OPTIMAL SENSOR SELECTION IN HETEROGENEOUS SENSOR NETWORKS

Marcelino Lázaro and Antonio Artés-Rodríguez

Department of Signal Theory and Communications
Universidad Carlos III de Madrid
Avda. Universidad 30, 28911, Leganés (Madrid) SPAIN
{mlazaro,antonio}@tsc.uc3m.es

ABSTRACT

We consider a large-scale, dense sensor network with two types of binary sensors with different discrimination performance and costs. We want to analyze what is the optimal proportion of both sensors in a target detection problem by assigning a given total cost per measurement. We obtain the hypothesis test, and we obtain the test performance using large deviation bounds. We found that the optimal proportion is "winner takes all" like. The sensor with the best performance/cost ratio is selected.

***Index Terms*—** Sensor networks, large deviations, cost-constrained.

1. INTRODUCTION

The fundamental performance limits in large-scale, randomly deployed, dense wireless sensor networks has received a great attention in the last years. Multiple contributions have been made to characterize the performance in several problems; the asymptotic performance for estimation [1, 2] and detection [3] problems are some examples.

Some efforts have also been made to determine the selection of the local detection rules (the kind of sensors to be used). Tsitsiklis [4] shown that when the number of sensors is arbitrarily large, the optimal binary decentralized detection is achieved by identical local detection rules, and this result has been recently extended [3] showing that using identical transmitter is also optimal.

One key aspect in networks with self-powered sensors is the energy consumption, which limits the time life of the network. The most energy demanding task, according to [5], is the wireless transmission. One alternative to economize energy is to employ censored transmission schemes. In [6] the censoring scheme is based on the idea that only sensors with positive detections try to transmit their positions. More elaborate censoring schemes has been proposed in the literature [7], but they do not apply in this setting because, as said before, the local decision could not be based on a likelihood ratio.

In this contribution we consider the problem of designing a binary network when different types of sensors, with different characteristics, are available and when a cost limitation is imposed. A censored transmission scheme is employed aiming to enlarge the network life. The simplest case, with two classes of sensors available, is analyzed. We do not restrict the local detectors to be based on a likelihood ratio, thus allowing the use of wide used non-parametric, learning-based local detectors, like in [8]. The behavior of the local detectors will be characterized using, for each class of sensors, 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, as proposed in [9]. The performance will be evaluated by means of a large deviation bound. The optimal solution will maximize this measure subject to the cost constraint.

This work has been partly supported by Ministerio de Educación y Ciencia of Spain (project 'DOIRAS', id. TIC2003-02602, and project 'MONIN', id. TEC2006-13514-C02-01), Comunidad de Madrid (project 'PRO-MULTIDIS-CM', id. S0505/TIC/0223), and the EC (Network of Excellence 'CRUISE', id. IST-4-027738).

2. PROBLEM STATEMENT AND NOTATION

We consider that two different classes of sensors, class-a, and class-b, are available to deploy a sensor network over a region $\mathcal{D} \in I\!\!R^2$, of area S. Sensors of both classes will be randomly deployed over $\mathcal{D}$ drawing a uniform distribution (for each class). The total cost dedicated to sensors is constrained to be a fixed amount C. In this case, if C^a and C^b are the costs per unit of sensors of class-a and class-b, respectively, and if ℓ^a and ℓ^b denote the number of sensors for each class in $\mathcal{D}$, the following constraint must be satisfied

$$\ell^a \cdot C^a + \ell^b \cdot C^b = C. \tag{1}$$

All sensors of a given class will apply the same binary detection rule, not necessarily based on a log-likelihood ratio (LLR) test. Given the region $\mathcal{D}$ and the position of a possible agent or target, $\boldsymbol{x}^t$, two hypothesis are defined: H_0, or null hypothesis if a target is not present at $\boldsymbol{x}^t$; H_1 or alternative hypothesis if a target is present. The probability of detection of an agent located at coordinates $\boldsymbol{x}^t$, for a sensor of class-j, $j \in \{a, b\}$, which is located at position $\boldsymbol{x}$, is denoted as

$$p_d^j(\boldsymbol{x}^t, \boldsymbol{x}, \alpha^j) = P(Y = 1 | \boldsymbol{X}^t = \boldsymbol{x}^t, \boldsymbol{X} = \boldsymbol{x}),$$

where α denotes the probability of false alarm (PFA) of the detector. This function depends on the nature of the specific detection process, but its general conditions are established in [9], where the joint probabilities of $\boldsymbol{X}$ and Y under each hypothesis are provided.

