DISTRIBUTED DETECTION WITH COLLISIONS IN A RANDOM, SINGLE-HOP WIRELESS SENSOR NETWORK

Gene T. Whipps[*†]   *Emre Ertin*[†]   *Randolph L. Moses*[†]

[*]U.S. Army Research Laboratory, Adelphi, MD 20783
[†]The Ohio State University, ECE Department, Columbus, OH 43210

Abstract

We consider the problem of distributed detection in a large network of sensors under network communication constraints. Sensor nodes are randomly deployed and follow a random sleep/wake schedule. When awake, sensor nodes perform local detection tests and communicate detections over a multiple-access channel to a fusion center. The fusion center can detect both successful communications and communication collisions in the channel. We show that the optimum fusion rule is a weighted sum of the number of detections received and the number of collisions detected by the fusion node. We derive analytical expressions that characterize the performance of the system. Simulation examples compare theoretical predictions with numerical results.

Index Terms— Distributed detection, random access sensor network

1. INTRODUCTION

This paper considers the problem of distributed detection in a large, random network of sensor nodes. We consider a localized event and focus on the case where sensor nodes communicate binary decisions over a single-hop wireless network to a common fusion node. The fusion node combines the received information in order to make a global decision on the presence or absence of a signal source.

This work is related to and draws from recent works in distributed detection. Niu *et al.* [1] show that, if local decisions are communicated perfectly, the fusion rule for independent and identically distributed (*i.i.d.*) binary observations, conditioned on the true hypothesis, simplifies to counting the number of detections. In [2], Niu *et al.* provide analytic and approximate expressions for the counting rule in a large, random sensor network and a random target location. Chang *et al.* [3] incorporate a random access protocol for a fixed number of sensor nodes, and they derive a fusion rule that is a weighted sum of the number of 1's and 0's successfully received at the fusion node. Similarly, Kapnadak *et al.* incor-

porate a random access protocol in [4], but for a random sensor network. In [5], Aldalahmeh *et al.* develop a real-time counting rule and provide an approximation to the system performance assuming that collisions are negligible. Similar to [2],[4], and [5], we consider a large, random sensor network and incorporate a random access protocol. In contrast, however, we assume a sensor node transmits only when it declares a detection (i.e., when a node decides a signal source is present). We show that the fusion rule is a weighted sum of the number of received detection reports and the number of message collisions.

The remainder of this paper is outlined as follows. In Section 2, we outline the sensor network model, including the local sensor operation, the network communication model, and the fusion node processing. In Section 3, we present numerical examples validating the analytic results and comparing with previous results under perfect communications. Finally, conclusions are given in Section 4.

2. SENSOR NETWORK MODEL

In this section we describe the sensor network model, including the local sensor operation, the network communication model, and the fusion node processing. We assume a large number of sensor nodes are randomly deployed independently and uniformly over a bounded, circular region. Each node awakens with some probability at each time interval, and if active, senses its environment, computes a local decision, and communicates any detections to the fusion node using a Slotted ALOHA protocol. When multiple nodes attempt to communicate over a shared wireless communications medium, transmitted messages are subject to errors due to collisions with other ongoing message transmissions. Therefore, the fusion node must take this into account when forming a decision rule.

2.1. Sensor Node Model

To conserve energy, each sensor node follows a random sleep/wake schedule according to a stationary Poisson process, independent of all other nodes. The number of active

nodes N during any period is modeled according to

$$\Pr\{N = n\} = e^{-\lambda}\frac{\lambda^n}{n!}, \quad n = 0, 1, 2, \ldots \quad (1)$$

where λ is the average number of active nodes in the circular region encompassing the sensor network.

Each active sensor node senses its environment and makes a local detection decision. Sensors collect noisy measurements of the environment for the purpose of detecting whether or not a source signal is present in the scene. Under hypothesis h_0, when no signal is present, sensors measure only noise. We assume the observations are independent and identically distributed (*i.i.d.*) given h_0 is the true state of nature. Under hypothesis h_1, in which a source is present, the sensor measures a signal component embedded in noise. The observations are modeled according to

$$h_j : f_{Z_i|H,D}(z|h_j, d_i), \quad (2)$$

for $j = 0, 1$ and $i = 1, 2, \ldots, N$, where $d_i = ||x_i - x_s||$ is the Euclidean distance between the signal source at x_s and the ith sensor at x_i. Under h_0, there is no dependence on the source signal, so we can write $f_{Z_i|H,D}(z|h_0, d_i) = f_{Z|H}(z|h_0)$. As a matter of notation, random variables are denoted with upper case letters and lower case letters represent specific outcomes.

