Impact of Correlated Primary Transmissions on the Design of a Cognitive Radio Inference Engine

Massimiliano Pesce, Marco Centenaro, Leonardo Badia, and Michele Zorzi
Department of Information Engineering, University of Padova
Via Gradenigo, 6/b, 35131 Padova, Italy
E-mail: {pescemas, marco.centenaro, badia, zorzi}@dei.unipd.it

***Abstract*—We consider sensing for cognitive network users, in particular focusing on a scenario where a primary user (PU) and a secondary user (SU) operate on the same frequency band. The SU is interested in identifying transmission opportunities when the PU is silent. We investigate how this sensing performed by the SU can be improved through modeling the PU transmission pattern with increasing accuracy. In particular, we are interested in evaluating the impact of correlation in PU's transmissions. Therefore, we assume that the real behavior of the PU follows a Markov chain, used to model correlation in its activity, and we discuss how the maximum likelihood estimation of the SU can be subsequently improved by adding more information about this underlying process. In this way, the estimate can evolve into a maximum a-posteriori criterion, and furthermore knowledge about the whole Markov chain can be exploited. Also, we investigate the practical setup of training periods of variable length used to estimate the PU's parameters.**

***Index Terms*—Cognitive networks; Markov process analysis; maximum likelihood estimator; performance evaluation.**

I. Introduction

Cognitive radio is a paradigm that refers to the application of techniques mimicking rationality and intelligence to configure wireless transmission parameters [1]. In particular, a common scenario of application involves the activity of a licensed primary user (PU) on a given frequency band and the concurrent operation over the same band of a secondary user (SU). The SU should estimate the ongoing PU activities, to be able to opportunistically exploit the instants where the PU is inactive [2].

This kind of investigation has recently become very appealing to solve problems of access coordination, network overhead, and high density of wireless devices in the same area. More in general, advanced designs of cognitive engines are able to increase network capacity and coverage for cellular systems, thus resulting in a better resource utilization for the operator as well as an improved quality-of-service (QoS) for the end user, ultimately leading to higher revenues for the former and better satisfaction for the latter.

If the activity patterns of the PU are accurately monitored, the SU can exploit idle channel resources that are temporarily unoccupied, while allowing interference-free communication for the PU [3]. However, the goodness of the instantaneous decision to make is related to the accuracy in the overall representation made by the SU about the PU. A precise modeling of when and how the PU is active leads to more effective spectrum usage by the SU [10].

The basic sensing procedure for spectrum sensing is to employ the Neyman-Pearson criterion that translates into a maximum likelihood comparison between the two hypotheses "PU is transmitting" and "PU is silent," both of which can be tested against a channel observation at the receiver's side [4]. In the former case, the receiver can hear the PU signal and an additional noise term, while in the latter only noise is present.

The approach of comparing the likelihoods of the two hypotheses can be derived under several conditions [5]. In the simplest case, the statistics of the noise and the transmission parameters are fully known. It can be proven via a Bayesian approach that it would be equivalent, only more cumbersome, to extract these values from actual transmissions and initial estimates that become more and more refined.

For the sake of simplicity, we start by considering that all physical parameters have been properly estimated, so that the only unknown at the physical layer is whether the PU is actually active or not. We will further relax this assumption by considering an iterative parameter estimation. However, our main focus is not on physical layer parameters, but rather we consider an upper layer characterization of the PU, and in particular we discuss how a richer description of its access/traffic characteristics can improve the sensing by the SU. However, this evaluation is challenging because there are several descriptors that can be taken into account to characterize the primary activity.

A first step in this sense is to include the transmission probability of the PU in the analysis and make it available to the SU. This corresponds to translating a maximum-likelihood criterion into a maximum a-posteriori probability one. Indeed, if the SU perfectly knows a prior for the instantaneous PU activity, it can properly weigh the likelihood of the two hypotheses.

