LOCATION-AIDED DISTRIBUTED PRIMARY USER IDENTIFICATION IN A COGNITIVE RADIO SCENARIO

Pavle Belanovic, Sergio Valcarcel Macua, and Santiago Zazo

Escuela Técnica Superior de Ingenieros de Telecomunicación
Universidad Politécnica de Madrid

ABSTRACT

We address a cognitive radio scenario, where a number of secondary users performs identification of which primary user, if any, is transmitting, in a distributed way and using limited location information. We propose two fully distributed algorithms: the first is a direct identification scheme, and in the other a distributed sub-optimal detection based on a simplified Neyman-Pearson energy detector precedes the identification scheme. Both algorithms are studied analytically in a realistic transmission scenario, and the advantage obtained by detection pre-processing is also verified via simulation. Finally, we give details of their fully distributed implementation via consensus averaging algorithms.

Index Terms— cognitive radio, distributed systems, wireless sensor networks, detection, consensus.

1. INTRODUCTION

In modern wireless networks, radio spectrum is a precious resource. Cognitive radio is one method of making ad-hoc use of unoccupied spectrum in order to increase the efficiency of its use. At the core of this approach lies the problem of detecting, and identifying, active primary users by a network of secondary users. We study the identification of which, if any, primary user is transmitting, by a network of secondary nodes without a fusion center and with only elementary location information. In a network of decision makers, distributed detection has been thoroughly studied and different solutions have been proposed. The problem is to decide what information the agents should share, and to find optimal fusion rules to combine the local outputs. Decentralized binary detection [1, 2, 3, 4, 5, 6] proposes a parallel architecture in which every node sends a summary of its own observations (e.g. quantized values, test outputs or hard decisions) to a fusion center in charge of making the final decision. Recently, completely distributed implementations, in which there is no fusion center so the nodes have to collaborate with each other to converge to the global solution [7], have also appeared; paying special attention to on-line algorithms in which nodes collaborate and detect in the same timescale [8, 9, 10].

The M-ary hypothesis testing, in particular with no prior knowledge of the probability distributions of the alternative hypotheses, has received much less attention. A number of decentralized approaches, which rely on a fusion center, have been proposed. For instance, [11] applies a blind algorithm after estimating the prior probabilities of the hypotheses; while in [12] the M-ary detection problem is converted into a sequence of binary detection problems. A fully distributed scheme based on belief propagation has been proposed in [13], but it requires knowledge of the prior probabilities in order to maximize the posterior distribution.

In this paper we introduce two fully distributed algorithms for transmission detection and primary user identification (M-ary hypothesis) when the only prior knowledge is that of the noise distribution. Nevertheless, we make the assumption that rudimentary location information is available: each secondary node knows its attenuation factor from each primary user. This assumption is reasonable in practical scenarios, because the nodes can easily learn the attenuations though calibration (indoor or outdoor, static only), fingerprinting (indoor, static or dynamic), GPS location and a propagation model (outdoor, static or dynamic), or any other method, all of which are beyond the scope of this paper.

2. PROBLEM DEFINITION

Let us assume a cognitive radio scenario, where $P$ primary users and $S$ secondary users share the same geographic area. Each primary user may transmit at any time (though we assume at most one primary user is transmitting at any moment) using a random "bursty" transmission. Each such transmission by a primary node $p$ is modeled as a signal $s_p$ which alternates between an active and a passive state, whose lengths are Poisson random variables, with parameters $\lambda q$ and $(1 - \lambda)q$, such that $\lambda$ is the activity factor and $q$ is the expected number of samples in each cycle. During the active state, $s_p \sim N(0, \sigma_p^2)$, and in the passive state $s_p = 0$. For each transmission, the primary node selects random $\sigma_p^2$ and $\lambda$, unknown to the secondary nodes.

The transmitted signal is then propagated to all the secondary users. Here we are not concerned with any propagation model in particular. Instead, we assume a static model of the received signal $x_s$ at the secondary node $s$ as an attenuation of the transmitted signal $s_p$ in AWGN, $x_s = \alpha_p s_p + n$, where $n \sim N(0, \sigma_n^2)$. We assume the realizations of $\alpha$ are iid at all the secondary nodes, and each node estimates $\sigma_n^2$ perfectly. Globally, the attenuation is given by a matrix $\mathbf{A} = [\alpha_p]_{p \times S}$, where each coefficient is assumed static and derived by any means, e.g. geometric model, measurements, fingerprinting technique, etc. We assume that each node $s$ has complete knowledge of its column of $\mathbf{A}$, but not of any other nodes’ attenuations. Other than this, no further location information is required by any node: neither its own location, nor of any other (primary orsecondary) node.

