Spectrum sensing by higher-order SVM-based detection

Angelo Coluccia, Alessio Fascista, and Giuseppe Ricci
Dipartimento di Ingegneria dell'Innovazione
Università del Salento, via Monteroni, 73100 Lecce, Italy.
e-mail: name.surname@unisalento.it.

***Abstract*—A novel spectrum sensing algorithm based on support vector machine is proposed. The idea is to map the received signals into a multi-dimensional feature space obtained from well-known spectrum sensing statistics and their higher-order combinations. The approach has been implemented and validated on a software-defined radio testbed. Experimental results have shown the receiver operating characteristic (ROC) curve of the proposed detector can outperform classical spectrum sensing approaches without requiring knowledge of the noise variance.**

***Index Terms*—spectrum sensing, support vector machine, detection, software-defined radio**

I. Introduction

There is an increasing demand for higher data rates in wireless communications, which however is more and more problematic to meet due to the scarcity of the spectrum. Traditionally, licensed spectrum is allocated over relatively long time periods and is intended to be used only by licensees. There is evidence today that spectral resources are underutilized in certain portions or "spectrum holes"; this is fostering the development of cognitive radio (CR) technologies [1], in order to reuse unused spectrum in an opportunistic manner.

Cognitive radio systems typically involve (licensee) primary users (PU), and secondary users (SU) who seek to opportunistically use the spectrum when the primary users are idle. The impact on the primary system (interference) must be kept at a minimal level, thus an effective spectrum sensing algorithm is of fundamental importance to decide whether a particular slice of the spectrum is available or not.

Spectrum sensing is also of great interest in passive radios, which have enormous applications in security and defense. In particular, localization of a non-cooperative target is preferably carried out in stealth mode, i.e., using only passive technologies instead of active ones such as conventional radar systems. A wireless sensor network (WSN) can be used to this aim, equipped with a passive radio which listens over a particular frequency and tries to detect a transmission.[1]

Classical approaches to the problem of spectrum sensing typically try to exploit the statistical properties of the received signal: signal processing tools are used to estimate specific characteristics, captured by e.g. the autocorrelation function or eigenvalue distribution, that can spotlight the presence of a structured signal representative of the PU transmission. To this aim, the analog signal is sampled and converted to the digital domain for processing, and novel techniques are available today for making such a process parsimonious (i.e., using a reduced number of samples compared to the standard Nyquist-rate prescription, see e.g. [4], [5]). At the same time, several machine learning approaches have been recently attempted to improve the detection capabilities of spectrum sensing, in particular adopting k-means, support vector machines (SVM), k-nearest neighbors (KNN), and other techniques [6]–[12]. Such approaches use as feature vector different quantities extracted from the samples: energy levels [7], eigenvalues or their ratios [8], [11], and other heuristic features obtained from ad-hoc signal statistics [9].

In this work, a novel spectrum sensing approach is proposed. It is based on SVM and it does not require to know the statistics of the noise, which is one of the main limitation of classical spectrum sensing based on energy detection. The idea is to nonlinearly map the received signals into a multi-dimensional feature space obtained from well-known spectrum sensing statistics and their higher-order combinations; then, SVM is adopted to perform binary classification through a hyperplane that ideally separates the data in such a space according to the hypothesis actually in force. SVM has several advantages including best linear separation of two clusters, potential to overcome the curse of dimensionality using kernels, and low-complexity [9]. The proposed approach of extending the feature vector through higher-order combinations of its components can be justified on the basis of Cover's theorem on the separability of patterns [13], which states that a classification problem, cast in a high-dimensional space nonlinearly, is more likely to be linearly separable than in a low-dimensional space [14], [15]. More specifically, we consider up to third-order combinations of three well-known spectrum sensing decision statistics, based on energy levels and eigenvalues.

The experiments carried out via a software-defined receiver reveal that the proposed approach is more effective than classical spectrum sensing techniques to detect (also weak) signals, without requiring knowledge of the noise power.

[1]As a practical example, the project *"SafeShore"* (*http://safeshore.eu*) [2] is aimed at detecting small targets such as drones (which can potentially carry explosives or can be used for smuggling); for this task, passive radio is used to detect the control signal from a remotely-piloted drone, which in turn is useful to "filter out" false alarms due e.g. to birds, indistinguishable through other complementary detection technologies such as LIDAR [3].