In addition, in this paper we are also interested in evaluating the impact of correlation within the PU's activity, and therefore we further introduce a 2-state Markov chain to represent the PU activity [6], which keeps into account the memory in the transmission process. Markov models are well established in the literature as a simple but elegant way to characterize these situations.

For this reason, we focus on the application of cognitive engines to a two-state Markov scenario, where we are able to describe the average transmission rate and correlation, and discuss whether this richer description can be beneficial. Specifically, we compare three cases. First, a basic ML estimation, with full characterization of the channel parameters and the PU's transmitting power, but no information at all (and therefore the assumption of equally distributed prior) on its actual transmitting probability. Second, a Maximum Posterior criterion where, in addition to all aspects mentioned before, also the PU's transmission probability is known, through a first-order description. Finally, we also include transmission correlation in a "Markovian" estimation, which means that the real underlying Markov process statistics is fully known by the SU and exploited to make the best possible decision.

Thus, we quantify the benefit brought by the improved description of the PU's activity pattern. Not only do we show that including PU's transmission correlation in the cognitive criterion is helpful to achieve a better estimate (and therefore an enhanced access operation) by the SU, but we also precisely state how much is the added value, and how critical is the parameter evaluation for this procedure.

The rest of the paper is organized as follows. We review related work in Section II. In Section III, the system model and the proposed inference engines are described. We show performance evaluations for the three approaches in Section IV. Finally, in Section V we draw the conclusions.

II. Related Work

The key feature of a cognitive radio is its ability to measure, sense, learn, and be aware of the radio channel characterization, also including the overall radio environment as well as the instantaneous spectrum availability, possibly including infrastructure and regulatory aspects as well [4].

Cognitive procedures can be employed towards different goals, including channel identification and estimation, interference detection, synchronization, user localization, and so on [7]. In this paper, we focus on a primary-secondary scenario, and assume that most of these evaluations, which involve a long-term estimation procedure, have already been performed, whereas the activity pattern of the PU and its impact on SU's transmissions are unknown.

Most of the contributions in the literature deal with characterizations of such patterns and interrelationships. For example, [8] investigates the correlation between PU's and SU's activity from an information-theoretic perspective, and [9] pushes this analysis into considering the interaction among multiple secondary users due to mutual interference. Furthermore, [10] argues that the PU's activity is in practice not smooth and burst-free, and introduces a metric to account for correlation statistics.

Further models of PU's transmission are discussed in [6]. Actually, there are indeed several similar models to represent the PU activity that all loosely translate into a 2-state Markov chain, or an equivalent formulation, where the PU is idling or transmitting depending on its activity in the previous time unit.

On the other hand, the direct estimation of the PU's activity is based on energy detection mechanisms [11], which translates in a decision based on likelihood (or log-likelihood) values, so as to exploit the Neyman-Pearson criterion. Most of the times, collaborative improvements of this decision are suggested by letting other independent nodes participate in the energy detection [12].

In reality, our investigation in the present paper is complementary to all of these contributions, since we are not interested in determining how the PU's Markov chain is obtained, but rather our goal is to determine how the knowledge of this underlying correlation can be exploited to improve the energy detection mechanism. We also do not aim at refining the signal acquisition or the energy detection per se, but we are interested in giving a quantitative assessment of the achievable gain by including correlation in the likelihood terms.

III. System Model

Consider a network scenario in which two transmitter-receiver pairs, $\{T_1, R_1\}$ and $\{T_2, R_2\}$, operate on the same wireless channel. In each pair, the receiver's role is just passive, while what really matters is the transmitter activity. Thus, we can consider a "user" to be either the whole pair, or just the transmitter, and we regard the pair $\{T_1, R_1\}$ as a PU, while $\{T_2, R_2\}$ is seen as a SU aiming at transmitting opportunistically exploiting the PU's inactivity periods. The time axis is slotted and every time slot is considered either busy if T_1 transmits a data packet, or idle when T_1 remains silent.