A censored transmission scheme is considered, similar to the one proposed in [10]. The extension of the scenario proposed in [10] to deal with two classes of sensors can be summarized as follows:

- The exploration of $\mathcal{D}$ can produce, potentially, the following data set:

$$\{(\boldsymbol{x}_i^a, y_i^a), (\boldsymbol{x}_k^b, y_k^b) : i = 1, \dots, \ell^a,\ k = 1, \dots, \ell^b\}$$

  with $\boldsymbol{x}_i^j \in \mathcal{D}, y_i^j \in \{0, 1\}$.

- Each pair $(\boldsymbol{x}_i^j, y_i^j)$ represents the reading of a sensor of class-j located at coordinates $\boldsymbol{x}_i^j$ that can detect $(y_i^j = 1)$, or not $(y_i^j = 0)$ a specific target.
- As in [10], to reduce the energy consumption, and thus enlarging the time life of the network, we assume that a parameter p_s, which defines the probability of sensing, can be dynamically tuned at all sensors. At every sensing instant (automatic or beacon driven), each sensor independently decides to sense with a probability p_s. We denote $\ell_s^j \leq \ell^j$ the number of sensors of each class that sense.
- For the sake of saving energy, only sensors with a positive detection $(Y = 1)$ try to transmit their position, $\boldsymbol{x}_i^j, j \in \{a, b\}$, to the fusion center. As in [10], to model the effect of medium access and transmission errors, we consider that each sensor that tries to transmit has a probability of transmission error, p_e. At the fusion center only the positions of the $\ell_t^j \leq \ell_s^j \leq \ell^j$ sensors of class-j that achieved a successful transmission are available. In the following, these positions will be labeled as $\{\boldsymbol{x}_i^j, i = 1, \cdots, \ell_t^j\}, j \in \{a, b\}$.

It is also interesting to remark the information that is not available at the fusion center:

- The number of sensors of each class, ℓ_n^j, that, after sensing, obtain a negative detection $(Y = 0)$, and the positions of such sensors.
- The number of sensors of each class, ℓ_e^j, that, after sensing and obtaining a positive detection $(Y = 1)$, fail to transmit their position to the fusion center, and the positions of such sensors.

In this scenario, we try to determine the optimal number of sensors of each class, ℓ^a, and ℓ^b, if a cost constraint as (1) is established.

3. HYPOTHESIS TEST FOR DETECTION

The hypothesis test for the detection problem stated above is based on the following vector of observations

$$\theta = [\boldsymbol{x}_1^a, \cdots, \boldsymbol{x}_{\ell_t^a}^a, \boldsymbol{x}_1^b, \cdots, \boldsymbol{x}_{\ell_t^b}^b, \ell_t^a, \ell_t^b],$$

which includes, in this order, the positions of the ℓ_t^a sensors of class-a that achieved a successful transmission, the positions of the ℓ_t^b sensors of class-b that achieved a successful transmission, and the number of sensors of each class that are available at the fusion center, ℓ_t^a and ℓ_t^b, respectively.

The likelihood ratio is

$$\Gamma = \frac{f_{\boldsymbol{\Theta}|H}(\boldsymbol{\theta}|H_1)}{f_{\boldsymbol{\Theta}|H}(\boldsymbol{\theta}|H_0)},$$

where the pdf of the observations under hypothesis H_u is

$$f_{\boldsymbol{\Theta}|H}(\boldsymbol{\theta}|H_u) = \prod_{i=1}^{\ell_t^a} f_{X|H,Y}^a(\boldsymbol{x}_i^a|H_u, 1) \prod_{k=1}^{\ell_t^b} f_{X|H,Y}^b(\boldsymbol{x}_k^b|H_u, 1) \cdot f_{L_t^a, L_t^b|L^a, L^b, H}(\ell_t^a, \ell_t^b|\ell^a, \ell^b, H_u).$$