978-9-0827-9703-9/19/$31.00 ©2019 IEEE

II. Spectrum Sensing Techniques

A. Problem formulation

The presence or absence of a PU signal can be regarded as a binary hypothesis testing problem. Denote by $s(t)$ the complex envelope of the (zero-mean) signal that comes from a primary TX and received at the RX. $\mathbb{H}_1$ and $\mathbb{H}_0$ correspond to the presence/absence of such a signal, respectively; thus, the complex envelope of the received signal is modeled as

$$y(t) = \begin{cases} w(t) & \mathbb{H}_0 \\ s(t) + w(t) & \mathbb{H}_1 \end{cases} \tag{1}$$

where $w(t)$ denotes the thermal (white) Gaussian noise. According to the level of information about the primary signal, different kinds of detection schemes can be used for spectrum sensing. When the received signal is known, the matched filter can be used for spectrum sensing, which is the optimal detector for an AWGN environment. The main advantage of matched filter detection is the short sensing time to achieve a good performance, because signal coherence is exploited. When the signals are unknown, other techniques should be used. Three different algorithms, in particular *Covariance Matrix-Based Detector* [16], [17], *Maximum-Minimum Eigenvalue Detector* [18]–[21] and *Energy Detector* [20], [22] are considered in this work; these approaches have different requirements and advantages/disadvantages.

We first introduce the discrete model for the received signal. The sampled received signal with a certain sampling frequency f_s and sampling period $T = 1/f_s$ will be

$$y(nT) = \begin{cases} w(nT) & \mathbb{H}_0 \\ s(nT) + w(nT) & \mathbb{H}_1 \end{cases}.$$

To ease the notation, we use the shorthands $y(n) = y(nT)$, $s(n) = s(nT)$, and $w(n) = w(nT)$. The signals can be thus represented by vectors, i.e.,

$$\begin{aligned} \mathbf{y} &= [y(1)\ y(2)\ \cdots\ y(N_s)]^\top \\ \mathbf{s} &= [s(1)\ s(2)\ \cdots\ s(N_s)]^\top \\ \mathbf{w} &= [w(1)\ w(2)\ \cdots\ w(N_s)]^\top \end{aligned} \tag{2}$$

where $^\top$ denotes transposition and N_s is the number of samples. The problem finally becomes

$$\mathbf{y} = \begin{cases} \mathbf{w} & \mathbb{H}_0 \\ \mathbf{s} + \mathbf{w} & \mathbb{H}_1 \end{cases}. \tag{3}$$

B. Energy Detector

Energy detection is a major and basic method. Unlike other methods, energy detection does not need any information about the signal to be detected since it uses only the energy E as detection statistics. However, energy detection is extremely sensitive to noise power uncertainty, because the method relies on the accurate knowledge of the noise power [20], which is difficult to obtain in practice since it depends on many factors including temperature, environment, and frequency.

C. Covariance Matrix-Based Detection

To overcome the limitation of the energy detector, a Covariance Matrix-Based Detector [16], [17] has been proposed, which does not require knowledge of σ^2. Let us assume that $w(k)$ is a stationary process satisfying $E[w(k)] = 0$, $E[w(k)w^*(h)] = \sigma_w^2 \delta_{kh}$, where * denotes complex conjugate, δ_{kh} is the Kronecker symbol (1 if $k = h$ and 0 otherwise), and $E[\cdot]$ is the statistical mean (expectation). Considering only L of the N_s consecutive samples, i.e.,

$$\begin{aligned} \mathbf{y}(n) &= [y(n)\ y(n+1)\ \cdots\ y(n+L-1)]^\top \\ \mathbf{s}(n) &= [s(n)\ s(n+1)\ \cdots\ s(n+L-1)]^\top \\ \mathbf{w}(n) &= [w(n)\ w(n+1)\ \cdots\ w(n+L-1)]^\top \end{aligned} \tag{4}$$

(L is also called *smoothing factor*), the statistical covariance matrices of the signal and noise are given by

$$\mathbf{R}_y = E\left[[\mathbf{y}(n) - E[\mathbf{y}(n)]]\,[\mathbf{y}(n) - E[\mathbf{y}(n)]]^H\right] \tag{5}$$

$$\mathbf{R}_s = E\left[[\mathbf{s}(n) - E[\mathbf{s}(n)]]\,[\mathbf{s}(n) - E[\mathbf{s}(n)]]^H\right] \tag{6}$$

where ^{H} denotes conjugate transposition (Hermitian), and

$$\mathbf{R}_y = \mathbf{R}_s + \sigma_w^2 \mathbf{I}_L \tag{7}$$

under $\mathbb{H}_1$, with $\mathbf{I}_L \in \mathbb{R}^{L\times L}$ the identity matrix of size L.