A. Wireless Channel Model

We consider the wireless medium as an additive white Gaussian noise (AWGN) channel with constant and known power attenuation.[1] For the sake of analytical tractability in the derivations of the proposed estimation methods, we assume that the instantaneous noise power N is distributed according to a chi-squared distribution, i.e., $N \sim \chi^2(0, P_\mathrm{N})$, where P_N is the noise power variance. It is worth remarking that this choice, although only an approximation of the real channel conditions, is made here to gain insights on the design of the proposed estimation method.[2]

The instantaneous value of the useful signal power S is, instead, deterministic, since it depends only on the transmit power and the modulation scheme used by T_1. In the following, for the sake of simplicity, we assume that its power value is constant and equal to P_S. Note that a similar analysis could be performed in the presence of different power levels for the signal, and an iterative decision mechanism could be applied to jointly interpret the power level of the primary (which is however not strictly required for the cognition process we are interested in) and the noise. Thus, we focused on a constant P_S to get a simpler analysis that provides the same insight.

[1] Further channel models, e.g., fading channels, will be considered as future work.

[2] We will validate this heuristic by means of numerical simulations showing the benefits of this model.

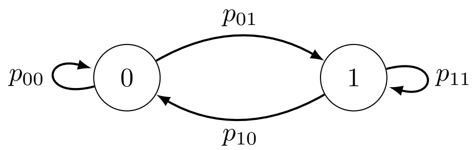

Fig. 1: Graphical model of the Markov chain which describes the transmission process of the primary link $\{T_1, R_1\}$.

Therefore, in our scenario, the signal-to-noise ratio (SNR) Γ is defined as

$$\Gamma = \frac{P_\mathrm{S}}{P_\mathrm{N}} \ . \tag{1}$$

According to these assumptions, the total input power at the receiver P_rx is distributed according to a chi-squared distribution, as well. In particular, P_rx has zero mean and variance P_N in case only noise is transmitted; on the other hand, it has mean P_S and variance P_N when useful data are sent over the channel.

B. Primary Transmission Model

We assume that the PU traffic pattern is *bursty*, i.e., T_1 typically remains silent for some periods of time and then transmits sequences of data packets (*bursts* of data). The behavior of the PU can be modeled using the two-state Markov chain depicted in Fig. 1, in which state `0` corresponds to the idle channel state and state `1` to the busy channel state.

The transition matrix of the Markov chain is

$$\boldsymbol{P} = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} \ . \tag{2}$$

The steady-state probabilities of the primary chain π_0 and π_1 are defined as

$$\pi_0 = \frac{p_{01}}{p_{01} + p_{10}} \ , \qquad \pi_1 = 1 - \pi_0 \ . \tag{3}$$

Since we are interested in modeling a bursty primary transmitter, we consider the *average burst length* B as

$$B = \frac{1}{p_{01}} \ . \tag{4}$$

We remark that the primary Markov chain is fully described by choosing the values of parameters B and π_1.

C. Secondary Transmission Model

In every time slot, T_2 senses the channel and decides whether it carries just noise N (denote this event with $\mathcal{N}$) or the sum of useful signal power S and noise power N (denote this event with $\mathcal{S}$). Three techniques to estimate the state of the primary chain are considered in the following, according to different levels of knowledge of the PU's behavior.

a) Maximum Likelihood (ML) estimation: With just knowledge of the average noise power value P_N and the useful signal power P_S, T_2 uses a threshold-based detection. Let us define threshold λ as follows [5]:

$$\lambda = \frac{(P_\mathrm{N} + P_\mathrm{S}) + P_\mathrm{N}}{2} = P_\mathrm{N} + \frac{P_\mathrm{S}}{2} \ . \tag{5}$$