Given this model, we tackle the problem of identifying the trans-
mitting primary user, if any, by means of a distributed algorithm that does not rely on the availability of a fusion center serving the network of secondary nodes. In other words, the nodes must cooperatively decide among \(P + 1\) hypotheses \(\{\mathcal{H}_0, \mathcal{H}_1, \ldots, \mathcal{H}_P\}\), where \(\mathcal{H}_0\) represents no transmission from any primary. To this end, we propose two suitable algorithms, presented in the following sections.

3. IDENTIFICATION SCHEME

The first scheme we propose performs direct identification based on each node's performance energy sampling, where \(W\) integration windows, each of length \(L\) samples, produces an energy estimate \(y_s[w] = \frac{1}{L} \sum_{l=1}^{L} (x_s[l])^2\), with \(w \in \{1, 2, \ldots, W\}\). Using the knowledge of the noise statistic, we generate a new variable \(z_s = y_s - \sigma^2_s\), distributed \(^1\) as

\[
z_s \sim \begin{cases} \mathcal{N}(0, \frac{2\sigma^4_s}{L}) & \mathcal{H}_0 \\ \mathcal{N} \left( \frac{\beta_p \sigma^2_p + \sigma^4_s}{\alpha_{ps}^2}, \sigma^2_{\alpha_{ps}} \right) & \mathcal{H}_p \end{cases}
\]

where \(\sigma_{ps} = \sigma_p \sigma_s\) and \(p \in \{1, 2, \ldots, P\}\). Each node \(s\) then constructs \(P + 1\) hypotheses to test, by compensating its own received distribution of \(W\) samples of \(z_s\) exactly \(P + 1\) times. The first compensation represents \(\mathcal{H}_0\), i.e., the possibility that \(z_s\) contains only noise energy, and is constructed simply by using the raw data itself (no compensation). The following \(P\) compensations are performed by multiplying the received distribution by a compensation factor \(\beta_p \sigma^2_p, p \in \{1, 2, \ldots, P\}\), i.e., one compensated distribution for each possible primary node. The factor \(\beta_p\) serves to normalize each of the hypotheses, relative to \(\mathcal{H}_0\), so that later on their variances will be directly comparable. Hence, \(\beta_p = (\| \sigma_p \|)^{-1/2}\), where \(\sigma_p = [\sigma_{p1}, \sigma_{p2}, \ldots, \sigma_{pP}]^T\).

Therefore, assuming that the hypothesis \(\mathcal{H}_p\) is true (shown in bold), each node \(s\) has a set of compensated distributions

\[
\mathcal{N} \left( \frac{\beta_p \sigma^2_p + \sigma^4_s}{\alpha_{ps}^2}, \sigma^2_{\alpha_{ps}} \right) \quad \mathcal{H}_p
\]

Hence, there are \(S\) distributions for each hypothesis, one per node.

Another way of seeing this is that each node estimates, in isolation, the product of the only two parameters common to all, \(\sigma^2_L\), since \(E(\sigma_{ps}^2 | y_s, \sigma^2_p) = \sigma^2_L\) \(\forall s\), when primary \(p\) is transmitting. We note that for the correct hypothesis \(p\), though all the nodes agree perfectly in the mean, they do not in the variance, due to the different attenuations \(\alpha_{ps}\), which do not get canceled out.

Estimating which hypothesis is true, \(j\), across the \(S\) secondary nodes is the next challenge, and the first to use cooperation among the secondary nodes. An intuitive approach would be to choose the hypothesis with a maximum sum of distances among the \(S\) distributions. Remembering that in the correct hypothesis, the variance in all the nodes does not match, prevents us from using Bhattacharyya, Mahalanobis, or any distance metrics that take variance into consideration. In other words, we disregard the variance information in each node, and use only on the means. Hence, we opt for the Euclidean metric, such that

\[
j = \arg \min_j \sum_{m=1}^{S} \frac{|\mu_m|_{|j} - \mu_n|_{|j}|}{\sigma_m}
\]

It is of course easy to see that the function reaches its minimum, being 0, when \(j = p\) and is strictly greater otherwise. Unfortunately this problem formulation cannot be used in a distributed scenario because it requires the knowledge of the entire matrix \(A\).

