Performance Analysis of Distributed Detection in a Random Sensor Field

Ruixin Niu, *Member, IEEE*, and Pramod K. Varshney, *Fellow, IEEE*

***Abstract*—For a wireless sensor network (WSN) with randomly deployed sensors, the performance of the counting rule, where the fusion center employs the total number of detections reported by local sensors for hypothesis testing, is investigated. It is assumed that the signal power decays as a function of the distance from the target. For both the case where the total number of sensors is known and the wireless channels are lossless, and the case where the number of sensors is random and the wireless channels have nonnegligible error rates, the exact system level probability of detection is derived analytically. Some approximation methods are also proposed to attain an accurate estimate of the probability of detection, while at the same time to reduce the computation load significantly. To obtain a better system level detection performance, the local sensor level decision threshold is determined such that it maximizes the system level deflection coefficient.**

***Index Terms*—Counting rule, decision fusion, distributed detection, wireless sensor networks.**

I. Introduction

DUE to their great potential in various applications, such as battlefield surveillance, security, traffic, and environmental monitoring, wireless sensor networks (WSNs) have attracted significant attention. Many aspects of WSNs, including routing protocols, network structures, distributed data compression and transmission, and collaborative signal processing, have recently been investigated and covered in the literature [1], [2]. In this paper, we focus on distributed detection, a fundamental task that a WSN needs to accomplish. There are already numerous publications on the conventional distributed detection and decision fusion problem. In [3], the optimum decision fusion rule has been obtained under the conditional independence assumption. Decision fusion with correlated observations has been studied in [4]–[6]. Many papers focused on the problem of distributed detection with constrained system resources [7]–[10]. Specifically, these publications have proposed solutions to optimal bit allocation (or sensor selection) given a communication constraint.

However, in most of these publications, it has been assumed that the local sensors' detection performances, namely either the local sensors' signal-to-noise ratios (SNRs) or their probabilities of detection and false alarm, are known to the fusion center. For a dynamic target and a WSN with passive sensors, it is very difficult to estimate local sensors' performances via experiments because these performances are time-varying as the target traverses the wireless sensor field. For a large WSN consists of densely deployed low-cost and low-power sensors, in our previous work [11], [12], we have proposed a counting rule that uses the total number of detections ("1"s) transmitted from local sensors as the fusion statistic. We have assumed that either the total number of sensors in the region of interest (ROI) [11] or its expected value [12] is very large, so that performance evaluation based on the central limit theorem (CLT) can be carried out. However, the assumption of a large number of sensors is not always true. In [13], a saddlepoint approximation technique has been used to provide accurate estimates of the error probabilities for the distributed detection problem, even for a system with a small number of sensors. However, the saddlepoint approximation technique works only for the approximation of the density or tail probability of the sum of independent and identically distributed (i.i.d.) random variables (RVs). It is not applicable to the problem formulated in this paper, where under hypothesis H_1, the test statistic is the sum of detections reported by local sensors, which are not i.i.d. Another assumption we have made in [11] and [12] was that the ROI is very large and the signal power decays very fast as the distance from the target increases, so that the system level detection performance was taken as approximately invariant to the target's location. Here we relax both of the above mentioned assumptions made in our previous work, derive the exact analytical expression for the detection performance, and propose some approximations via Binomial distributions and Demoivre-Laplace approximation, which require much less computation load and yet yield fairly accurate results. Furthermore, we investigate the detection performance for a more realistic scenario where the total number of sensors is random and the wireless channels between sensors and the fusion center are noisy.

In Section II, basic assumptions regarding the WSN are made and the signal decay model is introduced. In Section III, for a WSN with a known number of sensors and ideal communication channels between sensors and the fusion center, the system level detection performance is derived analytically. Several approximations to the exact system performance are also proposed, whose accuracies are compared in terms of the total variation distance (TVD). In Section IV, the assumptions of known number of sensors and perfect communication channels are relaxed. The number of sensors is assumed either to be a Binomial RV or a Poisson RV, and the noisy wireless channels are modeled as independent binary symmetric channels (BSCs). The system level detection performance and its approximations are derived. The problem of designing the local sensor level

Manuscript received July 21, 2006; revised May 10, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Mounir Ghogho.

The authors are with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244 USA (e-mail: rniu@ecs.syr.edu; varshney@ecs.syr.edu).

Digital Object Identifier 10.1109/TSP.2007.906770

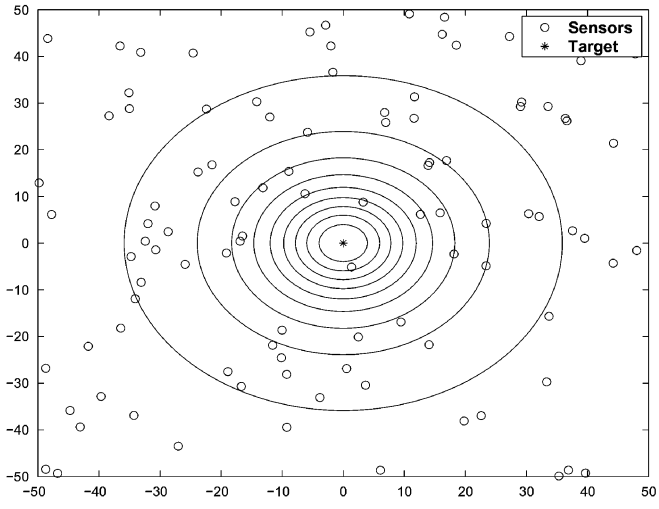


Fig. 1. A sensor deployment example.

threshold is investigated in Section V. Concluding remarks are provided in Section VI.

II. Problem Formulation

As shown in Fig. 1, a total of N sensors are randomly deployed in the ROI, which is a square with area b^2. The locations of sensors are unknown to the WSN, and they are i.i.d. and follow a uniform distribution in the ROI:

$$f(x_i, y_i) = \begin{cases} 1/b^2 & -\frac{b}{2} \le x_i, y_i \le \frac{b}{2} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

for $i = 1, \ldots N$, where (x_i, y_i) are the coordinates of sensor i. The location of the target, represented by its coordinates (x_t, y_t), is independent of the positions of sensors, and follows the same uniform distribution within the ROI as that described in (1).

We assume that the signal power emitted by the target decays isotropically as a function of the distance

$$P_i = g(d_i) \tag{2}$$

where P_i is the signal power measured at sensor i, and d_i is the Euclidean distance between the target and local sensor i

$$d_i = \sqrt{(x_i - x_t)^2 + (y_i - y_t)^2} \tag{3}$$

and function $g(\cdot)$ models how signal power decays as the distance from the target increases.

We assume that noises at the local sensors are i.i.d. and follow a certain distribution. All the sensors use an identical local threshold τ for a likelihood ratio test, and to obtain a local decision. Under hypothesis H_1 (target presence), the probability of detection at sensor i is a function of its receiver's SNR, which is a function of the signal strength P_0 at a reference distance d_0, the distance from the target, and the threshold τ. We denote it as $p_{d_i}(P_0, d_i, \tau)$, or $p_{d_i}(P_0, \sqrt{(x_i - x_t)^2 + (y_i - y_t)^2}, \tau)$. Under hypothesis H_0 (target absence), each sensor has the same probability of false alarm, and we denote it as $p_{\text{fa}}(\tau)$. Note that the detection performance analysis technique proposed in this paper is quite general, and its application is not restricted to a specific signal decay model or a specific noise model.

In the examples that we provide later in this paper, we used the following $g(\cdot)$:

$$g(x) = P_0 \begin{cases} 1 & 0 < x \le d_0 \\ d_0^n / x^n & d_0 < x \end{cases} \tag{4}$$

where P_0 is the signal power measured at a reference distance d_0, n is the signal decay exponent. By adopting the model described in (4), we prevent the receiver amplifier from saturation when the target is very close to the sensor. In the examples, we set $n = 2$ and $d_0 = 1$. We do not specify the type of the passive sensors and the power decay model of (4) is quite general. For example, for a spherical acoustic wave radiated by a simple source [14], the signal power decays at a rate inversely proportional to the square of the distance.