If the received power r is greater than λ then the slot is considered to be busy, otherwise it is assumed to be idle. Denoting with $\mathcal{R}_j^{(\mathrm{ML})}$ the decision region of event $j \in \{\mathcal{N}, \mathcal{S}\}$ and with $\hat{t}$ and t the estimated and the actual primary chain state, respectively, the detection success probability can be computed as follows:

$$\begin{aligned} P_\mathrm{succ}^{(\mathrm{ML})} &= \mathbb{P}[\hat{t} = t] \\ &= \sum_{j \in \{\mathcal{N}, \mathcal{S}\}} \mathbb{P}[r \in \mathcal{R}_j^{(\mathrm{ML})} | \mathrm{tx} = j] \cdot p_j \\ &= \sum_{j \in \{\mathcal{N}, \mathcal{S}\}} \int_{\mathcal{R}_j^{(\mathrm{ML})}} f_{r|\mathrm{tx}}(x|j) \cdot p_j \ \mathrm{d}x \\ &= \sum_{j \in \{\mathcal{N}, \mathcal{S}\}} p_j \cdot \int_{\mathcal{R}_j^{(\mathrm{ML})}} f_{r|\mathrm{tx}}(x|j) \ \mathrm{d}x \\ &= p_\mathcal{N} \cdot \int_0^\lambda f_{r|\mathrm{tx}}(x|\mathcal{N}) \ \mathrm{d}x \\ &\quad + p_\mathcal{S} \cdot \int_\lambda^{+\infty} f_{r|\mathrm{tx}}(x|\mathcal{S}) \ \mathrm{d}x \ , \end{aligned} \tag{6}$$

where $p_\mathcal{N} = \pi_0$ and $p_\mathcal{S} = \pi_1$.

b) Maximum Posterior Estimation (MPE): Assuming that T_2 knows also the steady-state probabilities of the primary chain, i.e., π_0 and π_1, T_2 can choose the threshold λ in such a way that it maximizes (6), i.e.,

$$\lambda^\star = \arg\max_\lambda \ P_\mathrm{succ}^{(\mathrm{ML})}(\lambda) \ . \tag{7}$$

Therefore, it is $P_\mathrm{succ}^{(\mathrm{MPE})} = P_\mathrm{succ}^{(\mathrm{ML})}(\lambda^\star)$.

c) Markovian Estimation (ME): If the complete transition matrix $\boldsymbol{P}$ is known to T_2, then the decision regions can be designed maximizing the following success probability:

$$\begin{aligned} P_\mathrm{succ}^{(\mathrm{ME})} &= \sum_{j \in \{\mathcal{N}, \mathcal{S}\}} \sum_{k \in \{\mathcal{N}, \mathcal{S}\}} \mathbb{P}[r \in \mathcal{R}_n^{(\mathrm{ME})} | \mathrm{tx} = k] \cdot p_{jk} \cdot \pi_j \\ &= \sum_{j,k \in \{\mathcal{N}, \mathcal{S}\}} \int_{\mathcal{R}_{jk}^{(\mathrm{ME})}} f_{r|\mathrm{tx}}(x|n) \cdot p_{jk} \cdot \pi_j \ \mathrm{d}x \\ &= \sum_{j,k \in \{\mathcal{N}, \mathcal{S}\}} p_{jk} \cdot \pi_j \cdot \int_{\mathcal{R}_{jk}^{(\mathrm{ME})}} f_{r|\mathrm{tx}}(x|n) \ \mathrm{d}x \, . \end{aligned} \tag{8}$$

In this case two distinct thresholds, λ_0 and λ_1, must be set for the cases in which the chain describing the PU's activity was previously in state `0` or `1`, respectively.

