One-Class based learning for Hybrid Spectrum Sensing in Cognitive Radio

M. Jaber*, A. Nasser*†, N. Charara*, A. Mansour† and K. C. Yao‡
* ICCS-Lab, Faculty of Science, American University of Culture and Education (AUCE), Beirut, Lebanon
† LABSTICC, CNRS, UMR 6285, ENSTA Bretagne, 2 Rue François Verny, 29806 Brest, France
‡ LABSTICC, CNRS, UMR 6285, UBO, 6 Avenue le Gorgeu, 29238 Brest, France

***Abstract*—The main aim of the Spectrum Sensing (SS) in a Cognitive Radio system is to distinguish between the binary hypotheses H_0: Primary User (PU) is absent and H_1: PU is active. In this paper, Machine Learning (ML)-based hybrid Spectrum Sensing (SS) scheme is proposed. The scattering of the Test Statistics (TSs) of two detectors is used in the learning and prediction phases. As the SS decision is binary, the proposed scheme requires the learning of only the boundaries of H_0-class in order to make a decision on the PU status: active or idle. Thus, a set of data generated under H_0 hypothesis is used to train the detection system. Accordingly, unlike the existing ML-based schemes of the literature, no PU statistical parameters are required. In order to discriminate between H0-class and elsewhere, we used a one-class classification approach that is inspired by the Isolation Forest algorithm. Extensive simulations are done in order to investigate the efficiency of such hybrid SS and the impact of the novelty detection model parameters on the detection performance. Indeed, these simulations corroborate the efficiency of the proposed one-class learning of the hybrid SS system.**

I. Introduction

The ever increasing demand on the wireless technologies pushed the communication community to tackle the problem of the spectrum scarcity. Cognitive Radio is one of the proposed solutions, which aims at utilising the frequency spectrum efficiently. The efficient use is based on sharing the spectrum between Unlicensed users, namely known as Secondary User (SU) and Primary User (PU). SU could access the frequency channel only when PU is absent. Thus, sensing the PU status, whether it is absent or active becomes an essential function of SU. To do so, Spectrum Sensing (SS) is responsible to verify the primary channel status by deriving a Test Statistic on the received signal such as Energy Detector (ED) [1], Cumulative Power Spectral Density (CPSD) detector [2], Cyclostationary detector [3], [4], etc [5].
In classical SS, TS is compared to a predefined threshold in order to make a decision on the PU status. When the TS is above a certain threshold PU is considered as active. In fact, this approach predetermines that the statistical distribution of TS is known, which is not always possible due to the unstable ans unknown statistical properties of the noise, the PU signal, and the channel.

To overcome the statistical problems of the classical SS and improve its performance, several works have been published proposing the adoption of the Machine Learning (ML) and the neural networks' techniques in order to make decision on the PU channel occupancy [6]–[12]. The main aim of the proposed works is to learn a statistical model for both statuses: the first one is H_0 when PU is assumed to be absent, and H_1 when PU is assumed to be active.
In [8], ML techniques such as the K-Means and Support-Vector Machine (SVM) are used to distinguish between the H_0 and H_1 hypotheses in a cooperative SS. Two low-dimension probability vectors related to both H_0 and H_1 of ED are used in order to train the system. SVM is used in order to set the threshold curve between H_0 and H_1 clusters. However, the proposed algorithm requires the pre-knowledge of the probability density function of ED under both H_0 and H_1, which is not always available since the PU signal statistical parameters are not always known.
In [9], [10], Artificial Neural Network (ANN) have been proposed in order to perform a hybrid SS. ANN is trained using the TSs of two detectors related to H_0 and H_1(in [9] ED and Cyclostationary Detector are used, and in [10] ED and likelihood ratio statistics are used). However, the application of the ANN requires the statistical parameters of the PU signal, which may be hard to be known in a CR context where the SU may deal with a great variety of primary signals.
Moreover, all the aforementioned works assumed two classes of PU activities: the first one lies with H_0 and the second is related to H_1. The latter class is related to the statistical properties of PU signal, such as the energy, the cyclo-stationary features, the sampling rate, etc. These statistical properties are not always available neither stable. Regarding the stability, PU may vary its transmit power, this may impact the position of the classifier hyper-plane line in ML techniques when distinguishing between the H0 and H1 classes. Furthermore, SU may deal with a great variety of signal types, as CR is based on the dynamic spectrum allocation. This variety makes the model learning process with all PU signals of the accessible channels very expensive.

