Decentralized detection for censored binary observations with statistical dependence

Marcelino Lázaro

ACCEPTED VERSION

(Submitted September 23, 2015. Accepted January 7, 2016)

Published in *Signal Processing*, Elsevier, ISSN: 0165-1684

Volume 123, June 2016, Pages 112-121

D.O.I.: 10.1016/j.sigpro.2016.01.005

Decentralized detection for censored binary observations with statistical dependence

Marcelino Lázaro

Universidad Carlos III de Madrid

Av. Universidad 30, Leganes 28911 MADRID, Spain

Abstract

This paper analyzes the problem of distributed detection in a sensor network of binary sensors. In particular, statistical dependence between local decisions (at binary sensors) is assumed, and two complementary methods to save energy have been considered: censoring, to avoid some transmissions from sensors to fusion center, and a sleep and wake up random schedule at local sensors. The effect of possible failures in transmission has been also included, considering the probability of having a successful transmision from a sensor to the fusion center. In this scenario, the necessary statistical information has been identified, the optimal decision rule at the fusion center has been obtained, and some examples have been used to analyze the effect of statistical dependence in a simple network with two sensors.

Keywords: Wireless sensor networks, correlation, censoring

1. Introduction

Joint advances in complementary disciplines such as material science, semiconductor technologies, networking, communications, or statistical signal processing, have allowed the ubiquitous deployment of wireless sensor networks (WSN) for numerous applications (see, for instance, [1] and references therein). Evolution of these networks has required an important development in many

Email address: mlazaro@tsc.uc3m.es (Marcelino Lázaro)

different research areas [2], and in particular the detection problem has received a great attention. This is one of the canonical problems in sensor networks: to decide between several possible hypothesis from a set of available observations. Statistical detection theory [3]-[5] provides the optimal solution to this problem under different criteria when conditional distributions for observations given each hypothesis are available. In the context of sensor networks, the detection problem typically becomes decentralized, because sensors usually process local observations to produce some quantization that is sent to a fusion center, instead of transmitting the observation value itself. Decentralized detection has also received a great attention in the literature, analyzing a great varity of network configurations and different related problems: optimal rules at the fusion center, at the local sensors, performance analysis, etc. Examples of interesting contributions can be found for instance in [6]-[11], although the literature on this topic is vast.

Initially, most results were based on the conditional independence assumption, i.e., conditional distributions of different observations were independent given each hypothesis. However, this assumption does not hold in many real applications, and different works have considered the decentralized detection problem under statistical dependence ([12]-[17] are again just examples of some interesting contributions). Unfortunatelly, in most cases analysis is much more involved under dependence, even in the case of networks with a small number of sensors, as it was perfectly illustrated in [18] or [19].

In networks with self-powered sensors, energy consumption becomes a crucial aspect for network lifetime [20], and therefore has to be taken into account at the different network levels [21], [22]. "Sleep" and "wake up" schedule of local sensor nodes is a technique that has demonstrated to increase the lifetime of the network [23, 24]. Using this technique, a given sensor will not sense, process its observation and transmit its decision at some "sleep" sampling instants or periods of time, which are specified according to some deterministic or random schedule of sensor activity. Depending on sensor characteristics, during these "sleep" intervals the sensor node can go to a sleeping mode or just can keep in an idle state that consumes a low amount of energy.

As one of the more energy demanding taks is wireless transmission [25], censored transmission schemes have been used to save energy. In a censored network, the transmission of the available information is avoided when under different criteria the local information is considered less informative or useful. In some cases transmission is prevented when the likelihood ratio of a given node is out of certain bounds [26, 27, 28]. Sequential tests have also been proposed at the local nodes [29] to only transmit when the certainty about the decision is high enough. However, these techniques are not useful for some widely used, non-parametric, learning-based local detectors, like in [30], or in commercial devices implementing unknown detection rules where likelihood measures of local observations are not available. In these cases, censoring has to be based on local decisions of sensors. In networks with binary sensors (sensors providing a binary local decision), a binary censoring scheme can be used, allowing transmission only in the case of a positive local detection, and preventing it in the case of a negative local detection [31].

In this paper a sensor network of binary sensors, with dependence in local decisions, will be considered, as in [13]. Unlike in [13], transmission from sensor nodes to the fusion center will not be considered ideal, but a given probability of failure in each transmission will be considered. Moreover, with the purpose of saving energy to increase the network lifetime, a sleep and wake up random schedule is combined with a binary censoring mechanism. This scenario was analyzed in [31] and [32] under the assumption of conditional independence, and will be extended here to include dependence in local decisions.

The paper is organized as follows. The problem will be stated in Section 2. Section 3 will provide the conditional distributions for observation, which will be used to obtain the optimal test at the fusion center. The statistical information that is necessary to know at the fusion center to implement the optimal decision will be identified in Section 4. Some examples will be used in Section 5 to analyze how dependence affects the performance of the proposed detection rule. Finally, Section 6 discusses some of the main conclusions that can be extracted from this work.

2. Problem Statement

This paper is focused on a decentralized detection problem that can be formulated as a binary hypothesis test when a network of ℓ binary sensors is deployed over a region $\mathcal{D}$. The problem is to decide between two possible hypotheses, the null hypothesis H_0 and the alternative hypothesis H_1, from the local binary decisions that are collected from the set of ℓ sensors in $\mathcal{D}$ by a fusion center. An example of this kind of problem is the one treated in [32], where a fusion center has to decide if a target is or is not present at a specific position $\boldsymbol{z} \in \mathcal{D}$ or inside a sub-region $\mathcal{D}' \subset \mathcal{D}$. In this example the two hypotheses are therefore defined as follows:

- Null hypothesis, H_0, if a target is not present (at $\boldsymbol{z}$ or in region $\mathcal{D}'$).
- Alternative hypothesis, H_1, if a target is present (at $\boldsymbol{z}$ or in region $\mathcal{D}'$).

It is assumed that all local sensors apply some binary detection rule, not necessarily identical for all sensors, and not necessarily based on a likelihood ratio test. The local detection rules may not be known at the fusion center.

As in [31] and [32], it will be assumed that the fusion center knows the number of sensors that are deployed in the network and the position of every sensor (as will be discussed in Section 3, the knowledge of sensor positions is not necessary in this paper; it is included here to maintain the same conditions of previous works [31] and [32]). Set $\mathcal{X}$ will contain the positions of the ℓ sensors

$$\mathcal{X} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_\ell\} \tag{1}$$

The exploration of $\mathcal{D}$, which can be either automatic or beacon driven, can potentially produce the following data set:

$$\left\{\{\boldsymbol{x}_j, y_j\}_{j=1}^{\ell}\right\} \tag{2}$$

with $\boldsymbol{x}_j \in \mathcal{D}$ being the position of the j-th sensor and $y_j \in \{0, 1\}$ being its binary local decision. Positions can be transmitted from sensors to the fusion center, but in static networks they can also be obtained at an initial calibration phase during the deployment of the network. In this case, in order to know the position of a given sensor, only a simple identification tag has to be transmitted (position will be implicit), and sensors do not have to incorporate a GPS chip or any other location device.

Vectors $\mathbf{x}$ and $\mathbf{y}$ will be used to group sensor positions and their local decisions, respectively

$$\mathbf{x} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_\ell], \ \mathbf{y} = [y_1, y_2, \cdots, y_\ell] \tag{3}$$

A many-to-one parallel transmission strategy to send this information to the fusion center is assumed. Not all information in (2) will be available at the fusion center to perform the hypothesis test. The reason is twofold. First, transmission failures can happen. Second, some techniques to economize power are used to prevent the sensors to sense and transmit at every sampling time. When a sensor tries to transmit to the fusion center, this transmission can be succesfull or can fail. The probability of a successful transmission will be denoted as p_t. In this work p_t will be considered to be equal for all sensors, although the extension to include different values for each sensor is straightforward.

Two particular mechanisms are used for power control: sleep and wake up schedule, and censoring. These techniques will be the same used in [31] and [32]

- A random sleep and wake up schedule is used locally, at sensor level. A sensor will only sense and process the local measure to obtain a local binary decision with a given probability p_s. This schedule is independent for each individual sensor in the network. Parameter p_s will tradeoff between performance and energy saving. Here, this parameter will be considered to be equal for all sensors, although the extension to include different values for each sensor is straightforward.