If the signal $\mathbf{s}(n)$ is not present, $\mathbf{R}_s = \mathbf{0}$; hence, the off-diagonal elements of $\mathbf{R}_y$ are (ideally) zeros. If there is a signal and the samples are correlated (as usually happens in real signals), $\mathbf{R}_y$ is not a diagonal matrix; hence some of the off-diagonal elements of $\mathbf{R}_y$ should be nonzeros. Denote as r_{ij} the element of $\mathbf{R}_y$ at the ith row and jth column, and let

$$T_1 = \frac{1}{L}\sum_{i=1}^{L}\sum_{j=1}^{L} |r_{ij}|, \quad T_2 = \frac{1}{L}\sum_{i=1}^{L} |r_{ii}|. \tag{8}$$

From the discussion above, it follows that if there is no signal $T_1/T_2 = 1$, while if there is a signal $T_1/T_2 > 1$. Thus, the T_1/T_2 ratio can be used to detect the presence of the signal. In practice, the statistical covariance matrix can only be calculated using a limited number of signal samples.

In this paper, we use the following estimator

$$\hat{\mathbf{R}}_y = \begin{bmatrix} \lambda(0) & \lambda(1) & \cdots & \lambda(L-1) \\ \lambda^*(1) & \lambda(0) & \cdots & \lambda(L-2) \\ \vdots & \vdots & \ddots & \vdots \\ \lambda^*(L-1) & \lambda^*(L-2) & \cdots & \lambda(0) \end{bmatrix} \tag{9}$$

where

$$\lambda(l) = \frac{1}{N_s}\sum_{m=1}^{N_s - l} \left(y(m) - \overline{y}\right)\left(y(m+l) - \overline{y}\right)^* \tag{10}$$

is the sample autocovariance of the received signal as defined in [23] with $\overline{y} = \frac{1}{N_s}\sum_{m=1}^{N_s} y(m)$ its sample mean.[2] The final decision statistic is given by (8) in which the estimate (9) is used in place of $\mathbf{R}_y$.

[2]We have omitted normalization since it cancels out in the ratio statistics.

D. Maximum-Minimum Eigenvalue Detector

This algorithm is based on the same considerations done for the covariance matrix-based detector. It consists in evaluating the ratio between the maximum and minimum eigenvalues and compare it to a threshold [18]–[21].

The rationale for such a procedure is that, by looking at eq. (7), the covariance matrix under the two hypotheses is

$$\mathbf{R}_y = \begin{cases} \sigma_w^2 \mathbf{I}_L & \mathbb{H}_0 \\ \mathbf{R}_s + \sigma_w^2 \mathbf{I}_L & \mathbb{H}_1 \end{cases} \tag{11}$$

hence the maximum and minimum eigenvalues are given by

$$(\lambda_{max}, \lambda_{min}) = \begin{cases} (\sigma_w^2, \sigma_w^2) & \mathbb{H}_0 \\ (\rho_{max} + \sigma_w^2, \rho_{min} + \sigma_w^2) & \mathbb{H}_1 \end{cases} \tag{12}$$

where ρ_{max} and ρ_{min} are the maximum and minimum eigenvalues of $\mathbf{R}_s$, respectively. Then, the ratio $\lambda_{max}/\lambda_{min}$, which in fact coincides with the condition number of the matrix $\mathbf{R}_y$, should be 1 under $\mathbb{H}_0$ and larger than 1 under $\mathbb{H}_1$.

Again, in practice, the statistical covariance matrix (hence its eigenvalues) can only be estimated using a limited number of signal samples. In this paper we adopt (9) as estimator also for the maximum-minimum eigenvalue detector.

E. Discussion

In principle, the energy detector should be able to detect a weak signal since it assumes a perfect "knowledge" of the environment, i.e., the noise power level σ_w^2. However, having such a knowledge is difficult in real scenarios, since it may vary according to different parameters such as temperature, quality of the electronics, and frequency. Approaches such as the maximum-minimum eigenvalue and covariance matrix-based detection appear more interesting since they do not require any *a priori* knowledge. However, although they offer good performance in case of strong signals, such methods may not be able to deal with weak sources.[3]

In this work, we tackle the issues of the different thresholds and noise uncertainty present in these algorithms and design a novel method to detect the absence or presence of signals. As we will see in the next section, we propose the adoption of a machine learning technique (SVM) to create a better detector that does not require knowledge of the noise power, and is able to detect also weak signals.