We assume a homogeneous medium that attenuates the source signal with distance in a manner that does not depend on the orientation relative to the signal source. For any given location of the signal source, the distances between the signal source and the sensors, denoted by D_i, are *i.i.d.* Thus, $f_{D_1,\ldots,D_N}(d_1, \ldots, d_N) = \prod_{i=1}^{N} f_D(d_i)$. It follows that the sensor observations are *i.i.d.* under h_1, and the joint distribution of the observations is

$$f_{Z_1,\ldots,Z_N|H}(z_1, \ldots, z_N|h_1) = \prod_{i=1}^{N} E_D(f_{Z_i|H,D}(z_i|h_1, d))$$

$$= \prod_{i=1}^{N} f_{Z|H}(z_i|h_1), \quad (3)$$

where $E_D(\cdot)$ represents the expectation with respect to the distance D between the signal source and a sensor. For the current problem, it was reported in [6] that the local tests in a parallel decision fusion system are likelihood ratio tests provided the local observations are independent, conditioned on each hypothesis. Accordingly, the test at the ith sensor decides between $H = h_0$ and $H = h_1$ according to

$$\Lambda_i(Z_i) \underset{h_0}{\overset{h_1}{\gtrless}} \tau_i, \quad (4)$$

where τ_i is the threshold and $\Lambda_i(z) = f_{Z|H}(z|h_1)/f_{Z|H}(z|h_0)$ is the likelihood ratio at the ith sensor. Although the likelihood ratio function is identical at each sensor (i.e., $\Lambda_i = \Lambda$), the local thresholds τ_i are not necessarily identical in optimal fusion of binary decisions [6]. However, we assume each sensor shares a common threshold τ. This choice dramatically simplifies the design and deployment of the sensor network, and it was shown in [7] to be asymptotically optimal.
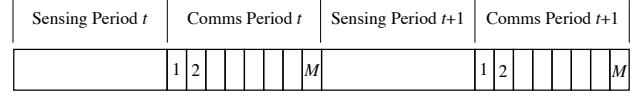

Fig. 1. Sensing and communications periods for all active sensor nodes.

The sensor-level performance is characterized by the probabilities of detection and false alarm. These probabilities are given by

$$\bar{p}_{i,j} = \Pr\{\Lambda(Z) > \tau|h_j\} \equiv \bar{p}_j, \quad (5)$$

for $j = 0, 1$ and $i = 1, \ldots, N$, with $\bar{p}_{i,0}$ representing the probability of false alarm and $\bar{p}_{i,1}$ representing the probability of detection for the ith sensor. Alternatively, $\bar{p}_{i,1}$ represents the average probability of detection at a sensor node, over all possible locations of the sensor node (hence the overbar).

Local decisions, denoted by $\hat{h}_i \in \{h_0, h_1\}$, are communicated to the fusion node. The fusion node then combines received information to make a global decision as to the presence or absence of a signal source.

2.2. Network Communication and Fusion

Sensor nodes are either awake or asleep during an activity period, which is formed by a sensing period followed by a communicating period. The sensing and communicating periods are synchronized so that each active sensor node observes the environment and makes a local decision during the sensing period. Then, nodes attempt to transmit their decisions to a common fusion node during the communications period. Fig. 1 graphically depicts the sensing and communicating periods. Note that there could be overlap between a communications period and the next sensing period, assuming one does not interfere with the other and so that similar periods do not overlap (e.g., Sensing Period t does not overlap with Sensing Period $t + 1$). Since we model the system as being time-invariant, we need only model a single activity period.

Since the sensor observations are *i.i.d.* given the true hypothesis, it is straightforward to show that the number of sensor nodes with detections, denoted by X, has a probability mass function given by

$$\Pr\{X = n|h_j\} = e^{-\lambda_j}\frac{\lambda_j^n}{n!}, \quad n = 0, 1, 2, \ldots \quad (6)$$

where $\lambda_j = \lambda\bar{p}_j$. In [1], Niu *et al.* show that, if the local detections are communicated perfectly, the fusion rule simplifies to counting the number of detections. In this paper, we derive the fusion rule for the case of sensors communicating through a delay-constrained media access control (MAC).