In SS, as stated above, only two hypotheses are available. In other words, when the hypothesis H_0 is eliminated, then H_1 is surly considered as the active state. Motivated by this fact, in our work the detection decision is based on a learning stage of only H_0-class. The main contributions of this paper can be presented as follows:

1) SS is performed without any need for the statistical parameters of the primary signal.
2) Hybrid Spectrum Sensing is applied, where the model is trained using records of two detectors under H_0. This fact enhances the accuracy of the sensing per-

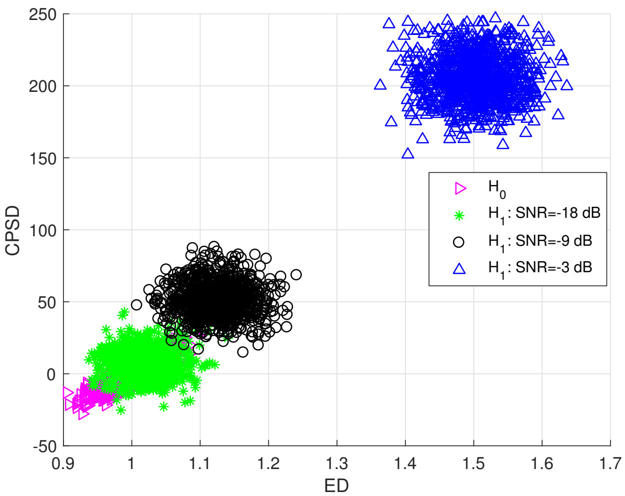


Fig. 1. The Scattering of (ED, CPSD) for N=1000 samples, 1000 trials and different values of SNR.

formance compared to using the data of only one detector.

3) A novelty detection approach based on Isolation Forest (iForest) [13] is adopted. We assume that we only observe examples of one class (H_0) and the second hidden class (H_1) is considered as the novelty element class.

4) The performance of the one-class hybrid detection is discussed according to the parameters of the learning technique and how they impact the detection performance.

II. System Model

The decision in SS is binary where two hypotheses must be distinguished:

$$H_0 : \text{PU is absent} \quad (1)$$

$$H_1 : \text{PU is active} \quad (2)$$

In classical SS, a detection method is applied on the received signal in order to outcome a TS. This TS leads SU to decide on the PU activity by comparing it to a predefined threshold. Accordingly, two classes of TS values have to be defined: H_0-class and H_1-class related to the hypotheses H_0 and H_1 respectively. In fact, H_0-class depends only on the system parameters such as the noise and the hardware imperfections, in other words it is independent from the PU signal as the received signal $y(n)$ can be presented as follows:

$$y(n) = w(n) \text{ under } H_0 \quad (3)$$

$$y(n) = s(n) + w(n) \text{ under } H_1 \quad (4)$$

where $w(n)$ is an Additive White Gaussian Noise (AWGN) and $s(n)$ is the PU signal to be detected, including the channel effects. Knowing that TS is a function of $y(n)$, H_0-class becomes independent of the PU signal as shown in (3).
In our system model, we assume that the SS is hybrid, i.e. SU uses more than one detector. The results of the detectors must be combined in order to make a decision on the PU activity. Figure 1 shows the scattering of two detectors: Energy Detector (ED) and Cumulative Power Spectral Density (CPSD) detector that are defined respectively as follows [1], [2]:

$$v_1 = \frac{1}{N}\sum_{1}^{N} |y(n)|^2 \quad (5)$$

$$v_2 = \frac{2}{N^2\sigma_w^2}\sum_{k=1}^{N/2}\left(\frac{N}{2} - k + 1\right)\frac{|Y(m)|^2 + |Y(-m+1)|^2}{2} \quad (6)$$