Here, L_t^j and L^j, $j \in \{a, b\}$ are the random variables modeling the number of sensors (of each class) that achieved a successful transmission and the total number of deployed sensors (of each class), respectively. It is straightforward to obtain $f_{X|H,Y}^j(\boldsymbol{x}_i^j|H_u, 1)$ by using the definition of the probability of detection $p_d^j(\boldsymbol{x}^t, \boldsymbol{x}_i, \alpha^j)$

$$f_{X|H,Y}^j(\boldsymbol{x}_i^j|H_1, 1) = \frac{p_d^j(\boldsymbol{x}^t, \boldsymbol{x}_i^j, \alpha^j)}{\int_{\mathcal{D}} p_d^j(\boldsymbol{x}^t, \boldsymbol{x}, \alpha^j)\, d\boldsymbol{x}},$$

and

$$f_{X|H,Y}^j(\boldsymbol{x}_i^j|H_0, 1) = \frac{1}{S}.$$

The number of sensors of each class are independent, therefore probabilistically they are modeled as follows:

$$f_{L_t^a, L_t^b|L^a, L^b, H}(\ell_t^a, \ell_t^b|\ell^a, \ell^b, H_u) = f_{L_t^a|L^a, H}(\ell_t^a|\ell^a, H_u) \cdot f_{L_t^b|L^b, H}(\ell_t^b|\ell^b, H_u), \quad (2)$$

where the distribution for the number of sensors of each class is given by (see [10])

$$f_{L_a^j|L^j, H}(\ell_a^j|\ell^j, H_u) = \sum_{\ell_s = \ell_a^j}^{\ell^j} \binom{\ell_s}{\ell_a^j} \left(p_{t|u}^j\right)^{\ell_t^j} \left(1 - p_{t|u}^j\right)^{\ell_s - \ell_a^j} \cdot \binom{\ell}{\ell_s} p_s^{\ell_s} \left(1 - p_s\right)^{\ell^j - \ell_s}.$$

Given a probability of transmission error, p_e, the probability of a sensor of class$-j$ having a successful transmission is

$$p_t^j = (1 - p_e) \cdot p_{\mathcal{D}}^j,$$

where $p_{\mathcal{D}}$ is the probability of having a positive detection for a sensor in region $\mathcal{D}$. Obviously, this probability depends on the underlying hypothesis, i.e.

$$p_{t|1}^j = (1 - p_e) \cdot p_{\mathcal{D}|1}^j,$$

and

$$p_{t|0}^j = (1 - p_e) \cdot p_{\mathcal{D}|0}^j,$$

where

$$p_{\mathcal{D}|1}^j = E\left\{p_d^j(\boldsymbol{x}^t, \boldsymbol{x}, \alpha^j)\right\} = \frac{1}{S} \int_{\mathcal{D}} p_d^j(\boldsymbol{x}^t, \boldsymbol{x}, \alpha^j)\, d\boldsymbol{x},$$

and

$$p_{\mathcal{D}|0}^j = \alpha^j.$$

Finally, the decision is usually given in terms of this ratio

$$\gamma = \ln \Gamma \underset{H_1}{\overset{H_0}{\lessgtr}} \tau.$$

The threshold can be obtained, for instance, by using the Bayes criteria or by means of asymptotic gaussianity for the Neyman-Pearson criteria (similar to the one obtained in [9]).

4. PERFORMANCE ANALYSIS BY LARGE DEVIATION BOUNDS

To evaluate the performance of the hypothesis test, we will use the large deviation framework by using error exponents. These asymptotic measurements will be used to search the optimal proportion of each class of sensors in a cost-constrained environment. If ε_n is the probability of error (of some kind) obtained with n observations, the error exponent is defined as

$$\lim_{n \to \infty} -\frac{1}{n} \ln \varepsilon_n$$

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

$$\lim_{n\to\infty} -\frac{1}{n}\ln\beta_n = D(f_{\Theta|H}(\boldsymbol{\theta}|H_0)\|f_{\Theta|H}(\boldsymbol{\theta}|H_1)),$$

where $D(f_{\Theta|H}(\boldsymbol{\theta}|H_0)\|f_{\Theta|H}(\boldsymbol{\theta}|H_1))$ denotes the Kullback-Leibler (KL) divergence [11] between the probability density functions of the observations under each hypothesis. We will use the notation $D(H_0\|H_1)$ for short.

