DECENTRALIZED DETECTION IN SENSOR NETWORKS USING RANGE INFORMATION

Antonio Artés-Rodríguez

Dept. of Signal Theory and Communications Universidad Carlos III de Madrid Avda. de la Universidad 30 28991 Leganés, Madrid, Spain antonio@ieee.org

ABSTRACT

We consider the problem of binary distributed detection in the context of large-scale, dense sensor networks. We propose to model the probability of detection in each sensor, p_d , as a function of the distance between the sensor and the source or target to be detected. We derive the Bayesian fusion rule under that model. We also derive, using the asymptotic gaussianity of the log-likelihood ratio, the Neyman-Pearson fusion rule. The performances of both tests is analyzed using large deviation bounds on the error probability and a parametric approximation to p_d . The main conclusions of the analysis of these bounds are that, for designing efficient tests in terms of energy comsumption, 1) the sensors must be grouped in areas of the order of the range of the local detectors, and, 2) the sensor must be configured to achieve the best local discrimination between hypothesis, independently of the configuration of the network.

1. INTRODUCTION

In 1986, Chair and Varshney [1] determined the optimum Bayes decision fusion rule for the binary distributed detection problem when the local detection rule is known. Two years later, Tsitsiklis [2] shown that when the number of sensor is arbitrarily large, the optimal binary decentralized detection is achieved by identical local detection rules. From that, many author have proposed and analyzed different optimal fusion rule under a variety of criteria (see [3] and the references therein for a review on this topic).

On the other hand, the recent advances in the development of small and low powered devices that integrate sensing, processing, and wireless communication capabilities [4] make possible the large-scale, dense and random sensor networks where the above algorithms are valuable tools. Nevertheless, some of the hypothesis of the original problem has to be revised. First, the conditional probability distribution under each hypothesis in each sensor can be different due to the spatial variation of the signal (or source, or target) to be detected and, second, if the sensors are identical as suggested in [2], the Probability of False Alarm (PFA) and Probability of Missdetection (PM) must vary from sensor to sensor according to a its position. Authors in [1] have been considered different PFA and PM for each sensor, but no further analysis was given. In this paper we propose a model in which the probability of detection (including false alarms) of the sensor varies as a function p_d of the distance between the sensor and the source or target to be detected. Following this model, we analyze the simple problem of detecting a target knowing its position and the position of the sensors under both Bayes and Neyman-Pearson (NP) tests. After deriving the log-likelihood ratio (LLR), we will obtain the threshold for the NP test using the asymptotic gaussianity of the LLR. The performances of both tests are analyzed using large deviation bounds on the error probability and a parametric approximation of p_d . We also provided, using these bounds, rules for designing the test for the exploration of a given spatial area.

The paper is organized as follows. The statement of the problem and the notation used in the paper is established in Section 2. The LLR tests are derived in Section 3. Sections 4 and 5 are devoted to the analysis of, respectively, asymptotic gaussianity and large deviation bounds. The optimization of the test is done in Section 6, and the conclusions end the paper.

2. PROBLEM STATEMENT AND NOTATION

We consider a random deployment of sensor with density ρ_s sensors per area unit over an area $\mathcal{D} \in \mathbb{R}^2$. Each sensor applies the same binary detection rule, not necessarily based on a LLR test.

The exploration of the area \mathcal{D} gives as a result the data set $\{(\boldsymbol{x}_i, y_i) : i = 1, \ldots, l, \boldsymbol{x}_i \in \mathcal{D}, y_i \in \{0, 1\}\}$, when each pair (\boldsymbol{x}_i, y_i) represents a successful reading of a sensor located at coordinates \boldsymbol{x}_i that can detect $(y_i = 1)$ or not $(y_i = 0)$ a target.

The probability of a positive detection (Y = 1) in a sensor located at coordinates \boldsymbol{x} when a target is present at coordinates \boldsymbol{x}^t is denoted as $p_d(\boldsymbol{x}^t, \boldsymbol{x}, \alpha)$, where α is the probability of false alarm (PFA) of the sensor when no target is present. In other words, $p_d(\boldsymbol{x}^t, \boldsymbol{x}, \alpha) = Pr(Y = 1 | \boldsymbol{X}^t = \boldsymbol{x}^t, \boldsymbol{X} = \boldsymbol{x})$ when the PFA of the detector is equal to α . $p_d(\boldsymbol{x}^t, \boldsymbol{x}, \alpha)$ has the following properties:

1.
$$p_d(\boldsymbol{x}^t, \boldsymbol{x}, \alpha) \ge \alpha$$

2. $p_d(\boldsymbol{x}^t, \boldsymbol{x}, \alpha) = p_d(\|\boldsymbol{x}^t - \boldsymbol{x}\|_2^2, \alpha)$
3. $p_d(\boldsymbol{x}^t, \boldsymbol{x}, \alpha) \ge p_d(\boldsymbol{x}^t, \boldsymbol{x}', \alpha) \Leftrightarrow \|\boldsymbol{x}^t - \boldsymbol{x}\|_2 \le \|\boldsymbol{x}^t - \boldsymbol{x}'\|_2$
4. $\lim_{\|\boldsymbol{x}^t - \boldsymbol{x}\|_2 \to \infty} p_d(\boldsymbol{x}^t, \boldsymbol{x}, \alpha) = \alpha$

Given \mathcal{D} , we define two hypothesis, H_0 or null hypothesis for the case when no target is present, and H_1 or alternative hypothesis for the case when a target is present.

This work were done while the author was visiting Cornell University. This visit was supported by a grant of the Secretaría de Estado de Educación y Universidades.

• Under hypothesis H_0 , the joint pdf of X and Y is

$$f_{\boldsymbol{X},Y|H_0}(\boldsymbol{x}, y|H_0) = \rho\left(\alpha\,\delta[y-1] + (1-\alpha)\,\delta[y]\right)$$

where ρ = (∫_D dx)⁻¹ and δ is the Kronecker function.
Under hypothesis H₁, the joint pdf of X and Y is

$$f_{\boldsymbol{X},Y|H_1}(\boldsymbol{x}, y|H_1) = \rho(p_d(\boldsymbol{x}^t, \boldsymbol{x}, \alpha) \,\delta[y-1] \\ + \left(1 - p_d(\boldsymbol{x}^t, \boldsymbol{x}, \alpha)\right) \,\delta[y])$$

We assume that samples in $\{(x_i, y_i) : i = 1, ..., l, x_i \in D, y_i \in \{0, 1\}\}$ are conditionally (under H_0 or H_1) independents.

When necessary, we can assume the following parametric approximation to p_d , that we called the "spanish hat" model:

$$p_d(\boldsymbol{x}^t, \boldsymbol{x}, \alpha) = \begin{cases} (1 - \beta) & \text{if } \|\boldsymbol{x}^t - \boldsymbol{x}\|_2 < r_0 \\ \alpha & \text{otherwise} \end{cases}$$

where r_0 is the range of the sensor. This model considers a constant probability of missdetection when the target is located inside the range of the sensor and a constant false alarm probability outside the range of the sensor. This simple model is provided to gain some insight into the performance analysis.

3. HYPOTHESIS DETECTION PROBLEMS

Given $\{(x_i, y_i) : i = 1, ..., l, x_i \in D, y_i \in \{0, 1\}\}$, the loglikelihood ratio between both hypothesis can be computed as

$$\lambda = \sum_{i=1}^{l} \ln \Gamma_i$$

where

$$\begin{split} \Gamma_i &= \frac{p_d(\boldsymbol{x}^t, \boldsymbol{x}_i, \alpha) \, \delta[y_i - 1] + (1 - p_d(\boldsymbol{x}^t, \boldsymbol{x}_i, \alpha)) \, \delta[y_i]}{\alpha \, \delta[y_i - 1] + (1 - \alpha) \, \delta[y_i]} \\ &= \begin{cases} \frac{1 - p_d(\boldsymbol{x}^t, \boldsymbol{x}_i, \alpha)}{\alpha} & \text{if } y_i = 0\\ \frac{p_d(\boldsymbol{x}^t, \boldsymbol{x}_i, \alpha)}{\alpha} & \text{if } y_i = 1 \end{cases} \end{split}$$

Under Bayes criteria, the threshold τ is easily set as

$$\tau = \ln \frac{\pi_0 (C_{10} - C_{00})}{\pi_1 (C_{01} - C_{11})}$$

but the resulting Bayes risk is hard to determine due to the non trivial partition of the input space, except for the single sensor (l = 1) case.

