Distributed Detection and Fusion in a Large Wireless Sensor Network of Random Size

Ruixin Niu

Department of Electrical Engineering and Computer Science, Syracuse University, 335 Link Hall, Syracuse,
NY 13244-1240, USA
Email: rniu@ecs.syr.edu

Pramod K. Varshney

Department of Electrical Engineering and Computer Science, Syracuse University, 335 Link Hall, Syracuse,
NY 13244-1240, USA
Email: varshney@ecs.syr.edu

For a wireless sensor network (WSN) with a random number of sensors, we propose a decision fusion rule that uses the total number of detections reported by local sensors as a statistic for hypothesis testing. We assume that the signal power attenuates as a function of the distance from the target, the number of sensors follows a Poisson distribution, and the locations of sensors follow a uniform distribution within the region of interest (ROI). Both analytical and simulation results for system-level detection performance are provided. This fusion rule can achieve a very good system-level detection performance even at very low signal-to-noise ratio (SNR), as long as the average number of sensors is sufficiently large. For all the different system parameters we have explored, the proposed fusion rule is equivalent to the optimal fusion rule, which requires much more prior information. The problem of designing an optimum local sensor-level threshold is investigated. For various system parameters, the optimal thresholds are found numerically by maximizing the deflection coefficient. Guidelines on selecting the optimal local sensor-level threshold are also provided.

Keywords and phrases: wireless sensor networks, distributed detection, decision fusion, deflection coefficient.

1. INTRODUCTION

Recently, wireless sensor networks (WSNs) have attracted much attention and interest, and have become a very active research area. Due to their high flexibility, enhanced surveillance coverage, robustness, mobility, and cost effectiveness, WSNs have wide applications and high potential in military surveillance, security, monitoring of traffic, and environment. Usually, a WSN consists of a large number of low-cost and low-power sensors, which are deployed in the environment to collect observations and preprocess the observations. Each sensor node has limited communication capability that allows it to communicate with other sensor nodes via a wireless channel. Normally, there is a fusion center that processes data from sensors and forms a global situational assessment.

In a typical WSN, sensor nodes are powered by batteries, and hence have a very frugal energy budget. To maintain

longer lifetimes of the sensors, all aspects of the network should be energy efficient. In [1], a data-centric energy efficient routing protocol is proposed. By using existing wireless local area network (WLAN) technologies, in [2], authors present a cluster-based ad hoc routing scheme for a multihop sensor network. In [3], an on-demand clustering mechanism, passive clustering, is presented to overcome two limitations of ad hoc routing schemes, namely limited scalability and the inability to adapt to high-density sensor distributions.

Many other important aspects of WSNs have been investigated too, such as distributed data compression and transmission, and collaborative signal processing [4, 5]. In a WSN, detection, classification, and tracking of targets require collaboration between sensor nodes. Distributed signal processing in a sensor network reduces the amount of communication required in the network, lowers the risk of network node failures, and prevents the fusion center from being overwhelmed by huge amount of raw data from sensors. In this paper, we focus on distributed target detection, one of the most important functions that a WSN needs to perform. There are already many papers on the conventional

distributed detection problem. In [6, 7], optimum fusion rules have been obtained under the conditional independence assumption. Decision fusion with correlated observations has been investigated in [8, 9, 10, 11]. There are also many papers on the problem of distributed detection with constrained system resources [12, 13, 14, 15, 16, 17, 18]. More specifically, these papers have proposed solutions to optimal bit allocation under a communication constraint.

However, most of these results are based on the assumption that the local sensors' detection performances, namely either the local sensors' signal-to-noise ratio (SNR) or their probability of detection and false alarm rate, are known to the fusion center. For a dynamic target and passive sensors, it is very difficult to estimate local sensors' performances via experiments because these performances are time varying as the target moves through the wireless sensor field. Even if the local sensors can somehow estimate their detection performances in real time, it will be very expensive to transmit them to the fusion center, especially for a WSN with very limited system resources. Usually a WSN consists of a large number of low-cost and low-power sensors, which are densely deployed in the surveillance area. Taking advantage of these unique characteristics of WSNs, in our previous paper [19], we proposed a fusion rule that uses the total number of detections ("1"s) transmitted from local sensors as the statistic.

In [19], we assumed that the total number of sensors in the region of interest (ROI) is known to the WSN. However, 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 random variable (RV). For example, in a battlefield or a hostile region, many microsensors can be deployed from an airplane to form a WSN. Data are transmitted from sensors to an access point, which could be an airplane that flies over the sensor field and collects data from the sensors. The total number of sensors within the network and the total number of sensors that can communicate with the access point (the flying airplane) at a particular time are RVs. In this paper, the results presented in [19] are extended to this more general situation. The performance of the fusion rule proposed in [19] will be analyzed with this extra uncertainty about the total number of sensors.

In Section 2, basic assumptions regarding the WSN are made, the signal attenuation model is provided, and the fusion rule based on the total number of detections from local sensors is introduced. In addition, it is shown that the proposed fusion rule can be adapted well to a large network with multiple-layer hierarchical structure. Analytical methods to determine the system-level detection performance are presented in Section 3. There, asymptotic detection performance is studied. In addition, the proposed fusion rule is compared to the likelihood-ratio (LR) based optimal fusion rule, which requires much more prior information. Simulation results are also provided to confirm our analyses. In Section 4, the problem of designing an optimum local sensor-level threshold is investigated, and the optimum
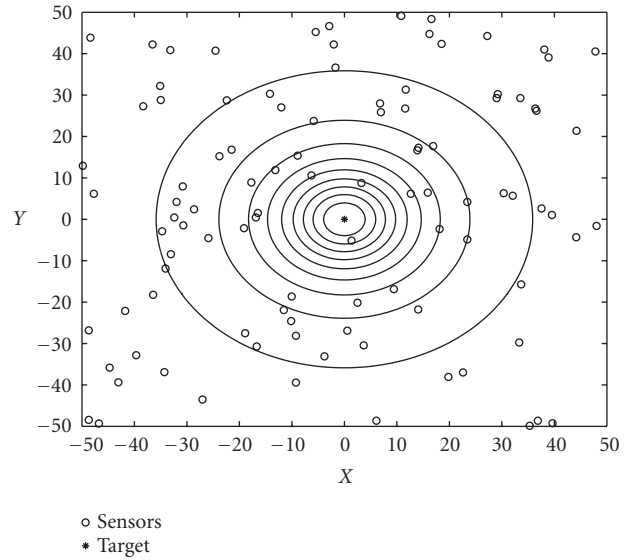

FIGURE 1: The signal power contour of a target located in a sensor field.

thresholds for various system parameters are found numerically. Conclusions and discussion are provided in Section 5.

2. SYSTEM MODEL AND DECISION FUSION RULE

2.1. Problem formulation

As shown in Figure 1, a total of N sensors are randomly deployed in the ROI, which is a square with area b^2. N is an RV that follows a Poisson distribution:

$$p(N) = \frac{\lambda^N e^{-\lambda}}{N!}, \quad N = 0, \ldots, \infty. \tag{1}$$

The locations of sensors are unknown to the WSN, but it is assumed that they are independent and identically distributed (i.i.d.) and follow a uniform distribution in the ROI:

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

for $i = 1, \ldots, N$, where (x_i, y_i) are the coordinates of sensor i.