For the test under analysis, the divergence is

$$D(H_0||H_1) = \sum_{\ell_t^a=0}^{\ell^a}\sum_{\ell_t^b=0}^{\ell^b} f_{L_t^a,L_t^b|L^a,L^b,H}(\ell_t^a,\ell_t^b|\ell^a,\ell^b,H_0) \cdot \left\{ \ln\frac{f_{L_t^a,L_t^b|L^a,L^b,H}(\ell_t^a,\ell_t^b|\ell^a,\ell^b,H_0)}{f_{L_t^a,L_t^b|L^a,L^b,H}(\ell_t^a,\ell_t^b|\ell^a,\ell^b,H_1)} - \left(\ell_t^a+\ell_t^b\right)\cdot\ln S - \ell_t^a\frac{1}{S}\int_{\mathcal{D}}\ln f^a_{X|H_1,Y}(\boldsymbol{x}|H_1,1)\,d\boldsymbol{x} - \ell_t^b\frac{1}{S}\int_{\mathcal{D}}\ln f^b_{X|H_1,Y}(\boldsymbol{x}|H_1,1)\,d\boldsymbol{x}\right\}. \quad (3)$$

5. OPTIMAL NUMBER OF SENSORS OF EACH CLASS UNDER COST CONSTRAINTS

Now, we want to obtain the optimal number of sensors of each class in the cost constrained problem defined by equation (1) by using the asymptotic performance measurement obtained in the previous section. This means that we look for the couple ℓ^a and ℓ^b that maximizes the divergence (3) subject to the constraint (1).

The divergence (3) can be divided in four terms, associated to the four terms into the brackets. Taking into account the independence of the number of vector for each class, which implies equation (2), it can be seen that the last three term are proportional to, respectively

$$E\left[\ell_t^a+\ell_t^b|L^a=\ell^a,L^b=\ell^b,H=H_0\right], \quad (4)$$

$$E\left[\ell_t^a|L^a=\ell^a,H=H_0\right], \quad (5)$$

and

$$E\left[\ell_t^b|L^b=\ell^b,H=H_0\right], \quad (6)$$

where, abusing of notation, the operator $E[\bullet|F=f]$ denotes expected value given $F=f$. Constraint (1) implies a linear relationship between ℓ^a and ℓ^b,

$$\ell^b = \frac{C-\ell^a\cdot C^a}{C^b}.$$

Expectation (5) is linear with the number of sensors of class-a, and therefore, linear with the numbers of sensors of class-a too. Expectation (6) is linear with the number of sensors of class-b, and therefore, linear with the number of sensors of class-a too. And expectation (4) is proportional to the total number of sensors, and, because of the independence, linear with the number of sensors of each class.

Finally, the first term of (3) can be rewritten as

$$\sum_{\ell_t^a=0}^{\ell^a}\sum_{\ell_t^b=0}^{\ell^b} f_{L_t^a|L^a,H}(\ell_t^a|\ell^a,H_0)\cdot f_{L_t^b|L^b,H}(\ell_t^b|\ell^b,H_0) \cdot\ln\frac{f_{L_t^a|L^a,H}(\ell_t^a|\ell^a,H_0)\cdot f_{L_t^b|L^b,H}(\ell_t^b|\ell^b,H_0)}{f_{L_t^a|L^a,H}(\ell_t^a|\ell^a,H_1)\cdot f_{L_t^b|L^b,H}(\ell_t^b|\ell^b,H_1)} = \sum_{\ell_t^a=0}^{\ell^a} f_{L_t^a|L^a,H}(\ell_t^a|\ell^a,H_0)\cdot\ln\frac{f_{L_t^a|L^a,H}(\ell_t^a|\ell^a,H_0)}{f_{L_t^a|L^a,H}(\ell_t^a|\ell^a,H_1)} + \sum_{\ell_t^b=0}^{\ell^b} f_{L_t^b|L^b,H}(\ell_t^b|\ell^b,H_0)\cdot\ln\frac{f_{L_t^b|L^b,H}(\ell_t^b|\ell^b,H_0)}{f_{L_t^b|L^b,H}(\ell_t^b|\ell^b,H_1)}.$$