However, it is also easy to show that this sum of distances is proportional to the sample variance of the set of \(S\) compensated means. Hence, the problem reduces to finding the hypothesis with minimum variance the \(S\) nodes. This problem is easily tackled in a distributed fashion using averaging consensus algorithms [14], following the idea of constructing the sample covariance matrix shown in [15]. Hence, our proposed algorithm is shown in Algorithm 1, and illustrated in Fig. 1, where the detection block is not active.

Algorithm 1 Identification algorithm at node \(s\)

1: INPUT \(x_s, \sigma_{ps}, y_s, L, W\)
2: \(y_s[w] \leftarrow L^{-1}x_s[w]\) (for each of the \(W\) elements of \(y_s\))
3: \(z_s \leftarrow y_s - \sigma^2_s\)
4: \(\mu_s \leftarrow W^{-1}y_s\) (sample mean of \(z_s\))
5: \(\beta_p = \frac{1}{\beta} \sum_{i=1}^{S} \alpha_{pi}^{-4} \quad \text{consensus loop}
6: \(m_s = [\mu_1, \mu_2, \ldots, \mu_S]^{T}\)
7: \(m_* = \frac{1}{S} \sum_{i=1}^{S} m_i \quad \text{consensus loop}
8: \(v_s \leftarrow \text{diag}(m_s - m_*) \times (m_s - m_*)^{T}\)
9: \(v_* \leftarrow \frac{1}{S} \sum_{i=1}^{S} v_i \quad \text{consensus loop}
10: j \leftarrow \arg \min_j v_j \quad \text{index of the minimum element}
11: \text{OUTPUT } j

4. DETECTION PRE-PROCESSING

In a low SNR regime the mode of failure of the scheme presented in Section 3, and indeed any identification scheme, is that of choosing \(\mathcal{H}_0\) even when a signal is present, simply because this signal is too weak to identify a particular primary node transmitting. Hence, it makes sense to perform an (optimal or nearly-optimal) detection step first, detecting the activity of any primary node, followed then by an identification procedure similar to that of Section 3, but this time with only \(P\), rather than \(P + 1\) hypotheses. This is illustrated in Fig. 1.

If each node \(s\) had the knowledge of all the parameters of its \(z_s\), the optimal test [16] based on the Neyman-Pearson criterion would be given by

\[
T(z_s) = \eta z_s^{T} + \theta z_s^{T} z_s
\]

Both \(\eta\) and \(\theta\) are functions of the statistics of \(z_s\), and since these are not known, the approach is not feasible. It is in principle possible to construct the generalized likelihood ratio test (GLRT) using the maximum likelihood (ML) estimates of \(\sigma^2_p\) and \(\lambda\). However, we found that in this cognitive radio scenario the GLRT performed poorly, since the estimates were very poor in the hypothesis \(\mathcal{H}_0\).

Therefore, we propose a sub-optimal approach \(T(z_s) = z_s^{T} z_s\).

\[
T(z_s) \sim \begin{cases} \mathcal{N} \left( \frac{2W \sigma^4_p}{L} \right), & \mathcal{H}_0 \\ \mathcal{N} \left( W \left( \sigma^4_p + \sigma^4_{ps} \lambda^2 \right), 2W \sigma^2_p + 2W \sigma^4_p \lambda^2 \right) & \mathcal{H}_p \end{cases}
\]
where \( \sigma_n^2 = \frac{1}{2}(\sigma_n^4 + \sigma_{ps}^4 + 2\sigma_{ns}^2\sigma_{ps}^2) \) for compactness. One obvious advantage is that this test does not depend on the estimates of \( \sigma_{ps}^2 \) and \( \lambda \).

The threshold is hence \( \gamma = \frac{2\sigma_n^4}{L} (\sqrt{2W}Q^{-1}(P_{fa}) + W) \), and is calculated by each node in isolation, for a defined probability of false alarm \( P_{fa} \), where \( Q^{-1}(\cdot) \) is the inverse Q-function. Typically each node would compare the local \( T(z_s) \leq \gamma \) producing a 1-bit detection decision, which are then combined globally (e.g. voting [17]).

Instead, we propose a weighted global test \( T_s = \frac{1}{S} \sum_{i=1}^{S} (T(z_s) - \gamma) \leq 0 \). Although the factor \( S^{-1} \) is quite unnecessary, it shows that this global value can also be derived in a distributed fashion via average consensus. If each node calculates its vote as a degree of confidence \( T(z_s) \leq \gamma \), simply the sign (+ or -) of the global average of the votes (available at all the nodes simultaneously) is the outcome of the global test. This weighing allows the nodes closer to any transmitting primary to exert a bigger influence, as desired.