- A censoring mechanism is used to reduce the number of transmissions. When a sensor is "awake", it will sense and process its local measure to obtain a binary decision. Censoring will be based on the local decision:

if local decision is positive ($y_j = 1$) the sensor will try to transmit (its position or its identification tag); if local decision is negative ($y_j = 0$) the sensor will not transmit to save power.

This scenario was analyzed in [31] and [32] under the conditional independence assumption for local binary decisions. Here the same scenario will be analyzed without this assumption, considering the case of conditional dependence of local decisions. Under the assumption of conditional independence under each hypothesis (see [31] or [32]), it was only necessary to know the conditional probabilities of a positive local decision for each isolated sensor under both hypothesis, $p_d(\boldsymbol{x}|H_k)$; under the independence assumption such a function is in general dependent on the distance of the sensor with target for H_1, and a constant probability of false alarm, α, for H_0. This fact allows the problem to be stated in terms of the joint distribution of sensor location and its local decision for a single sensor

$$f_{\boldsymbol{X},Y|H}(\boldsymbol{x},y|H_k) = f_{\boldsymbol{X}}(\boldsymbol{x})\left(p_d(\boldsymbol{x}|H_k)\delta[y-1] + (1-p_d(\boldsymbol{x}|H_k))\delta[y]\right) \tag{4}$$

where $f_{\boldsymbol{X}}(\boldsymbol{x})$ is the probability density function of sensor locations. If sensors are assumed to be randomly and uniformly located in $\mathcal{D}$, this function is

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{S_{\mathcal{D}}} \tag{5}$$

where $S_{\mathcal{D}}$ is the surface (or volume) of area $\mathcal{D}$, and when locations are known, and included in set $\mathcal{X}$

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{\ell} I_{\mathcal{X}}(\boldsymbol{x}) \tag{6}$$

where $I_{\mathcal{X}}(\boldsymbol{x})$ is the indicator function saying if $\boldsymbol{x}$ is contained in set $\mathcal{X}$.

In this paper some statistical dependence between local decisions will be considered. The statistical knowledge about this dependence will be in this case given by the joint probability of detection and joint probability of false alarm of every possible subset of n sensors. These probabilities are defined as the probability that all these n sensors obtain simultaneously a positive detection under H_1 and H_0, respectively. If the set $\mathcal{I}^n$ contains the indexes of n different sensors

$$\mathcal{I}^n = \{i_1, i_2, \cdots, i_n\} \tag{7}$$

with $i_k \neq i_j$ if $k \neq j$, and vector $\mathbf{x}_{\mathcal{I}^n}$ contains the locations of the n sensors with indexes contained in $\mathcal{I}^n$

$$\mathbf{x}_{\mathcal{I}^n} = [\boldsymbol{x}_{i_1}, \boldsymbol{x}_{i_2}, \cdots, \boldsymbol{x}_{i_n}] \tag{8}$$

the following joint conditional probabilities

$$p_d(\mathbf{x}_{\mathcal{I}^n}|H_k) = P(y_{i_1} = y_{i_2} = \cdots = y_{i_n} = 1|H_k) \tag{9}$$

denote the joint probabilities of detection (for H_1) and joint probabilities of false alarm (for H_0), which will be known for each $n \in \{1, 2, \cdots, \ell\}$ and for each possible set of n different indexes $\mathcal{I}^n$. In practice, these functions can be obtained analytically (based on the knowledge about the physical process or statistical model for measurements and local decision rules) or by means of empirical measurements. This allows to work with commercial devices providing a binary detection with an unknown decision rule. It is interesting to remark that for ℓ sensors, the number of joint probabilities is given by $2\left(2^\ell - 1\right)$, which limits the size of ℓ that can be treated in practice.

3. Hypothesis test at the fusion center

The optimal hypothesis test at the fusion center, under Neyman-Pearson and Bayes criteria, is given in terms of a likelihood ratio test (LRT) [5]. If Θ is the random variable denoting observations at the fusion center, and θ denotes its realization, the likelihood ratio test is defined as the ratio between conditional distributions of Θ given the hypothesis

$$\Lambda_\Theta(\theta) = \frac{f_{\boldsymbol{\Theta}|H}(\theta|H_1)}{f_{\boldsymbol{\Theta}|H}(\theta|H_0)} \tag{10}$$

The detection rule at the fusion center will be given by

$$u_{FC}(\theta) = \begin{cases} 0, & \text{if } \Lambda(\theta) < \lambda \\ \xi, & \text{if } \Lambda(\theta) = \lambda \\ 1, & \text{if } \Lambda(\theta) > \lambda \end{cases} \tag{11}$$

Threshold λ, and distribution of the Bernoulli randomization variable ξ for the NP test, will be obtained to fulfill the NP or Bayes criterion (here, a common notation for NP or Bayesian detection has been used to simplify the presentation because this is part of the well known theory of statistical decision; further details can be found in every classical book such as [3], [4] or [5]).

In this paper $\mathbf{X}$ is used to denote sensor positions, which are supposed to be known at the fusion center. In fact, it is not necessary to know the sensor positions, but it is only necessary for the fusion center to be able to identify the sensor which is associated to each received transmission (to make use of the appropriate statistical model in the likelihood ratio); therefore, it is straightforward to modify $\mathbf{X}$ to represent identification tags for each sensor in the network. The reason to include sensor positions in notation instead of identification tags is twofold: first, to mantain the same basic notation as in previous works [31] and [32] that analyzed the same scenarios under the conditional independence assumption; second, because in many real applications the probability of detection for a given sensor depends on its position (typically as a function of the distance with the target to be detected). The explicit inclusion of sensor position in notation can be helpful in these applications to make easier to understand the detection procedure.

3.1. Conditional distributions using binary censoring

In this section, conditional distributions defining the LRT will be obtained for a network where binary censoring, along with a sleep and wake up schedule, is used to economize energy. Distributions will be initially obtained for a network with only two sensors and will be later extended to an arbitrary number ℓ of sensors. For the case of two sensors, let $\mathbf{X}$ and $\mathbf{Y}$ be the random variables denoting the positions and local decisions, respectively, of both sensors

$$\mathbf{X} = [\boldsymbol{X}_1, \boldsymbol{X}_2],\ \mathbf{Y} = [Y_1, Y_2] \tag{12}$$

If $\mathbf{x}$ and $\mathbf{y}$ denote their possible realizations, $f_{\mathbf{X},\mathbf{Y}|H}(\mathbf{x}, \mathbf{y}|H_k)$ denotes the joint conditional distribution under hypothesis H_k, and $f_{\mathbf{Y}|\mathbf{X},H}(\mathbf{y}|\mathbf{x}, H_k)$ denotes the conditional distribution of local decisions given sensor positions. These ditributions (for $k \in \{0, 1\}$) are related with the joint probability of detection of the $\ell = 2$ sensors.

Let the random variable Θ denote the observation available at the fusion center, and θ will denote its realization. Because of the censoring mechanism that is used, the observation will be completely described by the positions of sensors having a successful transmission with the fusion center, because the binary decision $y = 1$ is implicit by the binary censoring strategy. Therefore, the alphabet of Θ will be given by

$$\theta \in \{[\,], [\boldsymbol{x}_1], [\boldsymbol{x}_2], [\boldsymbol{x}_1, \boldsymbol{x}_2]\}\,, \tag{13}$$

because depending on several factors, zero, one or the two sensors will successfully transmit to the fusion center at a given time. Conditional distributions for each possible value of observation, assuming that positions of the ℓ sensors are known at the fusion center, will be obtained as follows. When there is not any successful transmisión, i.e., $\theta = [\,]$

$$\begin{aligned} f_{\boldsymbol{\Theta}|H}(\theta|H_k) =& f_{\mathbf{Y}|\boldsymbol{X},H}([0,0]|\mathbf{x}, H_k) + f_{\mathbf{Y}|\boldsymbol{X},H}([1,0]|\mathbf{x}, H_k)(1 - p_s p_t) \\ &+ f_{\mathbf{Y}|\boldsymbol{X},H}([0,1]|\mathbf{x}, H_k)(1 - p_s p_t) + f_{\mathbf{Y}|\boldsymbol{X},H}([1,1]|\mathbf{x}, H_k)(1 - p_s p_t)^2 \end{aligned} \tag{14}$$

This expression is obtained taking into account that such observation can be produced in the following cases:

- Both sensors would have a null local decision even if they were "awake" (according to the sleep and wake up random schedule).
- Sensor 1 would have a positive local decision, but it was "asleep" or it was "awake" but its transmission failed, and at the same time sensor 2 would have a negative local decision in any case.
- Sensor 2 would have a positive local decision, but it was "asleep" or it was "awake" but its transmission failed, and at the same time sensor 1 would have a negative local decision in any case.