where N is the number of received samples, $Y(m)$ is the discrete Fourier transform of $y(n)$, and σ_w^2 is the noise variance.
Unlike H_0-class, the scattering of H_1-class values is a function of SNR as the H_1-class changes its position in the scattering space with the SNR (see figure 1). In order to distinguish between H_0- and H_1-class, the first idea which comes from the figure 1 is to use the available data to determine the optimal limit between these two classes. However, in a given scenario where the SNR is unstable, unknown, or relatively weak , set a threshold curve between H_0 and H_1 classes becomes hard to do. Note that other parameters such as the oversampling rate, the cyclic frequencies, the modulation type or order, etc. may impact the measure of the related TSs.

III. Proposed One-Class Hybrid Spectrum Sensing

Let $V = (v_1, v_2)$ be the vector of the evaluated values of the two detectors, ED and CPSD, used in the SS. H_0-class is trained from the values of V, when PU is absent. Accordingly, each value outside H_0-class has to be considered to be in H_1-class. Accordingly, instead of profiling the H_0 behavior, we aim to isolate the novel and unusual observations. These unusual observations are considered belonging to H_1-class. This approach may be sufficient to make a decision on the PU's activities:

$$V \in H_0\text{-class} \rightarrow \text{PU is absent} \quad (7)$$

$$V \notin H_0\text{-class} \rightarrow \text{PU is active} \quad (8)$$

Indeed, the main challenge becomes how to isolate the unusual observations of the so-called H_1-class. To achieve our goal, we assume that SU is capable to generate N_t trials of V under H_0. Based on these N_t values, the boundaries of H_0-class are virtually estimated. Note that no cooperation with PU is required at this stage as the H_0-class values are independent from the PU signal. As depicted in figure 1 H_0 data are gathered in a well distinguishable location is the scattering space. Here, H_1 instances become considered as *novelty* compared to H_0 instances, which represents the normal ones. Subsequently, one of the powerful techniques, that can distinguish between unusual and usual instances is the iForest [13].

A. One-Class based learning model

Isolation Forest (iForest) is a learning algorithm that isolates anomalies from the rest of normal instances, instead of profiling the normal behavior. This strategy is well adopted for the one-class training paradigm. Indeed, iForest introduces the use of anomaly score rather than the commonly used distance and density measures for the novelty detection [14], [15].
The iForest starts with a training phase, Binary trees (iTrees) are constructed using sub-samples of random instances. In

these trees, Partitions are generated by selecting a feature and then selecting a random split value between the selected feature's minimum and maximum value. iForset takes only two parameters, the number of trees and the subsampling size. To avoid problems due to tree algorithm randomness, the process is repeated several times and the average path length is calculated. The anomalies are those cases of short average path lengths on the iTrees. After several iterations the mean path length converges. Each algorithm for anomaly detection will calculate its data points and instances, and measure the confidence of the algorithm in their possible anomalies. In iForest, the leading and distinguishing insight is that anomalies remain closer to the root of the tree. The anomaly value is known as the path length $h(x)$, where x is the number of edges crossed from the root node. The anomaly score is defined as [13]:

$$s(x,n) = 2^{-\frac{E(h(x))}{c(n)}} \tag{9}$$

where $E(h(x))$ is the average path length of observation x, $c(n)$ is the average path length of unsuccessful search in a Binary Search Tree, and n is the number of external nodes.
An anomaly score is given to each observation and the following decision can be made on its basis: a score close to 1 indicates anomalies, score much less than 0.5 indicates regular observations. If a is score is equal to 0.5 then it does not have a clear anomaly.
The training stage of our one-class based learning process has been inspired from iForest. In this stage, we build binary trees using sub-samples of H_0 training set. At testing stage, we calculate the novelty score by the same way as an anomaly score (9) for each instance using the trained binary trees.