To both save transmission energy and reduce communication collisions, sensor nodes only transmit detections (i.e., when $\hat{h}_i = h_1$). Furthermore, the detection messages do not

need to be unique to the sensor nodes. In [4, 5], similar distributed detection models were considered, but both 0 and 1 local decisions are communicated to the fusion node. Here, messages to be transmitted represent only $\hat{h}_i = h_1$ decisions. In addition, we assume sensor nodes do not retransmit messages in the event of a collision, again to conserve energy. Since detection messages are identical in this work, collisions provide valuable information at the fusion node.

During the communication period, each sensor with a detection attempts to communicate a single detection message via Slotted ALOHA. The communications period is broken into $M < \infty$ equal-duration slots. On the receiving side, the fusion node does not know which slots will be utilized, so it must wait the duration of all M slots during the communications period before making a decision. Thus, the communications delay is directly proportional to M.

Nodes with detections randomly choose one out of M slots, independent of all other nodes, according to the same probability distribution. Let S_i represent the slot number chosen by the ith node. The slot numbers are then drawn according to $\Pr\{S_i = m\} = p_m$, for $m = 1, 2, \ldots, M$.

We assume the fusion node can detect collisions in any of the M slots. A message is considered to be received successfully in a slot if only one sensor node transmits in that slot. Otherwise, a slot is either unused or contains collisions. A slot with 2 or more transmissions is counted as 1 collision. The observations at the fusion node at the end of a communications period are then the total number of slots with detections, denoted N_1, and the total number of slots with collisions, denoted N_c.

We start by characterizing the probability mass function of the number of nodes attempting a transmission in each slot. Let T_m represent the number of sensor nodes that select slot m for transmitting a detection message. Since the slot numbers are selected *i.i.d.* across the nodes with detections, the probability of occupancy in slot m, conditioned on n nodes (with detections) and hypothesis h_j, is given by

$$\Pr\{T_m = k | X = n, h_j\} = \binom{n}{k} p_m^k (1 - p_m)^{n-k}, \quad (7)$$

for $k = 0, 1, \ldots, n$. Under this network model, the system designer can select the distribution from which nodes choose slot numbers. For a discrete uniform distribution (i.e., $p_m = 1/M$ for $m = 1, 2, \ldots, M$), the slot occupancies become identically distributed across slots. It is clear that the slot occupancies are also independent conditioned on the true hypothesis due to the conditional independence of the local decisions. Subsequently, the probability of occupancy in slot m, given h_j, is then

$$\Pr\{T_m = k | h_j\} = e^{-\frac{\lambda \bar{p}_j}{M}} \frac{\left(\frac{\lambda \bar{p}_j}{M}\right)^k}{k!}, \quad (8)$$

for $k = 0, 1, 2, \ldots$ and $m = 1, 2, \ldots, M$. Let $\pi_{0,j} = \Pr\{T_m = 0 | h_j\}$, $\pi_{1,j} = \Pr\{T_m = 1 | h_j\}$, and $\pi_{2,j} = \Pr\{T_m > 1 | h_j\}$ denote the conditional probabilities of an

unused slot, a successful transmission in a slot, and a collision in a slot, respectively. Note that $\pi_{2,j} = 1 - \pi_{0,j} - \pi_{1,j}$. Finally, for k successful transmissions and l collisions out of M slots, the joint conditional distribution of transmissions and collisions is given by

$$\Pr\{N_1 = k, N_c = l | h_j\} =$$
$$\frac{M!}{k! l! (M - k - l)!} (\pi_{0,j})^{M-k-l} (\pi_{1,j})^k (\pi_{2,j})^l, \quad (9)$$

for $k = 0, 1, \ldots, M$ and $l = 0, 1, \ldots, M - k$. The decision rule at the fusion node is then formed by the likelihood ratio test

$$\frac{\Pr\{N_1, N_c | h_1\}}{\Pr\{N_1, N_c | h_0\}} \underset{h_0}{\overset{h_1}{\gtrless}} \eta. \quad (10)$$

Substituting the right-hand-side of (9) into (10), and after some simplification, an equivalent test is given by

$$N_1 + \beta N_c \underset{h_0}{\overset{h_1}{\gtrless}} \eta', \quad (11)$$

where

$$\beta = \frac{\log\left(e^{\mu_1} - 1 - \mu_1\right) - \log\left(e^{\mu_0} - 1 - \mu_0\right)}{\log(\mu_1) - \log(\mu_0)}, \quad (12)$$

and $\mu_1 = \frac{\lambda \bar{p}_1}{M}$ and $\mu_0 = \frac{\lambda \bar{p}_0}{M}$.