In the examples, we also assume that at the local sensors, the noises are additive, i.i.d., and follow the standard Gaussian distribution:

$$n_i \sim \mathcal{N}(0, 1) \tag{5}$$

where n_i is the noise at the ith sensor. For a local sensor i, the binary hypothesis testing problem is

$$\begin{aligned} H_1&: s_i = a_i + n_i \\ H_0&: s_i = n_i \end{aligned} \tag{6}$$

where s_i is the received signal, a_i is the signal amplitude, which is

$$a_i = \sqrt{g(d_i)}. \tag{7}$$

The threshold τ and the false alarm rate p_{fa} satisfy the following relationship:

$$p_{\text{fa}} = \int_{\tau}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = Q(\tau) \tag{8}$$

where $Q(\cdot)$ is the complementary distribution function of the standard Gaussian. The probability of detection at local sensor i is, therefore

$$p_{d_i} = Q\Big(\tau - \sqrt{g(d_i)}\Big). \tag{9}$$

III. Known Number of Sensors and Perfect Channels

In this section, we assume that the total number of sensors N is known. Further, the wireless channels between the sensors and the fusion center are assumed to be perfect, with negligible error rates. The results derived in this section will form the basis for the more general results that we will obtain in the next section.

A. Decision Fusion

We denote the binary data from local sensor i as $I_i = \{0, 1\}$. I_i takes the value 1 when there is a detection; otherwise, it takes the value 0. Based on local decisions transmitted from sensors, the fusion center makes a final decision about a target's presence

or absence. For a binary decision fusion problem, the implementation of the optimal Chair-Varshney fusion rule [3] requires the knowledge of p_{d_i} and p_{fa_i}.

As long as the threshold τ is known, the probability of false alarm at each sensor is known ($p_{fa_i} = p_{\text{fa}}$), which can be derived from the probability density function (pdf) of the sensor noise. However, at each sensor, it is very difficult to calculate p_{d_i} since it is determined by each sensor's distance to the target (d_i) and the signal power emitted by the target (P_0). Without the knowledge of p_{d_i}s, the fusion center is forced to treat detections from every sensor equally. An intuitive choice is to use the total number of "1"s as a statistic. As proposed in [11], the counting rule makes a system level decision by first counting the number of detections made by local sensors and then comparing it with a threshold T

$$\Lambda_1 = \sum_{i=1}^{N} I_i \underset{H_0}{\overset{H_1}{\gtrless}} T. \tag{10}$$

B. Exact Performance Analysis

In this section, we derive the system performance measures of the counting rule, namely the probability of false alarm P_{fa} and probability of detection P_d at the fusion center.

1) Calculation of P_{fa}: At the fusion center, the probability of false alarm P_{fa} is

$$P_{\text{fa}} = Pr\{\Lambda_1 = \sum_{i=1}^{N} I_i \geq T \,|\, H_0\}. \tag{11}$$

Obviously, under hypothesis H_0, the total number of detections $\Lambda_1 = \sum_{i=1}^{N} I_i$ follows a Binomial (N, p_{fa}) distribution. For a given threshold T, the false alarm rate can be calculated as follows:

$$P_{\text{fa}} = \sum_{k=T}^{N} \binom{N}{k} p_{\text{fa}}^{k} (1 - p_{\text{fa}})^{N-k}. \tag{12}$$

2) Calculation of P_d: Under hypothesis H_1, I_is are not independent of each other, since they are all dependent on the target's coordinates (x_t, y_t), which are RVs. As a result, the distribution of $\Lambda_1 = \sum_{i=1}^{N} I_i$ can not be obtained by the saddlepoint approximation technique proposed in [13], since it requires I_is be i.i.d. We need a new method to derive the distribution of Λ_1.

In Section II, we have assumed that the noises at local sensors are i.i.d. and the locations of local sensors are i.i.d. Based on these assumptions, we show that under hypothesis H_1 and conditioned on the target's location (x_t, y_t), Λ_1 follows a Binomial distribution. This is summarized in the following theorem.

Theorem 1:

$$\Lambda_1 \,|\, (H_1, x_t, y_t) \sim \text{Binomial}\,(N, \bar{p}_d(x_t, y_t)) \tag{13}$$

where

$$\bar{p}_d(x_t, y_t) \triangleq \frac{1}{b^2} \int_{-\frac{b}{2}}^{\frac{b}{2}} \int_{-\frac{b}{2}}^{\frac{b}{2}} p_d(P_0, \sqrt{(x-x_t)^2 + (y-y_t)^2}, \tau)\, dx\, dy. \tag{14}$$

Proof: See Appendix I.

According to Theorem 1, we have the closed-form solution for the probability mass function (pmf) of $\Lambda_1 \,|\, (H_1, x_t, y_t)$

$$\begin{aligned} &\Pr(\Lambda_1 = k \,|\, (H_1, x_t, y_t)) \\ &\quad = \binom{N}{k} [\bar{p}_d(x_t, y_t)]^k [1 - \bar{p}_d(x_t, y_t)]^{N-k} \\ &\qquad\qquad (k = 0, \ldots, N) \end{aligned} \tag{15}$$

and it is straightforward to attain the pmf of $\Lambda_1 \,|\, H_1$

$$\begin{aligned} \Pr(\Lambda_1 = k \,|\, H_1) = \frac{1}{b^2} \binom{N}{k} \int_{-\frac{b}{2}}^{\frac{b}{2}} \int_{-\frac{b}{2}}^{\frac{b}{2}} \\ \times [\bar{p}_d(x_t, y_t)]^k [1 - \bar{p}_d(x_t, y_t)]^{N-k} dx_t dy_t \end{aligned} \tag{16}$$

and the system level probability of detection is, therefore

$$P_d = \sum_{k=T}^{N} \Pr(\Lambda_1 = k \,|\, H_1). \tag{17}$$

C. Performance Evaluation via Approximations

In the previous subsection, we have derived the exact formula to evaluate the system detection performance, namely (16) and (17), which involve the computationally intensive fourfold integration. In this section, some approximation methods are proposed to evaluate the performance, which require much less computation load than the exact evaluation of P_d.

Under hypothesis H_1, if sensors' locations $(x_i, y_i)s$, and the target's location (x_t, y_t) are known, then I_is are independent Bernoulli RVs, and the probability of success for I_i is $\Pr(I_i = 1) = p_{d_i}$, which is a constant. As a result, $\Lambda_1 = \sum I_i$ is the sum of N independent Bernoulli RVs, and its distribution is often called Poisson Binomial distribution [15], which has a very complicated structure if the p_{d_i}s are nonidentical. Many publications have addressed the problem of approximating Poisson Binomial distribution by either a Poisson or a Binomial distribution [15]–[18].

However, in this paper, under hypothesis H_1, Λ_1 does not follow a Poisson Binomial distribution, since sensors' locations $(x_i, y_i)s$, and the target's location (x_t, y_t) are unknown, and I_is are dependent on each other. Nevertheless, it is still interesting to find approximations of the distribution of Λ_1, based on Binomial or Poisson distribution. Through experiments, we have found that the Poisson distribution is a poor approximation for the distribution of $\Lambda_1 \,|\, H_1$, and we do not discuss it here.

1) Binomial I Approximation: Inspired by the fact that conditioned on (x_t, y_t), $\Lambda_1 \,|\, H_1$ is a Binomial distributed RV, here we use a Binomial (N, p_1) distribution to approximate the distribution of $\Lambda_1 \,|\, H_1$. The parameter of the Binomial distribution is

chosen such that its mean matches the mean of $\Lambda_1 \mid H_1$, namely $p_1 = \bar{p}_d$, where $\bar{p}_d$ is the expected value of p_d

$$\bar{p}_d = \frac{1}{b^4} \int_{-\frac{b}{2}}^{\frac{b}{2}} \int_{-\frac{b}{2}}^{\frac{b}{2}} \int_{-\frac{b}{2}}^{\frac{b}{2}} \int_{-\frac{b}{2}}^{\frac{b}{2}} p_d(P_0, \sqrt{(x-x_t)^2 + (y-y_t)^2}, \tau) \times dx\, dy\, dx_t\, dy_t. \quad (18)$$

Again, the evaluation of $\bar{p}_d$ requires a fourfold integration. However, given P_0 and τ, p_d is only a function of $d^2 = (x-x_t)^2 + (y-y_t)^2$. As a result, once the distribution of $v \triangleq d^2$ is known, the evaluation of $\bar{p}_d$ is reduced to a single-fold integration over v, instead of a fourfold integration. In [19], we have derived the distribution of v and we state it in the following theorem.