In order to understand the difficulties regarding the partition of the input space, lets analyze the case of two sensor (l = 2) with the threshold set to 0 ($\tau = 0$, that could corresponds with uniform costs and equally likely *a priori* probabilities). If $y_1 = y_2 = 1$, the log-likelihood is greater than zero ($\lambda > 0$) independently of the values of x_1, x_2 , and α . If $y_1 = y_2 = 0$, $\lambda < 0$ independently of the values of x_1, x_2 , and α . If $y_1 = 1$ and $y_2 = 0$ (or viceversa) λ can be less or greater than 0 depending on the values of x_1, x_2 , and α . In this case

$$\lambda = \ln \frac{p_d(\boldsymbol{x}^t, \boldsymbol{x}_1, \alpha)}{\alpha} + \ln \frac{1 - p_d(\boldsymbol{x}^t, \boldsymbol{x}_2, \alpha)}{1 - \alpha} \underset{H_0}{\overset{H_1}{\gtrless}} 0$$

induces a nonlinear partition of the space $x_1 \times x_2$.

Under the NP criteria, the determination of the threshold given the power of the test is basically as hard as the determination of the Bayes risk.

In the cases of interest the number of sensors is very high and, thus, under either Bayes or NP criteria, the solution of the problem becomes impossible to set and/or analyze. On the other hand, the high number of sensors opens the door of asymptotic statistics, that we will explore in two directions: first, using the asymptotic normality of λ to set the threshold of the NP test and to analyze its power and, second, using large deviation exponential bounds on the test error based for both Bayes and NP tests.

4. ASYMPTOTIC GAUSSIANITY

When the number of sensor, l, tends to infinity, the log-likelihood ratio, λ , tends to a normal random variable. We will first determine the means and variances of λ under the two hypothesis, and then determine the threshold and power of the NP test under gaussian statistics.

Lets denote $\gamma = \ln \Gamma$ and $\gamma_{H_i} = \gamma|_{H=H_i}$. The mean of γ_{H_0} is given by

$$E\{\gamma_{H_0}\} = -D(f_{\mathbf{X},Y|H_0} \| f_{\mathbf{X},Y|H_1})$$

where D is the Kullback-Leibler (KL) divergence [5]. We will denote $D(f_{X,Y|H_0}||f_{X,Y|H_1})$ as $D(H_0||H_1)$ for short. In our problem, $D(H_0||H_1)$ can be decomposed as

$$D(H_0 || H_1) = -H(\alpha) - \alpha P_1 - (1 - \alpha) P_0$$

where H is the binary entropy function, and

$$P_1 = \int_{\mathcal{D}} \rho \ln p_d(\boldsymbol{x}^t, \boldsymbol{x}_i, \alpha) \, d\boldsymbol{x}$$
$$P_0 = \int_{\mathcal{D}} \rho \ln(1 - p_d(\boldsymbol{x}^t, \boldsymbol{x}_i, \alpha)) \, d\boldsymbol{x}$$

Similarly,

$$E\{\gamma_{H_1}\} = D(f_{\mathbf{X},Y|H_1} \| f_{\mathbf{X},Y|H_0}) = D(H_1 \| H_0)$$

and

$$D(H_1||H_0) = -h(\mathbf{X}, Y|H_1) - \ln \alpha Pr(Y = 1|H_1) - \ln(1 - \alpha)(1 - Pr(Y = 1|H_1)) - \ln \mu$$

where h is the differential entropy. The variances of γ_{H_0} and γ_{H_1} are, respectively

$$E\{(\gamma_{H_0} + D(H_0 || H_1))^2\} = E\{\gamma_{H_0}^2\} - D^2(H_0 || H_1)$$
$$E\{(\gamma_{H_1} - D(H_1 || H_0))^2\} = E\{\gamma_{H_1}^2\} - D^2(H_1 || H_0)$$

The threshold τ of the NP test of level α_l is, according to the above

$$\tau = \sqrt{l \left(E\{\gamma_{H_0}^2\} - D^2(H_0 \| H_1) \right)} Q^{-1}(\alpha_l) - l D(H_0 \| H_1)$$

and its power, β_l , is given by

$$\beta_l = Q\left(\frac{l D(H_1 \| H_0) - \tau}{\sqrt{l \left(E\{\gamma_{H_1}^2\} - D^2(H_1 \| H_0)\right)}}\right)$$

where Q is the Marqum's function, and Q^{-1} its inverse.