After some mathematical manipulations

$$\sum_{\ell_t^j=0}^{\ell^j} f_{L_t^j|L^j,H}(\ell_t^j|\ell^j,H_0)\cdot\ln\frac{f_{L_t^j|L^j,H}(\ell_t^j|\ell^j,H_0)}{f_{L_t^j|L^j,H}(\ell_t^a|\ell^j,H_1)} = \ell^j\cdot\left\{p_{t|0}^j\cdot p_s\cdot\ln\frac{p_{t|0}^j}{p_{t|1}^j} + \left[(1-p_s)+(1-p_{t|0}^j)p_s\right]\cdot\ln\frac{(1-p_s)+(1-p_{t|0}^j)p_s}{(1-p_s)+(1-p_{t|1}^j)p_s}\right\},$$

i.e., linear with the number of sensors. The divergence (3) is linear with the number of sensors of both classes, independently of the shape of the probability of detection function. Therefore, the optimal solution of the cost-constrained problem will be one of the following possibilities:

$$\left(\ell^a=\frac{C}{C^a},\ell^b=0\right) \text{ or } \left(\ell^a=0,\ell^b=\frac{C}{C^b}\right),$$

i.e., to choose all sensors of the same class. The class to be selected will depend on the performance/cost ratio of each class. Figures 1 and 2 show the divergence in a cost-constrained problem for two probabilities of detection, the "spanish hat",

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

and the exponential

$$p_d(\boldsymbol{x}^t,\boldsymbol{x},\alpha) = \alpha + (1-\alpha-\beta)e^{-\theta\|\boldsymbol{x}^t-\boldsymbol{x}\|_2^2}, \quad (7)$$

where in both cases α denotes the probability of false alarm, and β the probability of miss-detection. Figure 1 considers the following normalized costs: $C=100$, $C^a=2$, and $C^b=1$. The parameters for "spanish hat" and exponential are $\alpha^a=0.05$, $\beta^a=0.05$, $\alpha^b=0.25$, $\beta^b=0.25$ (all these are common), and $r_o^a=1$, $\theta^a=1$, $r_o^b=0.75$, $\theta^b=1.33$. It can be seen how in this case the divergence is linearly growing with ℓ^a for both probabilities of detection. Therefore, in this case the best option is to choose all sensors of class-a (50 sensors in this case).

Figure 2 considers the following normalized costs: $C=100$, $C^a=4$, and $C^b=1$. Now, class-a sensors are notably better than the previous ones, but also more expensive. The parameters for "spanish hat" and exponential are $\alpha^a=0.01$, $\beta^a=0.01$,

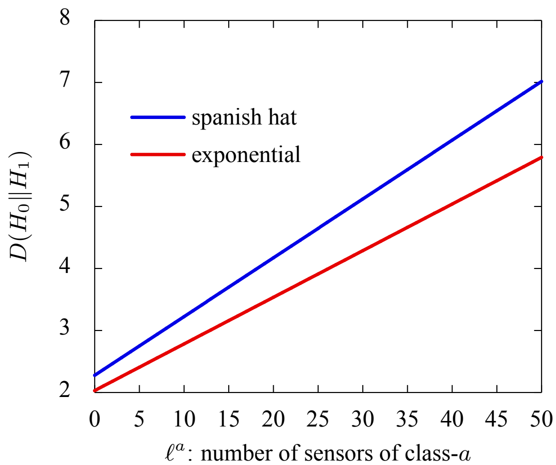

Fig. 1. $D(H_0||H_1$ for normalized costs $C = 100$, $C^a = 2$, and $C^b = 1$. Parameters for "spanish hat" and exponential: $\alpha^a = 0.05$, $\beta^a = 0.05$, $r_o^a = 1$, $\theta^a = 1$, $\alpha^b = 0.25$, $\beta^b = 0.25$, $r_o^b = 0.75$, $\theta^b = 1.33$.

$\alpha^b = 0.15$, $\beta^b = 0.15$ (all these are common), and $r_o^a = 1$, $\theta^a = 1$, $r_o^b = 0.75$, $\theta^b = 1.33$. It can be seen how in this case the divergence is linearly decreasing with ℓ^a for both probabilities of detection. Therefore, in this case the best option is to choice all sensors of class-b (100 sensors in this case).