Theorem 2: Assuming a local sensor's two coordinates and the target's two coordinates, namely x_i, y_i, x_t, and y_t are i.i.d. and follow a uniform distribution within the interval $[-b/2, b/2]$, the pdf of the square of their distance ($v = d^2$) is

$$f_V(v) = \begin{cases} \frac{\pi}{b^2} + \frac{v}{b^4} - \frac{4\sqrt{v}}{b^3}, & 0 < v \le b^2 \\ h(v), & b^2 < v \le 2b^2 \\ 0, & \text{otherwise} \end{cases} \quad (19)$$

where

$$h(v) \triangleq \frac{2}{b^2} \arcsin\left(\frac{2b^2 - v}{v}\right) - \frac{v}{b^4} + \frac{4\sqrt{v - b^2}}{b^3} - \frac{2}{b^2}.$$

Proof: See [19].

Given $f_V(v)$, the average p_d can be easily calculated as follows:

$$\bar{p}_d = \int_0^{2b^2} p_d(P_0, \sqrt{v}, \tau) f_V(v)\, dv. \quad (20)$$

2) Binomial II Approximation: The Binomial I approximation method only matches the mean of $\Lambda_1 \mid H_1$. It will be of interest to study the approximation that matches both the mean and variance of $\Lambda_1 \mid H_1$. We derive the variance of $\Lambda_1 \mid H_1$ and its bounds, and provide them in Proposition 1.

Proposition 1: The variance of $\Lambda_1 \mid H_1$ is

$$\mathrm{Var}(\Lambda_1 \mid H_1) = N\bar{p}_d - N^2\bar{p}_d^2 + (N^2 - N)E\left[\bar{p}_d^2(x_t, y_t)\right] \quad (21)$$

which satisfies the following inequality

$$N\bar{p}_d - N\bar{p}_d^2 \le \mathrm{Var}(\Lambda_1 \mid H_1) \le (N^2 - N)E\left(p_d^2\right) + N\bar{p}_d - N^2\bar{p}_d^2 \quad (22)$$

where

$$E(p_d^2) = \int_0^{2b^2} p_d^2(P_0, \sqrt{v}, \tau) f_V(v)\, dv. \quad (23)$$

Proof: See Appendix II.

From (21), it is clear that the evaluation of $E[\bar{p}_d^2(x_t, y_t)]$, and hence that of $\mathrm{Var}(\Lambda_1 \mid H_1)$ requires a fourfold integration. Both the calculations of lower and upper bounds of $\mathrm{Var}(\Lambda_1 \mid H_1)$ require only a single-fold integration. However, experiments show that these bounds, especially the upper bound, are not very tight. As a result, we still need to calculate the true variance of Λ_1 through a fourfold integration. Similar to the method proposed in [18], matching a Binomial (M, p_2) RV's mean Mp_2 and variance $Mp_2(1-p_2)$ to $N\bar{p}_d$ and $\mathrm{Var}(\Lambda_1 \mid H_1)$ respectively, we have

$$M = \mathrm{Round}\left\{(N\bar{p}_d)^2 / [N\bar{p}_d - \mathrm{Var}(\Lambda_1 \mid H_1)]\right\}$$
$$p_2 = N\bar{p}_d / M. \quad (24)$$

We name this approximation the Binomial II approximation. Note that since according to Proposition 1

$$\mathrm{Var}(\Lambda_1 \mid H_1) \ge N\bar{p}_d(1 - \bar{p}_d)$$

it can be shown that $M \ge N$.

3) DeMoivre-Laplace Approximation: According to DeMoivre-Laplace Theorem [20], if $Mp_2(1-p_2) \gg 1$, then

$$\binom{M}{k} p_2^k (1-p_2)^{M-k} \approx \frac{1}{\sqrt{2\pi M p_2(1-p_2)}} e^{-(k - Mp_2)^2 / [2Mp_2(1-p_2)]}.$$

As a result, when M is large, the pmf of a Binomial (M, p_2) distribution can be approximated by the samples of a Gaussian pdf, whose mean and variance are $N\bar{p}_d$ and $\mathrm{Var}(\Lambda_1 \mid H_1)$, respectively

$$\Pr(\Lambda_1 = k \mid H_1) \approx \frac{1}{\sqrt{2\pi \mathrm{Var}(\Lambda_1 \mid H_1)}} e^{-(k - N\bar{p}_d)^2 / [2\mathrm{Var}(\Lambda_1 \mid H_1)]}. \quad (25)$$

Note that when using DeMoivre-Laplace Approximation (DLA), terms in (25) are normalized so that the sum of the pmf values is one.

4) Total Variation Distance: To measure the accuracy of the approximations, we adopt the total variation distance (TVD) [16] to quantify the difference of two distributions. Let r_1 and r_2 be two integer-valued RVs, which follow two distributions with pmfs f_1 and f_2, respectively. The TVD between these two distributions is defined as

$$d_{TV}(f_1, f_2) = \sup_A |\Pr(r_1 \in A) - \Pr(r_2 \in A)| \quad (26)$$

where A is any subset of the set of nonnegative integers and it is well known that [17]

$$d_{TV}(f_1, f_2) = \frac{1}{2} \sum_{k=0}^{\infty} |f_1(k) - f_2(k)|. \quad (27)$$

Since by definition, the TVD is the maximum absolute difference of probability mass on any subset A between two distributions, and P_d is equivalent to a complementary cumulative distribution function (see (17)), the absolute difference between

TABLE I
COMPARISON OF DIFFERENT APPROXIMATION METHODS

Approximation method	Computation load	Approximation Accuracy
Binomial I	1 single-fold integration	High for small N
Binomial II	1 single-fold integration and 1 four-fold integration	High for any N
DLA	1 single-fold integration and 1 four-fold integration	High for large N
Exact Evaluation	$N+1$ four-fold integrations	Exact results

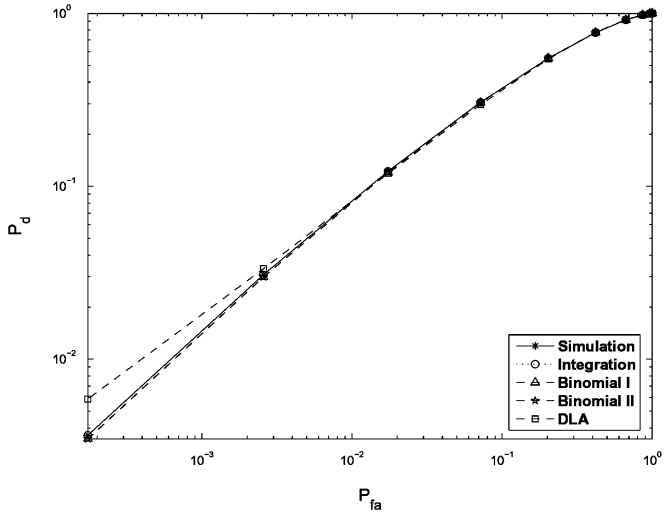


Fig. 2. The ROC curves obtained by simulation, fourfold integration, Binomial I, Binomial II, and DLA approximations. $n = 2, d_0 = 1, b = 50, N = 10, P_0 = 50$, and $\tau = 0.2$.

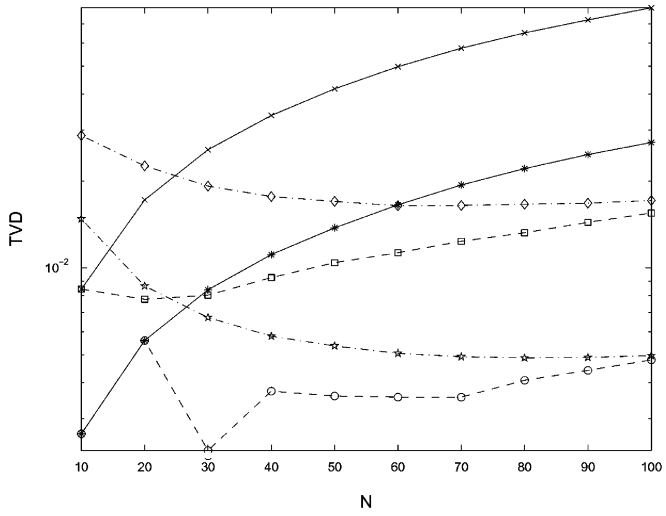


Fig. 3. The TVD between the exact pmf of $\Lambda_1 \mid H_1$, and its approximations based on Binomial I and II, and DLA. $P_0 = 25$. Solid line + x-mark: Binomial I, $b = 20$; solid line + star: Binomial I, $b = 40$; dashdot line + diamond: DLA, $b = 20$; dashdot line + pentagram: DLA, $b = 40$; dashed line + square: Binomial II, $b = 20$; dashed line + circle: Binomial II, $b = 40$.

the true P_d and its approximated value is also bounded by the TVD.