Now we will analyze the form that takes the most relevant parameter, $D(H_0||H_1)$, when using the "spanish hat" approximation of p_d . In order to simplify the analysis, lets assume that the area $\mathcal{D} \in \mathbb{R}^2$ is a circle of radius R centered on x^t . In that case, $D(H_0||H_1)$ is

$$D(H_0 \| H_1) = \begin{cases} \frac{r_0^2}{R^2} \left(\alpha \ln \frac{\alpha}{1-\beta} + (1-\alpha) \ln \frac{1-\alpha}{\beta} \right) & \text{if } R \ge r_0 \\ \left(\alpha \ln \frac{\alpha}{1-\beta} + (1-\alpha) \ln \frac{1-\alpha}{\beta} \right) & \text{if } R < r_0 \end{cases}$$
(1)

It is important to realize that the "spanish hat" model is a first order approximation of any function p_d . If we denote by D_{sh} de KL divergence under the "spanish hat" model, for any function $p_d^*(\mathbf{x}^t, \mathbf{x}_i, \alpha^*)$ that induces a divergence $D^*(H_0 || H_1)$, we can obtain a $D_{sh}(H_0 || H_1)$ such that $D_{sh}(H_1 || H_0) = D^*(H_1 || H_0)$ by setting the parameters of the "spanish hat" model as follows: $\alpha = \alpha^*$, and solving β and r_0 for equating P_0 and P_1 under the two p_d functions.

Using an equivalent procedure, different values of β and r_0 can be obtained by setting $D_{sh}(H_1||H_0) = D^*(H_1||H_0)$, and thus functions like $D(H_1||H_0) + D(H_0||H_1)$ that determines the power of the test can be lower and upper bounded, and so any function of moments of γ . Alternatively, instead of using multiple "spanish hat" functions for bounding any function of moments of γ , we can use a similar single function that takes constant values of p_d over concentric rings for equating a given p_d .

5. LARGE DEVIATION BOUNDS

As some authors point out [5], the estimates of the probability of error based on the central limit theorem give only good approximation only for a few standard deviations from the mean. A more powerful estimates of the probability of error are the large deviation bounds in the form of error exponents. If ϵ_l is the probability of error (of some kind) obtained with l observation, the error exponent is defined as

$$\lim_{l\to\infty}-\frac{1}{l}\ln\epsilon_l$$

In NP test, the best error exponent is given by the Stein's lemma, that applied to our problem says that for any $\alpha_n \in (0, 1)$

$$\lim_{l \to \infty} -\frac{1}{l} \ln \beta_l = D(H_0 || H_1)$$

a function that has been analyzed in the preceding section.

In Bayes tests (assuming that $C_{10} - C_{00} = C_{01} - C_{11}$), the best achievable error exponent is the Chernoff information, $C(f_{\boldsymbol{X},\boldsymbol{Y}|H_1}, f_{\boldsymbol{X},\boldsymbol{Y}|H_0})$ or $C(H_1, H_0)$ for short, defined as

$$C(f_{\mathbf{X},Y|H_1}, f_{\mathbf{X},Y|H_0}) = D(f_{\mathbf{X},Y|s_0} \| f_{\mathbf{X},Y|H_1})$$

= $D(f_{\mathbf{X},Y|s_0} \| f_{\mathbf{X},Y|H_0})$ (2)

where

$$\begin{split} f_{\boldsymbol{X},Y|s}(\boldsymbol{x},y|H_0) &= \\ & \frac{f_{\boldsymbol{X},Y|H_1}^s(\boldsymbol{x},y|H_1) f_{\boldsymbol{X},Y|H_0}^{1-s}(\boldsymbol{x},y|H_0)}{\sum \int f_{\boldsymbol{X},Y|H_1}^s(\boldsymbol{x}',y|H_1) f_{\boldsymbol{X},Y|H_0}^{1-s}(\boldsymbol{x}',y|H_0) \, d\boldsymbol{x}'} \end{split}$$

and s_0 the value of s such that (2) is satisfied.

The Chernoff information can also be obtained as minus the minimum of the cumulant generating function (cgf) of the log-likelihood ratio per sample under hypothesis H_0 (or H_1), i.e.,

$$C(H_1, H_0) = -\frac{1}{l} \min_{0 \le s \le 1} \mu_{\lambda,0}(s)$$

that in our problem takes the form

$$\mu_{\lambda,0}(s) = l \ln \left[\int_{\mathcal{D}} \rho \left(\frac{p_d(\boldsymbol{x}^t, \boldsymbol{x}, \alpha)^s}{\alpha^{s-1}} + \frac{(1 - p_d(\boldsymbol{x}^t, \boldsymbol{x}, \alpha))^s}{(1 - \alpha)^{s-1}} \right) d\boldsymbol{x} \right]$$