Noises at local sensors are i.i.d. and follow the standard Gaussian distribution with zero mean and unit variance:

$$n_i \sim \mathcal{N}(0, 1), \quad i = 1, \ldots, N. \tag{3}$$

For a local sensor i, the binary hypothesis testing problem is

$$\begin{aligned} H_1 &: s_i = a_i + n_i, \\ H_0 &: s_i = n_i, \end{aligned} \tag{4}$$

where s_i is the received signal at sensor i, and a_i is the amplitude of the signal that is emitted by the target and received at

sensor i. We adopt the same isotropic signal power attenuation model as that presented in [19]

$$a_i^2 = \frac{P_0}{1 + \alpha d_i^n}, \tag{5}$$

where P_0 is the signal power emitted by the target at distance zero, d_i is the distance between the target and local sensor i:

$$d_i = \sqrt{(x_i - x_t)^2 + (y_i - y_t)^2}, \tag{6}$$

and (x_t, y_t) are the coordinates of the target. We further assume that the location of the target also follows a uniform distribution within the ROI. n is the signal decay exponent and takes values between 2 and 3. α is an adjustable constant, and a larger α implies faster signal power decay. Note that the signal attenuation model can be easily extended to 3-dimensional problems. Our attenuation model is similar to that used in [20]. The difference is that in the denominator of (5), instead of d_i^n, we use $1 + \alpha d_i^n$. By doing so, our model is valid even if the distance d_i is close to or equal to 0. When d_i is large ($\alpha d_i^n \gg 1$), the difference between these two models is negligible.

In this paper, we do not specify the type of the passive sensors, and the power decay model adopted here is quite general. For example, in a radar or wireless communication system, for an isotropically radiated electromagnetic wave that is propagating in free space, the power is inversely proportional to the square of the distance from the transmitter [21, 22]. Similarly, when spherical acoustic waves radiated by a simple source are propagating through the air, the intensity of the waves will decay at a rate inversely proportional to the square of the distance [23].

Because the noise has unit variance, it is evident that the SNR at local sensor i is

$$\text{SNR}_i = a_i^2 = \frac{P_0}{1 + \alpha d_i^n}. \tag{7}$$

We define the SNR at distance zero as

$$\text{SNR}_0 = 10 \log_{10} P_0. \tag{8}$$

Assuming that all the local sensors use the same threshold τ to make a decision and with the Gaussian noise assumption, we have the local sensor-level false alarm rate and probability of detection:

$$p_{\text{fa}} = \int_\tau^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = Q(\tau), \tag{9}$$

$$p_{d_i} = \int_\tau^\infty \frac{1}{\sqrt{2\pi}} e^{-(t-a_i)^2/2} dt = Q(\tau - a_i), \tag{10}$$

where $Q(\cdot)$ is the complementary distribution function of the standard Gaussian, that is,

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt. \tag{11}$$

We assume that the ROI is very large and the signal power decays very fast. Hence, only within a very small fraction of the ROI, which is the area surrounding the target, the received signal power is significantly larger than zero. By ignoring the border effect of the ROI, we assume that the target is located at the center of the ROI, without any loss of generality. As a result, at a particular time, only a small subset of sensors can detect the target. To save communication and energy, a local sensor only transmits data ("1"s) to the fusion center when its signal exceeds the threshold τ.

2.2. Decision fusion rule

We denote the binary data from local sensor i as $I_i = \{0, 1\}$ ($i = 1, \ldots, N$). I_i takes the value 1 when there is a detection; otherwise, it takes the value 0.

We know that the optimal decision fusion rule is the Chair-Varshney fusion rule [6], and it is a threshold test of the following statistic:

$$\begin{aligned}
\Lambda_0 &= \sum_{i=1}^N \left[I_i \log \frac{p_{d_i}}{p_{\text{fa}_i}} + (1 - I_i) \log \frac{1 - p_{d_i}}{1 - p_{\text{fa}_i}} \right] \\
&= \sum_{i=1}^N I_i \log \frac{p_{d_i}(1 - p_{\text{fa}_i})}{p_{\text{fa}_i}(1 - p_{d_i})} + \sum_{i=1}^N \log \frac{1 - p_{d_i}}{1 - p_{\text{fa}_i}}.
\end{aligned} \tag{12}$$

This fusion statistic is equivalent to a weighted summation of all the detections ("1"s) that a fusion center receives. The decision from a sensor with a better detection performance, namely higher p_{d_i} and lower p_{fa_i}, gets a greater weight, which is given by $\log(p_{d_i}(1 - p_{\text{fa}_i})/p_{\text{fa}_i}(1 - p_{d_i}))$.

As long as the threshold τ is known, the probability of false alarm at each sensor is known ($p_{\text{fa}_i} = p_{\text{fa}}$) from (9). However, at each sensor, it is very difficult to calculate p_{d_i} since according to (10), p_{d_i} is decided by each sensor's distance to the target and the amplitude of the target's signal. To make matters worse, we do not even know the total number of sensors N because the fusion center only receives data from those sensors whose received signals exceed the threshold τ, as we have assumed in Section 2.1. An alternative scheme would be that each sensor transmits raw data s_i to the fusion center, and the fusion center will make a decision based on these raw measurements. However, the transmission of raw data will be very expensive especially for a typical WSN with very limited energy and bandwidth. It is desirable to transmit only binary data to the fusion center. 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 since the information about which sensor reports a "1" is of little use to the fusion center. As proposed in [19], the system-level decision is made by first counting the number of detections made by local sensors and then comparing it with a threshold T:

$$\Lambda = \sum_{i=1}^N I_i \underset{H_0}{\overset{H_1}{\gtrless}} T, \tag{13}$$
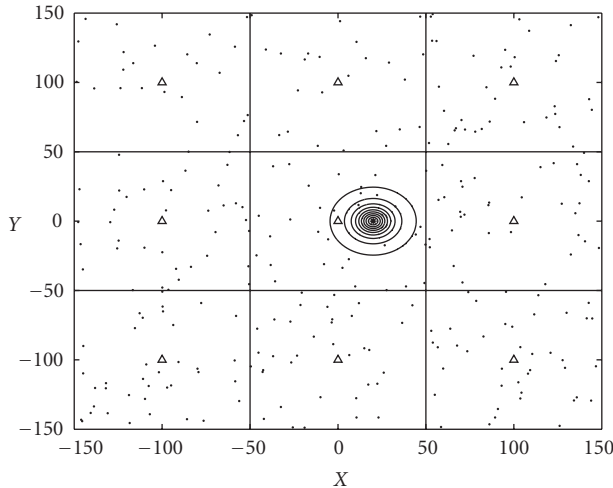
FIGURE 2: The signal power contour of a target located in a sensor field with nine cluster heads and their corresponding subregions. Points: sensors; triangles: cluster heads; star: target.

FIGURE 3: Three-layered hierarchical sensor network structure.

where $I_i = \{0, 1\}$ is the local decision made by sensor i. We also call this fusion rule the "counting rule."

2.3. Hierarchical network structure

In this paper, we focus on the application aspect of the WSN. Routing protocols and network structures are beyond the scope of this paper. In Sections 2.1 and 2.2, a very simple network structure is implied. That is, all the sensors in the ROI report directly to the fusion center. However, our analysis results, which are based on this simple assumption and will be presented later, are quite general and can be applied to various scenarios and network structures. In this section, we give an example to show how the proposed approach can be adapted to complicated and practical applications.