IV. Numerical Results

In this section, the efficiency of the proposed scheme is numerically evaluated under several scenarios related to the iForest parameters, and the effect of SNR. The evaluation is based on the probability of detection P_d and the probability of false alarm P_{fa}. P_d represents a True Negative decision of the iForest system, while P_{fa} stands for a False Positive decision. Increasing the detection PU signal is a main challenge for SU, and this requires an efficient detection performance, so a high True Negative accuracy is required. On the other hand, the SU data rate should be increased as possible in order to achieve a high spectral efficiency. For that reason, P_{fa} should be minimized as possible.
Without loss of generality, the data generated to learn the system is based on an AWGN noise with a zero mean and unit variance. The PU signal, which is not involved in the training stage, is assumed to be 16-QAM modulated with an over-sampling factor of 4. The TSs related to ED and CPSD are found based on 1000 received samples of $y(n)$ under both H_0 and H_1.

To figure out the effects of the subsampling on the performance of the hybrid ED-CPSD detector, figure 2 depicts the evolution of P_d, when the subsampling size varies. These results come out for $P_{fa} = 0.1$. For low SNR -24 dB to -15 dB, P_d increases with the size of the subsampling while this size is lower than 4000. Beyond this value P_d becomes constant. On the another hand, P_d is not affected

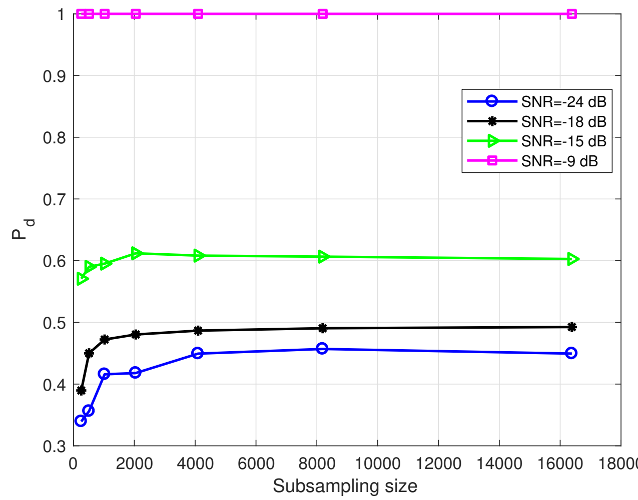


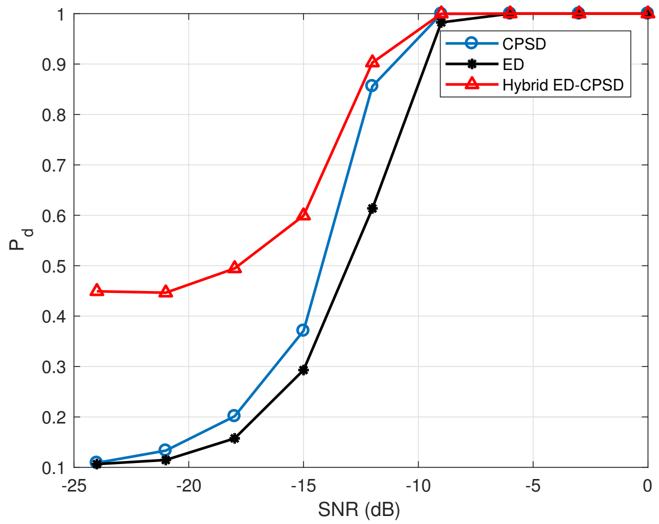
Fig. 2. The effect of the sub-sampling on P_d of the iForest-based one class HSS for various values of SNR. P_{fa} is fixed to 0.1.