D. Inference Engine Design

Consider a real transmission chain where T_1 transmits packets of length $L = 1024$ bit using a 64-PSK digital modulation. T_2 overhears a window of N_slots time slots and

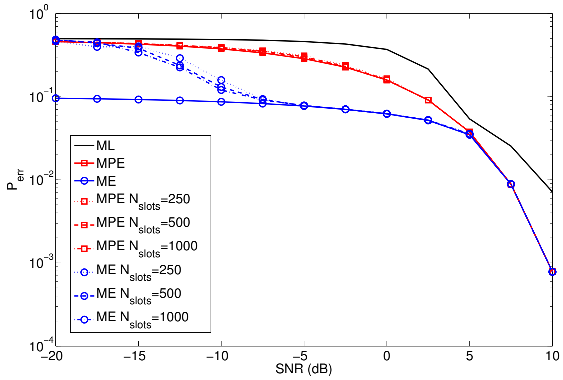

Fig. 2: Error probability vs. SNR for MPE and ME with different learning window sizes. $\pi_1 = 0.5$, $B = 10$.

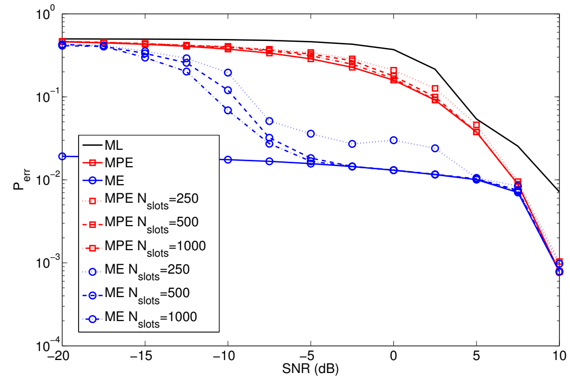

Fig. 3: Error probability vs. SNR for MPE and ME with different learning window sizes. $\pi_1 = 0.5$, $B = 50$.

for every slot decides whether it is busy (event $\mathcal{S}$) or idle (event $\mathcal{N}$), using the MLE threshold λ defined in (5). Then, T_2 is able to estimate the parameters of the primary chain, i.e., $\hat{\pi_0}$, $\hat{\pi}_1$, and $\hat{\boldsymbol{P}}$, using a frequentist approach.

IV. Performance Evaluation

In this section, we evaluate the performance of the proposed cognitive radio procedures.

A. Simulation Setup

We considered a transmission of a PU, whose activity pattern follows Fig. 1 and uses a 64-PSK modulation over a radio channel. We consider an SNR spanning from -20 to 10 dB, which would roughly correspond to a received power value considered within a range from -142 to -122 dBm for a setup where the noise power spectral density is that of the thermal noise and the bandwidth of the common channel shared by PU and SU is 160 kHz. Simulations have been carried out over Monte Carlo iterations using the parameters listed in Table I. We remark that different estimation lengths in terms of number of slots are considered for the simulations.

B. Simulation Results

In Fig. 2 we can observe how the knowledge of the Markov chain improves the overall performance according to the size of the learning window. We report two different approaches for computing the error probability, i.e., MPE and ME. For each

Parameter	Value
π_1	$\{0.5, 0.9\}$
B	$\{2, 10, 50\}$
N_{slots}	$\{250, 500, 1000\}$
Modulation	64-PSK
L_{pkt}	1024
N_{iters}	100

TABLE I: System parameters.

method there are four different curves: three are associated to three values of the learning window size while the solid unmarked one represents the perfect knowledge performance bound. Simulation data demonstrate that we can reduce the error probability (P_{err}) up to 5 times for low SNR values by applying the proposed Markov-based approach.

This, however, comes at the price of a higher sensitivity to the window size for the ME criterion in comparison with the MPE method. Even with a large learning window, i.e., 1000 slots, for low SNR values, the Markov approach does not reach the perfect knowledge bound. Moreover, also for higher SNRs and with small window lengths, the MPE approach achieves better performance.

Considering both Figs. 2 and 3 at the same time, the impact of the burst length (B) is worth noticing. With increasing values of B, ME shows its advantages with respect to MPE, but also its sensitivity to the learning process. In Fig. 3, the error probability for the Markov approach gets even lower, reaching, more or less, values ten times smaller.

