FULLY DISTRIBUTED SIGNAL DETECTION: APPLICATION TO COGNITIVE RADIO

Franck Iutzeler and Philippe Ciblat

Telecom ParisTech, 46 rue Barrault 75013 Paris, France
email: firstname.lastname@telecom-paristech.fr

ABSTRACT

Cooperative detection without fusion center has many applications such as spectrum sensing in cognitive radio or intrusion detection in ad hoc network. In this paper, we propose a new asynchronous fully-distributed cooperative algorithm which does not require any knowledge on the underlying nodes network. Discussion about the threshold choice and the respective duration of the sensing step and the gossip step is conducted.

***Index Terms*—** Cooperative detection, spectrum sensing, fully-distributed decision

1. INTRODUCTION

There are some applications where agents have to detect rapidly the presence or absence of a signal of interest. For instance, one can mention the spectrum sensing in cognitive radio or the intrusion detection in military mobile ad hoc networks. In order to make an accurate decision, these agents/nodes may have to cooperate with each other. The traditional way to fix this problem consists in providing hard or soft detection decisions to a fusion center. This centralized approach is clearly sensitive to fusion center failure. Moreover, in ad hoc networks context, fusion center election protocol and routing protocol have to be carried out which is costful in terms of overhead and time. Therefore, designing fully distributed decision algorithms is of great interest. Such algorithms rely on decision test function and decision threshold computed in a distributed way, *i.e.*, when only exchanges of local data with neighbor are allowed.

Cooperative detection has recently received a lot of attention (see [1] and references therein). Nevertheless, most works assume the existence of a fusion center and finally focus on the design of operations done at each node in order to help the fusion center to make the right decision. In the literature, only a few algorithms are fully distributed in the sense defined above [2, 3]. At each iteration of these algorithms, a spectrum sensing step is operated followed by a gossip step. These algorithms are well adapted to time-varying environments but they suffer from threshold distributed computation. Indeed, in [2], the threshold is chosen in an asymptotic regime and performance (especially, false alarm probability) are not ensured in finite time. In [3], the threshold is chosen assuming the absence of diffusion/gossiping step. Hence, threshold distributed computation remains an open issue.

We thus propose a new fully distributed signal decision algorithm based on a recently-developed gossiping algorithm [6] where sensing steps are followed by gossiping steps, and where the threshold is chosen adequately. In addition, thanks to the separation of both steps, we are able to optimizing their durations at the expense of less adaptivity compared to [2, 3].

This paper is organized as follows: in Section 2, we introduce the signal model. In Section 3, we remind some results of centralized cooperative detection. In Section 4, we propose a new fully-distributed cooperative detection algorithm. Threshold distributed computation is discussed, and ROC curves are derived. In Section 5, numerical results confirm our claims.

2. SYSTEM MODEL

We consider a network of K nodes collaborating to detect the presence or absence of a signal. The received signal at time n on node k writes $y_k(n)$. We assume that the duration of sensing is the same for all nodes and equal to N_s. Let $\mathbf{y}_k = [y_k(1) \cdots y_k(N_s)]^{\mathrm{T}}$ be the sensing data associated with node k, and where the superscript $(.)^{\mathrm{T}}$ stands for the transposition operator. The signal to (potentially) detect is denoted by $\mathbf{x}_k = [x_k(1) \cdots x_k(N_s)]^{\mathrm{T}}$ at node k where $x_k(n)$ corresponds to its value at time n. Finally, an additive noise can disturbed the detection and is denoted by $\mathbf{n}_k = [n_k(1) \cdots n_k(N_s)]^{\mathrm{T}}$ at node k where $n_k(n)$ corresponds to its value at time n. Let $\mathcal{N}(\boldsymbol{m}, \boldsymbol{\Sigma})$ be a Gaussian vector with mean $\boldsymbol{m}$ and covariance matrix $\boldsymbol{\Sigma}$. $\mathbf{n}_k$ is a i.i.d. Gaussian vector of distribution $\mathcal{N}(0, \sigma_k^2 \mathbf{I}_{N_s})$ where $\mathbf{I}_{N_s}$ stands for the identity matrix of size N_s. Throughout the paper, we assume that the statistics of $\mathbf{x}_k$ and $\mathbf{n}_k$ are known at node k.