- Both sensors would have a positive local decision, but they were "asleep" or they were "awake" but their transmissions failed.

Following the same reasoning, if only first sensor transmission has been received at the fusion center, i.e., $\theta = [\boldsymbol{x}_1]$

$$f_{\boldsymbol{\Theta}|H}(\theta|H_k) = f_{\mathbf{Y}|\boldsymbol{X},H}([1,0]|\mathbf{x},H_k)p_s p_t + f_{\mathbf{Y}|\boldsymbol{X},H}([1,1]|\mathbf{x},H_k)p_s p_t(1-p_s p_t) \tag{15}$$

and similarly, if the received transmission is only for second sensor, i.e., $\theta = [\boldsymbol{x}_2]$

$$f_{\boldsymbol{\Theta}|H}(\theta|H_k) = f_{\mathbf{Y}|\boldsymbol{X},H}([0,1]|\mathbf{x},H_k)p_s p_t + f_{\mathbf{Y}|\boldsymbol{X},H}([1,1]|\mathbf{x},H_k)p_s p_t(1-p_s p_t) \tag{16}$$

Finally, when fusion center receives the transmission of both sensors, i.e., $\theta = [\boldsymbol{x}_1, \boldsymbol{x}_2]$, conditional distribution is given by

$$f_{\boldsymbol{\Theta}|H}(\theta|H_k) = f_{\mathbf{Y}|\boldsymbol{X},H}([1,1]|\mathbf{x},H_k)(p_s p_t)^2 \tag{17}$$

Extension of these expressions for the general case of ℓ sensors is very simple. The vectors of random variables $\mathbf{X}$ and $\mathbf{Y}$ will now model positions and local decisions, respectively, for all sensors

$$\mathbf{X} = [\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_\ell],\ \mathbf{Y} = [Y_1, Y_2, \cdots, Y_\ell] \tag{18}$$

with $\mathbf{x}$ and $\mathbf{y}$ denoting the corresponding realizations

$$\mathbf{x} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_\ell],\ \mathbf{y} = [y_1, y_2, \cdots, y_\ell] \tag{19}$$

Observation at a given instant will contain the positions of the n sensors that after sensing and having a positive local decision had a successful transmission. Therefore an instance of this observation can be written as

$$\theta = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n] \tag{20}$$

where, without lack of generality and for the sake of a simpler notation, we have considered that the n sensors with a successful transmission are the ones corresponding to the first n positions in $\mathbf{x}$. In this case, the conditional distribution of the observation under hypothesis H_k is

$$f_{\boldsymbol{\Theta}|H}(\theta|H_k) = \sum_{\mathbf{y}_{na}\in\mathcal{Y}_{NA}^{\ell-n}} f_{\mathbf{Y}|\boldsymbol{X},H}([\mathbf{1}_n, \mathbf{y}_{na}]|\mathbf{x}, H_k) \times (p_s p_t)^n (1-p_s p_t)^{||\mathbf{y}_{na}||^2} \quad (21)$$

where $\mathbf{1}_n$ denotes a vector of n ones, and the set $\mathcal{Y}_{NA}^{\ell-n}$ represents the set of all possible values for non available measures, which in practice includes the set of $2^{\ell-n}$ binary vectors of length $\ell-n$. Note that $||\mathbf{y}_{na}||^2$ accounts for the number of ones in $\mathbf{y}_{na}$, which will be in the the range from 0 to $\ell-n$.

3.2. Conditional distributions without censoring

Although this paper is focused on binary censoring, in this section conditional distributions of the observations will be obtained also for a network that does not use censoring, which will be used as a reference to compare performance in Section 5 for a network with two sensors. Without censoring, every sensor that is "awake", according to its local random sleep and wake up schedule, will try to transmit its local decision y_i to the fusion center, along with its position or identification tag. Therefore, the alphabet of the observation for a network with 2 sensors will be given in this case by

$$\theta \in \{[\,],\ [(\boldsymbol{x}_1, y_1)],\ [(\boldsymbol{x}_2, y_2)],\ [(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2)]\} \quad (22)$$

Taking into account that $y_1, y_2 \in \{0, 1\}$, the observation can take 9 possible values. Conditional distributions of the observations for observation $\theta = [\,]$ are

$$f_{\boldsymbol{\Theta}|H}(\theta|H_k) = (1-p_s p_t)^2 \quad (23)$$

because this observation happens when both sensors were "asleep" or their transmissions failed. For observation $\theta = [(\boldsymbol{x}_1, y_1)]$

$$f_{\boldsymbol{\Theta}|H}(\theta|H_k) = f_{\mathbf{Y}|\boldsymbol{X},H}([y_1, 0]|\mathbf{x}, H_k) p_s p_t (1-p_s p_t) + f_{\mathbf{Y}|\boldsymbol{X},H}([y_1, 1]|\mathbf{x}, H_k) p_s p_t (1-p_s p_t) \quad (24)$$

and for $\theta = [(\boldsymbol{x}_2, y_2)]$

$$f_{\boldsymbol{\Theta}|H}(\theta|H_k) = f_{\mathbf{Y}|\boldsymbol{X},H}([0, y_2]|\mathbf{x}, H_k) p_s p_t (1-p_s p_t) + f_{\mathbf{Y}|\boldsymbol{X},H}([1, y_2]|\mathbf{x}, H_k) p_s p_t (1-p_s p_t) \quad (25)$$

because in both cases the missing decision can be 0 or 1. Finally, for $\theta = [(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2)]$

$$f_{\boldsymbol{\Theta}|H}(\theta|H_k) = f_{\mathbf{Y}|\mathbf{X},H}([y_1, y_2]|\mathbf{x}, H_k)(p_s p_t)^2 \tag{26}$$

Extension of these expressions for the general case of ℓ sensors, when decisions for n sensors are available, is also simple. Assuming again for the sake of simplicity that available measures correspond to the first n elements of $\mathbf{x}$ and $\mathbf{y}$ in (19), the conditional distribution of the observation under hypothesis H_k is

$$f_{\boldsymbol{\Theta}|H}(\theta|H_k) = \sum_{\mathbf{y}_{na} \in \mathcal{Y}_{NA}^{\ell-n}} f_{\mathbf{Y}|\mathbf{X},H}([\mathbf{y}_n, \mathbf{y}_{na}]|\mathbf{x}, H_k) \times (p_s p_t)^n (1 - p_s p_t)^{\ell-n} \tag{27}$$

where $\mathbf{y}_n$ denotes a vector with the n received local decisions, and the set $\mathcal{Y}_{NA}^{\ell-n}$ represents the set of all possible values for non available measures, which in practice includes the set of $2^{\ell-n}$ binary vectors of length $\ell - n$.

4. Necessary statistical information

Expressions obtained for conditional distributions of the observations, required to implement the optimal NP or Bayes detection rule in the proposed scenario, allow to identify the information that is necessary to have at the fusion center to perform an optimal NP or Bayes test:

- Parameter p_s defining the sleep and wake up local schedule for sensors.
- Parameter p_t modeling the probability of a successful transmission.
- The conditional distributions of local binary decisions

$$f_{\mathbf{Y}|\mathbf{X},H}(\mathbf{y}|\mathbf{x}, H_k),\ k \in 0, 1 \tag{28}$$

  for all the 2^ℓ possible values of binary vector $\mathbf{y}$.

The following proposition states that these conditional distributions are implicit in the set of joint probabilities of detection for all possible sets of sensors, similarly to the result provided in [13] for an scenario without censoring, sleep and wake up mechanism or transmission failures.