Using the "spanish hat" model and a circular $\ensuremath{\mathcal{D}}$ as before, the cgf is

$$\mu_{\lambda,0}(s) = l \begin{cases} \ln\left(\frac{r_0^2}{R^2} \left(\frac{(1-\beta)^s}{\alpha^{s-1}} + \frac{\beta^s}{(1-\alpha)^{s-1}}\right) + \left(1 - \frac{r_0^2}{R^2}\right) \right) \text{ if } R \ge r_0\\ \ln\left(\frac{(1-\beta)^s}{\alpha^{s-1}} + \frac{\beta^s}{(1-\alpha)^{s-1}}\right) & \text{ if } R < r_0 \end{cases}$$
(3)

That achieves its minimum respect to s when

$$s = s_0 = \frac{\ln \frac{1-\alpha}{\alpha} + \ln \ln \frac{1-\alpha}{\beta} - \ln \ln \frac{1-\beta}{\alpha}}{\ln \left(\frac{1-\beta}{\beta} \frac{1-\alpha}{\alpha}\right)}$$
(4)

6. OPTIMIZATION OF THE TESTS

The sensors are assumed to be battery powered, and the wireless transmissions from sensors to the fusion center is the most energy consuming operation [4]. For elongating the life of the sensor network, a reasonable criteria is to read the minimum number of sensors to achieve a probability of error in the problem of detecting the target less or equal a given arbitrarily small value, ϵ^1 . Also, as the power-related quantity that is assumed to be constant is the number of deployed sensor per area unit, the above criteria must be transformed to the minimum number of sensors per area unit.

Assuming that l is large enough or, equivalently, that the sensors are densely deployed, the number of read sensor to achieve a probability of error less or equal to ϵ , l, is

$$l \geq \frac{\ln \epsilon}{D}$$

where $D = D(H_0||H_1)$ for NP tests, and $C(H_1, H_0)$ for Bayes tests. The number of read sensor per area unit to achieve a probability of error less or equal to ϵ , $l\rho$, is

$$l\rho \ge \frac{\rho \ln \epsilon}{D} \tag{5}$$

By minimizing the right side of (5), we want to answer the following questions:

- 1. Given the sensors, are they any optimum configuration of the exploration area D?.
- Given D, and supposing that we can tune the sensors for different values of α following the ROC curve of the sensor, what is the optimum value of α?.
- 3. Are the solutions of the above questions independents?

¹For the NP tests, ϵ is the power of the test, and for Bayes tests, ϵ is the mean probability of error.

The analysis will be performed using the "spanish hat" model and circular exploration areas around the target.

For NP tests, the minimum of (5) is equivalent to the maximum of

$$\frac{D(H_0||H_1)}{\rho} = \begin{cases} \pi r_0^2 \left(\alpha \ln \frac{\alpha}{1-\beta} + (1-\alpha) \ln \frac{1-\alpha}{\beta} \right) & \text{if } R \ge r_0 \\ \pi R^2 \left(\alpha \ln \frac{\alpha}{1-\beta} + (1-\alpha) \ln \frac{1-\alpha}{\beta} \right) & \text{if } R < r_0 \end{cases}$$
(6)

and the first conclusion is that the exploration area, \mathcal{D} , must cover, at least, the range of the sensor, having no penalty for exploring big areas (other than the managing larger amount of data). This conclusion is maintained using more realistic p_d functions, as shown in Figure 1, where all the shown functions have the same effective range $r_e = \sqrt{\frac{\int (p_d(\boldsymbol{x}^t, \boldsymbol{x}, \alpha) - \alpha) d\boldsymbol{x}}{\pi(p_d(\boldsymbol{x}^t, \boldsymbol{x}^t, \alpha) - \alpha)}}$, that here takes value equal to 1.



Fig. 1. Different p_d functions (a) and its corresponding values of Eq. 6 as a function of R.

Considering now a fixed value of R, and considering also that r_0 , the range of the sensor, is an intrinsically fixed parameter of the sensor, finding the maximum of (6) is equivalent to finding the maximum of the binary discrimination [6] between α and $(1 - \beta)$,

$$L_b(\alpha, 1-\beta) = \alpha \ln \frac{\alpha}{1-\beta} + (1-\alpha) \ln \frac{1-\alpha}{\beta}$$
(7)

Obviously, $L_b(\alpha, 1 - \beta)$ can be infinite if α or β (or both) are zero, but we can only achieve the combinations α and β that allows the ROC curve of the sensor. Depending on the ROC curve, the maximum of (7) can reached in any point of the ROC curve. Nevertheless, one conclusion than we can obtain here, and that is also the answer to the third question, is that the value of α , the PFA of the sensor, is fixed locally for obtaining the maximum discrimination, independently of the rest of the parameters of the sensor network.