Let $Y = N_1 + \beta N_c$ and define the set $\Gamma = \{(k, l) | l \in \{0, 1, \ldots, M - k\}, k \in \{0, 1, \ldots, M\}\}$. Note that Y is a discrete random variable. Therefore, the test in (11) is a randomized test. Then, at the fusion node, the probability of false alarm is

$$P_{fa} = \sum_{\substack{(k,l) \in \Gamma \\ k + \beta l > \eta'}} \Pr\{N_1 = k, N_c = l | h_0\}$$
$$+ \gamma \sum_{\substack{(k,l) \in \Gamma \\ k + \beta l = \eta'}} \Pr\{N_1 = k, N_c = l | h_0\}, \quad (13)$$

and, similarly, the probability of missed detection is

$$P_m = \sum_{\substack{(k,l) \in \Gamma \\ k + \beta l \leq \eta'}} \Pr\{N_1 = k, N_c = l | h_1\}$$
$$- \gamma \sum_{\substack{(k,l) \in \Gamma \\ k + \beta l = \eta'}} \Pr\{N_1 = k, N_c = l | h_1\}, \quad (14)$$

where γ is the randomization parameter of the fusion rule. For a given λ, $\bar{p}_0$ (and thus τ), and M, the probaaility masses in (9) under both hypothesis (and thus β) are known. Since Γ is a finite point set, it is straightforward to find η' and γ that satisfies $P_{fa} = \alpha$, for some $\alpha \in (0, 1)$.

3. NUMERICAL RESULTS

In this section, we explore three simulation examples to highlight the detector performance at the fusion node. In the first example, the analytical performance is validated against Monte Carlo samples of the test statistic. For the second example, the analytical performance is evaluated versus
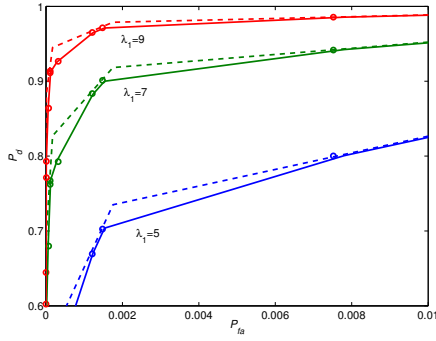
Fig. 2. ROC curves for $\lambda_1 = 5$ (blue), 7 (green), 9 (red) with $\lambda_0 = 0.5$ and $M = 5$. Solid lines represent the analytical ROC, circles represent sample estimates, and dashed lines represent the ROC under perfect communications.
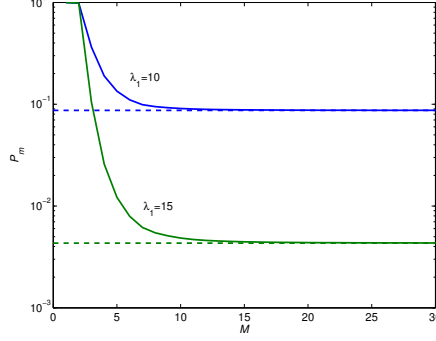
Fig. 3. P_m versus M for $\lambda_1 = 10$ (blue) and 15 (green), with $\lambda_0 = 0.5$ and fixed $P_{fa} = 10^{-5}$. Curves represent P_m for the decision rule with collisions (solid) and the counting rule under perfect communications (dashed).

Fig. 4. P_m versus λ for $M = 5$ (blue) and 15 (green), with $\lambda_0 = 0.5$ and fixed $P_{fa} = 10^{-5}$. Curves represent P_m for the decision rule with collisions (solid) and the counting rule under perfect communications (dashed).
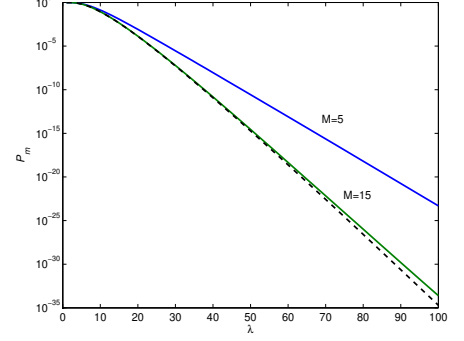
the number of communications slots M for fixed averaged number of nodes λ. In the third example, the analytical performance is evaluated versus λ for fixed M (i.e., a delay constraint). In each example, the results are also compared to the case for $M \to \infty$, where the probability of a collision in a slot goes to zero. In which case, the fusion rule is the counting (of detections) rule of [1] with probabilities of miss detection and false alarm evaluated from (6) with an appropriate randomization parameter.