The hypothesis test dealing with our problem can thus be written as follows

$$\left\{ \begin{array}{ll} \mathcal{H}_0 : & \forall k, \;\; \mathbf{y}_k = \mathbf{n}_k \\ \mathcal{H}_1 : & \forall k, \;\; \mathbf{y}_k = \mathbf{x}_k + \mathbf{n}_k \end{array} \right. \tag{1}$$

In the context of cognitive radio (which is our main application of interest), we prefer to consider hypothesis test tar-

The work has been partially funded by DGA (French Defense Agency).

geting a fixed probability of detection and so a variable false alarm probability that we wish to minimize rather than the standard Neyman-Pearson approach. Indeed, in this context, the secondary users (the nodes in our framework) should not disturb the primary user (the signal to detect in our framework) up to a pre-defined probability. Moreover, a high false alarm probability only implies that the secondary users do not use the white spaces while they could. Therefore we would like to minimize this false alarm probability.

According to the approach developed in [4], one can prove that our optimal test (minimizing the false alarm probability given a target probability of detection) still boils down to the so-called Likelihood Ratio Test (LRT)

$$\Lambda(\mathbf{y}) := \log\left(\frac{p(\mathbf{y}|\mathcal{H}_1)}{p(\mathbf{y}|\mathcal{H}_0)}\right) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda \tag{2}$$

where $p(\mathbf{y}|\mathcal{H})$ is the probability density of $\mathbf{y}$ given the tested hypothesis $\mathcal{H}$ and where λ is chosen such that the target probability of detection, denoted by P_D^{target}, is ensured.

3. REVIEW ON CENTRALIZED COOPERATIVE SPECTRUM SENSING

Before going further, we remind some important results about centralized cooperative spectrum sensing. We focus, on the one hand, on an energy-based detector (when the sought signal is unknown) and, on the other hand, on a training-based detector (when the sought signal is known and thus corresponds to a training sequence [3]).

3.1. Energy-based detector

When the sought signal is unknown, it is usual to assume $\mathbf{x}_k$ is an zero-mean i.i.d. Gaussian vector with covariance matrix $\gamma_k^2\mathbf{I}_{N_s}$. Then the Signal-to-Noise Ratio (SNR) at node k is equal to $\text{SNR}_k := \gamma_k^2/\sigma_k^2$ and is assumed to be known at node k. Assuming independence of the received signals at different nodes (this assumption is reasonable since even if the same signal is transmitted by the primary user, the random wireless channel leads to independent received signals between nodes), the test given in Eq. (2) and achieved at the fusion center can be decomposed as follows: $\Lambda(\mathbf{y}) = \sum_{k=1}^K \Lambda_k(\mathbf{y}_k)$ with $\Lambda_k(\mathbf{y}_k) = \log(p(\mathbf{y}_k|\mathcal{H}_1)/p(\mathbf{y}_k|\mathcal{H}_0))$. As $\mathbf{x}_k \sim \mathcal{N}(\mathbf{0}, \gamma_k^2\mathbf{I}_{N_s})$ and $\mathbf{n}_k \sim \mathcal{N}(\mathbf{0}, \sigma_k^2\mathbf{I}_{N_s})$, we obtain the following test by removing the constant terms

$$T(\mathbf{y}) := \frac{1}{K}\sum_{k=1}^{K} \frac{\|\mathbf{y}_k\|_2^2}{\gamma_k^2+\sigma_k^2}\text{SNR}_k \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta \tag{3}$$

where η must be chosen such as $\mathbb{P}(T(\mathbf{y}) > \eta|\mathcal{H}_1) = P_D^{\text{target}}$.

In order to compute the threshold η, we need to exhibit the probability density of T. Unfortunately, due to unequal SNRs, T is not χ^2-distributed. In [7], it is advocated that the density of T can be approximated with a Gamma distribution, denoted by $\Gamma(\kappa, \theta)$, whose the probability density function is equal to $g_{\kappa,\theta}$ defined by

$$g_{\kappa,\theta}(x) = \frac{1}{\Gamma(\kappa)\theta^\kappa} x^{\kappa-1}\mathrm{e}^{-x/\theta}, \quad x \geq 0, \tag{4}$$

and 0 otherwise. The terms κ and θ are chosen in order to match the mean and the variance of T. In the following, we denote its cumulative distributive function by $G_{\kappa,\theta}$ and its inverse by $G_{\kappa,\theta}^{(-1)}$. After some algebraic manipulations, we obtain that $T \sim \Gamma(\kappa_T, \theta_T)$ with

$$\kappa_T = \frac{KN_s}{2} \cdot \frac{(\frac{1}{K}\sum_{k=1}^K \text{SNR}_k)^2}{\frac{1}{K}\sum_{k=1}^K \text{SNR}_k^2} \tag{5}$$

and

$$\theta_T = \frac{2}{K} \cdot \frac{\frac{1}{K}\sum_{k=1}^K \text{SNR}_k^2}{\frac{1}{K}\sum_{k=1}^K \text{SNR}_k}. \tag{6}$$

