

# SPATIAL DIVERSITY AWARE DATA FUSION FOR COOPERATIVE SPECTRUM SENSING

Nuno Pratas<sup>(1,2)</sup> Neeli R. Prasad<sup>(1)</sup> António Rodrigues<sup>(1,2)</sup> Ramjee Prasad<sup>(1)</sup>

<sup>(1)</sup> Aalborg University, Denmark, nup@es.aau.dk, np@es.aau.dk, prasad@es.aau.dk

<sup>(2)</sup> Instituto Superior Técnico, Instituto de Telecomunicações, Portugal, ar@lx.it.pt

## ABSTRACT

Studies have shown that when data fusion schemes are used in cooperative spectrum sensing, there is a significant gap between the available resources and the ones perceived by the network.

In this paper a cluster based adaptive counting rule is proposed, where the local detectors that experience similar signal conditions are grouped by the fusion center in clusters and where the data fusion is then done separately at each cluster.

The proposed algorithm uses the correlation between the binary decisions of the local detectors over an observation window to select the cluster where each local detector should go. It was observed that in the case where there is only one signal source, that the proposed algorithm is able to achieve the same level of performance when compared to the perfect clustering algorithm where full information about the signal conditions at each local detector is available.

**Index Terms**— Data Fusion, Cooperative Spectrum Sensing, Exposed Node Problem, Clustering

## 1. INTRODUCTION

The main appeal of a Cognitive Radio Network (CRN) is its ability to perform opportunistic access to frequency bands when they are vacant, which can only occur if accurate information about the surrounding environment is made available. The information about which channels are vacant is obtained from the Spectrum Sensing (SS) process, where the CRN node samples the monitored spectrum and then decides on the presence of an incumbent signal. The SS serves a dual purpose, to detect when channels are vacant and also to limit the interference that the CRN nodes may cause in the incumbent network. Therefore the detection performance of the SS scheme in use by the CRN can affect both the performance of the CRN as well as the one of the incumbent network.

The detection performance is affected by the channel conditions, which depend on the path loss, multipath, shadowing, local interference and noise uncertainty [1, 2]. The combination of these phenomena can result in regimes where the Signal to Noise Ratio (SNR) is below the detection threshold

The authors would like to acknowledge the FCT which supported this work under the grant SFRH/BD/36454/2007 and IT/LA.

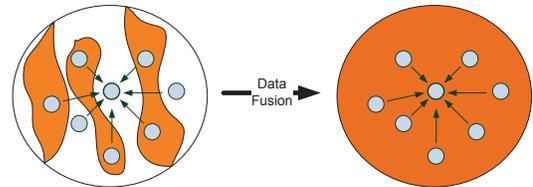
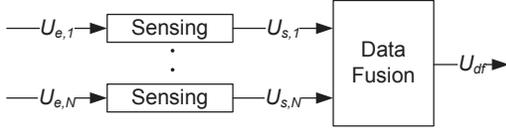


Fig. 1. Data fusion results in the the loss of spatial diversity

of the detector, leading to the hidden node problem. These regimes can be overcome through the use of a Cooperative Spectrum Sensing (CSS) scheme, [3, 4, 5, 6, 2]. The main idea behind the CSS is to enhance the detection performance by exploiting the spatial diversity in the observations of spatially separated CRN nodes.

The use of CSS schemes also brings a cooperation overhead, which refers to any extra effort that the CRN node needs to do to accomplish the CSS. A overhead not considered in the literature [7, 8] is that although performing the CSS using CRN nodes which are under correlated shadowing might decrease the detection performance, there is a drawback of not doing it so. In the CSS the local decisions of the CRN nodes are combined at the fusion center through the Local Decisions Data Fusion (LDDF) process. When the LDDF is performed over uncorrelated local decisions, then the decisions of far apart Local Detectors (LDs) will be combined. So, one will lose information about possible available spectrum opportunities, i.e. one loses information about the spatial diversity, which leads to the exposed node problem. This phenomenon is illustrated in Figure 1, where the coloured regions represent where the spectrum is occupied and non-coloured where the spectrum is available. After the LDDF occurs the CRN loses the information about the vacant regions.