Suppose that the sensor field is quite vast and the signal decays very fast as the distance from the target increases. As a result, only a tiny fraction of the sensors can detect the signals from the target, as illustrated in Figure 2. Most sensors' measurements are just pure noises. Since the local decisions from these sensors do not convey much information about the target, it is neither very useful nor energy efficient to transmit them to the fusion center. When the sensor network is very large, there is also the issue of scalability. One reasonable solution is to use a three-layered hierarchical network structure, as shown in Figure 3. Sensors that are close to each other will form a cluster and each cluster has its own cluster head or cluster master, which serves as the local fusion center and is supposed to have more powerful computation and communication capabilities. Each cluster is in charge of the surveillance of a subregion of the whole ROI, as shown in Figure 2. Instead of transmitting data to a faraway central fusion center, sensors will send data to their corresponding cluster head. Based on data transmitted from sensors located within a specific cluster/subregion, the corresponding cluster head will make a decision about target presence/absence within tha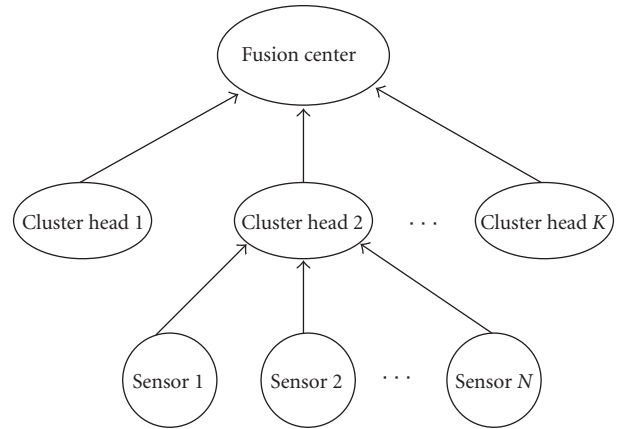t subregion. The decisions from cluster heads will be further transmitted to the fusion center to inform it if there is a target or event in specific subregions.

The theoretical analysis provided later in this paper can be used to evaluate the detection performance at the cluster-head level, as long as the assumptions made in Section 2.1 are still valid within each cluster/subregion.

3. PERFORMANCE ANALYSIS

In this section, the system-level detection performance, namely the probability of false alarm P_{fa} and probability of detection P_d at the fusion center, will be derived, and the analytical results will be compared to simulation results.

3.1. System-level false alarm rate

At the fusion center level, the probability of false alarm P_{fa} is

$$P_{\text{fa}} = \sum_{N=T}^{\infty} p(N) \Pr\{\Lambda \geq T | N, H_0\}. \tag{14}$$

Obviously, for a given N, under hypothesis H_0, Λ follows a binomial (N, p_{fa}) distribution. When N is large enough, $\Pr\{\Lambda \geq T | N, H_0\}$ can be calculated by using Laplace-De Moivre approximation [24]:

$$\Pr\{\Lambda \geq T | N, H_0\} = \sum_{i=T}^{N} \binom{N}{i} p_{\text{fa}}^i (1 - p_{\text{fa}})^{N-i}$$
$$\simeq Q\left(\frac{T - N p_{\text{fa}}}{\sqrt{N p_{\text{fa}}(1 - p_{\text{fa}})}}\right). \tag{15}$$

It is well known that the kurtosis of a Poisson distribution is $3 + (1/\lambda)$. As λ increases, the kurtosis of this Poisson distribution approaches that of a Gaussian distribution, and its distribution has a light tail. This can also be explained by the unique characteristic of the Poisson distribution. A Poisson RV with mean λ can be deemed as the summation of M i.i.d. Poisson RVs with mean $\lambda_0 = \lambda/M$. Therefore, a Poisson RV with a very large λ is a summation of a very large number (M) of i.i.d. Poisson RVs with a constant mean λ_0, and
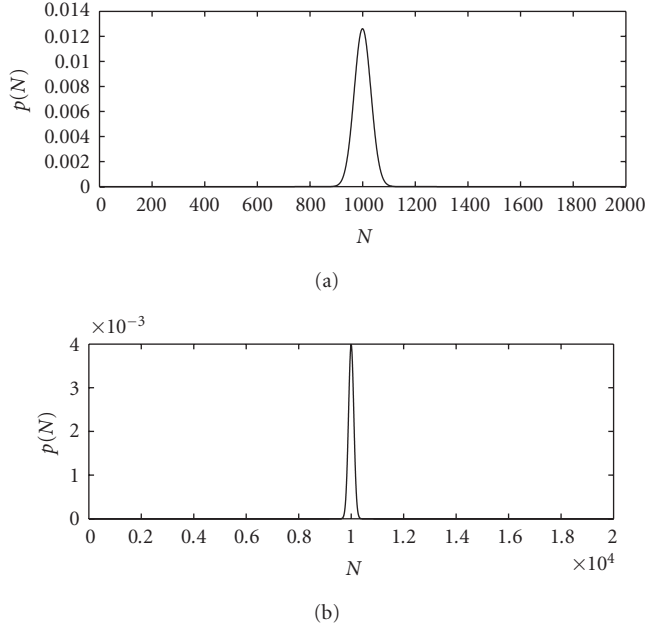
FIGURE 4: The probability mass function for Poisson distributions. (a): $\lambda = 1000$; (b): $\lambda = 10\,000$.

its distribution approaches a Gaussian distribution, according to the central limit theorem (CLT). As a result, when λ is large, the probability mass of N will concentrate around the average value (λ). This phenomenon is illustrated in Figure 4, where the probability mass function of N has been plotted for $\lambda = 1000$ and $\lambda = 10\,000$. Due to this characteristic of the Poisson distribution, using the fact that both the mean and variance of a Poisson RV are λ, we have the following approximation when λ is large:

$$\Pr\{\lambda - 6\sqrt{\lambda} \le N \le \lambda + 6\sqrt{\lambda}\} \simeq 1 \qquad (16)$$

or

$$\sum_{N_1}^{N_3} \frac{e^{-\lambda}\lambda^N}{N!} \simeq 1, \qquad (17)$$

where $N_1 = \lfloor \lambda - 6\sqrt{\lambda} \rfloor$ and $N_3 = \lceil \lambda + 6\sqrt{\lambda} \rceil$.

Hence, for a large λ, a "typical" N is also a large number. The probability that N takes a small value is negligible. For example, when $\lambda = 1000$, $\Pr\{N < 810\} = 2.4 \times 10^{-10}$; when $\lambda = 10\,000$, $\Pr\{N < 9400\} = 6.6 \times 10^{-10}$. Therefore, when λ is large enough, we have

$$
\begin{aligned}
P_{\text{fa}} &= \sum_{N=0}^{\infty} p(N) \sum_{i=T}^{N} \binom{N}{i} p_{\text{fa}}^i (1 - p_{\text{fa}})^{N-i} \\
&\simeq \sum_{N=N_2}^{N_3} \frac{\lambda^N e^{-\lambda}}{N!} Q\left(\frac{T - N p_{\text{fa}}}{\sqrt{N p_{\text{fa}}(1 - p_{\text{fa}})}}\right) \\
&= \sum_{N=N_2}^{N_3} \frac{\lambda^N e^{-\lambda}}{N!} Q\left(\frac{T - \mu_0}{\sigma_0}\right),
\end{aligned}
\qquad (18)
$$

where $N_2 = \max(T, N_1)$, $\mu_0 \triangleq N p_{\text{fa}}$, and $\sigma_0 \triangleq \sqrt{N p_{\text{fa}}(1 - p_{\text{fa}})}$. Note that for a large N, the Laplace-De Moivre approximation in (15) is valid, and this fact has been used in the derivation of (18). The significance of (17) also lies in the fact that the computation load in calculating P_{fa} or P_d (see (18) and (25)) is reduced significantly since in the computation, a summation of less than or equal to $12\sqrt{\lambda}$ terms is sufficient, rather than a summation of infinite number of terms.