III. SVM-based Spectrum Sensing: design

SVM, proposed by Vapnik in 1995 [24], is a supervised learning model with associated learning algorithms for classification. It provides a representation of the training sample as points in a space, mapped so that data are separated into a number of categories (classes). New data are classified according to which region of the space they fall in. Given a training set $\mathcal{T} = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \cdots, (\mathbf{x}_M, y_M)\}$, where $\mathbf{x}_i \in \mathbb{R}^n$, $y_i \in \mathbb{Y} = \{-1, 1\}$, $i = 1, \ldots, M$, the problem can be formalized as finding a real function $g(\mathbf{x})$ in $\mathbb{R}^n$ such that the *decision function* $f(\mathbf{x}) = \mathrm{sgn}(g(\mathbf{x}))$ will be able to predict the value of y for any $\mathbf{x}$, with $\mathrm{sgn}(\cdot)$ denoting the sign function [25]. Thus, solving a binary classification problem is equivalent to finding a criterion in order to separate the $\mathbb{R}^n$ space into two regions based on the training set $\mathcal{T}$. When $g(\mathbf{x})$ is restricted to be a linear function $g(\mathbf{x}) = \mathbf{w}^\top \mathbf{x} + b$, the corresponding method is referred to as a *linear classification machine*, with the hyperplane $\mathbf{w}^\top \mathbf{x} + b = 0$ separating $\mathbb{R}^n$ into two regions.

In this work SVM is applied to spectrum sensing, based on a particular vector of signal *features*. These features could be chosen in many different ways; the idea is to choose as feature vector the statistics of the different spectrum sensing algorithms discussed above, i.e.

$$\mathbf{x} = [x_1 \ x_2 \ x_3]^\top = \left[\frac{\lambda_{max}}{\lambda_{min}} \ \frac{T_1}{T_2} \ E\right]^\top \tag{13}$$

and their higher-order combinations, with the aim to promote stronger separability between the $\mathbb{H}_0$ and $\mathbb{H}_1$ regions; as mentioned in Sec. I, this approach can be justified on the basis of Cover's theorem on the separability of patterns [13]–[15]. In particular, we consider the feature vector extended up to the second order, i.e.,

$$\begin{aligned} \mathbf{x} =& [x_1 \ \cdots \ x_9]^\top \\ =& \left[\frac{\lambda_{max}}{\lambda_{min}} \quad \frac{T_1}{T_2} \quad E \right. \\ & \left. \frac{\lambda_{max}^2}{\lambda_{min}^2} \quad \frac{T_1^2}{T_2^2} \quad E^2 \quad \frac{\lambda_{max}}{\lambda_{min}}\frac{T_1}{T_2} \quad \frac{\lambda_{max}}{\lambda_{min}}E \quad \frac{T_1}{T_2}E \right]^\top \end{aligned}$$

as well as to the third order, i.e.,

$$\begin{aligned} \mathbf{x} =& [x_1 \ \cdots \ x_{19}]^\top \\ =& \left[\frac{\lambda_{max}}{\lambda_{min}} \quad \frac{T_1}{T_2} \quad E \right. \\ & \frac{\lambda_{max}^2}{\lambda_{min}^2} \quad \frac{T_1^2}{T_2^2} \quad E^2 \quad \frac{\lambda_{max}}{\lambda_{min}}\frac{T_1}{T_2} \quad \frac{\lambda_{max}}{\lambda_{min}}E \quad \frac{T_1}{T_2}E \\ & \frac{\lambda_{max}^2}{\lambda_{min}^2}\frac{T_1}{T_2} \quad \frac{\lambda_{max}^2}{\lambda_{min}^2}E \quad \frac{\lambda_{max}}{\lambda_{min}}\frac{T_1^2}{T_2^2} \quad \frac{T_1^2}{T_2^2}E \quad \frac{\lambda_{max}}{\lambda_{min}}E^2 \\ & \left. \frac{T_1}{T_2}E^2 \quad \frac{\lambda_{max}}{\lambda_{min}}\frac{T_1}{T_2}E \quad \frac{\lambda_{max}^3}{\lambda_{min}^3} \quad \frac{T_1^3}{T_2^3} \quad E^3 \right]^\top . \end{aligned}$$