One can then deduce that the optimal threshold given the target probability of detection is

$$\eta = G_{\kappa_T,\theta_T}^{(-1)}(1 - P_D^{\text{target}}).$$

3.2. Training-based detector

We now assume that each node k has the knowledge of the (possible) transmit signal $\mathbf{x}_k$. Typically, the signal $\mathbf{x}_k$ may decomposed as $h_k\mathbf{x}$ where h_k corresponds to the (known) channel fading between the node k and the sought transmitter and $\mathbf{x}$ is a training sequence [3]. Here, the signal power is $\gamma_k^2 = \|\mathbf{x}_k\|^2/N_s$.

Then the test given in Eq. (2) takes the following form

$$T(\mathbf{y}) := \frac{1}{K}\sum_{k=1}^{K} \frac{\mathbf{y}_k^{\mathrm{T}}\mathbf{x}_k}{\sigma_k^2} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta \tag{7}$$

As $\mathbf{x}_k$ is deterministic, T is Gaussian-distributed with mean m_T and variance ς_T^2 given by

$$m_T = N_s\left(\frac{1}{K}\sum_{k=1}^K \text{SNR}_k\right) \text{ and } \varsigma_T^2 = \frac{N_s}{K}\left(\frac{1}{K}\sum_{k=1}^K \text{SNR}_k\right).$$

As a consequence, the threshold is obtained as follows

$$\eta = \varsigma_T Q^{(-1)}\left(P_D^{\text{target}}\right) + m_T$$

where $Q^{(-1)}$ is the inverse of the Gaussian tail function.

4. FULLY DISTRIBUTED ALGORITHM

The purpose of this paper is to perform detection in a distributed way, i.e., without fusion center. Obviously, tests described in Eqs. (3)-(7) are not computable since a node may not have the contribution of the others. To overcome this problem, we propose to introduce a gossiping step in order to compute the involved averages. Prior to this gossiping step, the nodes perform a sensing step so that each node k can provide the term

$$t_k(\mathbf{y}_k) = \begin{cases} \|\mathbf{y}_k\|_2^2 \mathrm{SNR}_k/(\gamma_k^2+\sigma_k^2) & \text{if energy detector} \\ \mathbf{y}_k^{\mathrm{T}}\mathbf{x}_k/\sigma_k^2 & \text{if training detector.} \end{cases}$$

Let N_g be the duration of the gossiping step and $T = N_s + N_g$ be the total duration of the processing. Most gossiping algorithms for averaging can take the following form

$$\begin{bmatrix} T_1(\mathbf{y}) \\ \vdots \\ T_K(\mathbf{y}) \end{bmatrix} = \mathbf{P} \begin{bmatrix} t_1(\mathbf{y}_1) \\ \vdots \\ t_K(\mathbf{y}_K) \end{bmatrix}$$

where $T_k(\mathbf{y})$ is the final test function at node k, and where $\mathbf{P} = (p_{k\ell})_{k,\ell=1,\cdots,K}$ corresponds to the considered gossiping algorithm matrix after N_g iterations (see [5] for more details).

4.1. Energy-based detector

When an energy-based detection is carried out, the final test function at node k is

$$T_k(\mathbf{y}) = \sum_{\ell=1}^{K} p_{k\ell} \frac{\|\mathbf{y}_\ell\|_2^2}{\gamma_\ell^2+\sigma_\ell^2} \mathrm{SNR}_\ell \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_k,$$

where η_k is the threshold at node k.

Actually, the main issue in this paper is to find a distributive way for selecting a good threshold at any time, *i.e.*, ensuring the target probability of detection P_D^{target} as close as possible at any step of the algorithm[1].

