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 consumption, 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 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 \( \rho_s \) sensors per area unit over an area \( \mathcal{D} \subseteq \mathbb{R}^2 \). 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 \( \{ (x_i, y_i) : i = 1, \ldots, l, x_i \in \mathcal{D}, y_i \in \{0, 1\} \} \), when each pair \((x_i, y_i)\) represents a successful reading of a sensor located at coordinates \( x_i \), 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 \( x \) when a target is present at coordinates \( x^t \) is denoted as \( p_d(x^t, x, \alpha) \), where \( \alpha \) is the probability of false alarm (PFA) of the sensor when no target is present. In other words, \( p_d(x^t, x, \alpha) = Pr(Y = 1 | X = x^t, X = x) \) when the PFA of the detector is equal to \( \alpha \). \( p_d(x^t, x, \alpha) \) has the following properties:

1. \( p_d(x^t, x, \alpha) \geq \alpha \)
2. \( p_d(x^t, x, \alpha) = p_d(\|x^t - x\|_2^2, \alpha) \)
3. \( p_d(x^t, x, \alpha) \geq p_a(x^t, x') \Leftrightarrow \|x^t - x\|_2 \leq \|x^t - x'\|_2 \)
4. \( \lim_{\|x^t - x\|_2 \to \infty} p_d(x^t, x, \alpha) = \alpha \)

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

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

where $\rho = \int_D dx \rho x$ and $\delta$ is the Kronecker function.

• Under hypothesis $H_1$, the joint pdf of $X$ and $Y$ is

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

We assume that samples in $\{(x_i, y_i) : i = 1, \ldots, l, x_i \in \mathcal{D}, y_i \in \{0, 1\}\}$ 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:

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

where $r_0$ is the range of the sensor. This model considers a constant probability of misdetection 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 $\{\{(x_i, y_i) : i = 1, \ldots, l, x_i \in \mathcal{D}, y_i \in \{0, 1\}\}\}$, the log-likelihood ratio between both hypothesis can be computed as

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

where

$$\Gamma_i = \frac{p_d(x^i, x_i, \alpha) \delta[y_i - 1] + (1 - p_d(x^i, x_i, \alpha)) \delta[y_i]}{\alpha \delta[y_i - 1] + (1 - \alpha) \delta[y_i]}$$

$$= \begin{cases} \frac{1 - p_d(x^i, x_i, \alpha)}{1 - \alpha} & \text{if } y_i = 0 \\ \frac{p_d(x^i, x_i, \alpha)}{\alpha} & \text{if } y_i = 1 \end{cases}$$

Under Bayes criteria, the threshold $\tau$ 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 $\alpha$. If $y_1 = y_2 = 0$, $\lambda < 0$ independently of the values of $x_1$, $x_2$, and $\alpha$. If $y_1 = 1$ and $y_2 = 0$ (or viceversa) $\lambda$ can be less or greater than 0 depending on the values of $x_1$, $x_2$, and $\alpha$. In this case

$$\lambda = \ln \frac{p_d(x^i, x_1, \alpha)}{\alpha} + \ln \frac{1 - p_d(x^i, x_2, \alpha)}{1 - \alpha} \geq 0$$

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 $\lambda$ 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, $\lambda$, tends to a normal random variable. We will first determine the means and variances of $\lambda$ under the two hypothesis, and then determine the threshold and power of the NP test under gaussian statistics.

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

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

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

$$P_1 = \int_D \rho \ln p_d(x^i, x, \alpha) \, dx$$

$$P_0 = \int_D \rho \ln (1 - p_d(x^i, x, \alpha)) \, dx$$

Similarly,

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

and

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

$$- (1 - \alpha)(1 - Pr(Y = 1|H_1)) - \ln \rho$$

where $h$ is the differential entropy. The variances of $\gamma_{H_0}$ and $\gamma_{H_1}$ 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 $\tau$ of the NP test of level $\alpha_i$ is, according to the above

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

and its power, $\beta_i$, is given by

$$\beta_i = Q \left( \frac{l D(H_1 || H_0) - \tau}{l (E(\gamma_{H_1}^2) - D^2(H_1 || H_0))} \right)$$
where $Q$ is the Marquardt’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, let assume that the area $D \subset \mathbb{R}^d$ is a circle of radius $R$ centered on $x'$. In that case, $D(H_0||H_1)$ is

$$D(H_0||H_1) = \begin{cases} \frac{\pi}{4} \Big( \alpha \ln \alpha - (1 - \alpha) \ln \frac{1}{\alpha} \Big) & \text{if } R \geq r_0 \\ \frac{\pi}{4} \Big( \alpha \ln \alpha - (1 - \alpha) \ln \frac{1}{\alpha} \Big) & \text{if } R < r_0 \end{cases}$$

(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 $D_{ah}$ de KL divergence under the “spanish hat” model, for any function $p_d(x', x, \alpha^s)$ that induces a divergence $D^*(H_0||H_1)$, we can obtain a $D_{ah}(H_1||H_0) = D^*(H_1||H_0)$ by setting the parameters of the “spanish hat” model as follows: $\alpha = \alpha^s$, and solving $\beta$ and $\rho_0$ for equating $P_0$ and $P_1$ under the two $p_d$ functions.