First, we evaluate receiver operating characteristics (ROC) curves (P_d versus P_{fa}) for three different conditional means on the number of sensors nodes with detections. The mean under h_0 is fixed at $\lambda_0 = 0.5$ and the number of slots is set to $M = 5$. ROC curves are plotted in Fig. 2 for $\lambda_1 = 5$ (blue curves), 7 (green curves), 9 (red curves). For an appropriate signal model, the average number of detections λ_1 can be increased for a fixed number of false alarms λ_0 by increasing the local threshold τ while increasing the density of the sensor network [8]. As seen in Fig. 2, the performance is comparable to the case without collisions, with a small gap in performance. Collisions contain useful information since all of the messages are identical. Subsequently, the fusion node attempts to recover as much information from collisions as it does from received detection messages.

We next evaluate the probability of missed detection versus the number of message slots M for a fixed false alarm probability. The mean under h_0 is again fixed at $\lambda_0 = 0.5$ and the (global) false alarm probability is set to $P_{fa} = 10^{-5}$. Fig. 3 shows plots of P_m with the mean under h_1 set to $\lambda_1 = 10$ (blue curves) and 15 (green curves). It is clear from the figure that the gap in performance can be diminished by increasing the number of slots M. However, this is achieved at the expense of delay at the fusion node to make a global decision. This behavior is also evident by examining the occu-

pancy probability in (8). As M increases, the average number of occupants in any given slot goes to zero. Thus, collisions are reduced and the test statistic is dominated by the count of detections.

Finally, we evaluate the probability of missed detection versus the average number of active nodes λ for a fixed false alarm probability. Again, the mean under h_0 is fixed at $\lambda_0 = 0.5$ and the (global) false alarm probability is set to $P_{fa} = 10^{-5}$. Fig. 4 shows plots of P_m with the $M = 5$ (blue curve) and 15 (green curve). Despite a fixed number of message slots, this simulation shows that performance improves with increasing average number of nodes, while the average number of detections under h_0 (an thus transmission attempts) remains constant. The average number of collisions increases under h_1 while remaining constant under h_0, but the weighting of collisions in the test statistic also increases.

4. CONCLUSION

We have derived the fusion rule for the case of binary sensor nodes communicating through a random-access MAC on a single-hop wireless network. There is a chance that some detection messages are not received properly due to collisions with other messages since the communications medium is shared and randomly accessed. The fusion node receives potentially a subset of the detection messages over M slots and records the number of slots with any collisions. The fusion node then combines this information to make a global decision on the presence or absence of a signal source.

It was shown that the loss in performance is negligible, despite collisions, relative to the case of "ideal" communications. Additionally, the gap in performance can be reduced by increasing the number of slots M, but at the expense of added delay.

5. REFERENCES

[1] R. Niu, P. K. Varshney, and Q. Cheng, "Distributed detection in a large wireless sensor network," *Special Issue on the Seventh Int. Conf. on Information Fusion-Part I*, vol. 7, no. 4, pp. 380 – 394, 2006.

[2] R. Niu and P.K. Varshney, "Performance analysis of distributed detection in a random sensor field," *Signal Processing, IEEE Transactions on*, vol. 56, no. 1, pp. 339 –349, jan. 2008.

[3] Tse-Yao Chang, Teng-Cheng Hsu, and P.Y.-W. Hong, "Exploiting data-dependent transmission control and mac timing information for distributed detection in sensor networks," *Signal Processing, IEEE Transactions on*, vol. 58, no. 3, pp. 1369 –1382, march 2010.

[4] V. Kapnadak and E.J. Coyle, "Optimal density of sensors for distributed detection in single-hop wireless sensor networks," in *Information Fusion (FUSION), 2011 Proceedings of the 14th International Conference on*, july 2011, pp. 1 –8.

[5] S. Aldalahmeh, M. Ghogho, and A. Swami, "Distributed detection of an unknown target in clustered wireless sensor networks," in *Signal Processing Advances in Wireless Communications (SPAWC), 2011 IEEE 12th International Workshop on*, june 2011, pp. 126 –130.

[6] R. R. Tenney and N. R. Sandell, "Detection with distributed sensors," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 17, pp. 501–510, 1981.

[7] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Mathematics of Control, Signals, and Systems (MCSS)*, vol. 1, pp. 167–182, 1988, 10.1007/BF02551407.

[8] G. T. Whipps, E. Ertin, and R. L. Moses, "Threshold considerations in distributed detection in a network of sensors," in *SPIE Defense, Security, and Sensing*. International Society for Optics and Photonics, 2012, vol. 8389.