D. Experimental Results

In Fig. 2, the ROC curves corresponding to simulation, fourfold integration, Binomial approximations, and DLA are plotted. In the simulation, 10^7 Monte Carlo runs have been used. For different Monte Carlo runs, we use independent target coordinates (x_t, y_t), and independent sensor coordinates (x_i, y_i), for $i = 1, \ldots, N$. Hence, a total of 10^7 independent (x_t, y_t) pairs have been used in the simulation. It is clear that the curves obtained by simulation, integration, and Binomial I and II distributions are indistinguishable. There are discrepancies between the true ROC curve and that corresponding to DLA approximation. Note that $N = 10$ is too small for DLA to be a good approximation of a Binomial distribution.

To compare the accuracies of different approximations, in Fig. 3, we show their TVDs relative to the true pmf of $\Lambda_1 \mid H_1$, as a function of N. It is clear that the accuracy of Binomial I degrades rapidly as N increases. As we know, the variance of a Binomial I distribution, $N\bar{p}_d(1-\bar{p}_d)$ is linear in N, and $\mathrm{Var}(\Lambda_1 \mid H_1)$ is a quadratic function of N, according to Proposition 1. As N increases, their difference, $(N^2 - N)[E[\bar{p}_d^2(x_t, y_t) - \bar{p}_d^2]]$, increases, and the Binomial I approximation becomes less accurate, due to the mismatch in variance.

On the other hand, the Binomial II approximation has the smallest TVD among the three approximations, because both its mean and variance have been matched to those of $\Lambda_1 \mid H_1$. The jumpy behavior of its TVD curves is due to the fact that M can only take an integer value and hence its variance is close but not identical to the true variance of $\Lambda_1 \mid H_1$ in most cases. As N increases, the accuracy of DLA improves and converges to that of the Binomial II quickly. This is because as N increases, $M \geq N$ increases too, and the DLA becomes a more accurate approximation of Binomial (M, p_2) distribution.

Another observation is that when the size of the ROI b is large, the TVDs corresponding to all the approximations are small. This can be explained as follows. When the ROI is large, only within a small fraction of the ROI surrounding the target, received signal strength is significantly larger than zero. As a result, for most (x_t, y_t) pairs, except those near the border of the ROI, $\bar{p}_d(x_t, y_t)$ in (14) is almost a constant, independent of (x_t, y_t), and $\bar{p}_d(x_t, y_t) \approx \bar{p}_d$. In such cases, (16) becomes

$$\Pr(\Lambda_1 = k \mid H_1) \approx \binom{N}{k} \bar{p}_d^k (1-\bar{p}_d)^{N-k}$$

and Binomial distribution is a very good approximation to the exact distribution of $\Lambda_1 \mid H_1$.

In summary, when N is small, Binomial I approximation is a good choice, considering its accuracy and its small amount of computation involving only a single-fold integration. When N is large, however, the Binomial II and DLA approximations are better choices, since they need only one fourfold integration, compared with $N+1$ fourfold integrations required by the exact evaluation of the pmf of $\Lambda_1 \mid H_1$. The accuracies of different approximation methods, and their requirements for computation load are summarized in Table I.

IV. Random Number of Sensors and Noisy Channels

In a typical WSN, due to the limited energy and/or bandwidth budget for each sensor node, increasing power and/or employing powerful error correction codes may not always be feasible. Furthermore, in a hostile environment, the power of transmitted signal should be kept to a minimum to attain a low probability of intercept/detection (LPI/LPD). As a result, the data transmitted over the wireless channel between a sensor and the fusion center may endure errors. We assume that the local decisions are transmitted through binary symmetric channels (BSCs) [21] to the fusion center. The BSCs have been used to model the noisy channels in the context of distributed detection problems by many researchers since this simple model effectively characterizes the loss of information caused by nonideal channels. The channels corresponding to different sensors are assumed i.i.d., with the same crossover probability p_c

$$p_c \triangleq \Pr\{I_i' \neq I_i\} \tag{28}$$

where $I_i' = \{0,1\}$ is the ith sensor's decision received at the fusion center, which may not be the same as I_i, due to the nonzero channel error rate. Note that as long as $p_c < 1/2$, it is very easy to show that I_i' is a maximum likelihood estimator (MLE) of I_i. The statistic that the counting rule uses at the fusion center becomes

$$\Lambda_2 = \sum_{i=1}^{N} I_i'. \tag{29}$$

As shown later, the nonideal channels only change the probabilities of "1"s and "0"s received at the fusion center, and lead to performance degradation. We keep using the same decision fusion rule, the counting rule.

A. Binomial Distributed Random Number of Sensors

In many applications, the sensors are deployed randomly in and around the ROI, and oftentimes some of them are out of the communication range of the fusion center, malfunctioning or out of battery. Therefore, at a particular time, the total number of sensors that work properly in the ROI is a RV. Here we assume that this RV follows a Binomial distribution. We assume that in a network consisting of L sensors, each sensor successfully sends a binary local decision to the fusion center with a probability p_b. As a result, at a particular time, the number of sensors (N) that can communicate with the fusion center successfully is a Binomial (L, p_b) RV

$$N \sim \text{Binomial}\,(L, p_b). \tag{30}$$

Given this assumption and the BSC communication channels, we derive the distribution of $\Lambda_2 \,|\, (H_1, x_t, y_t)$, as summarized in the following theorem.

Theorem 3: Under hypothesis H_1, and conditioned on the target's coordinates x_t and y_t, Λ_2 follows a Binomial distribution, namely

$$\Lambda_2 \sim \text{Binomial}\,(L, p_b\mu(x_t, y_t)) \tag{31}$$

where

$$\mu(x_t, y_t) \triangleq p_c + (1 - 2p_c)\bar{p}_d(x_t, y_t)$$

and $\bar{p}_d(x_t, y_t)$ has been defined in Theorem 1.

Proof: See Appendix III.

With Theorem 3, it is easy to show that

$$\begin{aligned}\Pr(\Lambda_2 = k \,|\, H_1) &= \frac{1}{b^2}\binom{L}{k}\int_{-b/2}^{b/2}\int_{-b/2}^{b/2}[p_b\mu(x_t, y_t)]^k \\ &\quad \times [1 - p_b\mu(x_t, y_t)]^{L-k}dx_t dy_t\end{aligned} \tag{32}$$

for $k = 0, \ldots, L$. Similar to the proof of Theorem 3, we can show that

$$\Lambda_2 \,|\, H_0 \sim \text{Binomial}\,(L, p_b[p_c + (1 - 2p_c)p_{\text{fa}}]). \tag{33}$$

Since the evaluation of (32) requires fourfold integrations, once again we are interested in its approximations. Due to the similarity between the pmf of $\Lambda_2 \,|\, H_1$ and $\Lambda_1 \,|\, H_1$, we use techniques that have been used in Section III-C, namely the Binomial I and II approximations and DLA. These approximations need to match the mean and/or variance of the $\Lambda_2 \,|\, H_1$, which are provided without proof as follows:

Proposition 2: The mean and variance of $\Lambda_2 \,|\, H_1$ are

$$E(\Lambda_2 \,|\, H_1) = Lp_b[p_c + (1 - 2p_c)\bar{p}_d] \tag{34}$$

and

$$\begin{aligned}\text{Var}(\Lambda_2 \,|\, H_1) &= Lp_bE[\mu(x_t, y_t)]\{1 - Lp_bE[\mu(x_t, y_t)]\} \\ &\quad + L(L-1)p_b^2E[\mu^2(x_t, y_t)]\end{aligned} \tag{35}$$

where

$$E[\mu(x_t, y_t)] = p_c + (1 - 2p_c)\bar{p}_d \tag{36}$$

and

$$\begin{aligned}E[\mu^2(x_t, y_t)] &= p_c^2 + 2p_c(1 - 2p_c)\bar{p}_d \\ &\quad + (1 - 2p_c)^2E\left[\bar{p}_d^2(x_t, y_t)\right].\end{aligned} \tag{37}$$

To compare different approximation methods, in Fig. 4, we show their TVDs relative to the exact pmf of $\Lambda_2 \,|\, H_1$, as a function of L. It is clear that when the size of the sensor network L is small, both Binomial I and II approximations are very accurate. When L is large, however, the DLA approximation, which matches the mean and variance of $\Lambda_2 \,|\, H_1$ exactly, is a better choice.

