection and performance of distributed sensor systems," *IEEE Trans. Aerospace Elect. Sys.*, vol. 23, pp. 24–30, Jan. 1987.

[4] I.Y. Hoballah and P.K. Varshney, "Distributed Bayesian signal detection," *IEEE Trans. Inform. Theory*, vol. 35, pp. 995–1000, Sept. 1989.

[5] S.C.A. Thomopoulos, R. Viswanathan, and D.K. Bougoulias, "Optimal distributed decision fusion," *IEEE Trans Aerospace Elec. Syst.*, vol. 25, pp. 761–765, Sep. 1989.

[6] T.M. Duman and M. Salehi, "Decentralized detection over multiple-access channels," *IEEE Trans. Aeros. Elec. Systems*, vol. 34, pp. 469–476, Apr. 1998.

[7] B. Chen and P.K. Willett, "On the optimality of likelihood ratio test for local sensor decisions in the presence of nonideal channels," *IEEE Trans. Information theory*, vol. 51, pp. 693–699, Feb. 2005.

[8] B. Liu and B. Chen, "Joint source-channel coding for distributed sensor networks," in *Proc. 2004 Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2004.

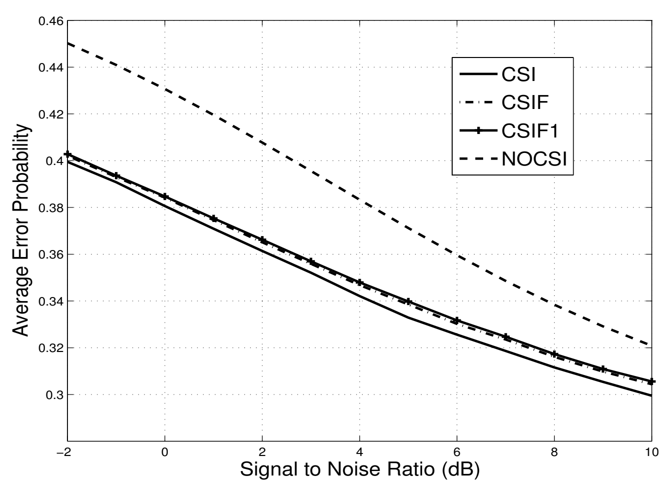

Fig. 2. Average error probability versus channel SNR.