Research Article

Learning-Based Spectrum Sensing for Cognitive Radio Systems

Yasmin Hassan, Mohamed El-Tarhuni, and Khaled Assaleh

Department of Electrical Engineering, American University of Sharjah, P.O. Box 26666, Sharjah, UAE

Correspondence should be addressed to Mohamed El-Tarhuni, mtarhuni@aus.edu

This paper presents a novel pattern recognition approach to spectrum sensing in collaborative cognitive radio systems. In the proposed scheme, discriminative features from the received signal are extracted at each node and used by a classifier at a central node to make a global decision about the availability of spectrum holes for use by the cognitive radio network. Specifically, linear and polynomial classifiers are proposed with energy, cyclostationary, or coherent features. Simulation results in terms of detection and false alarm probabilities of all proposed schemes are presented. It is concluded that cyclostationary-based schemes are the most reliable in terms of detecting primary users in the spectrum, however, at the expense of a longer sensing time compared to coherent based schemes. Results show that the performance is improved by having more users collaborating in providing features to the classifier. It is also shown that, in this spectrum sensing application, a linear classifier has a comparable performance to a second-order polynomial classifier and hence provides a better choice due to its simplicity. Finally, the impact of the observation window on the detection performance is presented.

1. Introduction

In the past few years, there have been remarkable developments in wireless communications technology leading to a rapid growth in wireless applications. However, this dramatic increase in wireless applications is severely limited by bandwidth scarcity. Traditionally, fixed spectrum assignments, in which frequency bands are statically assigned to licensed users are employed. The static spectrum allocation prevents from assigning vacant spectrum bands to new users and services. Further, spectrum occupancy measurements have shown that some licensed bands are significantly underutilized. For example, the Spectral Policy Task Force reported that radio channels are typically occupied 15% of the time [1]. Hence, the limitation in the available spectrum bands occurs mainly due the underutilization of available spectrum resulting from the inefficient static allocation techniques. This underutilization of available spectrum resources has led regulatory bodies to urge the development of dynamic spectrum allocation paradigms, called cognitive radio (CR) networks.

A CR network senses the operating environment for vacant spectrum opportunities and dynamically utilize the available radio resources [2, 3]. In CR technology, unlicensed (secondary) users are allowed to share the spectrum originally assigned to licensed (primary) users. Hence, frequency bands that are legally assigned to primary users are exploited by secondary users when primary users are idle. However, primary users have the right to occupy their assigned bands whenever needed. Consequently, secondary users should be aware of the variations in the surrounding environment and should be ready to adjust their operating parameters accordingly in order to make a productive usage of the spectrum [4].

Secondary users in CR networks are restrained by the condition of providing adequate protection to primary users. Hence, secondary users need to employ efficient spectrum sensing techniques that ensure the quality of service for primary users and exploit all dynamic spectrum sharing chances. That is to say, in order to facilitate dynamic spectrum access in licensed bands, effective spectrum sensing algorithms need to be developed whereby high reliability along with efficient utilization is achieved.

Spectrum sensing approaches that are commonly considered in CR applications include energy detection, cyclostationary feature detection, and coherent detection [2, 4–6]. Based on the prior knowledge a secondary user has about

primary users, a specific technique would be more appropriate. For instance, if a priori information about a primary user signal is known by secondary users, coherent detection can be utilized. Coherent detection uses features such as synchronization messages, pilots, preambles, midambles, and spectrum spreading sequences. When these patterns are known at the CR network, sensing is performed by correlating the incoming signal with the known patterns [6]. Coherent sensing based on pilot detection was implemented experimentally in [7]. On the other hand, when CRs have very limited information about the primary signal, energy detection is used. Another reason for using energy detection in spectrum sensing applications is the low complexity involved. However, the performance of energy detection in terms of the ability to detect primary signals is degraded, especially in low signal-to-noise ratio (SNR) conditions.

Another approach to spectrum sensing is based on cyclostationary detection to sense the presence of a primary user by exploiting cyclostationary features exhibited by the statistics of the primary signal [8]. In cyclostationary detection, the spectral correlation function (SCF) of a modulated signal is analyzed to decide on the presence of primary signal in the target spectrum. Cyclostationary feature detection based on multicycle detection has been proposed in [9, 10], where the cyclostationarity is detected at multiples of the cycle frequency. In orthogonal frequency division multiplexing (OFDM) systems, a cyclic prefix is intentionally inserted as a guard interval, which could be used to detect cyclostationarity of incumbent primary signals [11, 12]. Furthermore, the OFDM waveform could be modified in order to generate specific signatures at certain frequencies [13] such that the cyclic features created by these signatures are then extracted via cyclostationary detection to achieve an effective signal identification mechanism.

In order to preserve the quality of service for primary users, the interference caused by secondary users needs to be maintained below an acceptable level. Hence, reliable spectrum sensing needs to be performed by secondary users to detect the presence of a primary user, especially under shadowing and fading effects. Collaboration among spatially displaced secondary users is, hence, required to mitigate such effects without requiring excessively long detection times. In this case, several CR nodes utilize the spatial diversity gain provided by cooperative spectrum sensing to achieve better performance in fading environments [4, 9, 10, 14, 15].

In this work, we propose a collaborative spectrum sensing approach in CR applications. Specifically, we utilize classification techniques used in pattern recognition applications to identify the available and busy bands in the radio spectrum. Previously, pattern recognition techniques were used mainly in signal classification for determining type of modulation rather than spectrum sensing [16–18]. The proposed pattern recognition scheme represents a centralized cooperative CR network, whereby the decision of spectrum availability is made at a central node after collecting spectral sensing information from all collaborating users. Sensing information is subjected to a classifier model that outputs a global decision regarding the availability of the target spectrum band. Polynomial classifiers are proposed in this work as classifier models, in which first- and second-order expansions are investigated. Three spectrum sensing techniques are implemented to provide informative features to the classifier about the surrounding environment. Spectrum sensing techniques used for feature extraction can be classified into parametric and nonparametric. Nonparametric detection includes energy detection where the cognitive network does not have a priori knowledge on the primary users' signals. On the other hand, in parametric detection, cyclic features characterizing primary signals and prior knowledge of synchronizing preamble patterns are utilized. The parametric detection schemes include coherent detection and cyclostationary feature detection.

Many of the collaboration techniques in the prior work implement maximum ratio combining, likelihood ratio test, or hard decision rules, such as AND logic operation and one-out-of-n rule [4, 5, 19, 20]. Cooperative sensing based on energy detection has been proposed in [4], in which linear combination of local test statistics from multiple users is utilized in the decision making. The performance of a cyclostationary-based spectrum sensing cooperative CR system was considered in [20, 21], where binary decisions with different fusion rules of the secondary user's decisions using cyclic detectors were compared. Moreover, multiple user single-cycle detectors are proposed to accommodate secondary user collaboration [9], where different cyclic frequencies are utilized by different users and combined to make a global decision. In [10], the summation of local tests statistics of secondary users is employed as the fusion rule when multicycle detection is performed by CRs. Finally, cooperation based on hard decision rules was investigated with coherent detection in [7].