Once again, by assuming that T_k is well approximated by a Gamma distribution, we obtain that

$$\eta_k = G^{(-1)}_{\kappa_k,\theta_k}\left(1 - P_D(k)^{\mathrm{target}}\right)$$

where $P_D(k)^{\mathrm{target}}$ is the target probability of detection associated with node k, and

$$\kappa_k = \frac{N_s}{2} \cdot \frac{\left(\sum_{\ell=1}^K p_{k\ell}\mathrm{SNR}_\ell\right)^2}{\sum_{\ell=1}^K p_{k\ell}^2 \mathrm{SNR}_\ell^2} \tag{8}$$

and

$$\theta_k = 2 \cdot \frac{\sum_{\ell=1}^K p_{k\ell}^2 \mathrm{SNR}_\ell^2}{\sum_{\ell=1}^K p_{k\ell}\mathrm{SNR}_\ell}. \tag{9}$$

Then, we obtain the Receiver Operating Characteristic (ROC) curve is equal to

$$P_{FA}(k) = 1 - G_{\kappa'_k,\theta'_k}\left(G^{(-1)}_{\kappa_k,\theta_k}\left(1 - P_D(k)\right)\right) \tag{10}$$

[1] Actually, the decision is made before the convergence of the gossip algorithm. Indeed, assuming a primary user is present, T_k could be above the threshold whereas the gossip has not still converged to the consensus.

with

$$\kappa'_k = \frac{N_s}{2} \cdot \frac{\left(\sum_{\ell=1}^K p_{k\ell}\frac{\mathrm{SNR}_\ell}{1+\mathrm{SNR}_\ell}\right)^2}{\sum_{\ell=1}^K p_{k\ell}^2 \left(\frac{\mathrm{SNR}_\ell}{1+\mathrm{SNR}_\ell}\right)^2}$$

and

$$\theta'_k = 2 \cdot \frac{\sum_{\ell=1}^K p_{k\ell}^2 \left(\frac{\mathrm{SNR}_\ell}{1+\mathrm{SNR}_\ell}\right)^2}{\sum_{\ell=1}^K p_{k\ell}\frac{\mathrm{SNR}_\ell}{1+\mathrm{SNR}_\ell}}.$$

Unfortunately, the terms involving $p_{k\ell}^2$ in Eqs. (8)-(9) prevent to obtain the threshold η_k in a distributed way at node k for ensuring the probability of detection $P_D(k)^{\mathrm{target}}$. To overcome this issue, we propose hereafter two approaches.

Approach 1: distributed with knowledge of K. Actually, on the centralized scheme, the threshold depends on the average of the SNR and the square SNR through Eqs.(5)-(6). A simple idea is to replace these exact averages with the averages obtained thanks to the considered gossip algorithm. It is clear that if N_g is large enough, the obtained thresholds will be close to those of the centralized case and also to those described in Eqs. (8)-(9) since $p_{k,\ell}$ is close to $1/K$ and so $p_{k\ell}^2$ can be well approximated by $p_{k\ell}/K$. As a consequence, the new threshold is

$$\eta_k^{(1)} = G^{(-1)}_{\kappa_k^{(1)},\theta_k^{(1)}}\left(1 - P_D(k)^{\mathrm{target}}\right)$$

with

$$\kappa_k^{(1)} = \frac{KN_s}{2} \cdot \frac{\left(\sum_{\ell=1}^K p_{k\ell}\mathrm{SNR}_\ell\right)^2}{\sum_{\ell=1}^K p_{k\ell}\mathrm{SNR}_\ell^2}$$

and

$$\theta_k^{(1)} = \frac{2}{K} \cdot \frac{\sum_{\ell=1}^K p_{k\ell}\mathrm{SNR}_\ell^2}{\sum_{\ell=1}^K p_{k\ell}\mathrm{SNR}_\ell}.$$

This algorithm is still not fully distributed since the knowledge of the number of nodes is required. Furthermore, the target probability of detection is not ensured since the real probability of detection, denoted by $P_D(k)^{(1)}$, is given by

$$P_D(k)^{(1)} = 1 - G_{\kappa_k,\theta_k}(G^{(-1)}_{\kappa_k^{(1)},\theta_k^{(1)}}(1 - P_D(k)^{\mathrm{target}})).$$

In contrast, we prove that the ROC curve is the following one

$$P_{FA}(k)^{(1)} = 1 - G_{\kappa'_k,\theta'_k}(G^{(-1)}_{\kappa_k,\theta_k}(1 - P_D(k)^{(1)}))$$

which is the same as in Eq. (10). Consequently, the ROC curve is not degraded due to our approximate threshold, but depends on the gossip algorithm. In addition, the operating point in the ROC curve can not fixed a priori.