The accuracies of different approximation methods, and their requirements for computation load are summarized in Table II.

B. Poisson Distributed Random Number of Sensors

1) Exact Performance: In addition to the case where the number of sensors follows a Binomial distribution, we investigate the case of Poisson random number of sensors. The Poisson point process has been adopted widely to model the randomly distributed sensors or wireless nodes [12], [22]. Here we assume

TABLE II
COMPARISON OF DIFFERENT APPROXIMATION METHODS

Approximation method	Computation load	Approximation Accuracy
Binomial I	1 single-fold integration	High for small L
Binomial II	1 single-fold integration and 1 four-fold integration	High for small or medium L
DLA	1 single-fold integration and 1 four-fold integration	High for large L
Exact Evaluation	$L+1$ four-fold integrations	Exact results

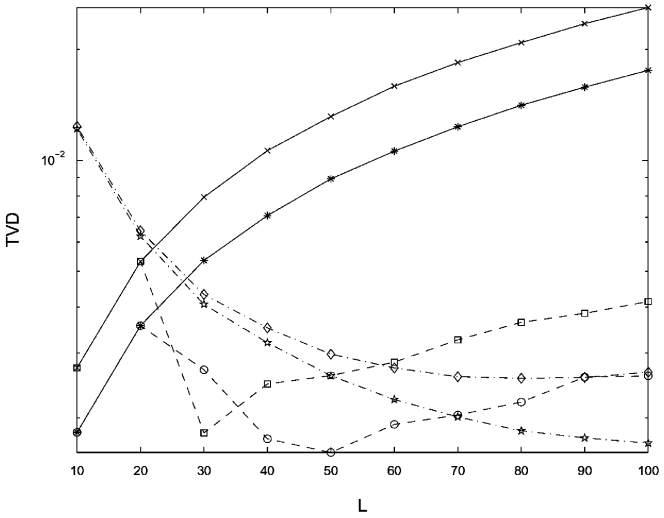


Fig. 4. The TVD between the exact pmf of $\Lambda_2 \mid H_1$, obtained through four-fold integration, and its approximations based on Binomial I and II, and DLA. $P_0 = 50, b = 50, p_b = 0.9$. Solid line + x-mark: Binomial I, $p_c = 0.01$; solid line + star: Binomial I, $p_c = 0.1$; dashdot line + diamond: DLA, $p_c = 0.01$; dashdot line + pentagram: DLA, $p_c = 0.1$; dashed line + square: Binomial II, $p_c = 0.01$; dashed line + circle: Binomial II, $p_c = 0.1$.

that the total number (N) of sensors within a ROI is a Poisson (λ) RV

$$f(N) = \frac{e^{-\lambda}\lambda^N}{N!} (N = 0, \ldots, \infty) \tag{38}$$

where λ is the average number of sensors deployed in the ROI. We denote the total number of detections received at the fusion center as $\Lambda_3 = \sum_{i=1}^{N} I_i'$.

Given the above assumptions and notations, we derive the distribution of $\Lambda_3 \mid (H_1, x_t, y_t)$, as summarized in the following theorem.

Theorem 4: Under hypothesis H_1, and conditioned on the target's two coordinates x_t and y_t, Λ_3 follows a Poisson distribution, namely:

$$\Lambda_3 \sim \text{Poisson}(\lambda\mu(x_t, y_t)) \tag{39}$$

where $\mu(x_t, y_t)$ has been defined in Theorem 3.

Proof: See Appendix IV.

With Theorem 4, it is easy to show that

$$\Pr(\Lambda_3 = k \mid H_1) = \frac{1}{b^2} \int\int \frac{e^{-\lambda\mu(x_t,y_t)} \left[\lambda\mu(x_t, y_t)\right]^k}{k!} dx_t dy_t \tag{40}$$

for $k = 0, \ldots, \infty$. Note that in practice, to get the pmf of $\Lambda_3 \mid H_1$, we only calculate a finite number of terms, which have nonnegligible values.

Similar to the proof of Theorem 4, we can show that under hypothesis H_0, Λ_3 follows a Poisson $(\lambda[p_c + (1-2p_c)p_{\text{fa}}])$ distribution.

2) Poisson Approximation: As we can see, the evaluation of (40) involves a fourfold integration. Since conditioned on $(x_t, y_t), \Lambda_3 \mid H_1$ is a Poisson distribution, here we try to approximate the pmf of $\Lambda_3 \mid H_1$ with a Poisson distribution, whose mean matches that of $\Lambda_3 \mid H_1$. In other words, we use a Poisson $(\lambda[p_c + (1-2p_c)\bar{p}_d])$ to approximate the pmf of $\Lambda_3 \mid H_1$, in which $\bar{p}_d$ has been defined in (20).

3) Gaussian Sample Approximation: Again, the Poisson approximation only matches the first moment, or the mean of $\Lambda_3 \mid H_1$. Similar to the proof of Proposition 1, we derive the variance of $\Lambda_3 \mid H_1$. We give it without proof in the following proposition.

Proposition 3: The variance of $\Lambda_3 \mid H_1$ is

$$\text{Var}(\Lambda_3 \mid H_1) = \lambda[p_c + (1-2p_c)\bar{p}_d] + \lambda^2(1-2p_c)^2 \left[E\left[\bar{p}_d^2(x_t, y_t)\right] - \bar{p}_d^2\right]. \tag{41}$$

Once the mean and variance of $\Lambda_3 \mid H_1$ are available, we use the samples of a Gaussian distribution to approximate the pmf of $\Lambda_3 \mid H_1$, since Gaussian pdf is simple to calculate, and its shape is completely characterized by its first two moments. Now we use the samples of a Gaussian $(\lambda[p_c + (1-2p_c)\bar{p}_d], \sqrt{\text{Var}(\Lambda_3 \mid H_1)})$ pdf to approximate the pmf of $\Lambda_3 \mid H_1$

$$\Pr(\Lambda_3 = k \mid H_1) \approx \frac{e^{-\{k-\lambda[p_c+(1-2p_c)\bar{p}_d]\}^2/[2\text{Var}(\Lambda_3 \mid H_1)]}}{\sqrt{2\pi\text{Var}(\Lambda_3 \mid H_1)}}. \tag{42}$$

Again, terms in (42) are normalized so that the sum of the pmf values is one.

4) Experimental Results: In Fig. 5, for two different channel error rates, we compare the ROC curves obtained through simulation, fourfold integrations, Poisson approximation, and Gaussian sample approximation. In the simulation, 10^7 Monte Carlo runs have been used. In each Monte Carlo run, we use independent target coordinates (x_t, y_t), an independent number of sensors N, and independent sensor coordinates (x_i, y_i), for $i = 1, \ldots, N$. As a result, a total of 10^7 independent (x_t, y_t) pairs have been used in the simulation. As expected, the ROC curve corresponding to a smaller channel error rate $(p_c = 0.01)$ is above that corresponding to a larger channel

TABLE III
COMPARISON OF DIFFERENT APPROXIMATION METHODS

Approximation method	Computation load	Approximation Accuracy
Poisson Approximation	1 single-fold integration	High for small λ
Gaussian Sample Approximation	1 single-fold integration and 1 four-fold integration	High for large λ
Exact Evaluation	O(λ) four-fold integrations	Exact result

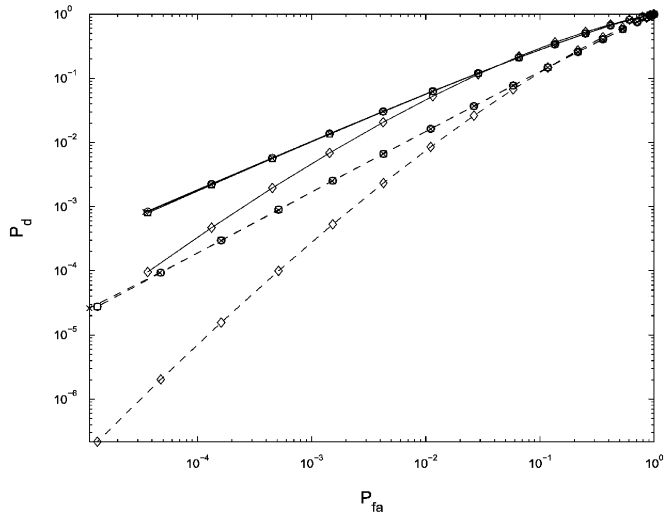


Fig. 5. The ROC curves obtained by simulation, four-fold integration, Poisson approximation, and Gaussian sample approximation. $n = 2, d_0 = 1, b = 50, \lambda = 10, P_0 = 50$, and $\tau = 0.2$. Solid line: $p_c = 0.01$; dashed line: $p_c = 0.4$. x-mark: simulation; circle: integration; square: Poisson approximation; diamond: Gaussian sample approximation.