The contributions of this paper are as follows. The problem of collaborative spectrum sensing in CR networks is investigated from a new perspective based on a pattern recognition approach. More specifically, polynomial classifiers are used in this work. The design, validation and evaluation of first- and second-order polynomial classifiers are presented. The parameters of these classifiers are optimized based on the signal strength of the individual secondary users in a collaborative manner. The performance in terms of false alarm rate and detection probability under low SNR conditions has been thoroughly examined and analyzed. Comprehensive performance evaluation of energy-based detection is provided. Finally, extensive simulations are performed to evaluate the performance of the proposed classifiers with parametric spectrum sensing schemes, where carrier frequency and synchronization preamble patterns are assumed to be known at the CR network. The results of this investigation were partially presented in [22, 23].

The rest of the paper is organized as follows: in Section 2, we introduce the signal model and the proposed cooperative spectrum sensing scheme. In Section 3, different feature extracting techniques are presented. The polynomial classifier structure is developed in Section 4. Simulation results and discussions are given in Section 5. Finally, Section 6 concludes the paper. All notations and symbols used in this paper are explained in Table 1.

TABLE 1: List of notations.

Notation/symbol	Explanation
$r_j[n]$	Received signal by user j
$x[n]$	Transmitted primary signal
g_j	Channel gain coefficients
$\eta_j[n]$	Zero-mean additive white Gaussian noise (AWGN) at user j's end
N	Number of users in cognitive network
$\mathbf{d}$	N-dimensional feature vector consisting of features extracted from different CRs
M	Number of samples observed to make a decision (observation window size)
$R_x^\alpha(\tau)$	Cyclic autocorrelation function
$S_x^\alpha[k]$	Spectral correlation density function
$R_{jc}[k]$	Cross-correlation between the received signal and the preamble sequence
$\boldsymbol{\varphi}(\mathbf{d})$	l-dimensional vector consisting of the monomials of a another vector $\mathbf{d}$
$\mathbf{w}_i$	The classifier model parameters for class i
$\mathbf{D}_{\text{train}}$	$K \times N$ matrix, where N is the dimensionality of the input feature vectors (provided by N CR users) and K is the number of feature vectors used in the training process
$\mathbf{M}$	A $(K \times l)$ matrix representing the polynomial expansion of elements in training data set $\mathbf{D}_{\text{train}}$
$\mathbf{t}_i$	An ideal target vector representing the ideal channel state (ON or OFF)
Y_d	Binary global decision on channel availability

2. Signal Model and System Description

We consider dynamic spectrum allocation in a collaborative CR network with the structure illustrated in Figure 1. The primary user and CR network are assumed to coexist within the same geographical area. The CR network consists of N users with a central node that detects the presence of primary signals and decides on the channel availability. CRs temporarily access the underutilized licensed frequency bands, without conflict with primary spectrum holders' usage.

The binary hypothesis test for spectrum sensing is formulated as

$$r_j[n] = \begin{cases} g_j x[n] + \eta_j[n] : & H_1 \\ \eta_j[n] : & H_0 \end{cases} \quad \text{for } j = 1, 2, \ldots, N, \quad (1)$$

where $r_j[n]$ represents the received signal by the jth CR user at the nth instant of time, and $x[n]$ denotes the primary user transmitted signal. H_1 represents the hypothesis of an occupied spectrum, while H_0 corresponds to an idle spectrum. The received signal at the jth user is corrupted by a zero-mean additive white Gaussian noise (AWGN), $\eta_j[n]$ with variance σ_j^2. The primary signal passes through a wireless channel to reach the jth CR user with a channel gain g_j. The wireless channel is modeled as a flat channel with slow fading. Each channel has a complex valued coefficient with Rayleigh distributed magnitude and uniformly distributed phase over the range $[0, 2\pi]$. The channel coefficients of different CRs in the network are assumed to be constant over a number of received signal symbols, that is, slow fading, and are also assumed to be independent and identically distributed.

In this paper, spectrum sensing in CR networks is formulated as a pattern recognition problem. Generally speaking, pattern recognition is used to classify a given set of data into several different categories. A pattern recognition system assigns an input signal to one of a number of known categories based on features derived to emphasize commonalities between those signals. A generic term that is used to describe input signals that need to be classified in a recognition system is *patterns*. Usually, patterns may not be useful for classification, and hence they need to be processed to acquire more useful input to the classifier [24, 25]. This processed information is called *features*. In supervised learning, a labeled training set of feature vectors is processed through the classification algorithm to determine the classifier model parameters. These parameters are used in predicting the class of new data that have not been seen during the learning phase. In this paper, supervised pattern recognition is utilized at the CR base station (CRBS) to classify available spectrum holes such that maximum detection is achieved with a desired false alarm rate.

In the proposed system, secondary users are constantly sensing the target spectrum band for primary signal presence. Within a secondary user receiver, discriminative features are extracted from the sensed signal. The extracted features from the difference secondary users are transmitted to the CRBS through a relatively low data rate control channel. This control channel is used for exchanging information between CRs and CRBS. At the CRBS, a decision about the spectrum availability is made based on a pattern recognition classifier that is previously trained. The block diagram of the proposed system is depicted in Figure 2 showing the signal flow of the CR inputs through feature extraction
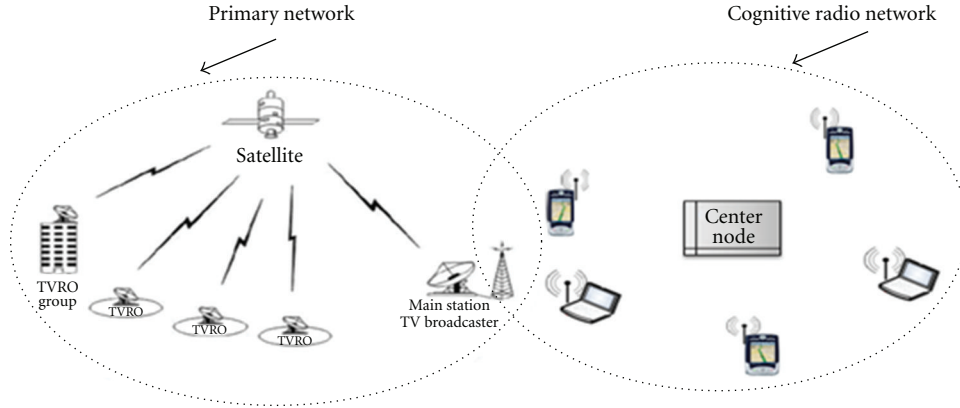
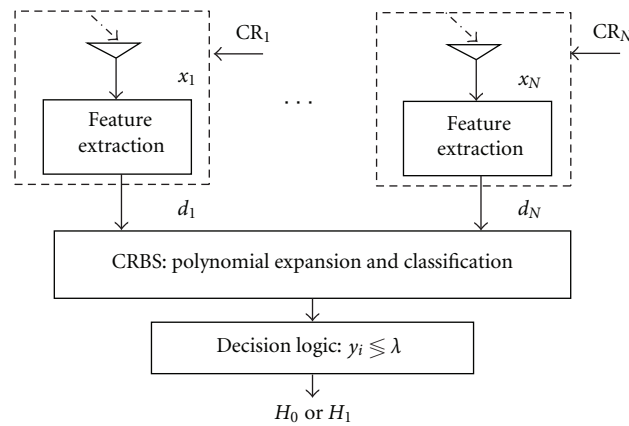FIGURE 1: An example of a centralized CR network.