3.2. System-level probability of detection

Because of the nature of this problem, different local sensors will have different p_{d_i}, which is a function of d_i as shown in (10). Therefore, under hypothesis H_1, the total number of detections (Λ) no longer follows a Binomial distribution. It is very difficult to derive an analytical expression for the distribution of Λ. Instead, we will obtain the P_d either through approximation or through simulation. In [19], through approximation by using the CLT, we derived the system level P_d when the number of sensors N is large:

$$\Pr\{\Lambda \ge T | N, H_1\} \simeq Q\left(\frac{T - N\bar{p}_d}{\sqrt{N\bar{\sigma}^2}}\right), \qquad (19)$$

where

$$\bar{p}_d = \frac{2\pi}{b^2} \int_0^{b/2} C(r) r\, dr + \left(1 - \frac{\pi}{4}\right)\gamma, \qquad (20)$$

$$\bar{\sigma}^2 = \frac{2\pi}{b^2} \int_0^{b/2} [1 - C(r)] C(r) r\, dr + \left(1 - \frac{\pi}{4}\right)\gamma(1 - \gamma), \quad (21)$$

$$C(r) = Q\left(\tau - \sqrt{\frac{P_0}{1 + \alpha r^n}}\right), \qquad (22)$$

$$\gamma = Q\left(\tau - \sqrt{\frac{P_0}{1 + \alpha(\sqrt{2}b/2)^n}}\right). \qquad (23)$$

Note that in [19], a different γ is used:

$$\gamma' = Q(\tau) = p_{\text{fa}}. \qquad (24)$$

γ used in this paper is slightly different from that used in [19], and it gives a more accurate approximation. But when the ROI is very large, meaning that b is large, the difference is really negligible. Interested readers can find the detailed derivations in [19]. Taking an average of (19) with respect to N, and similar to the derivation of (18), we have the system level P_d as

$$
\begin{aligned}
P_d &\simeq \sum_{N=N_2}^{N_3} \frac{\lambda^N e^{-\lambda}}{N!} Q\left(\frac{T - N\bar{p}_d}{\sqrt{N\bar{\sigma}^2}}\right) \\
&= \sum_{N=N_2}^{N_3} \frac{\lambda^N e^{-\lambda}}{N!} Q\left(\frac{T - \mu_1}{\sigma_1}\right),
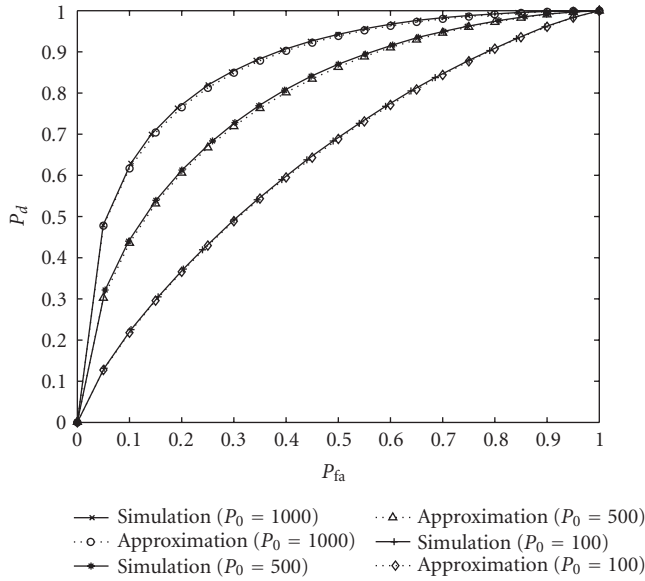\end{aligned}
\qquad (25)
$$

FIGURE 5: ROC curves obtained by analysis and simulations. $\lambda = 1000$, $n = 2$, $b = 100$, $\alpha = 200$, and $\tau = 0.77$, 0.73, 0.67 for $P_0 = 1000$, 500, 100, respectively.
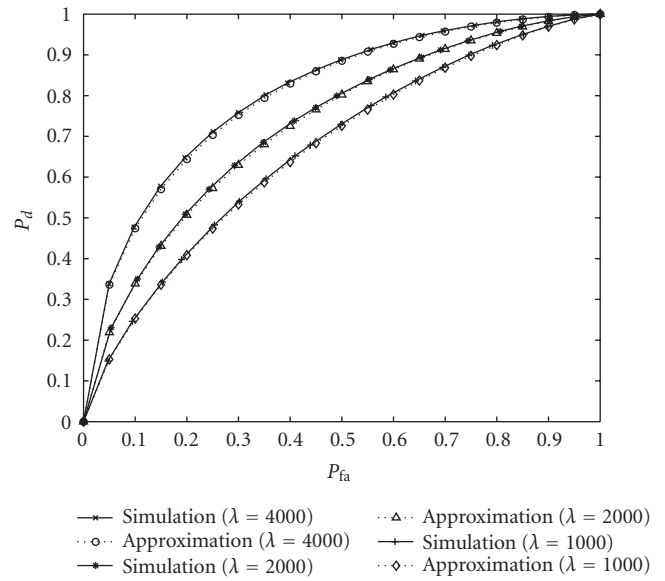

FIGURE 7: ROC curves obtained by analysis and simulations. $P_0 = 500$, $n = 3$, $b = 100$, $\alpha = 40$, and $\tau = 0.90$.
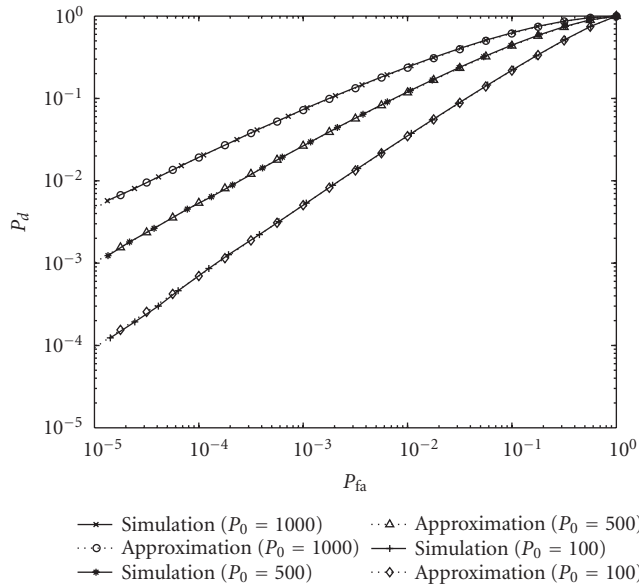

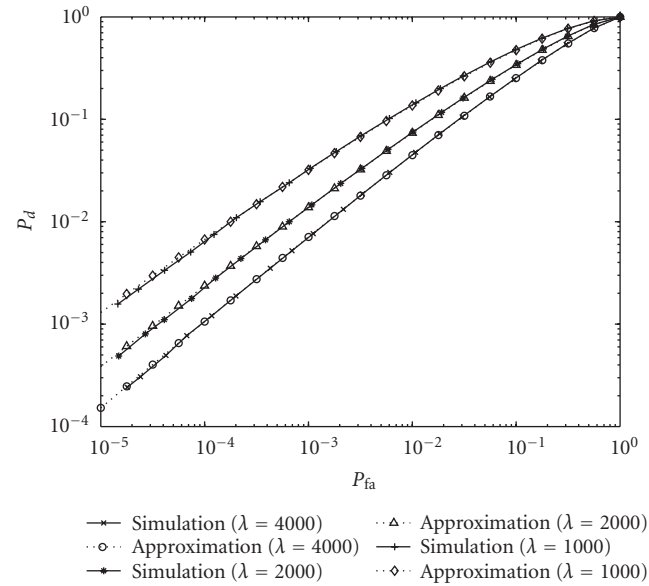FIGURE 6: ROC curves obtained by analysis and simulations. System parameters are the same as those listed in Figure 5.

FIGURE 8: ROC curves obtained by analysis and simulations. $P_0 = 500$, $n = 3$, $b = 100$, $\alpha = 40$, and $\tau = 0.90$.

where $\mu_1 \triangleq N\bar{p}_d$, and $\sigma_1 \triangleq \sqrt{N\bar{\sigma^2}}$. Again, we use the fact that for a large λ, a typical N is large. Therefore, the Gaussian approximation in (19) by using the CLT is still valid.