Approach 2: fully distributed. In this approach, the knowledge of the number of nodes will not be required anymore. Recently, new gossip algorithms, based on the sum-weight principle (see [6] and references therein) have been introduced in order to perform fast estimation of the average and the sum. Let $\mathbf{v} = [v_1, \cdots v_K]^{\mathrm{T}}$ be the vector whose component v_k is the value of the node k before gossiping (*e.g.*, SNR_k or SNR_k^2 for threshold computation, or t_k for test function computation). For computing simultaneously the average and the sum of $\mathbf{z}$, these algorithms rely on the three following variables, given by,

$$\mathbf{z} := \mathbf{Q}\mathbf{v},\ \mathbf{w}^{(1)} := \mathbf{Q}\mathbf{1},\ \mathbf{w}^{(e)} := \mathbf{Q}\mathbf{e}$$

where the matrix $\mathbf{Q}$ represents the gossip algorithm after N_g iterations (see [6] for more details), $\mathbf{1}$ is the K-sized vector whose elements are 1, $\mathbf{e}$ is the K-sized vector[2] whose the first component is equal to 1 and the others 0. Only the k-th component of all involved vectors is available at node k.

Then, each node k calculates the k-th component of $\mathbf{z}_p = \mathbf{z} \oslash \mathbf{w}^{(1)}$ and $\mathbf{z}_s = \mathbf{z} \oslash \mathbf{w}^{(e)}$ where $\oslash$ is the elementwise division. In [6], it is proven that $\mathbf{z}_p$ and $\mathbf{z}_s$ converge to the average and the sum of $\mathbf{v}$ respectively for large N_g. In addition, we remark that $\mathbf{z}_p = \mathbf{P}\mathbf{v}$ and $\mathbf{z}_s = \mathbf{S}\mathbf{v}$ with

$$\mathbf{P} = \mathrm{diag}\left(\mathbf{1} \oslash \mathbf{Q}\mathbf{1}\right)\mathbf{Q} \text{ and } \mathbf{S} = \mathrm{diag}\left(\mathbf{1} \oslash \mathbf{Q}\mathbf{e}\right)\mathbf{Q}. \quad (11)$$

Consequently, the final test function is computed with the gossip algorithm related to matrix $\mathbf{P}$ given in Eq. (11). The threshold is then obtained as follows

$$\eta_k^{(2)} = G^{(-1)}_{\kappa_k^{(2)}, \theta_k^{(2)}}\left(1 - P_D(k)^{\mathrm{target}}\right)$$

with

$$\kappa_k^{(2)} = \frac{N_s}{2} \cdot \frac{\left(\sum_{\ell=1}^{K} s_{k\ell}\mathrm{SNR}_\ell\right)^2}{\sum_{\ell=1}^{K} s_{k\ell}\mathrm{SNR}_\ell^2}$$

and

$$\theta_k^{(2)} = 2 \cdot \frac{\sum_{\ell=1}^{K} p_{k\ell}\mathrm{SNR}_\ell^2}{\sum_{\ell=1}^{K} s_{k\ell}\mathrm{SNR}_\ell}.$$

The algorithm is fully distributed since even the number of nodes is not required. Once again, the new threshold does not ensure the target probability of detection, and the ROC curve is still described by Eq. (10) but with $\mathbf{P}$ given by Eq. (11).

Finally, for both approaches, the ROC curves converge to the ROC curve related to the centralized case when N_g is large enough since the parameters κ_k, κ'_k, θ_k and θ'_k converge to those of the centralized case.

4.2. Training-based detector

When the training-based detector is implemented, the test function at node k after N_g gossiping iterations is

$$T_k(\mathbf{y}) = \sum_{\ell=1}^{K} p_{k\ell} \frac{\mathbf{y}_\ell^{\mathrm{T}}\mathbf{x}_\ell}{\sigma_\ell^2} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta_k,$$

[2]before gossipping, the second and third variables of each node must be initialized to match $\mathbf{1}$ and $\mathbf{e}$ respectively. For the second variable, each node is initialized to 1. For the third variable, only the first node is initialized to 1 whereas the others to 0. In cognitive radio context, the first node is the secondary user launching the sensing, *i.e.* wanting to access the medium.