FIGURE 2: Pattern recognition CR system model.

and classification leading to a decision about the spectrum availability at the CRBS.

The first step is spectrum sensing which involves input data acquisition by processing the signals received by antennas at different CR receivers. The received signal of the jth user is assumed to follow the mathematical model described in (1). These signals are transformed into multidimensional feature vectors that compactly characterize the sensed signals.

When secondary users in a CR network have no prior information about the transmitted primary signal, the energy of the received signal is used at the feature extraction stage and utilized by the classifier to discriminate between the noise only and primary signal present cases. On the other hand, if prior information, such as carrier frequency and synchronizing patterns, is known about the primary user's signal, feature extraction will be achieved by either exploiting cyclic features present in the signal or through coherent detection. Features extracted by any of the mentioned detection schemes will exhibit certain patterns when the spectrum is occupied by a primary user that are different from the patterns extracted when only noise is present in the spectrum. The difference between these patterns will be exploited as discriminative input data to

the pretrained classifier for decision making. The following section discusses the different feature extraction schemes used in this paper.

3. Feature Extraction Techniques

In this work, three different feature extraction schemes have been used, namely, noncoherent energy-based features, cyclostationary-based features, and coherent detection-based features.

3.1. Energy-Based Feature Extraction. Energy detection is one of the most commonly used techniques in spectrum sensing due to its low computational complexity and simple implementation. It does not require any prior knowledge of the primary users' signal; hence, it is considered as a nonparametric detection scheme. The classification system identifies spectrum availability relying on the energy of the received signal over an observation period. However, the task of detecting the signal becomes very challenging under low SNR levels and fading channels [5, 14]. As a preprocessing step, the received signal by the jth secondary user, $r_j(t)$, which follows the model specified in (1), is filtered according to the desired frequency band to obtain $y_j(t)$. The sampled
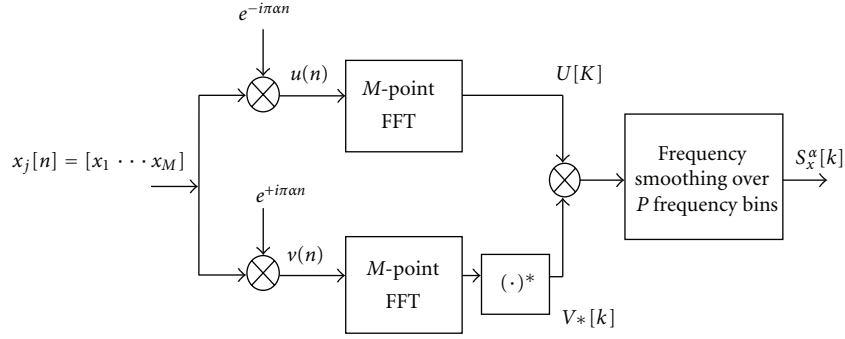
FIGURE 3: Cyclostationary feature detection estimator.

version of the signal is used to extract one representative energy feature d_j defined as

$$d_j = \frac{1}{M} \sum_{n=1}^{M} \left| y_j[n] \right|^2, \tag{2}$$

where M is the length of the observation period in samples. Accordingly, a feature vector for this observation period is constructed from the N users' signals such as $\mathbf{d} = [d_1 \cdots d_N]^T$, where T is the vector transpose operation. This feature vector will be presented to the classifier in the classification stage. A new set of energy features is obtained every observation window.

3.2. Cyclostationary-Based Feature Extraction. Most communication signals can be modeled as cyclostationary random processes, as they are usually characterized by built-in periodicities in their mean and autocorrelation. These underlying periodicities arise from the use of sinusoidal carriers, repeating spreading codes, pulse trains, or cyclic prefixes in signal transmission. On the other hand, noise signals do not exhibit such periodicity characteristics. Hence, the sensing of the spectrum availability can be based on the detection of the signal periodicity.

A binary phase shift keying (BPSK) that digitally modulated signal $x(t)$ with symbol duration T_0 has a cyclic autocorrelation function (CAF) defined as

$$R_x^\alpha(\tau) \triangleq \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} \left(x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi\alpha t} \right) dt, \tag{3}$$

where τ is a nonzero delay, and $\alpha \in \{0, \pm 1/T_0, \pm 2/T_0, \ldots\}$. $R_x^\alpha(\tau)$ represents the Fourier transform of the delay product $x(t + \tau/2)x^*(t - \tau/2)$ evaluated at the frequencies in α. The signal $x(t)$ is said to contain second order periodicity if and only if the Fourier transform of $R_x^\alpha(\tau)$ has discrete spectral lines at nonzero frequencies $\alpha \neq 0$ [26, 27].

Cyclostationarity of a signal leads to the presence of specific patterns in the spectrum of the signal, which can be examined using the so-called spectral correlation density function (SCD) [27, 28] defined as the Fourier transform of its CAF as

$$S_x^\alpha(f; \alpha) = \int_{-\infty}^{\infty} R_x^\alpha(\tau) e^{-j2\pi f \tau} d\tau. \tag{4}$$

The SCD is the cross-correlation function between the spectral translates of the signal at $f \pm \alpha/2$, for some cyclic frequency α. This function is smoothed using a frequency smoothing vector (window) W with P frequency bins. Since discrete-time samples of the received are used in estimating the SCD, it is possible to show that the SCD is given by

$$S_x^\alpha[k] = \frac{1}{P} \sum_{v=-(P-1)/2}^{(P-1)/2} X\left[k + \frac{\alpha}{2} + v\right] X^*\left[k - \frac{\alpha}{2} + v\right] W(v), \tag{5}$$

with

$$X[k] = \sum_{n=0}^{M-1} x[n] e^{-j2\pi kn/M}, \tag{6}$$

where $k = 0, 1, \ldots, M - 1$, and M is the number of samples over which the spectrum of the received signal is calculated (FFT length). The SCD estimation is implemented to obtain the cyclic feature extraction receiver structure as shown in Figure 3. The SCD of the received signal will depend on the presence or absence of the primary transmitted signal according to the following:

$$S_{r_j}^{\alpha_0}[k] = \begin{cases} \left| g_j \right|^2 S_x^{\alpha_0}[k] + S_{\eta_j}^{\alpha_0}[k] : & H_1, \\ S_{\eta_j}^{\alpha_0}[k] : & H_0, \end{cases} \tag{7}$$

where $S_{\eta_j}^{\alpha_0}[k]$ is the SCD of the AWGN only at the jth CR user, and $S_x^{\alpha_0}[k]$ is the SCD of the transmitted primary signal at some cyclic frequency $\alpha = \alpha_0$. Since the AWGN is a wide sense stationary process and does not possess second-order cyclostationarity, it will not have a peak at any cyclic frequency. On the other hand, the band-pass BPSK signal exhibits second-order cycle frequencies at $\alpha_0/2 = f_c + m/T_0$ [2, 9], for some integer m. Since the strongest spectral lines of a BPSK signal appear at $= f_c$, the strongest cyclic components

are observed at $\alpha_0/2 = f_c$ [9]. Therefore, the local decision variable for cyclostationary detection from jth CR user is chosen to be the value of the SCD in (7) evaluated at $k = 0$. This value is used as the jth element in the feature vector $\mathbf{d} = [d_1 \cdots d_N]^T$ that will be presented to the classifier in the classification stage, where $d_j = S_{r_j}^{\alpha_0}[k]$ with $k = 0$.