Proposition 1. *The collection of conditional joint probabilities of detection*

$$\mathcal{P}_d = \{p_d(\mathbf{x}_\mathcal{I}|H_k) : \mathcal{I} \subseteq \{1, 2, \cdots, \ell\}\} \tag{29}$$

where $\mathbf{x}_\mathcal{I}$ *denotes the vector containing the positions of sensors with indexes in set* $\mathcal{I}$*, as in* (8)*, along with the values of the parameter controlling the sleep and wake up schedule,* p_s*, and the probability of a successful transmission,* p_t*, contain the necessary statistical information to implement a Neyman-Pearson or a Bayes test at the fusion center. The conditional probability density functions defining the likelihood test are given by*

$$f_{\mathbf{Y}|\mathbf{X},H}(\mathbf{y}|\mathbf{x}, H_k) = \sum_{\mathcal{I}\subseteq\mathcal{A}_0} (-1)^{|\mathcal{I}|} p_d(\mathbf{x}_{\mathcal{A}_1\cup\mathcal{I}}|H_k) \tag{30}$$

where operator $|\cdot|$ *denotes cardinality of a set and*

$$\mathcal{A}_k = \{i \ : \ y_i = k\}, \ 1 \leq i \leq \ell, \ k \in \{0, 1\} \tag{31}$$

Proof. To demonstrate this proposition it is only necessary to remark the dependence of the likelihood ratio on p_s, p_t and conditional distributions $f_{\mathbf{Y}|\mathbf{X},H}(\mathbf{y}|\mathbf{x}, H_k)$. The expression for these conditional distributions from joint conditional probabilities of detection is equivalent to the expression of Theorem 1 in [13], using a slightly different notation (probabilities of detection are replaced in [13] by expectations over the product of local decisions). The same demonstration by induction can be replicated. □

As an example of the application of this proposition, for a system with $\ell = 3$ local detectors, all the necessary statistical information is conveyed by p_s, p_t and the following set of joint probabilities of detection

$$\begin{aligned}\mathcal{P}_d = \{&p_d(\boldsymbol{x}_1|H_k),\ p_d(\boldsymbol{x}_2|H_k),\ p_d(\boldsymbol{x}_3|H_k),\ p_d([\boldsymbol{x}_1, \boldsymbol{x}_2]|H_k),\\ &p_d([\boldsymbol{x}_1, \boldsymbol{x}_3]|H_k),\ p_d([\boldsymbol{x}_2, \boldsymbol{x}_3]|H_k),\ p_d([\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3]|H_k)\}\end{aligned} \tag{32}$$

for $k \in \{0, 1\}$.

A simple iterative procedure to compute the conditional distributions from joint conditional probabilities of detection is also proposed, as an alternative to the evaluation of the general expression given in the proposition. In equation (30), the calculation of conditional distributions for some different values of $\mathbf{y}$ share in many cases several of the elements in the addition. This iterative procedure allows to avoid repeating the addition of these common terms by choosing an appropriate order in the computation of conditional distributions. The procedure is based on the fact that conditional distribution for $f_{\mathbf{Y}|\boldsymbol{X},H}(\mathbf{y}|\mathbf{x},,H_k)$, where $\mathbf{y}$ has n positive detections (n ones in $\mathbf{y}$) can be written as the joint probability of detection for the n sensors with a positive detection, minus the addition of the conditional distributions $f_{\mathbf{Y}|\boldsymbol{X},H}(\mathbf{y}'|\mathbf{x},,H_k)$ for every $\mathbf{y}'$ sharing the n ones with $\mathbf{y}$ (i.e., considering all the $2^{\ell-n}$ possible combinations of zeros and ones at positions in $\mathbf{y}'$ where there are zeros at $\mathbf{y}$). The set $\mathcal{I}^\ell$ will denote the following set of ℓ ordered indexes

$$\mathcal{I}^\ell = \{1, 2, \cdots, \ell\} \tag{33}$$

Using this notation, the conditional joint probability density funcion $f_{\mathbf{X},\mathbf{Y}|H}(\mathbf{x},\mathbf{y}|H_k)$ for the case of all sensors providing a positive local decision is given by the corresponding conditional joint probability of detection of all sensors

$$f_{\mathbf{Y}|\boldsymbol{X},H}(\mathbf{1}_\ell|\mathbf{x},,H_k) = p_d(\mathbf{x}|H_k) = p_d(\mathbf{x}_{\mathcal{I}^\ell}|H_k) \tag{34}$$

The distribution for all sensors with positive local detection except the last sensor is obtained simply by

$$f_{\mathbf{Y}|\boldsymbol{X},H}([\mathbf{1}_{\ell-1},0]|\mathbf{x},H_k) = p_d(\mathbf{x}_{\mathcal{I}^{\ell-1}}|H_k) - p_d(\mathbf{x}|H_k). \tag{35}$$

It is straightforward to obtain the distribution for any $\mathbf{y}$ with $\ell-1$ positive local decisions and a single negative local decision just by reordering the set of sensor indexes.

Now, the distribution for all sensors with positive local detection except the last two sensors is obtained as

$$\begin{aligned} f_{\mathbf{Y}|\boldsymbol{X},H}([\mathbf{1}_{\ell-2},0,0]|\mathbf{x},H_k) &= p_d(\mathbf{x}_{\mathcal{I}^{\ell-2}}|H_k) - f_{\mathbf{Y}|\boldsymbol{X},H}([\mathbf{1}_{\ell-2},1,0]|\mathbf{x},H_k) \\ &- f_{\mathbf{Y}|\boldsymbol{X},H}([\mathbf{1}_{\ell-2},0,1]|\mathbf{x},H_k) - f_{\mathbf{Y}|\boldsymbol{X},H}([\mathbf{1}_{\ell-2},1,1]|\mathbf{x},H_k) \end{aligned} \tag{36}$$

The last three terms in the right hand side of (36) can be written in terms of the joint probabilities of detection, as it has been shown above. But the advantage of this iterative procedure is that now it is not necessary to compute again the additions leading to these terms.

It is straightforward to obtain the distribution for any situation with $\ell - 2$ positive local decisions and two negative local decisions just by reordering the set of sensor indexes. This procedure can be repeated to add a new negative local decision in each step to obtain all the joint conditional distributions.

5. Examples in a network with two sensors

This section will present some examples in a simple network with two sensors. In this case, the necessary statistical information will be p_s, p_t and

$$\mathcal{P}_d = \{p_d(\boldsymbol{x}_1|H_k),\ p_d(\boldsymbol{x}_2|H_k),\ p_d(\boldsymbol{x}_1, \boldsymbol{x}_2|H_k)\} \tag{37}$$

To quantify dependence, the correlation coefficient will be used, which for two random variables is defined as

$$\rho = \frac{\mathrm{cov}(Y_1, Y_2)}{\sqrt{\mathrm{var}(Y_1)\ \mathrm{var}(Y_2)}} \tag{38}$$

where $\mathrm{var}(\cdot)$ and $\mathrm{cov}(\cdot, \cdot)$ denote variance of a random variable and covariance between two random variables, respectively. By its own definition, this coefficient takes values in the range

$$-1 \leq \rho \leq +1 \tag{39}$$

it is null if Y_1 and Y_2 are independent, and it takes its maximum value $|\rho| = 1$ if Y_1 and Y_2 are linearly dependent. For binary random variables, this means that $\rho = 1$ for the maximum dependence given by $Y_1 = Y_2$. Taking into account that in this case Y_1 and Y_2 are binary variables, under hypothesis H_k this coefficient is given in terms of the set of joint probabilities of detection by

$$\rho_k = \frac{p_d(\boldsymbol{x}_1, \boldsymbol{x}_2|H_k) - p_d(\boldsymbol{x}_1|H_k)p_d(\boldsymbol{x}_2|H_k)}{\sqrt{p_d(\boldsymbol{x}_1|H_k)(1 - p_d(\boldsymbol{x}_1|H_k))p_d(\boldsymbol{x}_2|H_k)(1 - p_d(\boldsymbol{x}_2|H_k))}} \tag{40}$$

The following notation will be used in this section to simplify references to conditional distributions for local decisions

$$F^{(k)}_{y_1,y_2} \equiv f_{\mathbf{Y}|\boldsymbol{X},H}([y_1, y_2]|\mathbf{x}, H_k) \tag{41}$$

In this network with two sensors

$$F^{(k)}_{1,1} = p_d(\boldsymbol{x}_1, \boldsymbol{x}_2|H_k) \tag{42}$$

$$F^{(k)}_{1,0} = p_d(\boldsymbol{x}_1|H_k) - p_d(\boldsymbol{x}_1, \boldsymbol{x}_2|H_k) \tag{43}$$

$$F^{(k)}_{0,1} = p_d(\boldsymbol{x}_2|H_k) - p_d(\boldsymbol{x}_1, \boldsymbol{x}_2|H_k) \tag{44}$$

$$F^{(k)}_{0,0} = 1 + p_d(\boldsymbol{x}_1, \boldsymbol{x}_2|H_k) - p_d(\boldsymbol{x}_1|H_k) - p_d(\boldsymbol{x}_2|H_k) \tag{45}$$

To compare performance of different network configurations, the receiver operating characteristic (ROC) will be obtained and the area under the ROC (AUROC), often referred to simply as area under the curve (AUC), will be computed. ROC is the curve plotting the probability of detection of a classifier, p_D, as a function of its probability of false alarm, p_{FA} [5]. Different values of threshold decision λ in (11) will provide different values for p_{FA} and p_D, and therefore, different points in the ROC curve. The ROC curve defines the set of possible operating points (p_{FA}, p_D) that can be obtained modifying λ. For discrete observations, modification of λ will produce a discrete set of points, because the likelihood ratio $\Lambda_\Theta(\theta)$ has a limited number of possible values. Modification of statistical properties of randomization variable ξ allows a linear transition between these points, which makes the ROC to be piecewise linear. ROC curve has long been used to analyze the performance of a classifier. A detailed description of its properties can be found in [33].