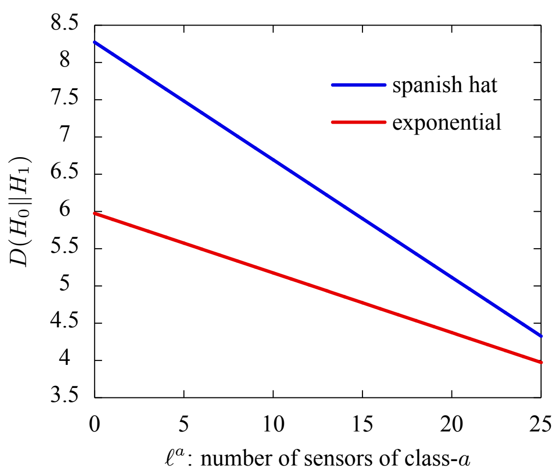

Fig. 2. $D(H_0||H_1$ for normalized costs $C = 100$, $C^a = 4$, and $C^b = 1$. Parameters for "spanish hat" and exponential: $\alpha^a = 0.01$, $\beta^a = 0.01$, $r_o^a = 1$, $\theta^a = 1$, $\alpha^b = 0.15$, $\beta^b = 0.15$, $r_o^b = 0.75$, $\theta^b = 1.33$.

6. CONCLUSIONS

We have shown that in a binary sensor network, with a censored transmission scheme, the optimal selection of sensors under a cost limitation is to choose all sensors from the same type. The error exponent is linear with the number of sensors of each class, independently of the probability of detection function. The class to be selected will depend on the performance/cost ratio of each class of sensors. Therefore, to select the best class it is only necessary to evaluate the divergence in (3) for two cases: only class-a sensors and only class-b sensors. This is equivalent to evaluate the simplest divergence expression derived for a single sensor type, presented in [10], for the maximum allowable number of sensors of each class considering the cost constraint, and to select the class with a higher divergence.

7. REFERENCES

[1] D. Marco, E. J. Duarte-Melo, M. Liu, and D. L. Neuhoff, "On the many-to-one transport capacity of a dense wireless sensor network and the compressibility of its data," in *International Workshop on Information Processing in Sensor Networks*, Berkeley, CA, Apr. 2003.

[2] H. El Gamal, "On the scaling laws of dense wireless sensor networks: the data gathering channel," *IEEE Transactions on Information Theory*, vol. 51, no. 3, pp. 1229–1234, Mar. 2005.

[3] J. F. Chamberland and V. V. Veeravalli, "Asymptotic results for decentralized detection in power constrained wireless sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 6, pp. 1007–1015, Aug. 2004.

[4] John N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Mathematics of Control, Signals, and Systems*, vol. 1, pp. 167–182, 1988.

[5] L. Doherty, B. A. Warneke, B. Boser, and K. S. J. Pister, "Energy and performance considerations for Smart Dust," *Int. Journal of Parallel and Distributed Sensor Networks*, vol. 4, no. 3, Dec. 2001.

[6] A. Artés-Rodríguez, M. Lázaro, and M. Sánchez-Fernández, "Decentralized detection in sensor networks with censored transmissions," in *Proceedings of the ICASSP 2005*, Philadelphia, USA, Mar. 2005.

[7] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: a low communication-rate scheme for distributed detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, no. 2, pp. 554–568, Apr. 1996.

[8] X. Nguyen, M. J. Wainwright, and M. I. Jordan, "Nonparametric decentralized detection using kernel methods," *IEEE Transactions on Signal Processing*, In press.

[9] Antonio Artés-Rodríguez, "Decentralized detection in sensor networks using range information," in *Proceedings of the ICASSP 2004*, Montreal, Canada, May 2004.

[10] A. Artés-Rodríguez, M. Lázaro, and M. Sánchez-Fernández, "Decentralized detection in sensor networks with censored transmissions," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Philadelphia, USA, Mar. 2005, vol. 4, pp. 817–820.

[11] T. M. Cover and J. A. Thomas, *Elements of information theory*, John Wiley & Sons, New York, NY, 1991.