As T_k is Gaussian distributed with mean m_k and variance ς_k^2, we have

$$\eta_k = \varsigma_k Q^{(-1)}\left(P_D^{\mathrm{target}}\right) + m_k$$

with

$$m_k = N_s \sum_{\ell=1}^{K} p_{k\ell}\mathrm{SNR}_\ell \quad \text{and} \quad \varsigma_k^2 = N_s \sum_{\ell=1}^{K} p_{k\ell}^2\mathrm{SNR}_\ell.$$

Once again, this algorithm can not be computed in a distributed way due to the presence of the terms $p_{k\ell}^2$ in the variance. To overcome this issue, previous proposed approaches can be applied straightforwardly.

5. NUMERICAL RESULTS

Except otherwise stated, the energy-based detector is carried out with $T = 128$, $N_s = N_g$ and $P_D^{\mathrm{target}} = 0.99$, and performance are averaged over random geographical graphs with $K = 10$ nodes. The SNRs at each node are exponentially-distributed with mean $\overline{\mathrm{SNR}}$. Only performance for the node exhibiting the smallest SNR realization will be plotted.

Hereafter, we always test the following algorithm configurations: i) the centralized one, ii) the pairwise gossip (PG) [5] with centralized threshold, iii) the pairwise gossip with approach 1 based threshold, iv) the broadcast sum-weight gossip (BWG) [6] with approach 2 based threshold.

In Fig. 1, we plot the ROC curve for the above-mentioned algorithm configurations. We remark that the ROC curves are

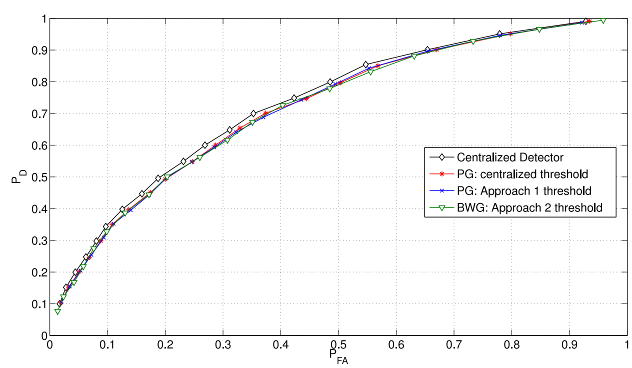

Fig. 1. ROC curve (P_D vs. P_{FA}).

very close to each other. In addition, when the same gossip algorithm is used, the ROC curve is identical regardless of the threshold technique computation.

In Figs. 2 and 3, we display empirical P_{FA} and P_D versus $\overline{\mathrm{SNR}}$ and N_s respectively So, the loss in false alarm probability for the fully-distributed approach is reasonable. Moreover its probability of detection is higher than the target one. We also remark that both steps (spectrum sensing and gossip) should have similar durations.

In Fig. 5, these algorithms have been evaluated when the hidden terminal practical configuration described in Fig. 4 has

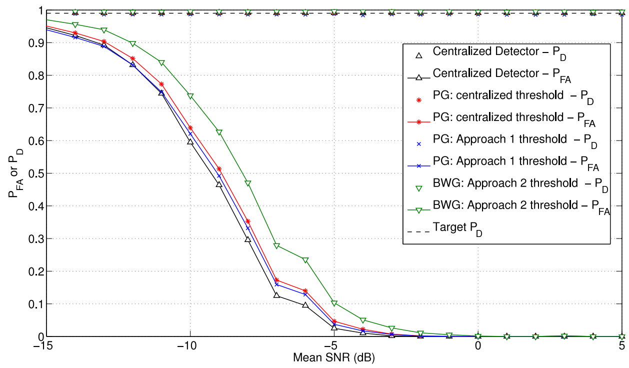

Fig. 2. P_{FA} and P_D versus $\overline{\text{SNR}}$.

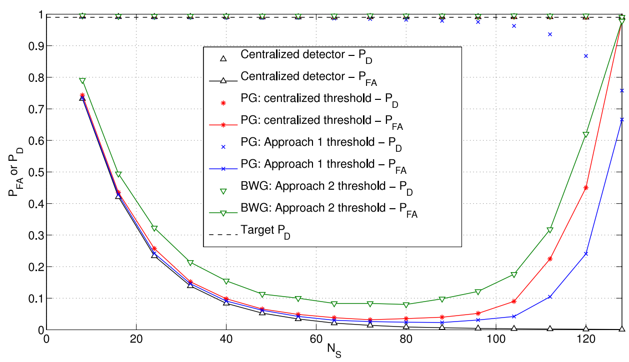

Fig. 3. P_{FA} and P_D versus N_s.