The motivation behind this paper is to provide a LDDF scheme which minimizes both the hidden and exposed node problem. The proposed LDDF scheme, groups LDs with similar signal conditions in clusters and then performs the LDDF separately at each of these clusters. The focus of the paper is on the part of the LDDF where the clustering occurs, while the data fusion method in place is the Adaptable Counting Rule (ACR), [9]. The clustering is achieved with the aid of the sample correlation measured from the LDs local decisions re-



**Fig. 2.** Capacity along the parallel data fusion chain

ceived at the fusion center. The performance of the proposed scheme is measured using the system perceived Capacity (C) and False Capacity (FC) metrics introduced in [8] and compared to the perfect clustering algorithm, where full information about the conditions at each of the LDs is available at the fusion center.

The remainder of this paper is organized as follows. In Section 2 it is depicted the system model as well as the metrics used to evaluate the proposed scheme performance. In Section 3 it is presented the proposed clustering mechanism together with a comparative performance evaluation. Finally, Section 4 concludes the paper with a recap of the contribution, the main results and an outlook on further development.

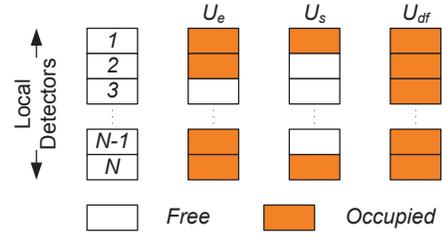
## 2. SYSTEM EVALUATION

### 2.1. Introduction

In Figure 2 the steps that constitute the Data Fusion Chain (DFC) are depicted, where  $U_{e,n}$ ,  $U_{s,n}$  and  $U_{df}$  quantify the perceived state of the sensed channel at each step of the data fusion chain. The  $U_{e,n}$  quantifies the experienced state of the channel targeted for sensing by the Local Detector (LD). The  $U_{s,n}$  quantifies the perceived state of the channel after sensing. Finally,  $U_{df}$  quantifies the perceived channel state after the data fusion. The values that each of these states can take, since binary decisions are considered, are,

$$U_{e,n}, U_{s,n}, U_{df} = \begin{cases} 1 & \text{if } H_0 \\ 0 & \text{if } H_1 \end{cases} \quad (1)$$

To illustrate the meaning of system perceived C, first consider that in a CSS session there are several LDs and that each of these experiences different signal conditions. Now if one considers that at the location of each of these LDs, the channel is deemed free for use if the experienced SNR,  $\gamma_{exp}$ , is below a given SNR threshold,  $\gamma_{thrs}$ , then it is expected that due to the mentioned varying channel conditions, some of the LDs will experience the same channel as free while other will experience it as occupied, when the  $\gamma_{exp} > \gamma_{thrs}$ . Note that what is meant by experienced channel state refers to the actual state of the sensed channel at a particular location, given by  $U_{e,n}$ , before the sensing takes place. Following the DFC in Figure 2, in Figure 3 is depicted the status of the perceived channel state at each step of the DFC, given by  $U_e$ ,  $U_s$  and  $U_{df}$ . Each of the figure's blocks represents a LD and its color the perceived channel state.



**Fig. 3.** Spectrum sensing capacity illustration

When comparing the experienced spectrum state,  $U_e$ , and sensed spectrum state,  $U_s$ , it can be seen that some of the LDs fail to detect that the channel is occupied, i.e. a missed detection occurs, while the other LDs judge the channel as occupied when it is not, i.e. a false alarm occurs. Both events have impact on the perceived system capacity, the missed detections because they cause the LD to perceive a channel as free when it is occupied, and the false alarm because the LD perceives the channel as occupied when it is free. So in the former, one assumes to have more resources than the ones available, while in the latter one misses the available resources. After the LDDF takes place,  $U_{df}$ , all LDs are assumed to perceive the channel state that resulted from the LDDF. From the example in Figure 3, after the LDDF all LDs are assumed to perceive the channel as occupied, although some of the LDs actually perceive the channel as free, causing a decrease of the system perceived capacity.