by the size of subsampling when the SNR is relatively high (-6 dB) as shown in figure 2 as for such value of SNR the classification becomes easier to the system due to the fact that H_0-instances and H_1-instances are practically separated. In order to show the efficiency of the one class based HSS, figures 3(a) and 3(b) show the variation of P_d in terms of SNR for a constant $P_{fa} = 0.1$ and $P_{fa} = 0.05$. Three detection scenarios are taken: ED, CPSD, and hybrid ED-CPSD. The hybrid detection outperforms both ED and CPSD for the both values of P_{fa}. The simulations of fig. 3(a) and 3(b) are based on a Number of trees and a subsampling size equal to 100 and 8192 respectively. For SNR=-15 dB and $P_{fa} = 0.1$ the hybrid ED-CPSD achieves $P_d = 0.6$ while P_d of ED and CPSD are 0.3 and 0.37 respectively. For a SNR of -12 dB, the P_d of hybrid ED-CPSD exceeds 0.9 while the classical ED achieves only 0.6. On the other hand, P_d showed by figure 3(a) is higher than the one of 3(b) due to the variation of P_{fa}. When P_{fa} is low, the contamination becomes more stringent which adversely impact the detection probability. However, when the contamination is less stringent P_{fa} and P_d increase accordingly.

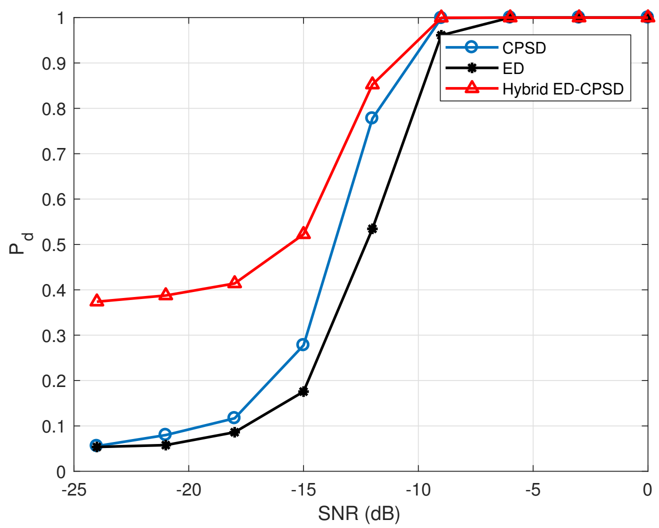
Figure 4 shows the impact of the number of trees on the accuracy convergence of the proposed iForest-based HSS. The accuracy is evaluated on the percentage of the True Positive and the True Negative relative to the overall of the validation instances. Standard deviation is evaluated based on the experiment outcomes of the iForest-based HSS for 100 iterations. The accuracy in terms on the number of trees is found for several SNRs. As it can be shown in figure 4, the accuracy average is constant relative to the number of trees for a given SNR. In contrast, the standard deviation increases as the number of trees decreases. However, for a relatively high SNR (i.e. SNR=-9 dB), the standard deviation is closed to zero. This is because distinguishing the novelty instances at such SNR becomes an easy task.

V. Conclusion

In this paper, SS in CR is performed using one-class based learning. Unlike existing ML based SS, in our work no pre-



(a) $P_{fa} = 0.1$



(b) $P_{fa} = 0.05$

Fig. 3. The evolution of P_d in terms of SNR for $P_{fa} = 0.1$ and $P_{fa} = 0.05$. Results ED, CPSD and HSS with ED-CPSD are presented. The Number of trees and the subsampling size are set to 100 and 8192 respectively.

information on the PU is required. HSS is adopted, where two detectors are used when performing the SS. The data gathered under H_0, i.e. PU is absent, is used to train the one-class model. iForest inspired technique was proposed to learn the H_0-class and to to detect the presence of the unusual H_1 observations. The obtained results of the HSS demonstrate that the proposed one-class scheme presents an efficient SS performance and enhances the SS compared to the non HSS.

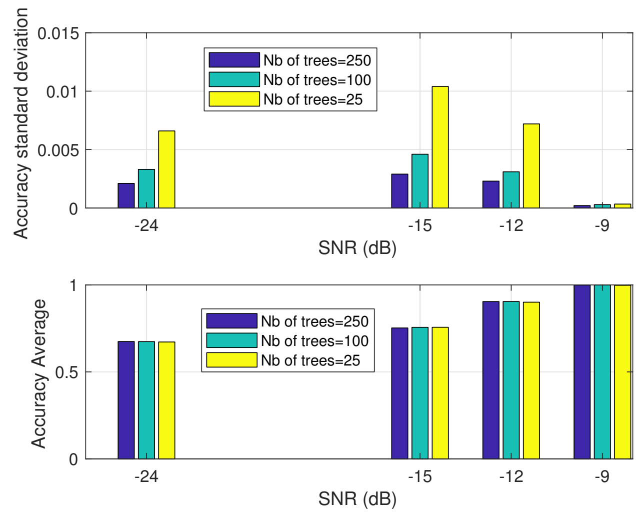


Fig. 4. Investigation on the effect of the number of trees on the performance of our iForest-based one class HSS for various values of SNR.