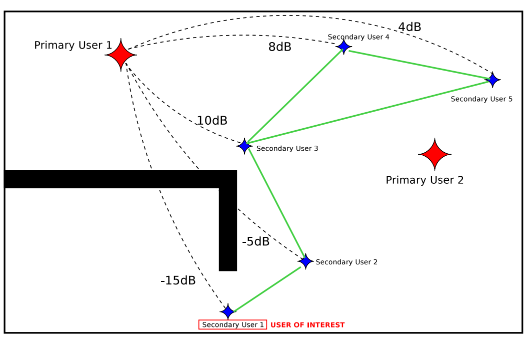

Fig. 4. Hidden terminal configuration

been simulated. Our proposed algorithms enable us to detect the hidden terminal quite quickly.

Finally, in Fig. 6, we compare our training-based algorithms to the diffusion LMS one described in [3]. Notice that N_s and N_g is incremented by 1 at each iteration of the diffusion LMS. We remark that our algorithms outperform the diffusion LMS. Actually, our block processing for the sensing step is much more efficient that the adaptive LMS one in [3]. Moreover, unlike diffusion LMS, our algorithms are asynchronous which simplifies the network management.

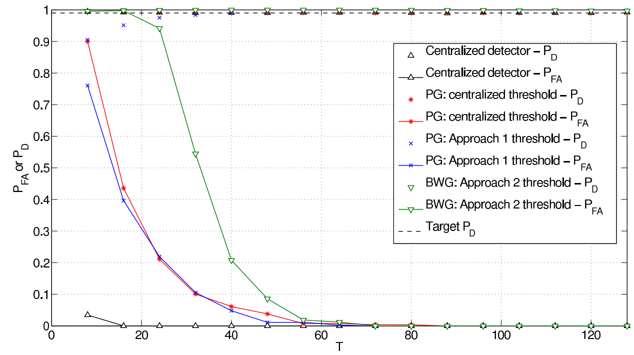

Fig. 5. P_{FA} and P_D versus T for the hidden terminal.

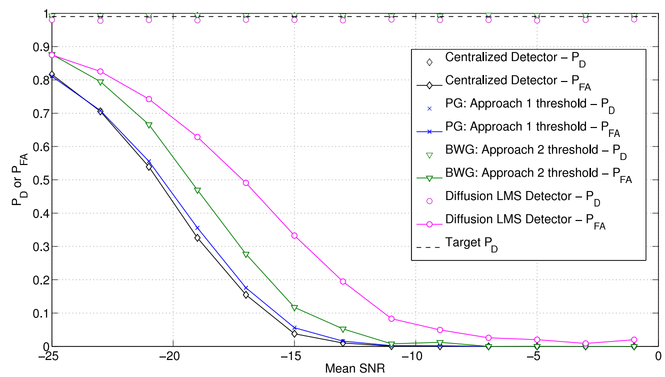

Fig. 6. P_{FA} and P_D versus $\overline{\text{SNR}}$ for proposed algorithms and diffusion LMS [3].

6. REFERENCES

[1] E. Axell, et al., *Spectrum sensing for cognitive radio*, IEEE Signal Processing Mag., vol. 29, no. 3, May 2012.

[2] P. Braca, et al., *Asymptotic optimality of running consensus in testing binary hypotheses*, IEEE Trans. Signal Processing, vol. 58, no. 2, Feb. 2010.

[3] F. Cattivelli, and A. Sayed, *Distributed detection over adaptive networks using diffusion adaptation*, IEEE Trans. on Signal Processing, vol. 59, no. 5, May 2011.

[4] H. Van Trees, *Detection, Estimation and Modulation Theory*, Wiley, 2004.

[5] S. Boyd, et al., *Randomized Gossip Algorithms*, IEEE Trans. on Information Theory, vol. 52, no. 6, June 2006.

[6] F. Iutzeler, et al., *Analysis of Sum-Weight-like algorithms for averaging in Wireless Sensor Networks*, accepted for publication at IEEE Trans. on Signal Processing. `http://arxiv.org/abs/1209.5912`

[7] P. Ciblat, et al., *Training Sequence Optimization for joint Channel and Frequency Offset estimation*, IEEE Trans. on Signal Processing, vol. 56, no. 8, Aug. 2008.