### 2.2. Performance Metrics

Several metrics are defined to measure the system perceived capacity at the different stages of the DFC. The capacity in the CSS context is the ratio of LDs that experience or perceive the channel state as free.

The potential capacity,  $C_r$ , is defined as,

$$C_r = \frac{\sum_{i=1}^N U_{e,i}}{N} \quad (2)$$

The post-sensing capacity,  $C_s$ , is defined as,

$$C_s = \frac{\sum_{i=1}^N U_{s,i} U_{e,i}}{N} \quad (3)$$

The post-data fusion capacity,  $C_{df}$ , is defined as,

$$C_{df} = U_{df} \frac{\sum_{i=1}^N U_{e,i}}{N} \quad (4)$$

The  $C_r$ ,  $C_s$  and  $C_{df}$  are the system's perceived capacity at three different points of the DFC, and the difference among them accounts for the probability of false alarm. But these metrics do not account for the effect of perceiving erroneously the channel state as free, i.e. they do not account for the occurrence of misdetections.

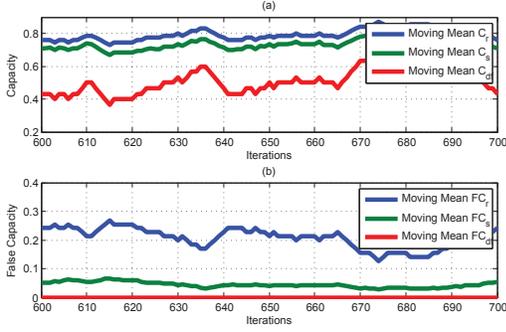


Fig. 4. Capacity and False Capacity metrics use illustration

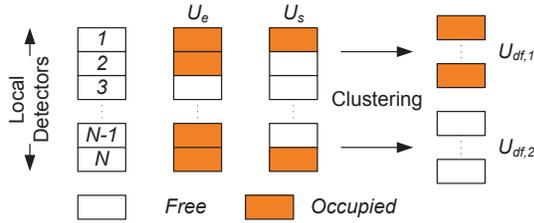


Fig. 5. Clustering of LDs illustration

The potential FC,  $FC_r$ , is defined as,

$$FC_r = 1 - C_r \quad (5)$$

The post-sensing false capacity,  $FC_s$ , is defined as,

$$FC_s = \frac{\sum_{i=1}^N U_{s,i} (1 - U_{e,i})}{N} \quad (6)$$

The post-data fusion FC is defined, given by  $FC_{df}$ , as,

$$FC_{df} = U_{df} \frac{\sum_{i=1}^N (1 - U_{e,i})}{N} \quad (7)$$

Through these metrics it is possible to characterize completely the perceived C and FC at each point of the DFC, and therefore to understand and quantify the capacity limits achieved by using different LDDF schemes as well on the case where the LDDF is not performed. In Figure 4 is depicted an example where the defined metrics are applied and where it can be seen that the ACR, [9], LDDF scheme minimizes both the C and FC.

### 2.3. Clustering motivation

The conclusion of the analysis found in [8], was that it might be possible to improve the network perceived C by gathering the LDs in different clusters, and then perform the data fusion individually at each cluster of LDs, as illustrated in Figure 5.

As a proof of concept consider the plot in Figure 6, where it is depicted the C and FC at the different levels of the parallel data fusion chain. The metrics of interest are the  $C_{df}$ ,