We need to discuss a final aspect. In classical detection theory, probability of false alarm (P_{FA}) is a design parameter set according to a maximum tolerable level; then, the design goal is typically to maximize the probability of detection (P_D). In the proposed SVM-based approach, once the features vector $\mathbf{x}$ is computed for a given training set $\mathcal{T}$, the optimal parameters $\mathbf{w}^*$ and b^* uniquely identify the separation hyperplane

[3] In addition to their own limitations, most detectors cannot detect a weak signal due to uncertainty on the noise process. In particular, the formulas available in the literature to compute the threshold that guarantees a nominal probability of false alarm, are often available only for white Gaussian noise; unfortunately, in many cases signals are filtered, hence the filtered noise could be not "white" anymore, and more in general non-linearities change the statistical distribution of the noise process. Whitening operations are proposed in [17] and [19], but this further complication is anyway not able to make the noise perfectly white and Gaussian in real experiments.

Figure 1. Experimental testbed with one RTL-SDR and one HackRF One connected to a Linux personal computer.

$\mathbf{w}^{*\top}\mathbf{x} + b^* = 0$ used for classifying new data. Consequently, P_D and P_{FA} are determined only on the basis of the specific training set $\mathcal{T}$, hence cannot be easily tuned. To tackle this issue, we propose to consider the parallel sheaf of hyperplanes with respect to the optimal hyperplane, parametrized in d:

$$\text{Shifted hyperplane: } g(\mathbf{x}) = \mathbf{w}^{*\top}\mathbf{x} + b^* + d. \qquad (14)$$

By using this heuristic, we expect that moving the hyperplane towards the $\mathbb{H}_1$ region (far away from the origin), P_D and P_{FA} will decrease simultaneously — of course not necessarily at the same pace — while moving it towards the $\mathbb{H}_0$ region (close to the origin) P_D and P_{FA} will increase simultaneously. Clearly, a trade-off must be considered.

IV. Experimental Testbed and Results

A. Experimental Testbed

The proposed SVM-based detector has been implemented in MATLAB and tested on real data acquired by RTL-SDR, a low cost (sub-20$) and easy-to-use USB device that receives RF radio signals in the range from 25 MHz to 1.75 GHz [26]. Originally, these devices were designed as DVB-T (Digital Video Broadcast-Terrestrial) receivers, but then it was discovered that they could be used as generic (receive only) SDRs by simply putting them into a different mode. The front end of the RTL-SDR receives RF signals, downconverts them to baseband, digitizes them, and finally outputs the samples of the baseband signal across its USB interface. The RTL-SDR can be interfaced with MATLAB & Simulink through specific support packages and libraries [26].

The acquisition has been performed in an observation time of $t_o = 1$ s with a sampling frequency $f_s = 240$ kHz, at a given frequency f_0. The resulting $N_s = 240$ ksamples, which represent the radiofrequency signal filtered in a bandwidth of ± 120 kHz around f_0, are stored in the vector form of (2). The first 4096 samples of each signal acquisition are discarded since they are the initial samples for setting up the RTL-SDR.

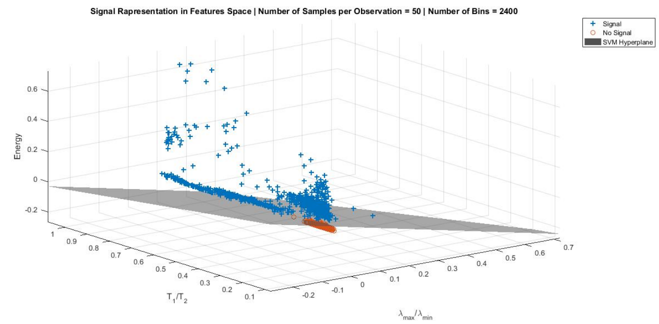

Figure 2. SVM training in 3D features space representation: red circles are data under $\mathbb{H}_0$, while blue crosses are data under $\mathbb{H}_1$; the learned hyperplane is shown in grey color.

Under $\mathbb{H}_1$ hypothesis, some signal has been transmitted using another SDR, called HackRF One [27], [28], interfaced with GNU Radio in Linux Xubuntu 16.10. The testbed is in Fig. 1.