3.3. Simulation results

The system-level P_d and P_{fa} can also be estimated by simulations. In Figures 5, 6, 7, and 8, the receiver's operative characteristic (ROC) curves obtained by using approximations in Sections 3.1 and 3.2 and those by simulations are plotted for various system parameters. The simulation results in Figures

5 and 7 are based on 10^5 Monte Carlo runs, and the simulation results in Figures 6 and 8 are obtained through 10^7 Monte Carlo runs. From these figures, it is clear that the results obtained by approximations are very close to those obtained by simulations, even when the system-level P_{fa} is very low (Figures 6 and 8).

3.4. Asymptotic analysis

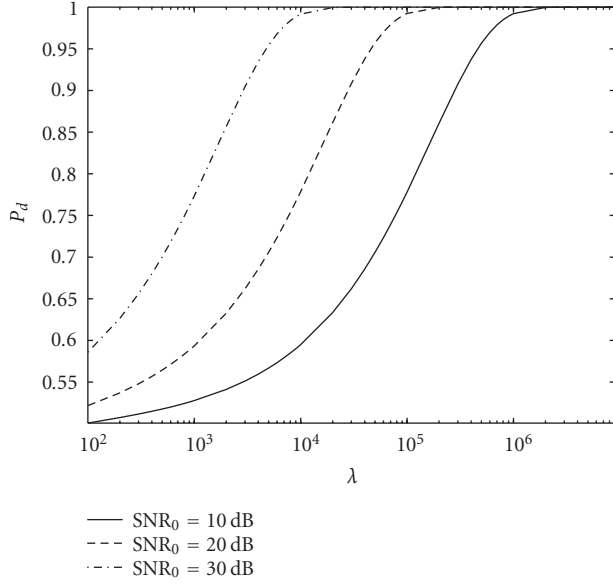It is useful to analyze the system performance when the average number of sensors λ is very large.

FIGURE 9: System-level probability of detection P_d as a function of λ. $n = 2$, $b = 100$, $\alpha = 200$, and $\tau = 0.5$.

FIGURE 10: System-level false alarm rate P_{fa} as a function of λ. $n = 2$, $b = 100$, $\alpha = 200$, and $\tau = 0.5$.

In (18), we know that

$$\max\left(T, \lfloor \lambda - 6\sqrt{\lambda} \rfloor\right) \leq N \leq \lceil \lambda + 6\sqrt{\lambda} \rceil. \quad (26)$$

Hence, as $\lambda \to \infty$, we have $N \to \lambda$, if $T \leq \lceil \lambda + 6\sqrt{\lambda} \rceil$. Assuming that the system-level threshold is in the form of $T = \beta\lambda$, we have

$$P_{\text{fa}} \simeq \sum_{N=N_2}^{N_3} \frac{\lambda^N e^{-\lambda}}{N!} Q\left(\frac{(\beta - p_{\text{fa}})\sqrt{\lambda}}{\sqrt{p_{\text{fa}}(1 - p_{\text{fa}})}}\right). \quad (27)$$

Similarly, from (25), we have

$$P_d \simeq \sum_{N=N_2}^{N_3} \frac{\lambda^N e^{-\lambda}}{N!} Q\left(\frac{(\beta - \bar{p}_d)\sqrt{\lambda}}{\sqrt{\bar{\sigma}^2}}\right). \quad (28)$$

Therefore, when $\lambda \to \infty$, if $\beta < p_{\text{fa}}$, $P_{\text{fa}} = P_d = 1$; if $p_{\text{fa}} < \beta < \bar{p}_d$, $P_{\text{fa}} = 0$ and $P_d = 1$; if $\beta > \bar{p}_d$, $P_{\text{fa}} = P_d = 0$. As a result, as long as β takes a value between p_{fa} and $\bar{p}_d$, as $\lambda \to \infty$, the WSN detection performance will be perfect with $P_d = 1$ and $P_{\text{fa}} = 0$. In Figures 9 and 10, P_d and P_{fa} as functions of λ are plotted. It is clear that the P_d converges to 1 and P_{fa} converges to 0, as λ increases. In this example, we set β such that $\beta = (p_{\text{fa}} + \bar{p}_d)/2$. Another conclusion is that when λ is large enough, even for a small SNR$_0$, the system can achieve a very good detection performance.

3.5. Optimality of the decision fusion rule

The proposed decision fusion rule (the counting rule) is actually a threshold test in terms of the total number of detections made by local sensors, and it is intuitive. It is important to compare the performance of this fusion rule to that of the
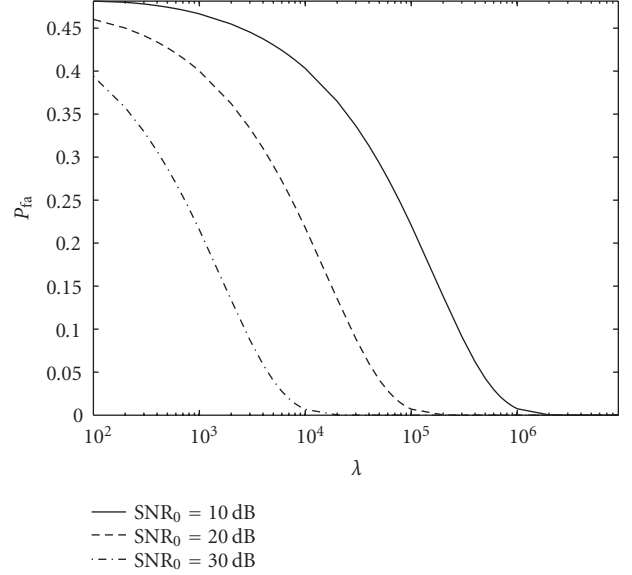
optimal decision fusion rule, which is also based on the total number of local detections from local sensors.

As we know, Λ in (13) is a lattice-type RV [24], which takes equidistant values from 0 to N. Hence, according to the CLT [24], for a large N, the probability $p_k = \Pr\{\Lambda = k|N\}$ equals the sample of the Gaussian density:

$$\Pr\{\Lambda = k|N\} \simeq \frac{1}{\sqrt{2\pi}\sigma} e^{-(k-\mu)^2/(2\sigma^2)} \quad (k = 0,\ldots,N). \quad (29)$$

Therefore, under hypothesis H_1, for a large λ, we have

$$\Pr\{\Lambda = k|H_1\} = \sum_{N=0}^{\infty} p(N) \Pr\{\Lambda = k|N, H_1\}$$
$$\simeq \sum_{N=0}^{\infty} \frac{\lambda^N e^{-\lambda}}{N!} \frac{1}{\sqrt{2\pi}\sigma_1(N)} e^{-[k-\mu_1(N)]^2/[2\sigma_1^2(N)]}, \quad (30)$$

where $\mu_1(N) = N\bar{p}_d$ and $\sigma_1(N) = \sqrt{N\bar{\sigma}^2}$. Similarly, under hypothesis H_0, for a large λ, we have

$$\Pr\{\Lambda = k|H_0\} \simeq \sum_{N=0}^{\infty} \frac{\lambda^N e^{-\lambda}}{N!} \frac{1}{\sqrt{2\pi}\sigma_0(N)} e^{-[k-\mu_0(N)]^2/[2\sigma_0^2(N)]}, \quad (31)$$

where $\mu_0(N) = N p_{\text{fa}}$ and $\sigma_0(N) = \sqrt{N p_{\text{fa}}(1 - p_{\text{fa}})}$. Now it is easy to show that the likelihood ratio of Λ is