Once the detection stage is performed in this distributed fashion, all the nodes can carry out the identification procedure (also distributed) as shown in Section 3, but this time with \( P \) rather than \( P + 1 \) hypotheses. This is shown in Algorithm 2 and illustrated in Fig. 1, where the detection block is active.

Algorithm 2: Identification algorithm with pre-detection at node \( s \)

1: **INPUT** \( x_s, \alpha_{ps} \forall p, \sigma_n, L, W, P_{fa} \)

2: \( \gamma \leftarrow \frac{2\sigma_n^4}{L} (\sqrt{2W}Q^{-1}(P_{fa}) + W) \) (NP threshold, local calc.)

3: \( T(z_s) \leftarrow z_s^Tz_s \) (simplified local test)

4: \( T_s \leftarrow \frac{1}{S} \sum_{i=1}^{S} (T(z_s) - \gamma) \) \( \leftarrow \) consensus loop

5: if \( T_s < 0 \) then

6: \( j = 0 \) (no transmission, \( \mathcal{H}_0 \))

7: else

8: \( m_s \rightarrow [\mu_1^s, \beta_1^s, \alpha_1^s, \ldots, \mu_P^s, \beta_P^s, \alpha_P^s]^T \)

9: end if

10: **OUTPUT** \( j \)

5. EXPERIMENTS

We verify, illustrate, and compare the functioning of the two proposed approaches via simulations. We define a simple scenario with four primary (\( P = 4 \)) and twenty secondary users (\( S = 20 \)) uniformly randomly located in a square area with sides of 200 m, choosing the primary users to be the most distant nodes. As discussed earlier, we assume iid zero-mean Gaussian noise at every secondary node, with the variance \( \sigma_n^2 \) perfectly estimated by the secondary users. Since every node has a different signal-to-noise ratio (SNR), depending on its attenuation \( \alpha_{ps} \), it is not possible to express the global results against SNR. Instead, we analyze the influence of \( \sigma_1 \) (keeping \( \sigma_n \) constant), or quite equivalently the transmitted-signal to receiver-noise ratio \( S/\eta N = 20 \log(\alpha_1/\sigma_n) \), which is a fictitious parameter. We ran \( 10^4 \) experiments for each value of \( S/\eta N, R \) in the range from 0 to 80 dB. In each experiment, we choose one of the \( P + 1 \) equally probable hypotheses. The identification scheme uses 100 integration windows (\( W = 100 \)) with two hundred samples each (\( L = 200 \)). The transmitter activity factor is 50\% (\( \lambda = 0.5 \)), with on average twenty samples per cycle (\( q = 20 \)).

In Fig. 2(a) we show the performance of both schemes in terms of a classic metric, the probability of detection. We see that at low to mid \( S/\eta N, R \) levels, the nearly-optimal detection pre-processing step brings about 10 dB of improvement. Both curves converge on the left to the value of the \( P_{fa} \), as usual, since both schemes fail in the same way. They are unable to separate the \( \mathcal{H}_0 \) and \( \mathcal{H}_p \) hypotheses, which at such low \( S/\eta N, R \) levels practically overlap completely.

On the other hand, in Fig. 2(b) we show the probability of identification, being the proportion of successful hypothesis identifications relative to the total number of experiments at that hypothesis, summed over all the hypotheses and normalized by \( (P + 1)^{-1} \). Again, we see a significant improvement won by the pre-detection step, of around 10 dB. As expected, both curves converge on the left to \( (P + 1)^{-1} \), which is 1/5 in this experiment.

6. CONCLUSIONS

In this work we studied the problem of identifying the active primary user by a distributed network of secondary nodes with very limited location information. We proposed two fully distributed algorithms, based on identification only, or identification with a detection pre-processing step. In the detection phase, we introduced a novel weighted global test, which allows the secondary nodes closer to the transmitter to exert a bigger influence. Both algorithms are implemented using averaging consensus to provide coordination among

![Fig. 1. Structure of the identification scheme with detection pre-processing](image)
the nodes. As expected, the nearly optimal detection step brings a compelling improvement of about 10 dB. Future work may include constructing hypotheses for multiple simultaneously transmitting primary users, and exploring the effect of imperfect knowledge (estimates) of the attenuation factors.

7. REFERENCES