B. Results

For the performance assessment we consider a training based on signals over different frequencies of the RTL-SDR. In particular, for the H_0 training set, acquisitions at the following frequencies have been performed in absence of detectable transmissions: 902 MHz (GSM), 433.9 MHz (remote control and key fobs), 569 MHz (TV broadcast UHF), 50 MHz (land mobile), 1500 MHz (amateur/land mobile) and 857 MHz (4G Uplink frequency). For the H_1 training set, conversely, the following frequencies have been used: 857 MHz (4G Uplink frequency), 816 MHz (4G Downlink frequency), 92.3 MHz (FM broadcast radio), 569 MHz with various transmission powers (FM signal transmitted with HackRF One), and 902 MHz (GSM frequency). The results are shown in Fig. 2.

The performance are evaluated by computing the receiver operating characteristics (ROC) curves, considering of course $\mathbb{H}_0$ and $\mathbb{H}_1$ signals not used for the SVM training. Several experiments have been conducted. Here we show one representative case: a signal at $f_0 = 902$ MHz for $\mathbb{H}_0$ with no transmission and a signal at $f_0 = 569$ MHz for $\mathbb{H}_1$ in which the transmitter has an attenuation factor so as to produce a weak signal. For evaluation purposes, to obtain the different combinations of P_D and P_{FA}, we adopt the heuristic in (14).

Fig. 3 shows the comparison between the proposed SVM-based approach and the three detectors presented in Sec. II. Remarkably, the obtained performance is better than the covariance-based detectors and almost attains the energy detector, without however requiring knowledge of the noise power; notice that, to obtain an estimate of the latter, the training data were used hence the threshold corresponds to the nominal P_{FA}. Fig. 3 also shows the comparison between the extended and the classical SVM-based approaches[4]. It is worth noticing that the extended SVM-based approach outperforms

[4]We cannot represent the feature space since it is not possible to visualize more than three dimensions.

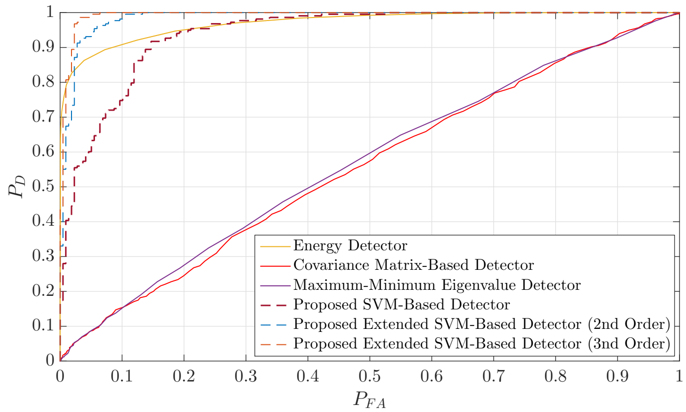

Figure 3. ROC curves comparison between the proposed SVM-based approaches and state-of-the-art competitors.

the energy detector, which additionally exploits knowledge of the noise power level. As a whole, SVM-based detectors can guarantee better performance without requiring such an information. Results for other frequencies are not shown since they are similar to the reported one, which as said represents a case of weak signal.

V. Conclusion

We have addressed the design of a novel spectrum sensing scheme based on SVM with feature space obtained from well-known spectrum sensing statistics and their higher-order combinations. Using this algorithm, the presence of a transmission over a certain frequency can be accurately detected without knowledge of the noise power. The proposed algorithm has been evaluated in real experiments through a software-defined radio testbed, showing that it can outperform classical spectrum sensing approaches also for weak signals.

Acknowledgement

The authors wish to thank Eng. Lucio Villani for his help with testbed development and experiments.

References

[1] Y.-C. Liang, K.-C. Chen, G. Y. Li, and P. Mahonen. Cognitive radio networking and communications: an overview. *IEEE Transactions on Vehicular Technology*, 60(7):3386–3407, Jun 2011.

[2] G. De Cubber, R. Shalom, A. Coluccia, O. Borcan, R. Chamrád, T. Radulescu, E. Izquierdo, and Z. Gagov. The safeshore system for the detection of threat agents in a maritime border environment. In *IARP Workshop on Risky Interventions and Environmental Surveillance*, Les Bons Villers, Belgium, 18-19 May 2017.

[3] A. Coluccia et al. Drone-vs-bird detection challenge at IEEE AVSS2017. In *14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pages 1–6, Lecce, Italy, 29 Aug-1 Sep 2017.