Using an equivalent procedure, different values of $\beta$ and $\rho_0$ can be obtained by setting $D_{ah}(H_1||H_0) = D^*(H_1||H_0)$, 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 $\gamma$. Alternatively, instead of using multiple “spanish hat” functions for bounding any function of moments of $\gamma$, 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 small 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 $\epsilon$ 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 $\alpha_n \in (0, 1)$

$$\lim_{l \to \infty} \frac{1}{l} \ln \beta_0 = D(H_0||H_1)$$

a function that has been analyzed in the preceding section.

In Bayes tests (assuming that $C_{i0} = C_{0i} = C_{01} - C_{11}$), the best achievable error exponent is the Chernoff information, $C(f_X, Y|H_1, f_X, Y|H_0)$ or $C(H_1, H_0)$ for short, defined as

$$C(f_X, Y|H_1, f_X, Y|H_0) = D(f_X, Y|H_0) - D(f_X, Y|H_1) = D(f_X, Y|H_0) - D(f_X, Y|H_1)$$

(2)

where

$$f_X, Y|H_0(x, y|H_0) = \frac{f_X, Y|H_0(x, y|H_1) f_Y^{|-s} H_0(x, y|H_0)}{\sum f_X, Y|H_0(x', y|H_1) f_Y^{|-s} H_0(x', y|H_0) dx'}$$

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.,

$$C(H_1, H_0) = -\min_{0 \leq s \leq 1} \mu_{\lambda_0}(s)$$

that in our problem takes the form

$$\mu_{\lambda_0}(s) = l \ln \left[ \int_D \rho \left( \frac{p_d(x', x, \alpha)^s}{\alpha^{s-1}} + \frac{1 - p_d(x', x, \alpha)^s}{(1 - \alpha)^{s-1}} \right) dx \right]$$

Using the “spanish hat” model and a circular $D$ as before, the cgf is

$$\mu_{\lambda_0}(s) = l \left( \ln \left( \frac{\beta^s}{\alpha^s} + \frac{\beta^s}{(1 - \alpha)^{s-1}} \right) + \ln \left( \frac{\beta}{\alpha^{s-1}} + \frac{\beta}{(1 - \alpha)^{s-1}} \right) \right)$$

if $R \geq \rho_0$

(3)

That achieves its minimum respect to $s$ when

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

(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, $\epsilon^1$. 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 $\epsilon$, $l_0$, is

$$l_0 \geq \frac{\ln \epsilon}{D}$$

(5)

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 $\epsilon$, $l_0$, is

$$l_0 \geq \frac{\ln \epsilon}{D}$$

(5)

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

1. Given the sensors, are there any optimum configuration of the exploration area $D'$?

2. Given $D$, and supposing that we can tune the sensors for different values of $\alpha$ following the ROC curve of the sensor, what is the optimum value of $\alpha$?

3. Are the solutions of the above questions independents?

\footnote{For the NP tests, $\epsilon$ is the power of the test, and for Bayes tests, $\epsilon$ 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

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

and the first conclusion is that the exploration area, $D$, 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 $r_e = \sqrt{\frac{\int (p_a(s) - p_a(t)) \, ds}{\int p_a(s) \, ds}}$, 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 $R$.](image)

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 $\alpha$ and $(1-\beta)$,

$$L_0(\alpha, 1-\beta) = \alpha \ln \frac{\alpha}{1-\beta} + (1-\alpha) \ln \frac{1-\alpha}{\beta}$$

(7)

Obviously, $L_0(\alpha, 1-\beta)$ can be infinite if $\alpha$ or $\beta$ (or both) are zero, but we can only achieve the combinations $\alpha$ and $\beta$ 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 $\alpha$, 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

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

(8)

where $s_0$ is as in (4).

For fixed values of $\alpha$, $\beta$ and $r_0$, (8) is a decreasing function of $R^2$ for $R \geq r_0$ and an increasing 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

$$C_b(\alpha, \beta) = \frac{(1-\alpha)^{\beta_0}}{\alpha^{\alpha_0-1}} + \frac{\beta}{\alpha}$$

(9)

that is a similar problem to finding the maximum of $L_0(\alpha, 1-\beta)$ 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 $\alpha$ and $\beta$ 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