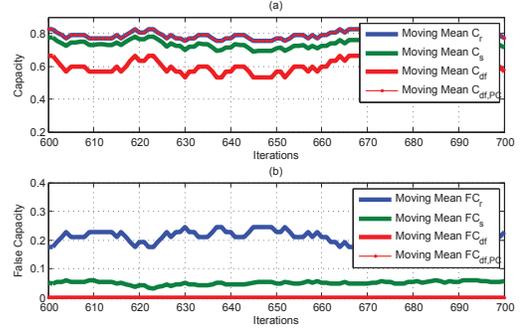


Fig. 6. Perfect clustering in data fusion with  $N = 40$

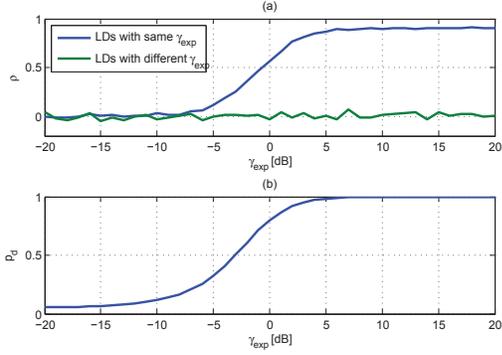
$C_{df,PC}$ ,  $FC_{df}$  and  $FC_{df,PC}$ , which measure the C and FC perceived by using the ACR LDDF and by dividing the LDs in clusters, respectively. In the latter case, the LDs where grouped into two clusters, and the division method was given by the  $\gamma_{thrs}$  where any LD with  $\gamma_{exp} < \gamma_{thrs}$  was assumed to be experiencing the channel as vacant. The motivation for defining such threshold, is that it is expected that below a certain  $\gamma_{thrs}$  it does not make sense to consider the channel to be occupied, since the amount of interference that the node associated to the LD would experience from the signal source can be neglected. This assumption is done from the CRN side, i.e. it does not consider the minimization of the interference of the CRN in the primary network and therefore this approach might not be applicable to scenarios where the goal is to minimize interference in the primary network. It should be noted that the purpose of this algorithm is to identify the maximum number of available opportunities for the CRN to use, and whether these will be used by the CRN will depend on the access control mechanism in place.

From the plot in Figure 6 it can be seen that it is worthwhile to group the LDs in clusters. The main challenge is to identify which information should be used as basis to perform the clustering process, since the information about the  $\gamma_{exp}$  is not available at the fusion center. An alternative source of information is the correlation observed between the LDs local decisions over time. The issue with this source of information is that the LDs are not perfect, i.e. their  $p_d < 1$  and their  $p_{fa} > 0$ , and therefore the LDs will most likely never be fully correlated. Therefore there is a need to define a correlation threshold which translates to the considered  $\gamma_{thrs}$ .

## 3. ALGORITHM AND RESULTS

### 3.1. Algorithm Description

Consider the plots in Figure 7, where it is depicted the variation of the Pearson correlation coefficient,  $\rho$ , in regards to the SNR experienced,  $\gamma_{exp}$ , by a pair of LDs in two different scenarios. In the first scenario, the blue curve, the pair of LDs have the same performance, i.e. they have the same  $p_d$ ,



**Fig. 7.** (a)  $\rho$  between the decisions of two LDs, with signal duty cycle of 0.5 and the LDs  $p_{fa} = 0.05$ , (b)  $p_d$  versus  $\gamma_{exp}$

which is dependent on the  $\gamma_{exp}$ . In the second scenario, the green curve, in one of the LDs the  $p_d$  is same in regards to the minimum  $\gamma_{exp}$ , while in the other LD the  $p_d$  varies with the  $\gamma_{exp}$ . The LD performance is plotted in Figure 7(b), using as reference the energy detector model presented in [3].

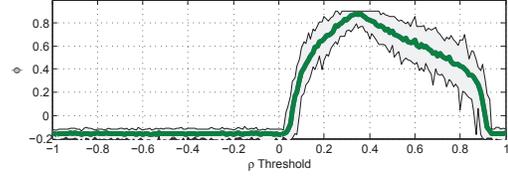
In the first scenario, it can be seen that the  $\rho$  is higher than 0.5 only when the  $\gamma_{exp} > 0dB$ , which from the plot in Figure 7(b) translates to a  $p_d > 0.8$ . This relationship leads to that by using the  $\rho$  it is possible to identify whether a LD is experiencing a SNR above a certain level, if there at least one other LD which is also experiencing the same level of SNR.