$$L(\Lambda) = \frac{\Pr\{\Lambda|H_1\}}{\Pr\{\Lambda|H_0\}}$$
$$\simeq \frac{\sum_{N=0}^{\infty} \lambda^N/[N!\sigma_1(N)] e^{-[\Lambda-\mu_1(N)]^2/[2\sigma_1^2(N)]}}{\sum_{N=0}^{\infty} \lambda^N/[N!\sigma_0(N)] e^{-[\Lambda-\mu_0(N)]^2/[2\sigma_0^2(N)]}}. \quad (32)$$
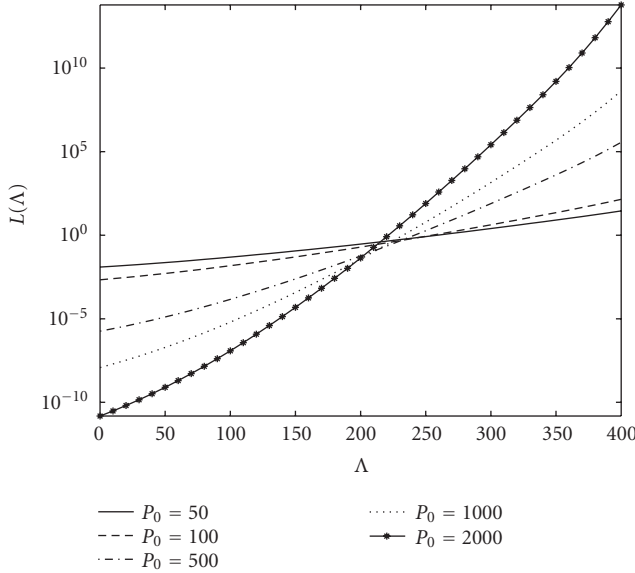
FIGURE 11: Function $L(\Lambda)$. $\lambda = 1000$, $n = 2$, $b = 100$, $\alpha = 200$, $\tau = 0.66, 0.67, 0.73, 0.77, 0.82$ for $P_0 = 50, 100, 500, 1000, 2000$, respectively.

FIGURE 12: Function $L(\Lambda)$. $n = 3$, $P_0 = 500$, $b = 100$, $\alpha = 40$, $\tau = 0.90$.

Hence, the optimal fusion rule at the fusion center is a likelihood ratio test:

$$L(\Lambda) \underset{H_0}{\overset{H_1}{\gtrless}} T_L. \tag{33}$$

Note that the implementation of the proposed counting rule for a Neyman-Pearson detector with a given system level P_{fa} requires only the knowledge of λ and τ (or p_{fa}) in order to find the system-level threshold T through (18). To choose an optimal local threshold τ, as we will see later in this paper, the knowledge of P_0 is required too. However, the counting rule can still be implemented without an optimal τ, and a good choice of τ based on some prior knowledge of P_0 can always be made. As a result, an exact knowledge of P_0 is not necessary for the implementation of the counting rule, even though it is needed in the evaluation of the system-level detection performance.

As for the implementation of the optimal fusion rule, we need to have the *exact* knowledge of α, P_0, and b to calculate $\bar{\sigma}^2$ and $\bar{p}_d$. Hence, the optimal fusion rule requires much more information, especially the knowledge of signal power P_0, which is unknown in most cases. Furthermore, because of its dependence on the exact knowledge of P_0, the optimal fusion rule is more sensitive to the estimation errors of P_0. Therefore, in this paper, the optimal fusion rule only has theoretical importance, and it is not very useful or robust in practical applications, where it is always difficult to estimate P_0.

As we can see from (32), $L(\Lambda)$ is a nonlinear transformation of Λ. The threshold tests of Λ and $L(\Lambda)$ will have identical detection performances if $L(\Lambda)$ is a monotonically increasing transformation of Λ.
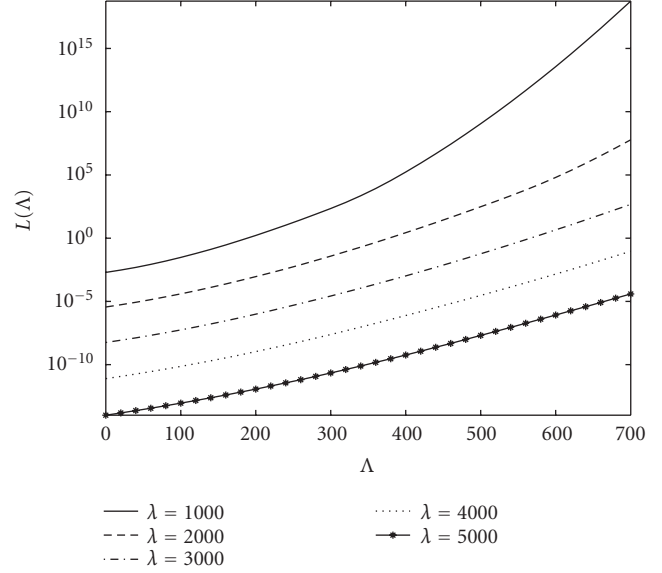
In Figures 11 and 12, L as a function of Λ is plotted for different system parameters. As we can see, in all the cases, $L(\Lambda)$ is a monotonically increasing function of Λ, meaning that the counting rule and the optimal fusion rule are equivalent in terms of detection performance. In addition to the cases shown in Figures 11 and 12, we have extensively investigated the relationship between L and Λ for various system parameters. For all the system parameters we have studied, $L(\Lambda)$ is a monotonically increasing function of Λ.

In Figure 13, the ROC curves obtained by simulations (based on 10^6 Monte Carlo runs) for the counting rule and the optimal fusion rule are shown. We can see that the ROC curves corresponding to the counting rule and those of the optimal fusion rule are indistinguishable.

4. THRESHOLD FOR LOCAL SENSORS

In addition to the ROC curve for performance comparison, one can also resort to the so-called deflection coefficient [25, 26], especially when the statistical properties of the signal and noise are limited to moments up to a given order. The deflection coefficient is defined as

$$D(\Lambda) = \frac{[E(\Lambda|H_1) - E(\Lambda|H_0)]^2}{\text{Var}(\Lambda|H_0)}. \tag{34}$$

In the case of $\text{Var}(\Lambda|H_1) = \text{Var}(\Lambda|H_0)$, this is in essence the SNR of the detection statistic. It is worth noting that the use of deflection criterion leads to the optimum LR receiver in many cases of practical importance [25]. For example, in the problem of detecting a Gaussian signal in Gaussian noise, an LR detector is obtained by maximizing the deflection measure. In the above sections, we have assumed that the threshold τ (or equivalently p_{fa}) is given. From (18), (20), (21),
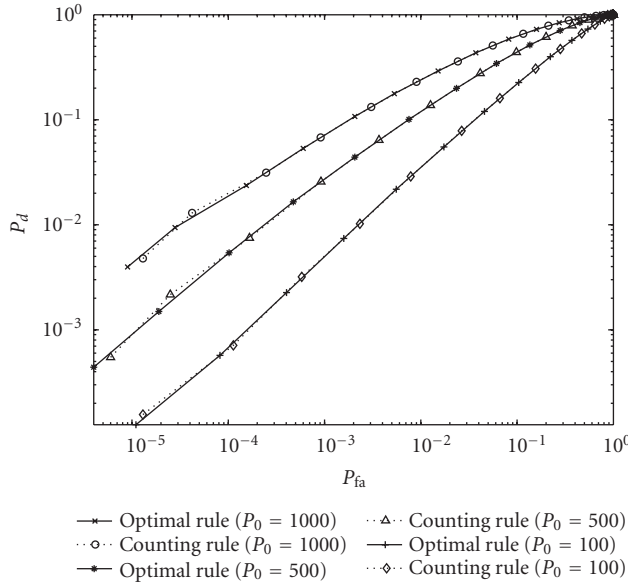
FIGURE 13: ROC curves for the counting rule and the optimal fusion rule. System parameters are the same as those listed in Figure 5.
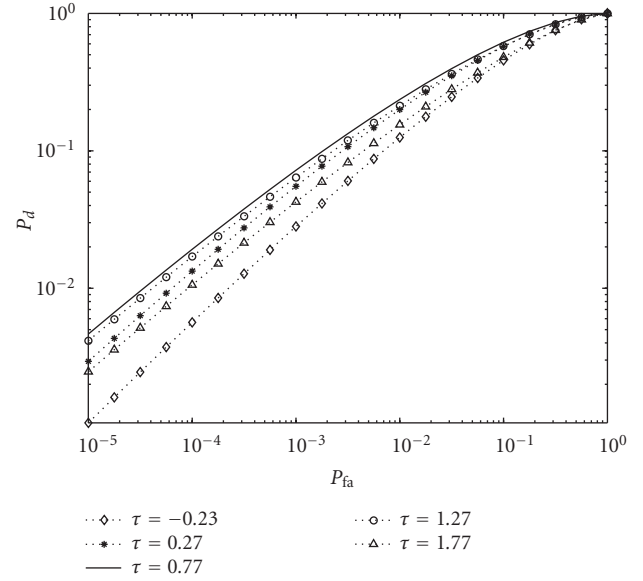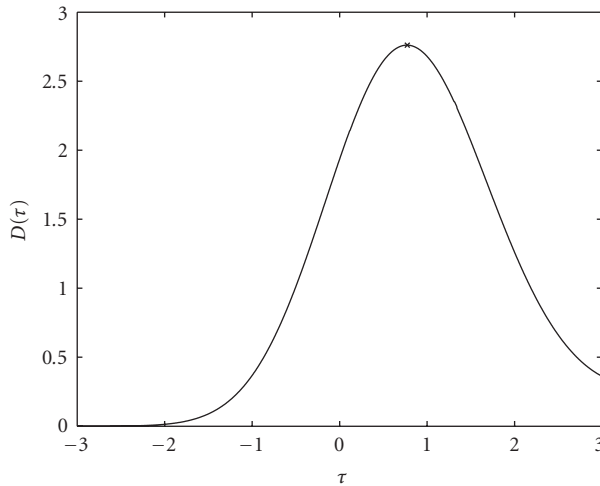