3.3. Coherent Based Feature Extraction.

Another sensing scheme investigated in this work is feature extraction using coherent detection. Coherent detection is performed by demodulating the primary user's signal, which requires a priori information of the primary signal such as packet format, control, or synchronization sequences. [6]. If the synchronizing preamble patterns are known at the CR network end, coherent sensing can be exploited by correlating the incoming signal with the known patterns. This is effectively correlating the signal with itself resulting in an autocorrelation function that peaks at zero delay when the primary user is present. Primary users are assumed to use a frame size of M bits with an L-bit synchronization preamble referred to as $x_p[n]$. Accordingly, the cognitive users will be acquiring data during the preamble period (i.e., L bits every M bits). The received signal from jth CR user, $r_j(t)$, is cross-correlated with the preamble sequence over the preamble length to obtain

$$R_{jc}[k] = \begin{cases} \left\{ |g_j|^2 \sum_{n=1}^{L} |x_p[n]|^2 + \mathrm{Re}\left\{ \sum_{n=1}^{L} x_p[n]\eta_j^*[n] \right\} \right\} : & H_1 \\ \mathrm{Re}\left\{ \sum_{n=1}^{L} x_p[n]\eta_j^*[n] \right\} & H_0. \end{cases}$$

(8)

Under hypothesis H_1, $R_{jc}[k]$ becomes an autocorrelation of the transmitted preamble having a strong peak at $k = 0$. However, under H_0, $R_{jc}[k]$ will not exhibit strong peaks. The peak value of $R_{jc}[k]$ is used as the jth element in the feature vector $\mathbf{d} = [d_1 \cdots d_N]^T$ that will be presented to the classifier in the classification stage, where $d_j = \max\{|R_{jc}[k]|\}$.

4. Polynomial Classifiers

In this paper, we use polynomial classifiers (PCs) as the classification models to decide on the spectrum availability. Polynomial classifiers have shown improved recognition performance with lower computational complexity as compared to other recognitions methods, such as neural networks and hidden Markov models [25, 29]. Furthermore, polynomial classifiers deal with simple mathematical operations such as multiplication and summation, which makes them suitable for practical implementation in digital signal processing algorithms. The principle of a polynomial classifier is the expansion of the input feature space into a higher dimensional space that is linearly separable [30].

Consider an input pattern $\mathbf{d} = [d_1 \cdots d_N]^T$, where N is the number of features and T represents the transpose operation. The rth order polynomial classifier first performs a vectorial mapping of the N-dimensional feature vector, $\mathbf{d}$, into an l-dimensional vector $\boldsymbol{\varphi}(\mathbf{d})$. The elements of $\boldsymbol{\varphi}(\mathbf{d})$ are monomials of the form [24, 25]

$$\begin{aligned} \boldsymbol{\varphi}(\mathbf{d}) = \big[&1, d_1, d_2, \ldots d_N, d_1^2, d_1 d_2, d_1 d_3, \ldots, d_1 d_N, d_1^3, d_1^2 d_2, \\ &\ldots d_1^2 d_N, \ldots, d_1^r, d_1^r d_2, \ldots d_1^r d_N, \\ &d_2^2, d_2 d_3, \ldots, d_2 d_N, d_2^3, d_2^2 d_1, d_2^2 d_3, \ldots d_2^2 d_N, \\ &\ldots, d_2^r, d_2^r d_1, \ldots, d_2^r d_N, \ldots \big]^T. \end{aligned}$$

(9)

Then, the output scorer y_i is obtained at the output layer after linearly combining the expansion terms $\boldsymbol{\varphi}(\mathbf{d})$ using

$$y_i = \mathbf{w}_i^T \boldsymbol{\varphi}(\mathbf{d}),$$

(10)

where $\mathbf{w}_i$ is the model (weights) of class i. The dimensionality of the expanded vector $\boldsymbol{\varphi}(\mathbf{d})$ can be expressed in terms of the polynomial order and the dimensionality of the input vector $\mathbf{d}$. The design of the classifier comprises of two stages, namely, training and testing.

4.1. Training.

The training process involves finding the optimal model parameters that best map a multidimensional input sequence to a corresponding one-dimensional target sequence. The model is designed to classify between two different classes, H_i for $i = \{0, 1\}$, corresponding to the binary hypotheses in (1). The multidimensional input sequence $\mathbf{D}_{\text{train}}$ is a $K \times N$ matrix, where N is the dimensionality of the input feature vectors (provided by N CR users) and K is the number of feature vectors used in the training process. The training matrix $\mathbf{D}_{\text{train}}$ is given by

$$\mathbf{D}_{\text{train}} = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1N} \\ \vdots & & \ddots & \vdots \\ d_{K1} & d_{K2} & \cdots & d_{KN} \end{bmatrix}.$$

(11)

The one-dimensional target vector $\mathbf{t}_i = [t_{i_1} \ t_{i_2} \cdots t_{i_K}]^T$ for $i = \{0, 1\}$ consists of K elements where $t_{i_z} = 1$ if the corresponding zth feature vector belongs to class i, and $t_{i_z} = 0$ if the corresponding zth feature vector does not belong to class i, for $z = 1, 2, \ldots, K$.

The training vectors are expanded into their polynomial terms as defined in (9) resulting in a model training $\mathbf{M}$ data set of size $(K \times l)$ that is defined by

$$\mathbf{M} = \big[\boldsymbol{\varphi}(\mathbf{d}_1) \ \boldsymbol{\varphi}(\mathbf{d}_2) \ \cdots \ \boldsymbol{\varphi}(\mathbf{d}_K) \big]^T.$$

(12)

Once training feature vectors are expanded into their polynomial basis terms, the polynomial classifier is trained to find an optimum set of weights, $\mathbf{w}_i^{\text{opt}}$, that minimizes the Euclidian distance between the ideal target vector $\mathbf{t}_i$ and the corresponding outputs of the classifier using the mean-squared error criterion to get

$$\mathbf{w}_i^{\text{opt}} = \arg\{\min_{\mathbf{w}} \|\mathbf{M}\mathbf{w} - \mathbf{t}_i\|_2\}.$$

(13)

The problem of (13) can be solved using the method of normal equation to explicitly obtain the optimal model for the two-class spectrum sensing problem as [25, 29]

$$\mathbf{w}_i^{\text{opt}} = \left(\mathbf{M}^T\mathbf{M}\right)^{-1}\mathbf{M}^T\mathbf{t}_i. \tag{14}$$