In Figs. 4 and 5 we report the same metrics as before, but with a higher π_1. In this case, the difference between MPE and ME is not so evident and the curves are close to each other. In particular, the full-knowledge criteria are perfectly overlapping in Fig. 4. This interesting situation turns out when $\pi_1 = 0.9$ and $B = 10$, i.e., when in the primary chain $p_{11} = p_{01}$. The one-step probabilities to reach state 1 are independent of the outgoing state, thus making the knowledge of the starting state worthless. When this condition is satisfied, the MPE criterion coincides exactly with the ME. However, incrementing B from 10 to 50, see Fig. 5, we force the system to be biased, altering the transition probabilities of the Markov chain. The ME approach takes advantage of this imbalance and this appears evident when looking at the gap between the two curves.

As a final remark, it is worth noticing that in all the plots the ML theoretic curve exhibits two different behaviors for high SNRs. This can be explained considering the definition of $P_{\text{succ}}^{(\text{ML})}$ in Eq. (6). The ML success probability increases

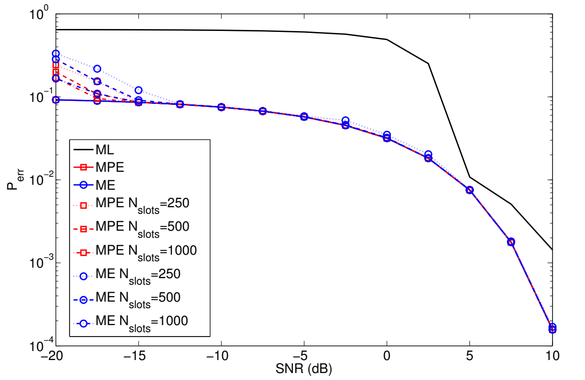

Fig. 4: Error probability vs. SNR for MPE and ME with different learning window sizes. $\pi_1 = 0.9$, $B = 10$.

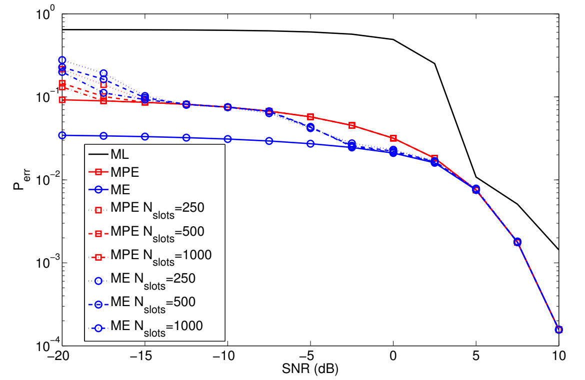

Fig. 5: Error probability vs. SNR for MPE and ME with different learning window sizes. $\pi_1 = 0.9$, $B = 50$.

as the values of the two integrals $\int_0^\lambda f_{P_{\rm rx|tx}}(x|\mathcal{N})\ dx$ and $\int_\lambda^{+\infty} f_{P_{\rm rx|tx}}(x|\mathcal{S})\ dx$ grow. However, the higher the SNR, the smaller the noise power and, therefore, the smaller the value of the ML threshold λ (see Eq. (5)). In particular, when λ is such that $\lambda \le P_S$, then $\int_\lambda^{+\infty} f_{P_{\rm rx|tx}}(x|\mathcal{S})\ dx = 1$, and remains costant for all the values of λ smaller than P_S. Thus, the variation of $P_{\rm succ}^{\rm (ML)}$ is given only by the term $p_\mathcal{N} \cdot \int_0^\lambda f_{P_{\rm rx|tx}}(x|\mathcal{N})\ dx$, which still varies according to the value of λ. The SNR threshold value such that $\lambda = P_S$ and the behavior of the curve changes can be obtained imposing

$$\hat{\lambda} = P_S \ , \ \hat{P}_N + \frac{P_S}{2} = P_S \ , \ \hat{P}_N = \frac{P_S}{2} \, , \tag{9}$$

yielding an SNR threshold

$$\hat{\Gamma} = \frac{P_S}{\hat{P}_N} = 2 \simeq 3 \text{ dB} \, . \tag{10}$$

Note that, because the total curve is a combination of two contributions, the breakpoint in the graphs appears at the SNR value $\Gamma = 5$ dB.