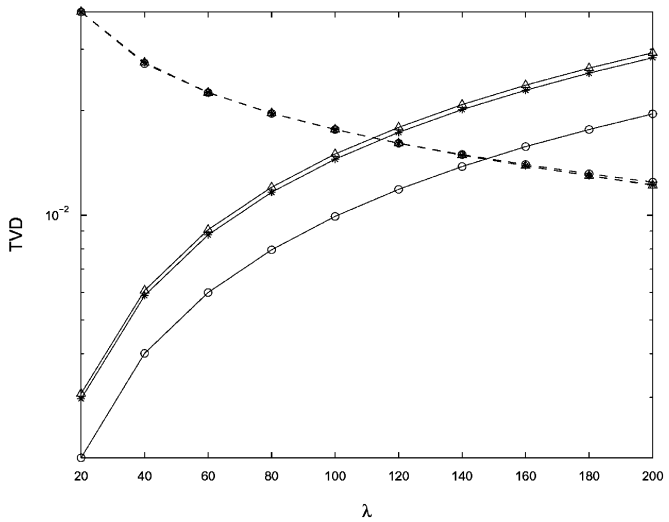


Fig. 6. The ROC curves obtained by simulation, fourfold integration, Poisson approximation, and Gaussian sample approximation. $n = 2, d_0 = 1, b = 50, P_0 = 50$, and $\tau = 0.2$. Solid line: Poisson approximation; dashed line: Gaussian sample approximation. Triangle: $p_c = 0.001$; star: $p_c = 0.01$; circle: $p_c = 0.1$.

error rate ($p_c = 0.4$), meaning that a system with a smaller channel error has a superior detection performance. It is clear that the curves obtained by simulation, integration, and Poisson approximation are indistinguishable. Note that in this particular case, the Gaussian sample approximation is not very accurate, and its corresponding ROC curve is significantly different from the true ROC curve in both cases of $p_c = 0.01$ and $p_c = 0.4$.

To compare the accuracies of different approximations, in Fig. 6, we show their TVDs relative to the true pmf of $\Lambda_3 \,|\, H_1$, as a function of λ. It is clear that the Poisson approximation becomes less accurate as λ increases. Another observation is that the larger p_c is, the better the Poisson approximation is. These phenomena can be explained as follows. According to Proposition 3, the difference between the variance of a Poisson ($\lambda[p_c + (1 - 2p_c)\bar{p}_d]$) RV and that of $\Lambda_3 \,|\, H_1$ is $\lambda^2(1-2p_c)^2[E[\bar{p}_d^2(x_t, y_t)] - \bar{p}_d^2]$, which is an increasing function of λ, and a decreasing function of p_c. A large λ or a small p_c leads to large mismatch in variance, making the Poisson distribution a less accurate approximation.

As we can see, as λ increases, the accuracy of the Gaussian sample approximation improves, and it is not sensitive to the variation of p_c.

In summary, for a small λ, the Poisson approximation is a very good choice, and it needs only a single-fold integration. For a large λ, the Gaussian sample approximation has better accuracy, and it requires only a fourfold integration, which is modest compared with the computation needed for the exact evaluation of the pmf of $\Lambda_3 \,|\, H_1$. When λ is large, and the approximation using samples of the Gaussian $(\lambda[p_c + (1 - 2p_c)\bar{p}_d], \sqrt{\mathrm{Var}(\Lambda_3 \,|\, H_1)})$ pdf is accurate, the number of non-negligible terms of the pmf of $\Lambda_3 \,|\, H_1$ can be deemed approximately as

$$\lambda[p_c + (1-2p_c)\bar{p}_d] + 7\sqrt{\mathrm{Var}(\Lambda_3 \,|\, H_1)} - \max\left(0, \lambda[p_c + (1-2p_c)\bar{p}_d] - 7\sqrt{\mathrm{Var}(\Lambda_3 \,|\, H_1)}\right)$$

which is in the order of λ, considering that $\mathrm{Var}(\Lambda_3 \,|\, H_1)$ is a quadratic function of λ. The accuracies of different approximation methods, and the computation load they require are summarized in Table III.

V. DECISION THRESHOLD AT LOCAL SENSORS

Since the local sensor level threshold τ affects the system level P_{fa} and P_d, it should be designed carefully to achieve a better system level detection performance. However, the calculation of ROC curves involves fourfold integrations, making ROC based optimization procedure computationally prohibitive. Therefore, We resort to the deflection coefficient [23], which is especially useful when the statistical properties of the signal and noise are limited to moments up to a given order. The deflection coefficient is defined as

$$D(\Lambda) = \frac{[E(\Lambda \,|\, H_1) - E(\Lambda \,|\, H_0)]^2}{\mathrm{Var}(\Lambda \,|\, H_0)} \tag{43}$$

where Λ is the test statistic used in a hypothesis testing problem. When $\mathrm{Var}(\Lambda \,|\, H_1) = \mathrm{Var}(\Lambda \,|\, H_0)$, this is in essence the SNR of the detection statistic. The use of deflection criterion leads to

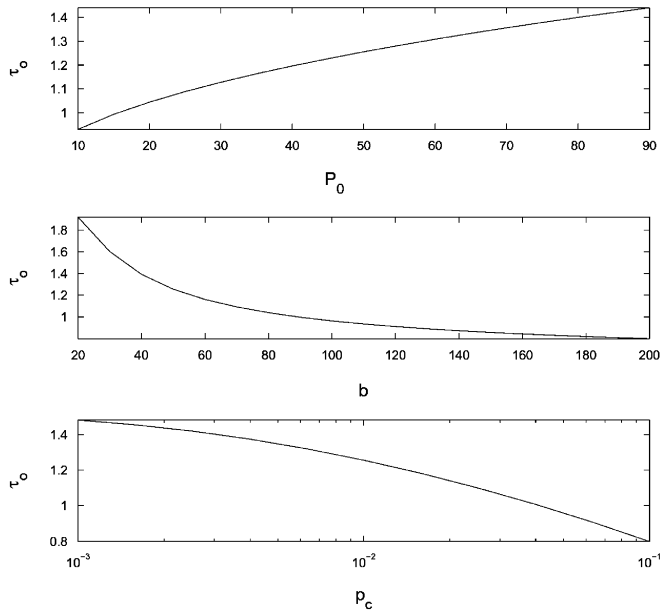


Fig. 7. Top figure: the optimal τ_o as a function of $P_0, n = 2, d_0 = 1, b = 50, p_c = 0.01$; middle figure: τ_o as a function of $b, n = 2, d_0 = 1, P_0 = 50$, and $p_c = 0.01$; bottom figure: τ_o as a function of $p_c, n = 2, d_0 = 1, P_0 = 50, b = 50$.

the optimum LR receiver in many cases of practical importance [23]. In the case of Poisson random number of sensors and noisy channels, the deflection coefficient of Λ_3 is

$$D(\Lambda_3) = \frac{\lambda(1-2p_c)^2(\bar{p}_d - p_{\text{fa}})^2}{p_c + (1-2p_c)p_{\text{fa}}}. \tag{44}$$

Note that the evaluation of the above formula requires only a single-fold integration. We design a sensor-level threshold τ that maximizes $D(\Lambda_3)$.