4.2. Testing. In the testing stage, novel feature vectors $\mathbf{d}_{\text{test}}$ are used to represent the testing data set. The features are initially expanded into their basis terms $\boldsymbol{\varphi}(\mathbf{d}_{\text{test}})$ and then presented to the trained models $\{\mathbf{w}_0^{\text{opt}}, \mathbf{w}_1^{\text{opt}}\}$ to obtain the corresponding set of scores $\{y_i\}$ as

$$\{y_i\} = \boldsymbol{\varphi}(\mathbf{d}_{\text{test}})\, \mathbf{w}_i^{\text{opt}} \quad \text{for } i = 0, 1. \tag{15}$$

Accordingly, we assign the testing feature vector to hypothesis H_i that satisfies [25]

$$Y_d = \arg_{i}\{\max\{y_i\}\}. \tag{16}$$

Ideally, the output from the classifier model, for a certain input feature vector, should be one when the spectrum is occupied and zero when the spectrum is idle as we apply it to the corresponding model $\mathbf{w}_i^{\text{opt}}$. However, when new input data are fed to the classifier, the output has values varying around one for hypothesis H_1 and values varying around zero for H_0 and vice versa. In order to achieve a desired level of constant false alarm rate, a threshold needs to be defined to separate the two classes instead of just comparing different models output scores.

An iterative algorithm is applied at the training stage to search for the threshold for different signal levels that achieves a specific false alarm rate as follows. First, the output score is computed by subjecting a validation data set (with known class labels) to the model $\mathbf{w}_1^{\text{opt}}$. The threshold is initialized to $\lambda = 0.5$, such that the global decision variable $Y_d = 1$ if $y_1 > \lambda$ and $Y_d = 0$ otherwise. The false alarm rate is then estimated by comparing the output decisions of all validation feature vectors to the ideal output $\mathbf{t}_1$. A false alarm will be declared when the output decision is one indicating that the spectrum is busy while the ideal output $\mathbf{t}_1$ is zero indicating that the actual spectrum is available. The threshold λ is incremented or decremented with a small value such that the desired false alarm rate is achieved with a specified accuracy, for example, a mean-squared error of less than 1%. The above steps are repeated for the validation data with different received SNR levels to form a lookup table that could be used when new test data is received. Note that the threshold setting operation, in addition to the training process, is performed offline. The training and validation data sequences are retrieved from a database that is maintained at the CRBS for offline training and validation.

5. Simulation Results

In this section, the performance of a first-order polynomial classifier (known also as linear classifier (LC)) and second-order polynomial classifiers (PCs) using the previously discussed feature extraction methods is evaluated. A band pass

BPSK primary signal is used when cyclostationary feature detection is utilized, while antipodal baseband signaling with $S(k) = \pm 1$ in the case of coherent detection. To emulate a more challenging and practical situation, we assume the distance between the CR network and the primary transmitter is relatively large; hence, the average received SNR_{avg} is in the low SNR range, that is, $\text{SNR}_{\text{avg}} \leq 0\,\text{dB}$. In addition, the jth CR receives a signal with a signal-to-noise ratio SNR_j that depends on the ith CR's proximity from the primary user. To account for signal shadowing, SNR_j follows a log-normal distribution with a variance $\sigma^2 = 4\,\text{dB}$ and a mean equivalent to SNR_{avg}. The small-scale channel variations follow a flat Rayleigh fading model. It is also assumed that the channel variation is relatively slow compared to the bit duration (slow fading model). We remark that the simulation parameters were used for illustrative purposes, and other values could be used without loss of generality.

5.1. Energy-Based Feature Extraction. Energy detection is performed at the various secondary users, and the extracted decision variables are provided to the recognition model at the CRBS. The probability of detection achieved by the LC and PC at different average received signals levels is presented in Figure 4. The results are obtained for a window size of 200 bits and a target false alarm (P_f) of 10%. The value of the false alarm rate was chosen to be consistent with the IEEE 802.22 requirements for CR networks [3]. It is interesting to notice that although the PC requires more memory and computational complexity to perform the expansion operation, it does not improve the detection probability performance compared to the LC. Hence, it is recommended to use an LC since it provides good performance with less required memory space and computational cost resulting in making faster decisions about the availability of the spectrum. Moreover, the advantage of cooperative sensing compared to single-radio-based sensing is demonstrated by the improvement in the detection performance as the number of secondary users contributing to signal classification is increased. For instance, a received SNR of around $-9\,\text{dB}$ is appropriate to reach a detection probability of 90% with three CRs, while a received signal with an average SNR of around $-6.5\,\text{dB}$ is required to achieve the same detection rate with one CR, resulting in a 2.5 dB gain which improves the ability to avoid interfering with weak primary users. It is notable from Figure 4 that the enhancement in performance diminishes as the number of receivers collaborating in global decision increases.

The probability of detection results of the energy detector for a received signal with $\text{SNR}_{\text{avg}} = -5\,\text{dB}$, $N = 3$ users, and $P_f = 10\%$ is depicted in Figure 5 for both the LC and PC as a function of the observation window size. It is evident that the detection performance is highly affected by the window size over which the local decision variables are estimated. As the window size increases, the data used for training and testing becomes more representative to the present signal in the spectrum, and hence the classifier's output score is more accurate. However, the larger the window size is, the longer
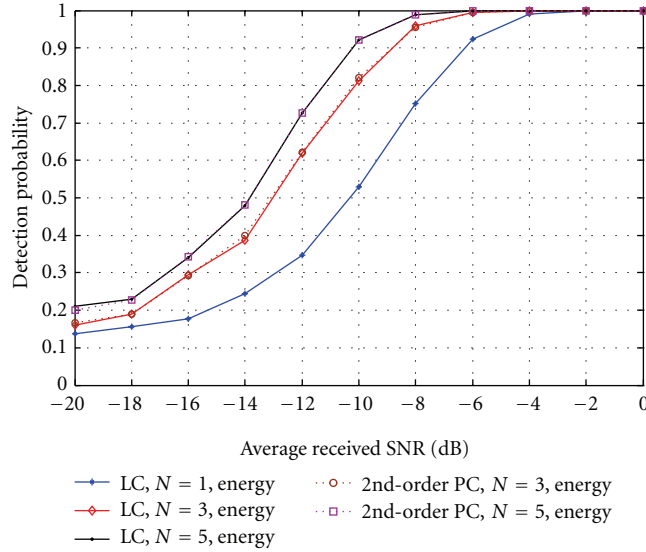
FIGURE 4: Detection performance for cooperative LC and PC with energy-based feature extraction.
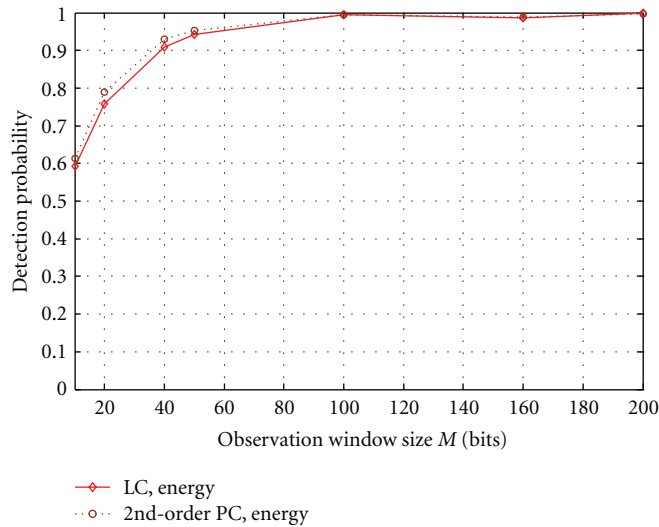

FIGURE 5: Impact of the window size on detection performance of energy-based schemes: $\text{SNR}_{\text{avg}} = -5$ dB, $N = 3$ users, and $P_f = 10\%$.

it takes for decision making of spectrum availability by the classifier at the base station. This yields a delay in spectral allocation when the spectrum is available, hence, resulting in lower spectral utilization.