V. Conclusions

We studied the impact of including awareness about correlation patterns of the primary activity in the sensing decision made by the secondary. The investigation was performed by assuming an underlying 2-state Markov process for the primary activity, and we compared three different decision criteria. In particular, our goal was to see, and quantitatively assess, the improvement brought by including correlation (i.e., average transmission burst length in the Markov process).

The results show that, when the primary activity is frequent, the benefit of including correlation is less evident. Actually, all three criteria perform almost identically if π_1 is high and correlation is mild. It is intuitive, indeed, that in these cases transmission opportunities for the secondary user are rare, and therefore the main decision criterion is simply related to physical parameters such as the power perceived on the channel. However, correlation is better kept into account when the primary user is active 50% of the time or less, and also when correlation is strong (i.e., when a high burst length significantly affects the transmission probability over subsequent slots).

The precise quantification of this last aspect is an interesting topic for future research, that can lead to adaptive cognitive criteria with increased realism and better performance. Thus, we believe this can be subject for future research.

References

[1] J. Mitola III and G. Q. Maguire Jr., "Cognitive radio: making software radios more personal," *IEEE Personal Communications*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[2] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201-220, Feb. 2005.

[3] G. Yuan, R. C. Grammenos, Y. Yang, and W. Wang, "Performance analysis of selective opportunistic spectrum access with traffic prediction," *IEEE Trans. Veh. Tech.*, vol. 59, no. 4, pp. 1949–1959, May 2010.

[4] T. Yücek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Comm. Surv.&Tutor.*, vol. 11, no. 1, pp. 116–130, Mar. 2009.

[5] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio: State-of-the-art and recent advances," *IEEE Sig. Proc. Mag.*, vol. 29, no. 3, pp. 101-116, Apr. 2012.

[6] Y. Saleem and M. H. Rehmani, "Primary radio user activity models for cognitive radio networks: A survey," *J. Netw. Comp. Appl.*, vol. 43, pp. 1-16, Aug. 2014.

[7] M. T. Masonta, M. Mzyece, and N. Ntlatlapa, "Spectrum decision in cognitive radio networks: A survey," *IEEE Comm. Surv.&Tutor.*, vol. 15, no. 3, pp. 1088–1107, July 2013.

[8] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proc. IEEE*, vol. 97, no. 5, pp. 894-914, Apr. 2009.

[9] A. Rabbachin, T. Q. S. Quek, H. Shin, and M. Z. Win, "Cognitive network interference," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 2, 480-493, Feb. 2011.

[10] B. Canberk, I. F. Akyildiz, and S. Oktug, "Primary user activity modeling using first-difference filter clustering and correlation in cognitive radio networks," *IEEE/ACM Trans. Netw.*, vol. 19, no. 1, pp. 170-183, Feb. 2011.

[11] D. Cabric, S. M. Mishra, and R. W. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," *Proc. IEEE Asilomar conf.*, vol. 1, pp. 772-776), Nov. 2004.

[12] W. Zhang, M. K. Ranjan, and K. b. Letaief, "Optimization of cooperative spectrum sensing with energy detection in cognitive radio networks," *IEEE Trans. Wirel. Commun.*, vol. 8, no. 12, pp. 5761-5766, Dec. 2009.

[13] A. Goldsmith, "Wireless Communications," Cambridge University Press, New York (NY), USA, 2005.