We investigate the relationship between the optimal threshold τ_o and system parameters such as P_0, the size of the ROI (b), and wireless channels' error rate (p_c), and show the results in Fig. 7. As we can see, τ_o is a monotonically increasing function of P_0. The intuition behind this phenomenon is that for a stronger P_0, a higher τ can reduce the sensor level false alarm rate p_{fa}, while at the same time it can still maintain a relatively high sensor level probability of detection. τ_o is a monotonically decreasing function of b. This is because for a large ROI, the received signal is weak at a larger fraction of sensors, and it is better to use a lower threshold to improve the chance of detecting the target. Also, τ_o is a monotonically decreasing function of p_c. The reason lies in the fact that for more noisy channels, more detections ("1"s) are likely to be erroneously transmitted to the fusion center. A lower threshold is needed to generate a larger number of detections to achieve the same detection performance as that in an ideal-channel case.

An example is provided to illustrate the benefit of the optimal threshold. The system-level ROC curves for different τ are plotted in Fig. 8. As we can see, the ROC curve corresponding to the optimal threshold $\tau_o(1.2557)$ is above those for other thresholds, meaning that τ_o provides the best system level performance.

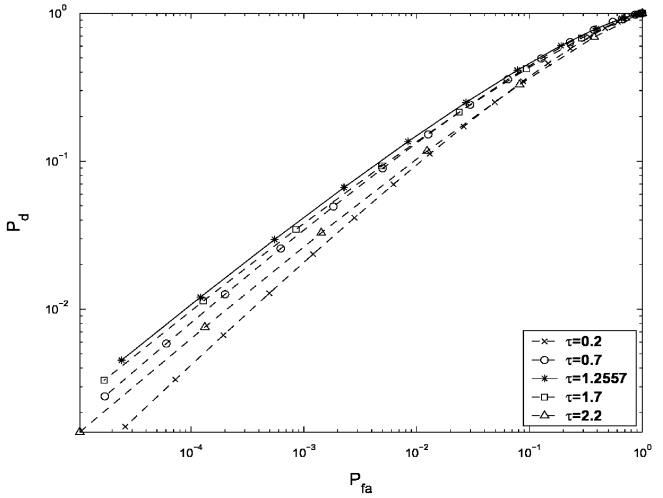


Fig. 8. The ROC curves of a WSN with different sensor-level decision threshold τ, obtained by fourfold integration. $n = 2, d_0 = 1, b = 50, P_0 = 50, \lambda = 20, p_c = 0.01$.

VI. Conclusions and Discussion

For a WSN with randomly deployed sensors, we have proposed a decision fusion rule that uses the total number of detections reported by local sensors as a test statistic. It is assumed that the received signal power is inversely proportional to the distance from the target. For both the case of a known number of sensors and perfect channels and the case of a random number of sensors and noisy channels, we have derived the system-level probabilities of detection and false alarm. The evaluation of the system-level probability of detection involves computationally intensive fourfold integrations. To reduce the computation burden, we have proposed several methods to approximate the system level probability of detection. Their accuracies have been compared in terms of the TVD. Our results show that for a WSN with a small number of sensors, or a small expected number of sensors, the approximation methods that match the mean of the true test statistic are good choices; for a WSN with a large number of sensors, or a large expected number of sensors, the approximation methods that match both the mean and variance of the true test statistic are good choices.

We have shown that the threshold at local sensors is a very important parameter to design, which can affect the system level performance significantly. Here, we have proposed to find the threshold by maximizing the system-level deflection coefficient. The optimal thresholds are calculated numerically for various system parameters. If the signal strength P_0 is high, the size of ROI b is small, and the channel error rate p_c is small, a higher local sensor level threshold τ should be chosen; otherwise, a lower local sensor level threshold τ should be employed to achieve a better performance.

Appendix I
Proof of Theorem 1

Starting from the assumption that the noises at local sensors are i.i.d. and the locations of local sensors are i.i.d., it is easy to

show that given (x_t, y_t) and under hypothesis H_1, I_is are i.i.d. RVs. Being a Bernoulli RV, I_i has the following pmf

$$f(I_i \mid H_1) = \begin{cases} p_{d_i}, & I_i = 1 \\ 1 - p_{d_i}, & I_i = 0. \end{cases} \tag{45}$$

Accordingly, given (x_t, y_t), the moment generating function (MGF) of I_i is

$$\begin{aligned} &M_{I_i \mid (H_1, X_t, Y_t)}(s) \\ &\quad = E\{E[e^{sI_i} \mid (x_i, y_i, x_t, y_t)]\} \\ &\quad = E[1 + (e^s - 1)p_{d_i}(x_i, y_i, x_t, y_t) \mid (x_t, y_t)] \\ &\quad = 1 + (e^s - 1)\bar{p}_d(x_t, y_t) \end{aligned} \tag{46}$$

where

$$\bar{p}_d(x_t, y_t) = \frac{1}{b^2} \int_{-\frac{b}{2}}^{\frac{b}{2}} \int_{-\frac{b}{2}}^{\frac{b}{2}} p_d(P_0, \sqrt{(x - x_t)^2 + (y - y_t)^2}, \tau)\, dx\, dy. \tag{47}$$

Note that in (46), the outer expectation is taken over x_i and y_i only. Because given (x_t, y_t), I_is are i.i.d., from (10), we have

$$\begin{aligned} &M_{\Lambda_1 \mid (H_1, X_t, Y_t)}(s) \\ &\quad = E\left[e^{s \sum_{i=1}^{n} I_i} \middle| (x_t, y_t) \right] \\ &\quad = \prod_{i=1}^{N} E[e^{sI_i} \mid (x_t, y_t)] \\ &\quad = [1 + (e^s - 1)\bar{p}_d(x_t, y_t)]^N \end{aligned} \tag{48}$$

which is nothing but the MGF of a Binomial $(N, \bar{p}_d(x_t, y_t))$ RV.

Appendix II
Proof of Proposition 1

From Theorem 1, it can be readily shown that

$$\mathrm{Var}[\Lambda_1 \mid (H_1, x_t, y_t)] = N\bar{p}_d(x_t, y_t)[1 - \bar{p}_d(x_t, y_t)]. \tag{49}$$

According to a characteristic of the conditional variance, we have

$$\begin{aligned} &\mathrm{Var}(\Lambda_1 \mid H_1) \\ &\quad = E\{\mathrm{Var}[\Lambda_1 \mid (H_1, x_t, y_t)]\} + \mathrm{Var}[E(\Lambda_1 \mid H_1, x_t, y_t)] \\ &\quad = E[N\bar{p}_d(x_t, y_t) - N\bar{p}_d^2(x_t, y_t)] + \mathrm{Var}[N\bar{p}_d(x_t, y_t)] \\ &\quad = N\bar{p}_d - N^2\bar{p}_d^2 + (N^2 - N)E\left[\bar{p}_d^2(x_t, y_t)\right]. \end{aligned} \tag{50}$$

Since

$$\mathrm{Var}[\bar{p}_d(x_t, y_t)] = E\left[\bar{p}_d^2(x_t, y_t)\right] - (\bar{p}_d)^2 \geq 0 \tag{51}$$

we have

$$E[\bar{p}_d^2(x_t, y_t)] \geq (\bar{p}_d)^2. \tag{52}$$

Similarly

$$E\left[p_d^2(x_i, y_i, x_t, y_t) \mid (x_t, y_t)\right] \geq \bar{p}_d^2(x_t, y_t). \tag{53}$$

Taking expectation with respect to (x_t, y_t) on both sides of the above inequality, we finally have

$$E\left[\bar{p}_d^2(x_t, y_t)\right] \leq E\left[p_d^2(x_i, y_i, x_t, y_t)\right]. \tag{54}$$

With (52), (54), and (50), it is easy to show that

$$\begin{aligned} N\bar{p}_d - N\bar{p}_d^2 &\leq \mathrm{Var}(\Lambda_1 \mid H_1) \\ &\leq (N^2 - N)E(p_d^2) + N\bar{p}_d - N^2\bar{p}_d^2. \end{aligned} \tag{55}$$

Appendix III
Proof of Theorem 3

Since the noises at local sensors are i.i.d., the locations of local sensors are i.i.d., and channels are i.i.d., it is easy to show that given (x_t, y_t) and under hypothesis H_1, I_i's are i.i.d. RVs, whose pmf is as follows:

$$f(I_i' \mid H_1) = \begin{cases} p_c + p_{d_i} - 2p_c p_{d_i}, & I_i' = 1 \\ 1 + 2p_c p_{d_i} - p_c - p_{d_i}, & I_i' = 0. \end{cases} \tag{56}$$

Following a similar procedure of the proof of Theorem 1, we can show that conditioned on N and (x_t, y_t), Λ_2 follows a Binomial $(N, p_c + (1 - 2p_c)\bar{p}_d(x_t, y_t))$ distribution. Hence, the corresponding MGF is

$$M_{\Lambda_2 \mid (N, X_t, Y_t, H_1)}(s) = [1 + (e^s - 1)\mu(x_t, y_t)]^N. \tag{57}$$

Taking expectation with respect to N on both sides of the above equation, we have

$$\begin{aligned} &M_{\Lambda_2 \mid (X_t, Y_t, H_1)}(s) \\ &\quad = \sum_{N=0}^{L} \binom{L}{N} p_b^N (1 - p_b)^{L-N} [1 + (e^s - 1)\mu(x_t, y_t)]^N \\ &\quad = [1 + (e^s - 1)p_b\mu(x_t, y_t)]^L \end{aligned} \tag{58}$$

which is the MGF of a RV with a Binomial $(L, p_b\mu(x_t, y_t))$ distribution. Note that we obtain the second equality in (58) according to Binomial Theorem.