A useful performance measure is the receiver operational characteristic (ROC) that represents the variation of the probability of detection with the false alarm probability at certain operational parameters. Since the LC provided a better choice with energy detection, its ROC is obtained as shown in Figure 6 when the primary signal is received at an $\text{SNR}_{\text{avg}} = -14$ dB and observation window size of $M = 200$ bits. It is observed that the detection probability deteriorates for low false alarm rates and improves when higher false alarm probability is tolerable. This behavior is expected since in order to achieve a low false alarm rate, the threshold level needs to be raised. Raising the

threshold level above classifier's output score, corresponding to occupied spectrum class, may lead to miss detecting primary signal's presence, and consequently causing more interference to the primary network's users. It can be noted that higher detection is accomplished with higher number of cooperating CRs.

5.2. *Cyclostationary Based Feature Extraction.* Simulation results for the proposed classification system when cyclostationary features are fed to the CRBS for spectrum sensing are presented in Figure 7. It is shown that cyclostationary feature detection can achieve very high detection probability even with low SNR values by using more cooperating CRs. For instance, a detection probability of about 90% is achieved with an average SNR of -18 dB when five CRs are used.
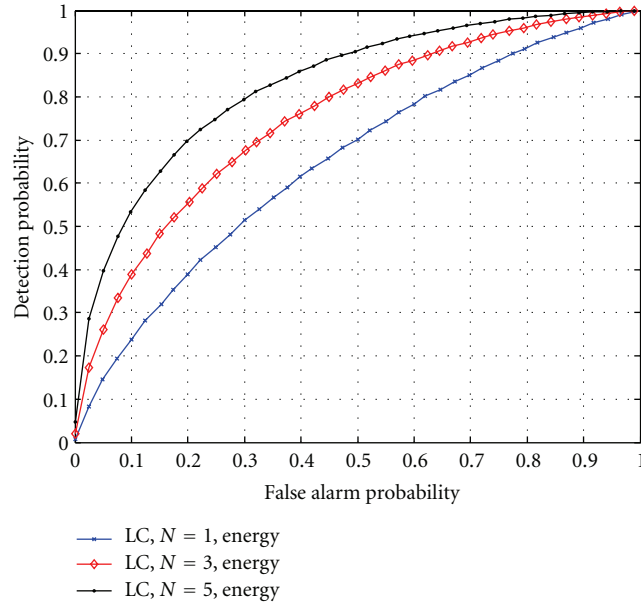
FIGURE 6: ROC curves of energy-based linear classifier scheme: $SNR_{avg} = -14\,dB$ and $M = 200\,bits$.
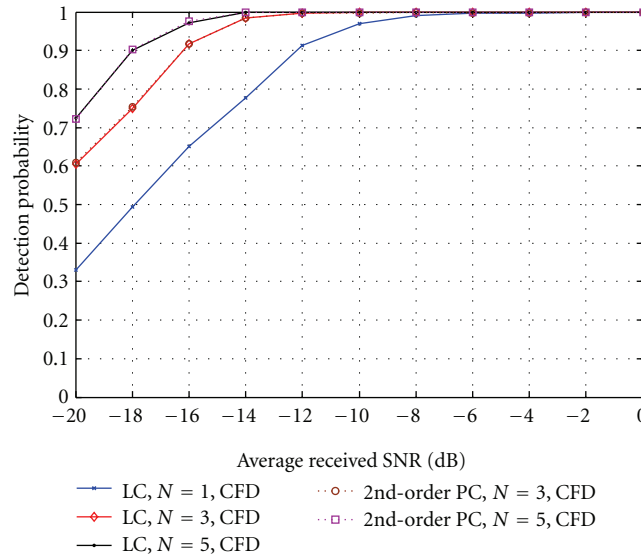

FIGURE 7: Detection performance for cooperative LC and PC with cyclostationary-based feature extraction: $P_f = 10\%$ and $M = 200\,bits$.

The results demonstrate improved performance compared to the energy-based detector. For instance, a gain of about 4 dB is observed when the number of CRs cooperating in making the decision increases from one to three at $P_d = 90\%$. As in the case of energy detection, it is observed that there is no significant performance improvement as the order of the classifier is increased from first order to second order. Furthermore, performance improvements due to increasing cooperative CRs saturate for higher number of users.

The detection performance of the cyclostationary detection scheme is improved by increasing the observation window size as illustrated in **Figure 8**. The detection results for the LC and PC are presented at $SNR_{avg} = -14\,dB$, $N = 3$ users, and $P_f = 10\%$. Increasing the observation window size from 20 to 200 bits results in improving the probability of detecting from about 70% to 98%, at a SNR_{avg} of $-14\,dB$. The ROC curve is shown in **Figure 9** for a $SNR_{avg} = -14\,dB$ and observation window of 200 bits indicating that using more cooperating radios results in better detection performance.

5.3. Coherent-Based Feature Extraction. For the coherent-based scheme, **Figure 10** shows the detection probability as the received primary signal's level is varied. Coherent
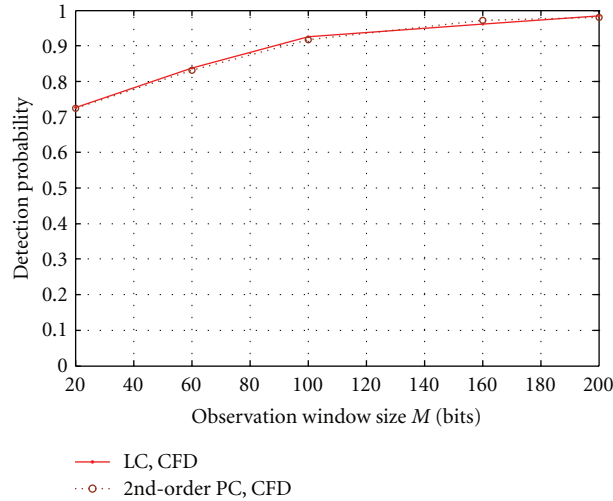
FIGURE 8: Impact of the window size on detection performance of cyclostationary-based schemes: $SNR_{avg} = -14$ dB, $N = 3$ users, and $P_f = 10\%$.
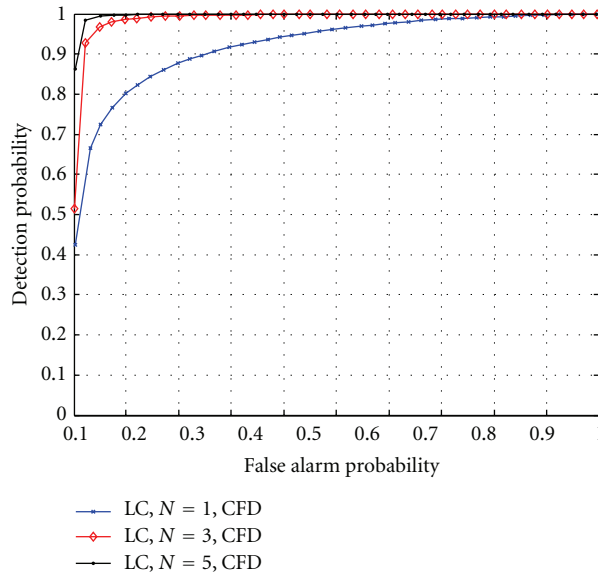