AUC, the measure of the area under the ROC curve, is commonly used as a figure of merit to compare classifiers [34], because it has some important statistical properties (see [33] for details).

For the scenario with binary censoring, taking into account that the observation can take one of the four possible values in (13), and therefore threshold λ can be in five different relative postions with respect to these points, the ROC will be given by five points: the extreme points $(0,0)$ and $(1,1)$, plus three additional points.

In a conventional detection network, it is reasonable to assume that

$$\Lambda_\Theta([\,]) < \Lambda_\Theta([\boldsymbol{x}_1]),\ \Lambda_\Theta([\,]) < \Lambda_\Theta([\boldsymbol{x}_2]) \tag{46}$$

and

$$\Lambda_\Theta([\boldsymbol{x}_1]) < \Lambda_\Theta([\boldsymbol{x}_1,\boldsymbol{x}_2]),\ \Lambda_\Theta([\boldsymbol{x}_2]) < \Lambda_\Theta([\boldsymbol{x}_1,\boldsymbol{x}_2]) \tag{47}$$

because a higher number of received transmissions (associated with possitive local decision, $y_i = 1$) will generally indicate a higher probability of hypothesis H_1. Situation $\Lambda_\Theta([\boldsymbol{x}_1]) < \Lambda_\Theta([\boldsymbol{x}_2])$ or $\Lambda_\Theta([\boldsymbol{x}_1]) > \Lambda_\Theta([\boldsymbol{x}_2])$ will depend on the relative detection performance of sensors at $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$. Under these assumptions, using the generic notation $(r_n^{(0)}, r_n^{(1)})$, $n \in \{1,2,3\}$ for points defining the ROC, these points will be given by

$$r_1^{(k)} = p^2 F_{1,1}^{(k)} \tag{48}$$

$$r_2^{(k)} = \begin{cases} p\left(F_{1,0}^{(k)} + F_{1,1}^{(k)}\right), & \text{if } \Lambda_\Theta([\boldsymbol{x}_1]) > \Lambda_\Theta([\boldsymbol{x}_2]) \\ p\left(F_{0,1}^{(k)} + F_{1,1}^{(k)}\right), & \text{if } \Lambda_\Theta([\boldsymbol{x}_1]) < \Lambda_\Theta([\boldsymbol{x}_2]) \end{cases} \tag{49}$$

$$r_3^{(k)} = p\left(F_{0,1}^{(k)} + F_{1,0}^{(k)} + 2F_{1,1}^{(k)}\right) - p^2 F_{1,1}^{(k)} \tag{50}$$

where $p = p_s \times p_t$ denotes the joint probability of being "awake" and having a successful transmission. Taking into account the possible relative locations of threshold λ in (11) with respect to the four possible values of $\Lambda_\Theta(\theta)$, these expressions can be easily obtained from probabilities associated to each possible observation under both hypothesis, which are given by (14)–(17). Without the reasonable constraints (46) and (47), the expressions for these points will change, although it is straightforward to obtain them given relative values of $\Lambda_\Theta(\theta)$ for all values of θ.

The ROC for the uncensored scenario can be obtained in a similar way. The ROC for the scenario where all sensors obtain their local decision and successfuly transmit it to the fusion center (scenario analyzed in [13]) will also be computed as a reference, because this scenario provides the best performance that can be achieved by the binary networks considered in this paper. It is convenient to remark that in this scenario, as all local decisions are available at the fusion center, the likelihood ratio will be given by

$$\Lambda_{\Theta}(\theta) = \frac{f_{\mathbf{\Theta}|H}(\theta|H_1)}{f_{\mathbf{\Theta}|H}(\theta|H_0)} = \frac{f_{\mathbf{Y}|\mathbf{X},H}([y_1, y_2]|\mathbf{x}, H_1)}{f_{\mathbf{Y}|\mathbf{X},H}([y_1, y_2]|\mathbf{x}, H_0)} \tag{51}$$

Therefore, in this section three scenarios will be compared

- Scenario analyzed in [13] without censoring and sleep and wake up schedule, and with a perfect transmission with the fusion center. Label "Ideal" will be used to identify this scenario.
- Scenario with binary censoring, sleep and wake up schedule, defined by parameter p_s, and possible failures in the transmission to the fusion center, modeled by parameter p_t defining the probability of a successful transmission. Label "Censored" will be used to identify this scenario.
- Scenario without binary censoring, but with sleep and wake up schedule and possible errors in the transmission to the fusion center, modeled by parameters p_s and p_t, respectively. Label "Uncensored" will be used to identify this scenario.

It is interesting to remark that these 3 scenarios are equivalent under condition

$$p_s = p_t = 1 \tag{52}$$

When all sensors are "awake" and successfully transmit to the fusion center, all the information (local decisions of all sensors) is available at the fusion center: this fact is evident in the "Uncensored" scenario, and in the "Censored" scenario sensors that have transmitted have a positive local decision and sensors that have not transmitted have a negative local decision.

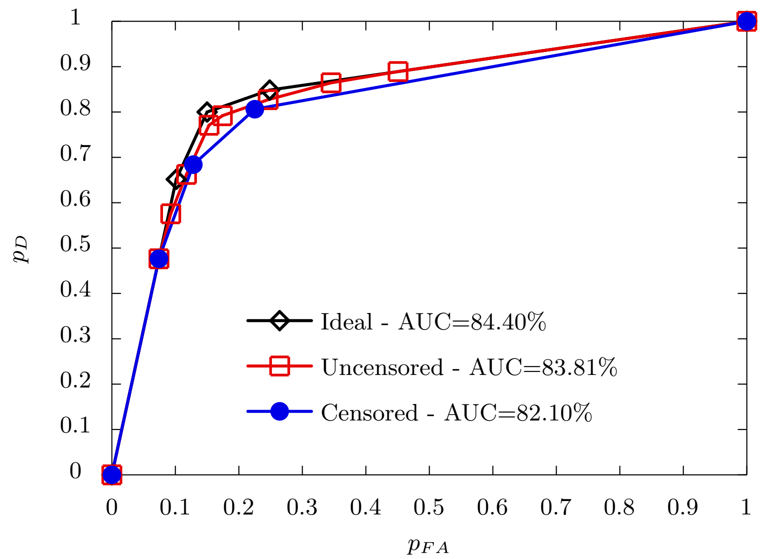

Figure 1: ROC and AUC for Example 1 with $p_s = 0.9$ and $p_t = 0.95$.

5.1. Example 1 - ROC and AUC as a function of p_s

In the first example, the following set of parameters has been used

$$p_d(\boldsymbol{x}_1|H_1) = 0.7,\ p_d(\boldsymbol{x}_2|H_1) = 0.8$$

$$p_d(\boldsymbol{x}_1|H_0) = 0.2,\ p_d(\boldsymbol{x}_2|H_0) = 0.15$$

$$\rho_1 = \rho_0 = \frac{1}{2}$$

Note that both hypotheses have the same degree of correlation. Note also that $p_d(\boldsymbol{x}_1|H_k)$ and $p_d(\boldsymbol{x}_2|H_k)$ along with ρ_k determine the value for $p_d(\boldsymbol{x}_1, \boldsymbol{x}_2|H_k)$ through (40). Figs. 1, 2 and 3 plot the ROC and the AUC (in %), when $p_t = 0.95$ for different values of p_s, 0.9, 0.75 and 0.5, respectively.