FIGURE 14: $D(\tau)$. $\lambda = 1000$, $n = 2$, $a = 100$, $\alpha = 200$, $\text{SNR}_0 = 30$ dB (or $P_0 = 1000$).

and (25), we know that both P_{fa} and P_d are functions of τ. Hence, τ is a parameter that can be designed to achieve a better system-level performance. In this paper, we will find the optimum local sensor-level threshold τ by maximizing the deflection coefficient. The deflection coefficient for the detection problem in this paper is derived and stated in the following theorem.

Theorem 1. *The deflection coefficient at the fusion center for the detection problem formulated in this paper is*

$$D(\tau) = \frac{\lambda\left[\bar{p}_d(\tau) - p_{\text{fa}}(\tau)\right]^2}{p_{\text{fa}}(\tau)}. \tag{35}$$

For the proof, see the appendix.

FIGURE 15: ROC curves for system with different τ. $\lambda = 1000$, $n = 2$, $b = 100$, $\alpha = 200$, $\text{SNR}_0 = 30$ dB (or $P_0 = 1000$).

The optimum τ can be found by maximizing $D(\tau)$ with respect to τ. As we can see in Figure 14, there exists an optimal $\tau(0.7694)$ that maximizes the deflection coefficient D. By employing this optimum τ_{opt}, a significant improvement in D can be achieved.

The system-level ROC curves for different τ are plotted in Figure 15. As we can see, the ROC curve corresponding to the optimal threshold $\tau_{\text{opt}}(0.77)$ is above those for other thresholds, meaning that τ_{opt} provides the best system level performance. In Figures 16 and 17, τ_{opt} and the corresponding optimal p_{fa} as functions of SNR_0 and α are shown. It is clear that τ_{opt} is a monotonically increasing function of SNR_0 and a monotonically decreasing function of α. This is because with a strong target signal (high SNR_0 and low α), by adopting a higher threshold, local sensors lower their false alarm rate, while at the same time they can still attain a relatively high probability of detection.

5. CONCLUSIONS

We have proposed and studied a decision fusion rule that is based on the total number of detections reported by local sensors for a WSN with a random number of sensors. Assuming that the number of sensors in a ROI follows a Poisson distribution, we have derived the system-level detection performance measures, namely the probabilities of detection and false alarm. We have shown that even at very low SNR, this fusion rule can achieve a very good system-level detection performance given that there are, on an average, a sufficiently large number of sensors deployed in the ROI. The average number of sensors needed for a prespecified system-level performance can be calculated based on our analytical expressions. Another important result is that the proposed fusion rule is equivalent to the optimal fusion rule, which
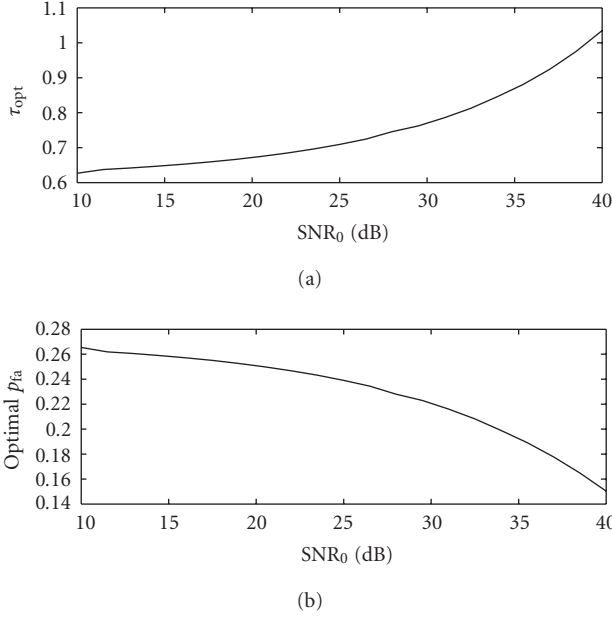
(a)

(b)

FIGURE 16: Optimal τ_{opt} and the corresponding optimal p_{fa} as functions of SNR_0. $\lambda = 1000$, $n = 2$, $b = 100$, $\alpha = 200$.

(a)

(b)
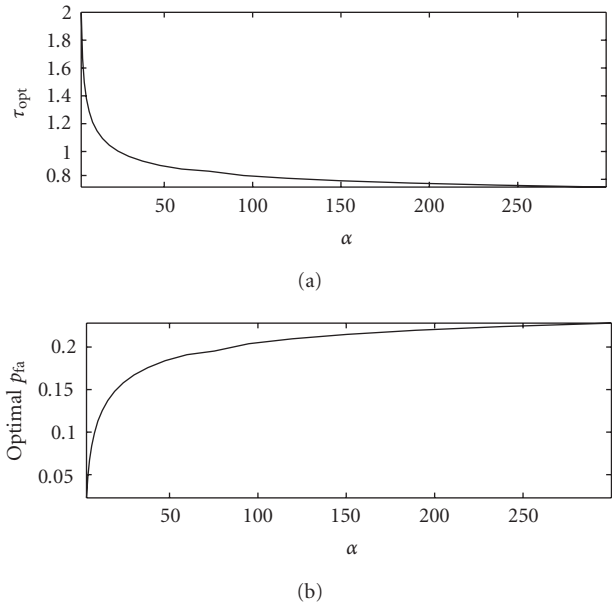
FIGURE 17: Optimal τ_{opt} and the corresponding optimal p_{fa} as functions of α. $\lambda = 1000$, $n = 2$, $b = 100$, $\text{SNR}_0 = 30$ dB.

requires much more prior knowledge of the system parameters, for all the different system parameters we have investigated.

We have also shown that a better system performance can be achieved if we choose an optimum threshold at the local sensors by maximizing the deflection coefficient. If SNR_0 is high, and α is small, a higher local sensor-level threshold τ should be chosen; otherwise, a lower τ should be employed to achieve a better performance.