FIGURE 9: ROC curves of cyclostationary-based linear classifier scheme: $SNR_{avg} = -14$ dB and $M = 200$ bits.

detection was simulated for a primary signal with a preamble size of $L = 16$ bits and a frame length of $M = 200$ bits. It is noticed that a gain of about 6 dB is achieved as the number of collaborating radios N increases from one to three, at a detection probability of 90%. The achieved gain, however, reduces to around 1.5 dB as N increases from three to five. Coherent detection provides reliable signal identification with $P_d > 98\%$, when the received signal level is above -10 dB and $N = 5$. We notice that the LC and PC perform comparably when coherent detection is utilized in feature extraction.

The ROC curve for coherent-detection-based sensing for various numbers of cooperative CRs is demonstrated in Figure 11. It is apparent that there is a performance variation as the different number of CRs collaborates in making the decision. The performance gap between various numbers of CRs shrinks as the false alarm probability increases.

Finally, Figure 12 shows the performance gain achieved using coherent detection LC scheme as the length of preamble sequence increases. The figure presents the SNR_{avg} required to obtain a specific detection probability and false alarm rate, as the preamble length increases. Longer preamble sequences result in lower values of the required SNR_{avg} indicating that a lower level of received primary signal is sufficient to achieve a certain detection rate as preamble length is increased.

5.4. Discussion. Among the three considered schemes, cyclostationary feature detection provides the best performance
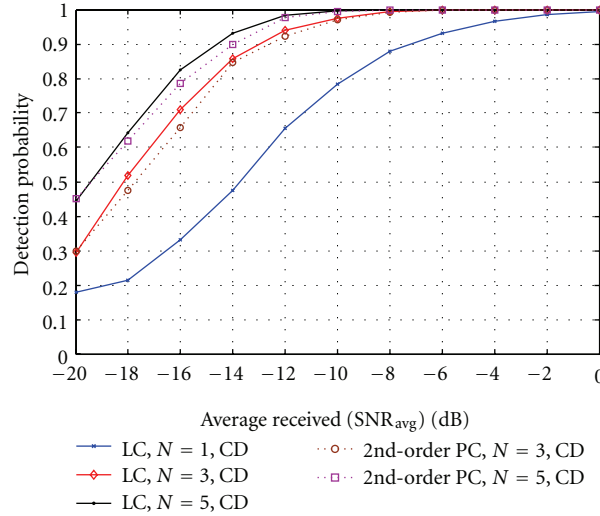
FIGURE 10: Detection performance for cooperative LC and PC with coherent-based feature extraction: $P_f = 10\%$ and preamble length $L = 16$ bits.
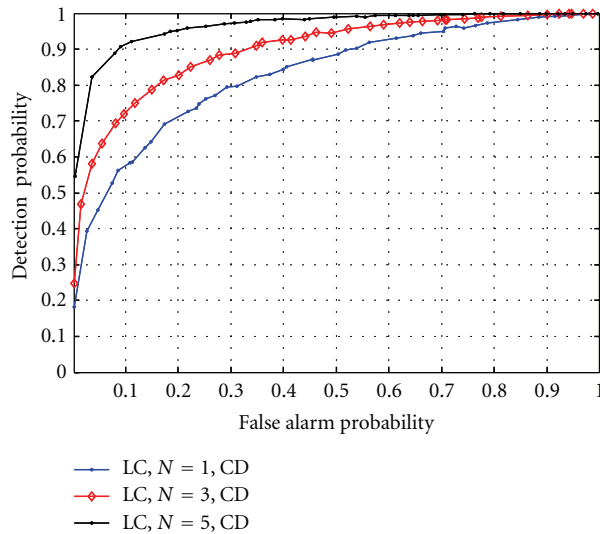

FIGURE 11: ROC curves of coherent-based linear classifier scheme: $SNR_{avg} = -14$ dB and $L = 16$ bits.

in terms of both detection and false alarm rates, while energy detection results in the poorest performer. The implementation of cyclostationary feature detection relies on the knowledge of carrier frequency and modulation type of the primary signal. The obvious drawback of cyclostationary detection is the high computational complexity required to extract cyclic features at CRs, as compared to other techniques. On the other hand, it provides high reliability to the CR network under low SNR conditions.

It has been shown that a longer sensing time can improve the detection performance considerably. However, detection improvement due to increasing sensing time is achieved at the expense of lowering the network's agility, since longer time is required to decide on the vacancy of the spectrum. This comment is very important when comparing

the performance of the cyclostationary scheme over coherent scheme. Specifically, the cyclostationary-based scheme achieves a better detection performance when compared to the coherent detection scheme. This is so because the former uses the entire frame in the process of decision making. On the other hand, the coherent-based scheme uses a shorter observation window (preamble) leading to a more timely decision making. Increasing the preamble length will improve the performance for the coherent-based scheme but at the expense of a reduction in the spectral efficiency of the primary user, while the cyclostationary-based scheme does not suffer from this drawback.

Finally, although energy detection represents the feature extraction scheme with least detection capability under low SNR conditions, representing severe fading and shadowing,
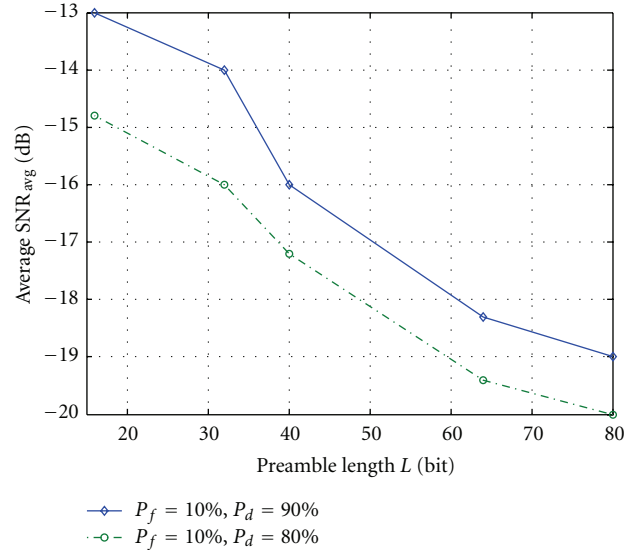
FIGURE 12: Impact of the preamble length on detection performance of coherent-based linear classifier scheme.

it is still an attractive technique under high SNR regimes due to its simplicity and minimum prior information requirements.