Appendix IV
Proof of Theorem 4

Following a similar procedure of the proof of Theorem 3, we can show that conditioned on N and (x_t, y_t), Λ_3 follows a Binomial $(N, p_c + (1 - 2p_c)\bar{p}_d(x_t, y_t))$ distribution. Hence, the corresponding MGF is

$$M_{\Lambda_3 \mid (N, X_t, Y_t, H_1)}(s) = [1 + (e^s - 1)\mu(x_t, y_t)]^N. \tag{59}$$

Taking expectation with respect to N on both sides of the above equation, we have

$$\begin{aligned} M_{\Lambda_3 \mid (X_t, Y_t, H_1)}(s) &= \sum_{N=0}^{\infty} \frac{e^{-\lambda}[\lambda + (e^s - 1)\lambda\mu(x_t, y_t)]^N}{N!} \\ &= e^{(e^s - 1)\lambda\mu(x_t, y_t)} \end{aligned} \tag{60}$$

which is the MGF of a RV with a Poisson $(\lambda\mu(x_t, y_t))$ distribution.

ACKNOWLEDGMENT

The authors are very grateful for the valuable comments and suggestions made by the anonymous reviewers.

REFERENCES

[1] S. Kumar, F. Zhao, and D. Shepherd, Eds., "Special issue on collaborative signal and information processing in microsensor networks," *IEEE Signal Process. Mag.*, vol. 19, no. 2, Mar. 2002.

[2] A. Sayeed, D. Estrin, G. Pottie, and K. Ramchandran, Eds., "Special issue on self-organizing distributed collaborative sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, Apr. 2005.

[3] Z. Chair and P. K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, no. 1, pp. 98–101, Jan. 1986.

[4] P. K. Willett, P. F. Swaszek, and R. S. Blum, "The good, bad, and ugly: Distributed detection of a known signal in dependent Gaussian noise," *IEEE Trans. Signal Process.*, vol. 48, no. 12, pp. 3266–3279, Dec. 2000.

[5] E. Drakopoulos and C. C. Lee, "Optimum multisensor fusion of correlated local decisions," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 27, no. 4, pp. 593–605, Jul. 1991.

[6] M. Kam, Q. Zhu, and W. S. Gray, "Optimal data fusion of correlated local decisions in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 7, pp. 916–920, Jul. 1992.

[7] C. Rago, P. K. Willett, and Y. Bar-Shalom, "Censoring sensors: A low-communication-rate scheme for distributed detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 4, pp. 554–568, Apr. 1996.

[8] C. T. Yu and P. K. Varshney, "Bit allocation for discrete signal detection," *IEEE Trans. Commun.*, vol. 46, no. 2, pp. 173–175, Feb. 1998.

[9] J. Hu and R. Blum, "On the optimality of finite-level quantization for distributed signal detection," *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 1665–1671, May 2001.

[10] J. Chamberland and V. V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 407–416, Feb. 2003.

[11] R. Niu, P. K. Varshney, and Q. Cheng, "Distributed detection in a large wireless sensor network," *Int. J. Inf. Fusion*, vol. 7, no. 4, pp. 380–394, Dec. 2006.

[12] R. Niu and P. K. Varshney, "Distributed detection and fusion in a large wireless sensor network of random size," *EURASIP J. Wireless Commun. Netw.*, vol. 2005, no. 4, pp. 462–472, Sep. 2005.

[13] S. A. Aldosari and J. M. F. Moura, "Detection in sensor networks: The Saddlepoint approximation," *IEEE Trans. Signal Process.*, vol. 55, no. 1, pp. 327–340, Jan. 2007.

[14] L. E. Kinsler and A. R. Frey, *Fundamentals of Acoustics*. New York: Wiley, 1962.

[15] L. Le Cam, "An approximation theorem for the poisson binomial distribution," *Pacific J. Math.*, vol. 10, no. 4, pp. 1181–1197, 1960.

[16] A. D. Barbour, L. Holst, and S. Janson, *Poisson Approximation, Oxford Studies in Probability*, 2nd ed. Oxford, U.K.: Clarendon , 1992.

[17] B. Roos, "Binomial approximation to the Poisson binomial distribution: The Krawtchouk expansion," *Theory Probab. Appl.*, vol. 45, no. 2, pp. 258–272, 2000.

[18] K. P. Choi and A. Xia, "Approximating the number of successes in independent trials: Binomial versus poisson," *Ann. Appl. Probab.*, vol. 12, no. 4, pp. 1139–1148, 2002.

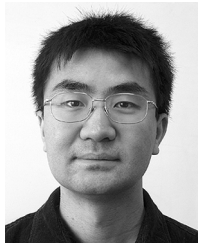
[19] R. Niu and P. K. Varshney, "Target location estimation in sensor networks with quantized data," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4519–4528, Dec. 2006.

[20] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1984.

[21] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

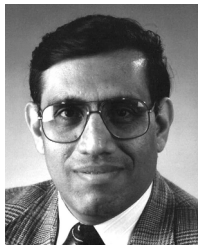
[22] Y. Sung, L. Tong, and A. Swami, "Asymptotic locally optimal detector for large-scale sensor networks under the poisson regime," *IEEE Trans. Signal Process.*, vol. 53, no. 6, pp. 2005–2017, Jun. 2005.

[23] B. Picinbono, "On deflection as a performance criterion in detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 3, pp. 1072–1081, Jul. 1995.



Ruixin Niu (M'04) received the B.S. degree from Xi'an Jiaotong University, Xi'an, China, in 1994, the M.S. degree from the Institute of Electronics, Chinese Academy of Sciences, Beijing, in 1997, and the Ph.D. degree from the University of Connecticut, Storrs, in 2001, all in electrical engineering.

He is currently a Research Assistant Professor with Syracuse University, Syracuse, NY. His research interests are in the areas of statistical signal processing and its applications, including detection, estimation, data fusion and communications. He received the Fusion 2004 Best Paper Award, from the Seventh International Conference on Information Fusion.



Pramod K. Varshney (F'97) was born in Allahabad, India, on July 1, 1952. He received the B.S. degree in electrical engineering and computer science (with highest honors), and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois at Urbana-Champaign in 1972, 1974, and 1976 respectively.

Since 1976, he has been with Syracuse University, Syracuse, NY, where he is currently a Distinguished Professor of electrical engineering and computer science. His current research interests are in distributed sensor networks and data fusion, detection and estimation theory, wireless communications, image processing, radar signal processing, and remote sensing. He has published extensively. He is the author of Distributed Detection and Data Fusion (New York: Springer-Verlag, 1997).

Dr. Varshney was a James Scholar, a Bronze Tablet Senior, and a Fellow while with the University of Illinois. He is a member of Tau Beta Pi and is the recipient of the 1981 ASEE Dow Outstanding Young Faculty Award. He was elected Fellow of the IEEE in 1997. He was the Guest Editor of the Special Issue on Data Fusion of the PROCEEDINGS OF THE IEEE, January 1997. In 2000, he received the Third Millennium Medal from the IEEE and Chancellor's Citation for exceptional academic achievement at Syracuse University. He serves as a distinguished lecturer for the AES society of the IEEE. He is on the editorial boards of the *International Journal of Distributed Sensor Networks* and the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He was the President of International Society of Information Fusion during 2001.