It can be seen how the ROC of "Censored" and "Uncensored" systems approach the "Ideal" ROC as p_s increases, and is progresively worse as p_s decreases. In fact, performance depends on the product of p_s and p_t, as it can be seen in Fig. 4, where the AUC is plotted as a funciton of $p_s \times p_t$. As this product goes to 0, both "Censored" and "Uncensored" ROC converge to the dummy detector with AUC equal to 50% (the one with ROC going through a

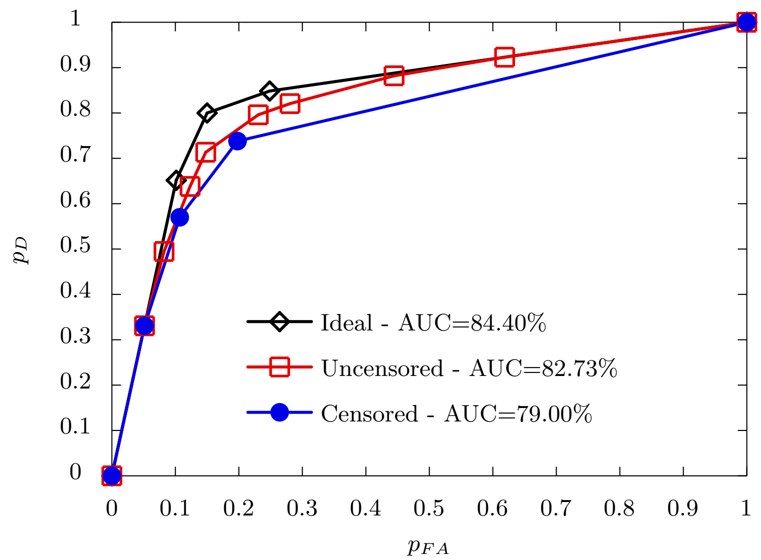

Figure 2: ROC and AUC for Example 1 with $p_s = 0.75$ and $p_t = 0.95$.

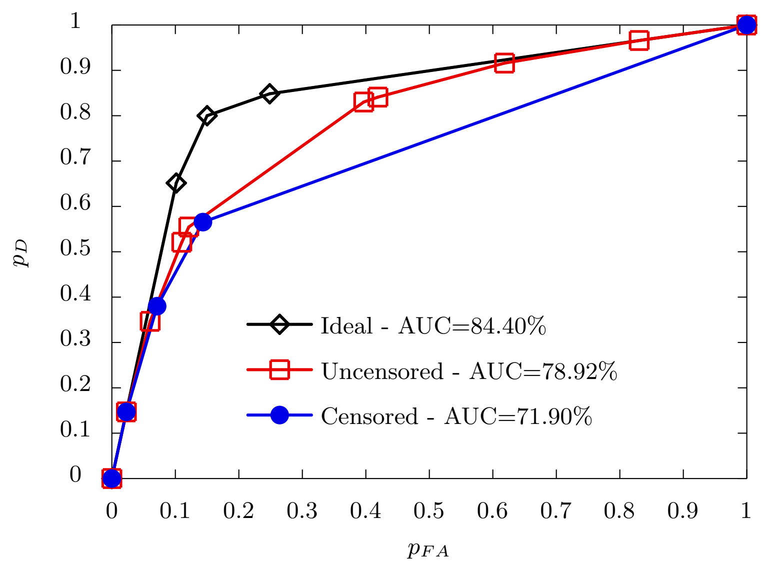

Figure 3: ROC and AUC for Example 1 with $p_s = 0.5$ and $p_t = 0.95$.

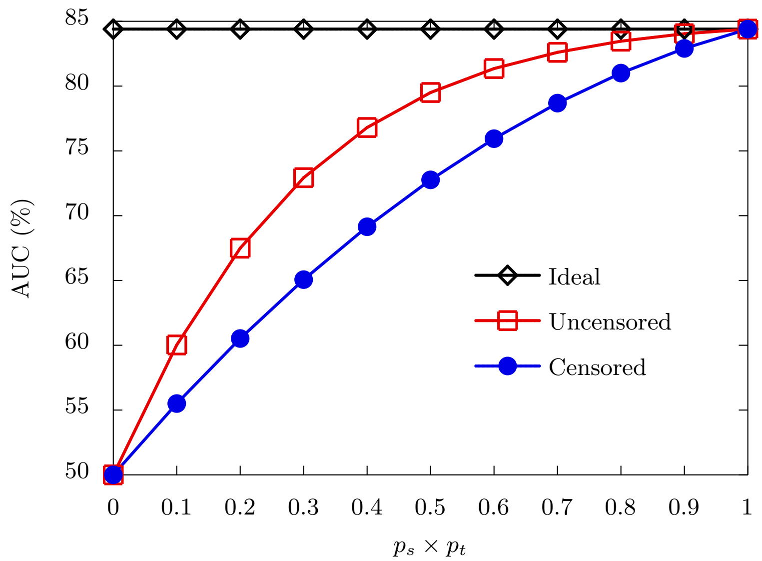

Figure 4: AUC as a function of the product $p_s \times p_t$ in Example 1.

straight line from $(0,0)$ to $(1,1)$ in the p_{FA} vs p_D plot). Obviously, the fusion center can not make useful decisions when it does not receive information about local decisions of the sensors. If the product goes to 1, both systems converge to the optimal solution, as it was previously discussed.

5.2. Example 2

In this second example, the effect of correlation in performance will be assessed. The following set of parameters has been used for different values of the correlation coefficient.

$$p_d(\boldsymbol{x}_1|H_1) = p_d(\boldsymbol{x}_2|H_1) = 0.75$$

$$p_d(\boldsymbol{x}_1|H_0) = p_d(\boldsymbol{x}_2|H_0) = 0.3$$

$$p_s = 0.9, \; p_t = 0.95$$

In this example $p_d(\boldsymbol{x}_1|H_k) = p_d(\boldsymbol{x}_2|H_k)$ to allow the correlation coefficient to range from 0 to 1 (the case $\rho_k = 1$, corresponding to the $y_1 = y_2$ situation, can only be obtained with this constraint in the probabilities of detection or false alarm for individual detectors).

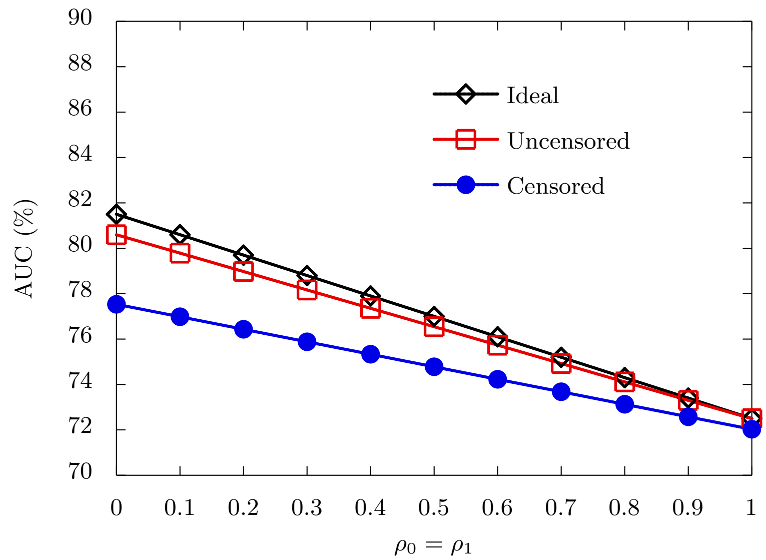

Figure 5: AUC as a function of the correlation coefficient for dependence under both hypothesis.

Fig. 5 plots AUC as a function of the correlation coefficient when this parameter is the same under both hypothesis, i.e., $\rho_0 = \rho_1$. In this case, AUC has a linear decrease as correlation coefficient increases for all methods. As local decisions become more correlated, each individual decision contains a lower information than in the case of conditional independent local decisions, which makes AUC decrease. Moreover, difference in AUC value decreases with respect to the ideal system as correlation increases.