References

[1] F. Digham, M.-S. Alouini, and K. Simon. On the Energy Detection of Unknown Signals Over Fading Channels. *IEEE Transactions on Communications*, 55(1):21 – 24, Jan. 2007.

[2] A. Nasser, A. Mansour, K. C. Yao, H. Abdallah, and H. Charara. Spectrum sensing based on cumulative power spectral density. *EURASIP Journal on Advances in Signal Processing*, 2017(1):38, 2017.

[3] M. Derakhshani, T. Le-Ngoc, and M. Nasiri-Kenari. Efficient cooperative cyclostationary spectrum sensing in cognitive radios at low snr regimes. *IEEE Transactions on Wireless Communications*, 10(11):3754 – 3764, Nov 2011.

[4] Abbass Nasser, Ali Mansour, K-C Yao, Mohamad Chaitou, and Hussein Charara. Spatial and time diversities for canonical correlation significance test in spectrum sensing. In *2016 24th European Signal Processing Conference (EUSIPCO)*, pages 1232–1236. IEEE.

[5] T. Yucek and H. Arslan. A Survey of Spectrum Sensing Algorithms for Cognitive Radio Applications. *IEEE Communication Surveys & Tutorials*, 11(1):116 – 130, First Quarter 2009.

[6] C. Clancy, J. Hecker, E. Stuntebeck, and T. O'Shea. Applications of machine learning to cognitive radio networks. *IEEE Wireless Communications*, 14(4):47–52, August 2007.

[7] K. M. Thilina, K. W. Choi, N. Saquib, and E. Hossain. Machine learning techniques for cooperative spectrum sensing in cognitive radio networks. *IEEE Journal on Selected Areas in Communications*, 31(11):2209–2221, November 2013.

[8] Y. Lu, P. Zhu, D. Wang, and M. Fattouche. Machine learning techniques with probability vector for cooperative spectrum sensing in cognitive radio networks. In *2016 IEEE Wireless Communications and Networking Conference*, pages 1–6, April 2016.

[9] M. R. Vyas, D. K. Patel, and M. Lopez-Benitez. Artificial neural network based hybrid spectrum sensing scheme for cognitive radio. In *IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pages 1–7, Oct 2017.

[10] Y. Tang, Q. Zhang, and W. Lin. Artificial neural network based spectrum sensing method for cognitive radio. In *2010 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM)*, pages 1–4, Sep. 2010.

[11] Z. Li, W. Wu, X. Liu, and P. Qi. Improved cooperative spectrum sensing model based on machine learning for cognitive radio networks. *IET Communications*, 12(19):2485–2492, 2018.

[12] C. Guo, M. Jin, Q. Guo, and Y. Li. Spectrum sensing based on combined eigenvalue and eigenvector through blind learning. *IEEE Communications Letters*, 22(8):1636–1639, Aug 2018.

[13] Fei Tony Liu, Kai Ming Ting, and Zhi hua Zhou. Isolation forest. In *In ICDM 08: Proceedings of the 2008 Eighth IEEE International Conference on Data Mining. IEEE Computer Society*, pages 413–422.

[14] Fei Tony Liu, Kai Ming Ting, and Zhi-Hua Zhou. Isolation-based anomaly detection. *ACM Trans. Knowl. Discov. Data*, 6(1), March 2012.

[15] Fei Tony Liu, Kai Ming Ting, and Zhi-Hua Zhou. Isolation forest. In *IEEE International Conference on Data Mining*, ICDM 08, pages 413 – 422, USA, 2008.