[4] F. Zabini and A. Conti. Inhomogeneous poisson sampling of finite-energy signals with uncertainties in R^d. *IEEE Transactions on Signal Processing*, 64(18):4679–4694, Sep 2016.

[5] A. Kipnis, Y. C. Eldar, and A. J. Goldsmith. Fundamental distortion limits of analog-to-digital compression. *IEEE Transactions on Information Theory*, 64(9):6013–6033, Sep 2018.

[6] 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*, Doha, Qatar, 1-6 Apr 2016.

[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, Nov 2013.

[8] O. P. Awe, Z. Zhu, and S. Lambotharan. Eigenvalue and support vector machine techniques for spectrum sensing in cognitive radio networks. In *2013 Conference on Technologies and Applications of Artificial Intelligence*, pages 223–227, Taipei, Taiwan, 6-8 Dec 2013.

[9] S. U. Jan, V. H. Vu, and I. S. Koo. Performance analysis of support vector machine-based classifier for spectrum sensing in cognitive radio networks. In *2018 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)*, pages 385–389, Zhengzhou, China, 18-20 Oct 2018.

[10] A. Tiwari, H. Chenji, and V. Devabhaktuni. Comparison of statistical signal processing and machine learning algorithms for spectrum sensing. In *2018 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6, Abu Dhabi, United Arab Emirates, 9-13 Dec 2018.

[11] H. Xue and F. Gao. A machine learning based spectrum-sensing algorithm using sample covariance matrix. In *2015 10th International Conference on Communications and Networking in China (ChinaCom)*, pages 476–480, Shanghai, China, 15-17 Aug 2015.

[12] Y. Xu, P. Cheng, Z. Chen, Y. Li, and B. Vucetic. Mobile collaborative spectrum sensing for heterogeneous networks: A bayesian machine learning approach. *IEEE Transactions on Signal Processing*, 66(21):5634–5647, Nov 2018.

[13] T.M. Cover. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Transactions on Electronic Computers*, pages 326–334, Jun 1965.

[14] Simon Haykin. *Neural Networks and Learning Machines (Third ed.)*. Upper Saddle River, New Jersey: Pearson Education Inc., 2012.

[15] Leo Breiman. Statistical modeling: The two cultures. *Statistical Science*, 2001.

[16] Y. Zeng and Y.-C. Liang. Covariance based signal detections for cognitive radio. In *2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, pages 202–207, Dublin, Ireland, 17-20 Apr 2007.

[17] Y. Zeng and Y.-C. Liang. Spectrum-sensing algorithms for cognitive radio based on statistical covariances. *IEEE Transactions on Vehicular Technology*, 58(4):1804–1815, Sep 2009.

[18] Y. Zeng and Y.-C. Liang. Maximum-minimum eigenvalue detection for cognitive radio. In *IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pages 1–5, Athens, Greece, 3-7 Sep 2007.

[19] Y. Zeng and Y.-C. Liang. Eigenvalue-based spectrum sensing algorithms for cognitive radio. *IEEE Transactions on Communications*, 57(6):1784–1793, Jun 2009.

[20] P. Alvarez, N. Pratas, A. Rodrigues, N. R. Prasad, and R. Prasad. Energy detection and eigenvalue based detection: An experimental study using GNU radio. In *14th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, pages 1–5, Le Quartz, Brest, France, 3-7 Oct 2011.

[21] F. Penna, R. Garello, D. Figlioli, and M. A. Spirito. Exact non-asymptotic threshold for eigenvalue-based spectrum sensing. In *4th International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, pages 1–5, Hanover, Germany, 22-24 Jun 2009.

[22] H. Urkowitz. Energy detection of unknown deterministic signals. *Proceedings of the IEEE*, 55(4):523–531, 1967.

[23] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel. *Time Series Analysis: Forecasting and Control (3rd ed.)*. Englewood Cliffs, NJ: Prentice Hall, 1994.

[24] C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, Feb 1995.

[25] N. Deng, Y. Tian, and C. Zhang. *Support vector machines: optimization based theory, algorithms, and extensions*. CRC press, 2012.

[26] R. W. Stewart, K. W. Barlee, D. S. W. Atkinson, and L. H. Crockett. *Software defined radio using MATLAB & Simulink and the RTL-SDR*. Strathclyde Academic Media, 2015.

[27] Great Scott Gadgets. Hackrf one, 2017.

[28] M. Ossmann. Hackrf one-an open source sdr platform, 2016.