Fig. 6 plots AUC for the case when dependence is only present under hypothesis H_1, while under H_0 local decisions are independent. In this case, both the “Ideal” and “Uncensored” systems exhibit the same behavior: AUC decreases as ρ_1 increases up to a given point; from this point, a higher correlation provides a better value for AUC. Again, as measures become more correlated under H_1, initially each measure has lower information than in the independent case, thus making AUC to decrease. However, as measures achieve a certain degree of correlation under H_1 but there are still independent under H_0, the own correlation vs not correlation of the observations under H_1 and H_0, respectively,

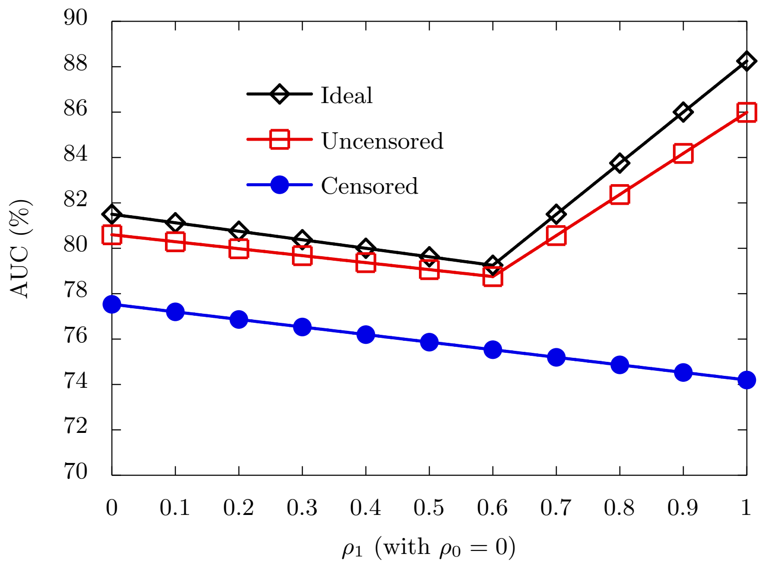

Figure 6: AUC as a function of the correlation coefficient with dependence only under H_1.

can be helpful to discriminate between both hypothesis. This behavior has also been seen in other kind of detection problems, for instance in the fusion of dependent local decision in tandem networks [35]. However, in this example the "Censored" system does not seem to have this behavior: AUC decreases linearly as correlation index increases, and there is not a point where slope changes sign.

Analyzing this case with more detail, it was observed that the point where the slope changes sign (and therefore the trend with correlation changes) moved accordingly to the quality of individual sensors. As this quality is improved ($p_d(\boldsymbol{x}_i|H_k)$ increases for H_1 and decreases for H_0) the point moved right, to higher values of correlation. When quality was lower ($p_d(\boldsymbol{x}_i|H_k)$ decreases for H_1 and increases for H_0), the point moved to the left, to lower values of correlation. This effect will be illustrated by using a couple of sensors that are "worse" than the ones in the initial setup, in particular

$$p_d(\boldsymbol{x}_1|H_1) = p_d(\boldsymbol{x}_2|H_1) = 0.7$$

$$p_d(\boldsymbol{x}_1|H_0) = p_d(\boldsymbol{x}_2|H_0) = 0.35$$

Fig. 7 plots the evolution of AUC as a function of ρ_1 using these sensors. It can

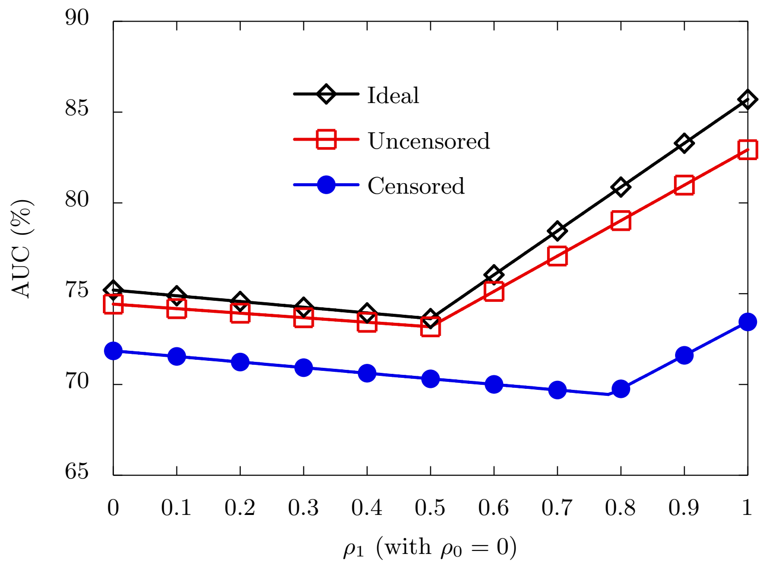

Figure 7: AUC as a function of the correlation coefficient with dependence only under H_1 for slightly "worse" individual sensors.

be seen that now all methods have basically the same kind of behavior, although a higher level of correlation is required to change the trend in the "Censored" method. As sensors with "better" individual performance are used, the value of correlation at which the trend of AUC changes moves to higher values, and at a given point there is no change in the negative slope for the "'Censored" method, as it happens in Fig. 6.

Finally, Fig. 8 plots the evolution of AUC when correlation is present only under H_0. The initial setup has been considered to obtain the picture. In this case, all methods exhibit the same kind of behavior, and the point where the trend with correlation changes is the same for the three methods.

6. Discussion

This paper has analyzed a distributed detection problem in a network with binary sensors with conditional dependence in local decisions. The necessary statistical information for optimal decisions has been identified. Expressions for the conditional distributions of observations have been obtained when a lo-

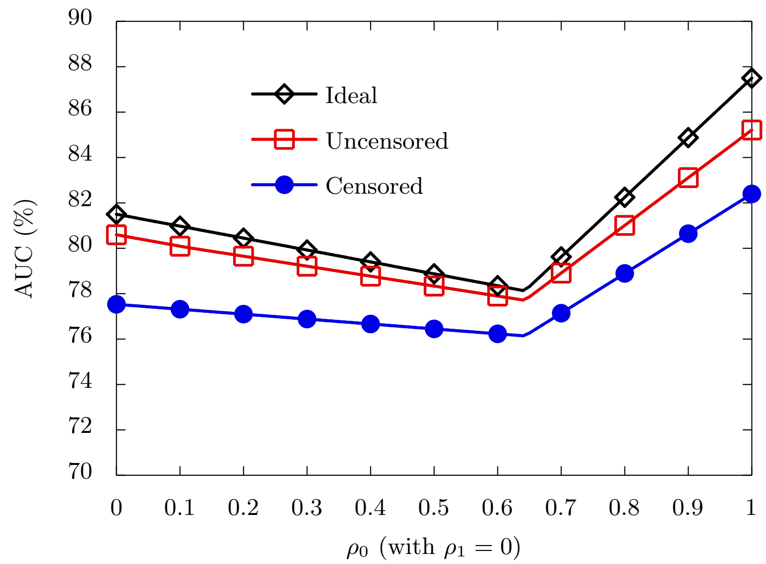

Figure 8: AUC as a function of the correlation coefficient with dependence only under H_0.

cal sleep and wake up schelude is used in combination with binary censoring, or without binary censoring. The effect of failed transmissions has been also considered. Performance of the proposed method has been analyzed in a simple network with two sensors by using the area under the ROC as a figure of merit. These experiments illustrate the different trends with correlation appearing when correlation is present just under one hypothesis or when it is present under both hypothesis.

Binary censoring and sleep and wake up schedules can clearly help to reduce the energy consumption in wireless sensor networks, thus increasing network lifetime. Obviously, there is a tradeoff between energy consumption and performance. Using large values for parameter p_s, which defines the local sleep and wake up random schedule, performance can be made to be close to the optimal perfomance without censoring. Lower values of p_s will save more energy at the price of a lower performance. The amount of energy that can be saved by using binary censoring combined with a local sleep and wake up schedule, as it is proposed here, will be highly dependent on the probability of hypothesis H_1, because this is related with the number of positive local detections (and the corresponding transmissions), and on the specific sensor characteristics: efficiency of sleep and wake up schedule, relative energy consumption due to sensing, processing local decision, and wireless transmission, etc. Therefore, design decision for parameter p_s will be strongly dependent on these characteristics and on the required performance. When required performance is time variant, it is also possible to dynamically modify parameter p_s to accomodate to the current performance requirements. The formulation used in this manuscript allows the designer to analyze the performance/lifetime tradeoff in practical binary networks an to select the optimal configuration taking into account the underlying constraints of the specific application, and the characteristics of the sensors that are used to implement the network.