In the second scenario, it can be seen that if the  $\gamma_{exp}$  by one of the LDs is low enough, then the  $\rho$  is always near zero. This is also observed in the first scenario, when the  $\gamma_{exp}$  by both LDs is low enough, where in the depicted scenarios low enough occurs when the  $\gamma_{exp} < -10dB$ . This occurs due to the mapping chosen in the local decisions in regards to the presence and absence of a signal, i.e. when the signal is absent,  $H_0$ , the local decision of the  $i^{th}$  detector is mapped as  $u_i = 0$ , while when the signal is present,  $H_1$ , the local decision of the  $i^{th}$  detector is mapped to  $u_i = 1$ , which is the opposite of the mapping considered in (1).

The clustering algorithm, should perform the clustering by using the correlation coefficients of the LDs decisions over time. For simplicity, it is considered that there is only one signal source, and therefore it is of interest to divide the LDs in two clusters. One cluster will include all the LDs where the  $\gamma_{exp} > \gamma_{thrs}$ , while the other will include the remaining LDs.

In Figure 8 is plotted, for the detector performance curve depicted in Figure 7(b), the evaluation of different  $\rho$  thresholds in regards to  $\phi$  coefficient, so to measure the performance of the classification mechanism which puts the LDs in one of two clusters. It can be observed that the  $\phi$  is maximized when the  $\rho \in [0.3, 0.4]$ , when all LDs are dimensioned according to the considered performance curve.

In Algorithm 1 is listed the algorithm which groups the LDs in clusters. Where  $|L_\rho|$  represents the number of ele-



**Fig. 8.**  $\rho$  threshold variation evaluated through  $\phi$

ments in the decrescent ordered list  $L_\rho$ . The  $SI$  is the sensing iteration counter, the  $SI_{min}$  is the minimum number of sensing iterations so that the obtained  $\rho$  is statistically significant.  $C_1$  is the LD cluster where are stored the LDs which are experiencing a level of correlation above  $\rho_{thrs}$ , which means that these LDs have a  $\gamma_{exp} \geq \gamma_{thrs}$ , while  $C_0$  is the cluster where all the uncorrelated LDs are stored. The cluster decision,  $C_{df,0}$ , for  $C_0$  is  $H_0$ , because when a LD is uncorrelated with all others LDs then it means that the  $\gamma_{exp} < \gamma_{thrs}$ .

---

**Algorithm 1** Single source clustering algorithm

---

```

if  $SI < SI_{min}$  then
    Obtain  $C_{df}$  using ACR algorithm [9]
else
    for all  $i, j$  LD pairs when  $i \neq j$  do
         $\rho_{ij} \leftarrow$  from local decisions  $\mathbf{u}_i$  and  $\mathbf{u}_j$  and  $L_\rho \leftarrow \rho_{ij}$ 
    end for
    Add all LDs to  $C_0$ 
    while  $|L_\rho| > 0$  do
        Remove  $\rho_{ij}$  from the head of  $L_\rho$ 
        if  $i, j \notin C_1$  then
            if  $\rho_{i,j} \geq \rho_{Thrs}$  then
                 $C_1 \leftarrow i, j$ 
                Remove  $i, j$  from  $C_0$ 
            end if
        else
            if  $\rho_{i,j} \geq \rho_{Thrs}$  then
                if  $i \notin C_1$  then
                     $C_1 \leftarrow i$ , remove  $i$  from  $C_0$ 
                else
                     $C_1 \leftarrow j$ , remove  $j$  from  $C_0$ 
                end if
            end if
        end if
    end while
    if  $|C_1| > 0$  then
        Obtain  $C_{df,1}$  using ACR algorithm [9]
    end if
    if  $|C_0| > 0$  then
         $C_{df,0} = H_0$ 
    end if
end if