APPENDIX

PROOF OF THEOREM 1

Under hypothesis H_0, we have

$$E(\Lambda|N, H_0) = N p_{\text{fa}}, \tag{A.1}$$

$$\text{Var}(\Lambda|N, H_0) = N p_{\text{fa}}(1 - p_{\text{fa}}). \tag{A.2}$$

Hence,

$$E(\Lambda^2|N, H_0) = \text{Var}(\Lambda|N, H_0) + E(\Lambda|N, H_0)^2$$
$$= N p_{\text{fa}}(1 - p_{\text{fa}}) + N^2 p_{\text{fa}}^2. \tag{A.3}$$

Since N is a Poisson RV, we have

$$E[N] = \lambda, \tag{A.4}$$

$$\text{Var}[N] = \lambda, \tag{A.5}$$

$$E(N^2) = \text{Var}(N) + [E(N)]^2 = \lambda + \lambda^2. \tag{A.6}$$

With (A.1) and (A.4), $E(\Lambda|H_0)$ can be derived as follows:

$$E(\Lambda|H_0) = E[E(\Lambda|N, H_0)] = E[N p_{\text{fa}}] = \lambda p_{\text{fa}}. \tag{A.7}$$

Given (A.3), (A.4), and (A.6), it is easy to show that

$$E(\Lambda^2|H_0) = E[E(\Lambda^2|N, H_0)]$$
$$= E[N p_{\text{fa}}(1 - p_{\text{fa}}) + N^2 p_{\text{fa}}^2]$$
$$= \lambda p_{\text{fa}}(1 - p_{\text{fa}}) + (\lambda + \lambda^2) p_{\text{fa}}^2$$
$$= \lambda p_{\text{fa}}(1 + \lambda p_{\text{fa}}). \tag{A.8}$$

Therefore,

$$\text{Var}(\Lambda|H_0) = E[\Lambda^2|H_0] - E[\Lambda|H_0]^2$$
$$= \lambda p_{\text{fa}}(1 + \lambda p_{\text{fa}}) - \lambda^2 p_{\text{fa}}^2 \tag{A.9}$$
$$= \lambda p_{\text{fa}}.$$

Under hypothesis H_1, according to [19], we have

$$E(\Lambda|N, H_1) = N \bar{p}_d. \tag{A.10}$$

Hence,

$$E[\Lambda|H_1] = E[E(\Lambda|N, H_1)]$$
$$= E[N \bar{p}_d] \tag{A.11}$$
$$= \lambda \bar{p}_d.$$

By substituting (A.7), (A.9), and (A.11) into (34), we finally get the deflection coefficient

$$D(\tau) = \frac{\lambda[\bar{p}_d(\tau) - p_{\text{fa}}(\tau)]^2}{p_{\text{fa}}(\tau)}. \tag{A.12}$$

REFERENCES

[1] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva, "Directed diffusion for wireless sensor networking," *IEEE/ACM Trans. Networking*, vol. 11, no. 1, pp. 2–16, 2003.

[2] H. Gharavi and K. Ban, "Multihop sensor network design for wide-band communications," *Proc. IEEE*, vol. 91, no. 8, pp. 1221–1234, 2003.

[3] T. J. Kwon, M. Gerla, V. K. Varma, M. Barton, and T. R. Hsing, "Efficient flooding with passive clustering—an overhead-free selective forward mechanism for ad hoc/sensor networks," *Proc. IEEE*, vol. 91, no. 8, pp. 1210–1220, 2003.

[4] S. Kumar, F. Zhao, and D. Shepherd, "Collaborative signal and information processing in microsensor networks," *IEEE Signal Processing Mag.*, vol. 19, no. 2, pp. 13–14, 2002.

[5] H. Gharavi and S. P. Kumar, "Special issue on sensor networks and applications," *Proc. IEEE*, vol. 91, no. 8, pp. 1151–1153, 2003.

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

[7] P. K. Varshney, *Distributed Detection and Data Fusion*, Springer, New York, NY, USA, 1997.

[8] 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 Processing*, vol. 48, no. 12, pp. 3266–3279, 2000.

[9] E. Drakopoulos and C.-C. Lee, "Optimum multisensor fusion of correlated local decisions," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 27, no. 4, pp. 593–606, 1991.

[10] M. Kam, W. Chang, and Q. Zhu, "Hardware complexity of binary distributed detection systems with isolated local Bayesian detectors," *IEEE Trans. Syst., Man, Cybern.*, vol. 21, no. 3, pp. 565–571, 1991.

[11] 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. 3, pp. 916–920, 1992.

[12] 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. 2, pp. 554–568, 1996.

[13] F. Gini, F. Lombardini, and L. Verrazzani, "Decentralised detection strategies under communication constraints," *IEE Proceedings—Radar, Sonar and Navigation*, vol. 145, no. 4, pp. 199–208, 1998.

[14] C.-T. Yu and P. K. Varshney, "Paradigm for distributed detection under communication constraints," *Optical Engineering*, vol. 37, no. 2, pp. 417–426, 1998.

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

[16] T. Kasetkasem and P. K. Varshney, "Communication structure planning for multisensor detection systems," *IEE Proceedings—Radar, Sonar and Navigation*, vol. 148, no. 1, pp. 2–8, 2001.

[17] J. Hu and R. S. Blum, "On the optimality of finite-level quantizations for distributed signal detection," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1665–1671, 2001.

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

[19] R. Niu, P. K. Varshney, M. H. Moore, and D. Klamer, "Decision fusion in a wireless sensor network with a large number of sensors," in *Proc. 7th IEEE International Conference on Information Fusion (ICIF '04)*, Stockholm, Sweden, June–July 2004.

[20] D. Li, K. D. Wong, Y. H. Hu, and A. M. Sayeed, "Detection, classification, and tracking of targets," *IEEE Signal Processing Mag.*, vol. 19, no. 2, pp. 17–29, 2002.

[21] N. Levanon, *Radar Principles*, John Wiley & Sons, New York, NY, USA, 1988.

[22] T. Rappaport, *Wireless Communications—Principles and Practices*, Prentice-Hall, Upper Saddle River, NJ, USA, 1996.

[23] L. E. Kinsler and A. R. Frey, *Fundamentals of Acoustics*, John Wiley & Sons, New York, NY, USA, 1962.

[24] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, NY, USA, 1984.

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

[26] S. M. Kay, *Fundamentals of Statistical Signal Processing II: Detection Theory*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1998.

Ruixin Niu received his B.S. degree from Xi'an Jiaotong University, Xi'an, China, in 1994, his M.S. degree from the Institute of Electronics, Chinese Academy of Sciences, Beijing, in 1997, and his Ph.D. degree from the University of Connecticut, Storrs, in 2001, all in electrical engineering. He is currently a Postdoctoral Research Associate with Syracuse University, Syracuse, NY. His research interests are in the areas of statistical signal processing and its applications, including detection, estimation, data fusion, communications, and image processing. He received the Fusion 2004 Best Paper Award in the Seventh International Conference on Information Fusion, Stockholm, Sweden, in June 2004.
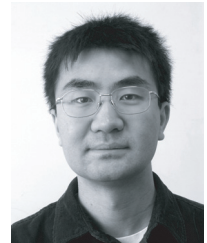
Pramod K. Varshney was born in Allahabad, India, on July 1, 1952. He received the B.S. degree in electrical engineering and computer science (with highest honors), and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois at Urbana-Champaign in 1972, 1974, and 1976, respectively. Since 1976, he has been with Syracuse University, Syracuse, NY, where he is currently a 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 has served as a Consultant to several major companies. He is the recipient of the 1981 ASEE Dow Outstanding Young Faculty Award. He was elected to the grade of Fellow of the IEEE in 1997 for his contributions in the area of distributed detection and data fusion. 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 was the President of International Society of Information Fusion during 2001.