One of the problems that appears with dependence in distributed detection problems is that in many cases it is difficult to know the statistical model of the underlying measures, which will define the optimal detection rule. In the case of binary sensors, the problem can be alleviatted, because the necessary statistical information is included in the joint probabilities of detection between different subsets of the sensors. These probabilities can be properly modeled in many practical scenarios if it is possible to have some initial calibration stage where probabilities of detection could be estimated from real measures. This can be feasible for instance in an application for indoor detection where the network can be deployed and some test procedures can be established to estimate the probabilities of detection.

Although this is beyond the scope of this paper, an interesting question can be how the fusion center can learn the statistical dependence of local decisions from received data. Introducing an adaptive mechanism to update the estimations of joint probabilities of detection can be helpful in nonstationary environments.

A final comment about p_s and p_t. It can be seen that in all expressions corresponding to conditional distributions of observations, the effect of these parameters is clearly coupled through their product. This is because both parameters produce the same effect: the lack at the fusion center of some data.

The reason to keep them separated in notation is to make explicit that this lack of information is produced by two very different causes: one is the inherent probability of failures in wireless transmissions, which is unavoidable in practice but undesirable; the other one is intentional, as p_s is a design parameter allowing to establish a tradeoff between performance and energy consumption.

References

References

[1] D. Puccinelli, M. Haenggi, Wireless sensor networks: applications and challenges of ubiquitous sensing, IEEE Circuits and Systems Magazine 5 (3) (2005) 19–29.

[2] C. Y. Chong, S. Kumar, Sensor networks: Evolution, opportunities, and challenges, Proceedings of the IEEE 91 (8) (2003) 1247–1256.

[3] H. L. Van Trees, Detection, estimation, and modulation Theory: Part I, John Wiley and Sons, New York, 1968.

[4] H. V. Poor, An introduction to signal detection and estimation, 2nd Edition, Springer Text in Electrical Engineering, Springer Verlag, New York, 1994.

[5] S. M. Kay, Fundamentals of statistical signal processing: detection theory, Prentice-Hall, Inc., Upper Saddle River, New Jersey, 1998.

[6] P. K. Varshney, Distributed detection and data fusion, Springer-Verlag, New York, NY, 1997.

[7] R. Viswanathan, P. K. Varshney, Distributed detection with multiple sensors: Part I - fundamentals, Proceedings of the IEEE 85 (1) (1997) 54–63.

[8] R. S. Blum, S. A. Kassam, H. V. Poor, Distributed detection with multiple sensors: Part II - advanced topics, Proceedings of the IEEE 85 (1) (1997) 64–79.

[9] J. F. Chamberland, V. V. Veeravalli, Decentralized detection in sensor networks, IEEE Transactions on Signal Processing 51 (2) (2003) 407–416.

[10] J. F. Chamberland, V. V. Veeravalli, Asymptotic results for decentralized detection in power constrained wireless sensor networks, IEEE Journal on Selected Areas in Communications 22 (6) (2004) 1007–1015.

[11] S. A. Aldosari, J. M. F. Moura, Detection in sensor networks: The saddlepoint approximation, IEEE Transactions on Signal Processing 55 (1) (2007) 327–340.

[12] V. Aalo, R. Viswanathan, On distributed detection with correlated sensors: two examples, IEEE Transactions on Aerospace and Electronic Systems 25 (3) (1989) 414–421.

[13] E. Drakopoulos, C. C. Lee, Optimum multisensor fusion of correlated local decisions, IEEE Transactions on Aerospace and Electronic Systems 27 (4) (1991) 593–606.

[14] V. Aalo, R. Viswanathan, Asymptotic performance of a distributed detection system in correlated Gaussian noise, IEEE Transactions on Signal Processing 40 (1) (1992) 211–213.

[15] M. Kam, Q. Zhu, W. S. Gray, Optimal data fusion of correlated local decisions in multiple sensor detection systems, IEEE Transactions on Aerospace and Electronic Systems 28 (3) (1992) 916–920.

[16] J. F. Chamberland, V. V. Veeravalli, How dense should a sensor network be for detection with correlated observations?, IEEE Transactions on Information Theory 52 (11) (2006) 5099–5106.

[17] Y. Sung, H. V. Poor, H. Yu, How much information can one get from a wireless ad hoc sensot network over a correlated random field?, IEEE Transactions on Information Theory 55 (6) (2009) 2827–2847.

[18] J. N. Tsitsiklis, M. Athans, On the complexity of decentralized decision making and decentralized decision, IEEE Transactions on Automatic Control 30 (5) (1985) 440–446.

[19] P. Willet, P. F. Swazek, R. S. Blum, The good, bad, and ugly: Distributed detection of a known signal in dependent Gaussian noise, IEEE Transactions on Signal Processing 48 (12) (2000) 3266–3279.

[20] J. C. Dagher, M. W. Marcellin, M. A. Neifeld, A theory for maximizing the lifetime of sensor networks, IEEE Transactions on Communications 55 (2) (2007) 323–332.

[21] V. Raghunathan, C. Schurgers, S. Park, M. B. Srivastava, Energy-aware wireless microsensor networks, IEEE Signal Proccesing Magazine 19 (2) (2002) 40–50.

[22] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, A survey on sensor networks, IEEE Communication Magazine 40 (8) (2002) 102–114.

[23] S. Bhattacharya, G. Xing, C. Lu, G. C. Roman, B. Harris, O. Chipara, Dynamic wake-up and topology maintenance protocols with spatiotemporal guarantees, in: Proceedings of the 4th International Conference on Information Processing in Sensor Networks, Los Angeles, CA, 2005, pp. 28–34.

[24] Q. Cao, T. Abdelzaher, T. He, J. Stankovic, Towards optimal sleep scheduling in sensor networks for rare-event detection, in: Proceedings of the 4th International Conference on Information Processing in Sensor Networks, Los Angeles, CA, 2005, pp. 20–27.

[25] L. Doherty, B. A. Warneke, B. Boser, K. S. J. Pister, Energy and performance considerations for Smart Dust, International Journal of Parallel and Distributed Sensor Networks 4 (3) (2001) 121–133.

[26] C. Rago, P. Willett, Y. Bar-Shalom, Censoring sensors: a low communication-rate scheme for distributed detection, IEEE Transactions on Aerospace and Electronic Systems 32 (2) (1996) 554–568.

[27] P. Willett, L. Tong, One aspect to cross-layer design in sensor networks, in: Proceedings of the MILCOM, Vol. 1, Monterrey, CA, 2004, pp. 59–64.

[28] S. Appadwedula, V. V. Veeravalli, D. L. Jones, Energy-efficient detection in sensor networks, IEEE Journal on Selected Areas in Communications 23 (4) (2005) 693–702.

[29] S. Marano, V. Matta, P. Willett, L. Tong, Cross-layer design of sequential detectors in sensor networks, IEEE Transactions on Signal Processing 54 (11) (2006) 4105–4117.

[30] X. Nguyen, M. J. Wainwright, M. I. Jordan, Nonparametric decentralized detection using kernel methods, IEEE Transactions on Signal Processing 53 (11) (2005) 4053–4066.

[31] M. Lázaro, A. Artés-Rodríguez, M. Sánchez-Fernández, Decentralized detection in binary dense sensor networks: to transmit or not to transmit, in: Proceedings of the International Workshop on Wireless Ad-hoc Networks (IWWAN), London, UK, 2005.

[32] M. Lázaro, M. Sánchez-Fernández, A. Artés-Rodríguez, Optimal sensor selection in binary heterogeneous sensor networks, IEEE Transactions on Signal Processing 57 (4) (2009) 1577–1587.

[33] T. Fawcett, An introduction to ROC analysis, Pattern Recognition Letters 27 (8) (2006) 861–874.

[34] A. P. Bradley, The use of the area under the ROC curve in the evaluation of machine learning algorithms, Pattern Recognition 30 (7) (1997) 1145–1159.

[35] J. Plata, M. Lázaro, Closed-form error exponent for the Neyman-Pearson fusion of dependent local decisions in a one-dimensional sensor network, IEEE Transactions on Signal Processing 59 (3) (2010) 1239–1254.