For Bayes tests, the minimum of (5) is equivalent to the maximum of

$$\frac{C(H_1, H_0)}{\rho} = -\pi R^2 \begin{cases} \ln\left(\frac{r_0^2}{R^2} \left(\frac{(1-\beta)^{s_0}}{\alpha^{s_0-1}} + \frac{\beta^{s_0}}{(1-\alpha)^{s_0-1}}\right) + \left(1 - \frac{r_0^2}{R^2}\right)\right) & \text{if } R \ge r_0 \\ \ln\left(\frac{(1-\beta)^{s_0}}{\alpha^{s_0-1}} + \frac{\beta^{s_0}}{(1-\alpha)^{s_0-1}}\right) & \text{if } R < r_0 \end{cases}$$
(8)

where s_0 is as in (4).

For fixed values of α , β and r_0 , (8) is a increasing function of R^2 for $R \ge r_0$ and an decreasing function of R^2 for $R < r_0$. So, it achieves its maximum when $R = r_0$.

Considering now fixed values of R and r_0 , finding the maximum of (8) is equivalent to finding the minimum of the function

$$C_b(\alpha,\beta) = \frac{(1-\beta)^{s_0}}{\alpha^{s_0-1}} + \frac{\beta^{s_0}}{(1-\alpha)^{s_0-1}} \tag{9}$$

that is a similar problem to finding the maximum of $L_b(\alpha, 1 - \beta)$ in NP tests, in the sense that its minimum can be reached in any point of the ROC curve of the sensor. Also, as in NP tests, the values of α and β that minimizes (9) are independent of the rest of the parameter, and also, as before, the problem of finding the parameters that minimizes the needed number of sensor per area unit can be divided in two independent problems: the problem of determining the best exploration area, and the problem of determining the best local decision parameters. The evaluation of different p_d functions corroborates these statements, but no results are shown here due to space limitations.

7. CONCLUSIONS

In this paper we proposed to model the probability of detection of each sensor, p_d , as a function of the distance between the sensor and the source or target to be detected. Based on that model, we derived the LLR for the detection problem, the Bayesian fusion rule and, under the asymptotic gaussianity of the LLR, the Neyman-Pearson fusion rule. The probability of error of both test were analyzed using Stein's lemma (NP test) and Chernoff information (Bayes test), and a simple parametric approximation to p_d .

Using as a criteria of efficiency the minimum number of read sensors per area unit to achieve a probability of error less or equal that a given value, the analysis revealed two main facts. The first is that the minimum area of exploration must cover the range of the sensors. In the case of NP tests, there is no penalty for considering bigger areas. On the other hand, in the case of Bayes tests, considering bigger exploration areas with the same density of read sensors penalizes the error exponent. The second is that, for configuring the sensors before deploying or before the reading, the only information that it is needed is the function p_d and the type of global fusion rule (NP or Bayes). The rest of the parameters of the sensor network does not compromise the performances, as well as the local decision rule is configured for the greatest discrimination between hypothesis.

ACKNOWLEDGMENTS

The author thanks Lang Tong for their helpful comments and suggestions, and fruitful discussions.

8. REFERENCES

- Z. Chair and P.K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. AES*, vol. AES-22, no. 1, pp. 98–101, Jan. 1986.
- [2] J.N. Tsitsiklis, "Decentralized detection by a lare number of sensors," *Math. Contr., Sign., Sys.*, vol. 1, pp. 167–182, 1988.
- [3] P.K. Varshney, Distributed Detection and Data Fusion, Springer-Verlag, New York, NY, 1997.
- [4] B.E. Boser, L. Doherty, B.A. Warneke and K.S.J. Pister, "Energy and performance considerations for smart dust," *Int. J. of Par. Distr. Sys. and Netw.*, vol. 4, no. 3, pp. 121–133, 2001.
- [5] T.M. Cover and J.A. Thomas, *Elements of Information The*ory, John Wiley & Sons, New York, NY, 1991.
- [6] R.E. Blahut, Principles and Practice of Information Theory, Addison-Wesley, Reading, MA, 1987.