6. Conclusion

In this paper, pattern recognition models were proposed to tackle the problem of spectrum sensing in CR networks. The proposed classifier model is based on collaborative sensing, in which secondary users monitor channel usage in a given area and cooperate through a centralized node to provide the spectrum occupancy information. The cooperation between secondary users was achieved through first- and second-order polynomial classifiers that were modeled, trained, validated, and evaluated. Results indicate that both the linear and the polynomial classifiers provide high detection rates of primary signal over wireless fading channel and very small signal to noise ratios. Moreover, simulation results show that both classifiers perform comparably; consequently, the linear classifier is chosen as the best model for cooperative CR networks due to its lower complexity. Energy-, cyclostationary-, and coherent-based feature extraction techniques were compared. Simulation results demonstrated that cyclostationary detection constitutes the best candidate for feature extraction when information on primary signal is available, since it outperforms energy and coherent detection substantially. However, the remarkable detection capability of cyclostationary detection is achieved at the expense of higher implementation complexity.

References

[1] M. Naraghi and T. Ikuma, "Autocorrelation-based spectrum sensing for cognitive radios," *IEEE Transactions on Vehicular Technology*, vol. 59, pp. 718–733, 2010.

[2] E. Hossain and V. Bhrgava, *Cognitive Wireless Communication Network*, Springer, 1st edition, 2007.

[3] C. Cordeiro, K. Challpali, and D. Birru, "IEEE 802.22: an introduction to the first wireless standard based on cognitive radios," *Journal of Communications*, vol. 1, pp. 38–47, 2006.

[4] Z. Quan, S. Cui, and A. H. Sayed, "Optimal linear cooperation for spectrum sensing in cognitive radio networks," *IEEE Journal on Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 28–40, 2008.

[5] T. Yücek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Communications Surveys and Tutorials*, vol. 11, no. 1, pp. 116–130, 2009.

[6] D. Cabric, S. M. Mishra, and R. W. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," in *Proceedings of the 38th Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 772–776, November 2004.

[7] D. Cabric, A. Tkachenko, and R. W. Brodersen, "Spectrum sensing measurements of pilot, energy, and collaborative detection," in *Proceedings of the Military Communications Conference (MILCOM '06)*, pp. 1–7, October 2006.

[8] P. Wang, J. Fang, N. Han, and H. Li, "Multiantenna-assisted spectrum sensing for cognitive radio," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 4, pp. 1791–1800, 2010.

[9] M. Derakhshani, M. Nasiri-Kenari, and T. Le-Ngoc, "Cooperative cyclostationary spectrum sensing in cognitive radios at low SNR regimes," in *Proceedings of the IEEE International Conference on Communications (ICC '10)*, pp. 604–608, May 2010.

[10] J. Lundén, V. Koivunen, A. Huttunen, and H. V. Poor, "Collaborative cyclostationary spectrum sensing for cognitive radio systems," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4182–4195, 2009.

[11] N. Khambekar, C. Spooner, and V. Chaudhary, "Listen-while-talking: a technique for primary user protection," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '09)*, pp. 1–5, April 2009.

[12] N. Khambekar, D. Liang, and V. Chaudhary, "Utilizing OFDM guard interval for spectrum sensing," in *Proceedings of the IEEE*

Wireless Communications and Networking Conference (WCNC '07), pp. 38–42, March 2007.

[13] P. D. Sutton, K. E. Nolan, and L. E. Doyle, "Cyclostationary signatures in practical cognitive radio applications," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 1, pp. 13–24, 2008.

[14] D. Cabric, A. Tkachneko, and R. Brodersen, "Experimental study of spectrum sensing based on energy detection and network cooperation," Tech. Rep., Berkeley Wireless Research Center, Berkeley, Calif, USA, 1997.

[15] N. S. Shankar, C. Cordeiro, and K. Challapali, "Spectrum agile radios: utilization and sensing architectures," in *Proceedings of the 1st IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN '05)*, pp. 160–169, November 2005.

[16] K. Kim, I. A. Akbar, K. K. Bae, J. S. Um, C. M. Spooner, and J. H. Reed, "Cyclostationary approaches to signal detection and classification in cognitive radio," in *Proceedings of the 2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 212–215, April 2007.

[17] K. Assaleh, K. Farrel, and R. Mammone, "A new method of modulation classification for digitally modulated signals," in *Proceedings of the IEEE Military Communications Conference*, vol. 2, pp. 712–716, 1992.

[18] A. F. Cattoni, M. Ottonello, M. Raffetto, and C. S. Regazzoni, "Neural networks Mode classification based on frequency distribution features," in *Proceedings of the 2nd International Conference on Cognitive Radio Oriented Wireless Networks and Communications, CrownCom*, pp. 251–257, August 2007.

[19] T. Y. Yücek and H. Arslan, "Spectrum characterization for opportunistic cognitive radio systems," in *Proceedings of the IEEE Military Communications Conference*, pp. 1–6, 2006.

[20] B. Wang, *Dynamic spectrum allocation and sharing in cognitive cooperative networks [Ph.D. thesis]*, University of Maryland, 2009.

[21] T. Zhang, G. Yu, and C. Sun, "Performance of cyclostationary features based spectrum sensing method in a multiple antenna cognitive radio system," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '09)*, pp. 1–5, April 2009.

[22] Y. Hassan, M. El-Tarhuni, and K. Assaleh, "Comparison of linear and polynomial classifiers for co-operative cognitive radio networks," in *Proceedings of the IEEE 21st International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC '10)*, pp. 797–802, Istanbul, Turkey, September 2010.

[23] Y. Hassan, M. El-Tarhuni, and K. Assaleh, "Knowledge based cooperative spectrum sensing using polynomial classifiers in cognitive radio networks," in *Proceedings of the 4th International Conference on Signal Processing and Communication Systems (ICSPCS '10)*, Sydney, Australia, December 2010.

[24] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, Academic Press, San Diego, Calif, USA, 3rd edition, 2006.

[25] W. M. Campbell, K. T. Assaleh, and C. C. Broun, "Speaker recognition with polynomial classifiers," *IEEE Transactions on Speech and Audio Processing*, vol. 10, no. 4, pp. 205–212, 2002.

[26] W. Gardner, *Cyclostationarity in Communications and Signal Processing*, IEEE Press, 1st edition, 1994.

[27] W. A. Gardner, "Exploitation of spectral redundancy in cyclostationary signals," *IEEE Signal Processing Magazine*, vol. 8, no. 2, pp. 14–36, 1991.

[28] Y. Lin and C. He, "Subsection-average cyclostationary feature detection in cognitive radio," in *Proceedings of the IEEE International Conference on Neural Networks & Signal Processing*, pp. 604–608, 2008.

[29] T. Shanableh, K. Assaleh, and M. Al-Rousan, "Spatio-temporal feature-extraction techniques for isolated gesture recognition in arabic sign language," *IEEE Transactions on Systems, Man, and Cybernetics B*, vol. 37, no. 3, pp. 641–650, 2007.

[30] D. Specht, "Generation of polynomial discriminant functions for pattern recognition," *IEEE Transactions on Electronic Computers*, vol. 16, no. 3, pp. 308–319, 1967.