```

---

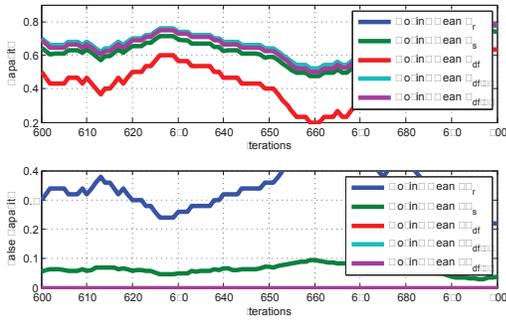


Fig. 9. Clustering comparison with  $N = 40$

### 3.2. Algorithm Results

In Figure 9 is depicted the comparison over time using the C and FC metrics when using Data Fusion (DF) with and without clustering, where the indices  $df$ ,  $PC$  and  $df$ ,  $CC$  refer to perfect clustering and the correlation based clustering, as defined by Algorithm 1 with a  $\rho_{thrs} = 0.35$ , respectively. As seen from the results the proposed correlation based clustering algorithm is able to achieve the performance level of the perfect clustering algorithm. This performance was achieved because a proper  $\rho_{thrs}$  was set, which as discussed before depends on the detection performance of the LDs.

## 4. CONCLUSIONS AND FUTURE WORK

In this paper the concept of Capacity and False Capacity in the Cooperative Spectrum Sensing context was introduced, and it was shown how these metrics can be used to measure the performance along the data fusion chain. Through these metrics it was observed that when using data fusion the information about the LDs spatial diversity is discarded, which leads to the aggravation of the exposed node problem. To overcome this limitation it was proposed an algorithm which groups LDs in a cluster, with the purpose of minimizing the information loss in regards to the LDs spatial diversity. These clusters are then created by grouping together LDs which are correlated in regards to their local decisions.

It was shown, that in the case where there is only one signal source, that the proposed algorithm was able to achieve the same level of performance, measured through C and FC when compared to the perfect clustering algorithm where full information about the conditions at each of the LDs was available at the fusion center.

The clustering algorithm can in principle be generalized for multiple signal sources, if one considers that LDs which experience similar conditions can be grouped together. The maximum number of cluster is expected to be  $2^{\text{signal sources}}$ , although limited by the number of LDs since there should be at least 3 LDs per cluster to take advantage of the LDDF. The generalization of the proposed algorithm and evaluation is left for future work.

## 5. REFERENCES

- [1] R. Tandra and A. Sahai, "Snr walls for signal detection," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 4–17, February 2008.
- [2] K.C. Chen and R. Prasad, *Cognitive Radio Networks*, John Wiley and Sons Ltd, April 2009.
- [3] D. Cabric, A. Tkachenko, and R. W. Brodersen, "Experimental study of spectrum sensing based on energy detection and network cooperation," in *Proceedings of the first international workshop on Technology and policy for accessing spectrum*, New York, NY, USA, 2006, TAPAS '06, ACM.
- [4] A. Ghasemi and E.S. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments," in *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005. (DySPAN 2005)*, November 2005, pp. 131–136.
- [5] S.M. Mishra, A. Sahai, and R.W. Brodersen, "Cooperative sensing among cognitive radios," in *IEEE International Conference on Communications, 2006. (ICC 2006)*, 2006, vol. 4, pp. 1658–1663.
- [6] E. Visotsky, S. Kuffner, and R. Peterson, "On collaborative detection of tv transmissions in support of dynamic spectrum sharing," in *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005. (DySPAN 2005)*, Nov. 2005, pp. 338–345.
- [7] Ian F. Akyildiz, Brandon F. Lo, and Ravikumar Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: A survey," *Physical Communication*, vol. 4, no. 1, pp. 40–62, 2011.
- [8] N. Pratas, N. Marchetti, N. R. Prasad, A. Rodrigues, and R. Prasad, "System capacity limits introduced by data fusion on cooperative spectrum sensing under correlated environments," *MTA Review*, vol. XX, no. 4, pp. 245–262, 2010.
- [9] Nuno Pratas, Nicola Marchetti, Neeli Prasad, António Rodrigues, and Ramjee Prasad, "Adaptive counting rule for cooperative spectrum sensing under correlated environments," *Wireless Personal Communications*, pp. 1–14, 10.1007/s11277-012-0519-4.