Wireless Netw (2014) 20:319–330
DOI 10.1007/s11276-013-0555-3

On cognitive processes in cognitive radio networks

Boris Oklander · Moshe Sidi

Published online: 9 June 2013

Abstract In this article we model the cognitive processes and evaluate their impact on the performance of cognitive radio networks (CRN). Operation of the cognitive radio nodes, can be characterized by two types of processes: communication processes such as packets transmission, and cognitive processes such as estimation of the network state and decision-making for dynamic resource allocation. We propose a continuous time Markov chain model of CRN that couples these processes into unified queueing framework and analyze it by means of the matrix-geometric approach. From the obtained results, we derive the performance measures of CRN such as average delay and throughput, and establish their dependencies on the underlying cognitive processes. Additionally, we design an efficient policy for accessing the vacant channels and managing the transmission-sensing trade-off, which arises when transmissions and sensing are mutually exclusive. The policy search is carried out by the stochastic optimization method of cross-entropy. The optimized policy leads to significantly enhanced performance of CRN.

Keywords Cognitive radio networks · Dynamic spectrum access · State estimation · Queueing analysis · Cross-entropy

1 Introduction

The requirement of additional bandwidth for wireless access in voice, video, multi-media and other high rate data applications is steadily increasing. CRN is a candidate to cope with a wide spectrum of challenges arising in the face of the growing communication demands. Although researchers and standardization bodies generally agree that CR should be able to sense the environment and autonomously adapt to changing operating conditions, there are different views concerning the levels of cognitive functionality [1]. This functionality is an important factor which distinguishes CRN as an unique class of wireless networks.

Different studies have addressed CRNs capability of opportunistic spectrum access [3, 4], in which spectrum bands licensed to primary users (PU) are shared with the cognitive users called secondary users (SU). It is well known that a significant part of the allocated spectrum is vastly underutilized [5, 6], and the CRN goal in this scheme is to improve spectrum utilization while avoiding interference with the PUs [7–9]. This requires management of the sensing-transmission tradeoff [10–13]. Additionally, there are works that treat CRN while taking sensing errors into account [14]. Most of the studies however treated different operational scenarios of CRN without addressing its cognitive behavior directly.

The cognitive behavior of a CRN can be represented by the cognition cycle [2]. It is the main control process which enables CR to stay aware of its communication environment and to adapt to its changing conditions. There are different views of what phases the cognition cycle consists [2, 4], but basically all the versions share the observation, orientation, decision and action (OODA) phases. During the observation phase, CR continuously senses the environment in order to collect the input information for the cognition cycle. In the orientation phase, CR uses the gathered information from its sensors to estimate the current network state. Next, given the estimated network

B. Oklander (✉) · M. Sidi
Department of Electrical Engineering, Technion–Israel Institute of Technology, Haifa, Israel
e-mail: oklander@tx.technion.ac.il

M. Sidi
e-mail: moshe@ee.technion.ac.il

conditions, CR enters the decision-making phase in which it applies some policy to decide on the course of action. Finally, CR completes the cognition cycle by entering the action phase, which carries out the chosen actions. In addition, machine learning can be structured into these phases of cognition cycle in order to update them in the face of changing environment's situation or user needs.

The cognition cycle implies strong correlation between the perception (sensing and estimation) and the action (transmissions). Essentially, the interdependence of perception and action is a fundamental principle governing CRs behavior. Perceiving both the environment and the inner states enables CR to intelligently adapt its actions in the face of the dynamically changing conditions of the network. CR controls the sensing process, and therefore it can deliberately modify its perception level by changing the resource allocation (e.g. varying the sensing rate). Both the perception and the action processes make use of CRs limited resources such as computation power, spectrum bandwidth etc. Therefore, by applying appropriate policy for resources allocation, CR should adaptively optimize its operating point.

It is not common to find studies that directly address the interdependent processes composing the cognition cycle. The main reason for this is the difficulty to design analytically tractable models for systems characterized by cognitive behavior. As for now, a substantial gap remains between the perception and action-taking models. In [15] state of the art protocols for medium access in cognitive radio networks are overviewed. The authors point out that the existing studies do not fully integrate both the spectrum sensing and the spectrum access in one framework which is required in order to maintain the capability of adaptation to the environment changes [16].

The authors of [17] derive a threshold strategy for the sequential channel sensing process aiming to maximize the aggregated throughput of CRN. While the model in [17] assumes independent transmissions over different channels, our model can deliberately utilize any number of channels it can observe simultaneously and therefore achieves higher degree of spectral agility at the expense of strong correlation between the channels. Additionally, our model embeds the CRs buffer, which allows a more accurate performance evaluation of CR in general and obtaining the delay performance in particular. The authors of [18] derive a queueing framework to study the performance of CRN accessing the spectrum in an opportunistic manner. Although this model allows an analytic study of CRN, it lacks the modeling of the cognition cycle as it neglects the phase of environment sensing and its state estimation. A basic version of cognition cycle model is given in our previous work [19]. However it lacks the sensing-transmission tradeoff and penalty for interference with PU.

This paper presents three significant contributions to the problem of modeling the CRN. Firstly, we enhance the model of [19] by introducing CRN with penalty for interfering PU. The penalty provides an incentive for CRN to enhance its perception level in order to avoid interference with PU, which is an essential requirement in any realistic scenario. Secondly, we introduce a decision-making process, which is responsible both for selecting the channels to be accessed and for managing the sensing-transmission tradeoff [10–13]. Thirdly, we propose a cross-entropy optimized policy for controlling the CRN. The task of policy optimization is rather hard due to the high complexity of the model. To overcome this problem we use the method of stochastic optimization of cross-entropy, which is an efficient tool at hand for the task of policy optimization [26]. The resulting policies reflect the intelligent behavior induced by the described above cognition cycle.

This paper is organized as follows. In Sect. 2 we present a stochastic model of the cognition cycle. In Sect. 3 the model is analyzed using Matrix-Geometric approach. The numerical analysis of the proposed model introduces insights into the performance of CRN. In particular, we establish the relations between different quantities such as input rate, environment state estimation rate and delay of SU. In Sect. 4 we use the cross-entropy method to optimize the allocation of the CRN resources. Section 5 summarizes the work.

2 Cognition cycle model

We regard the cognition cycle as an aggregation of interdependent processes through which CRN interacts with the communication environment. CRN access channels temporarily unoccupied by PU in order to transmit data. We denote by S_t the state of the environment at time t, which is actually the number of available channels for CRN access. In the case when CRN tries to access channels erroneously estimated as vacant, the transmissions fail. This penalty for interference with PU implies a significant incentive for CRN to allocate resources required for enhancing its perception level.

The perception process consists of sensing the environment and estimating its state. We denote by $\hat{S}_t$ the estimation of the environment state S_t. CRN observes the environment by sensing the network channels and it has some control over the observation process by deliberately tuning the sensing rate over time. For example, CRN could increase the sensing rate in order to keep track of rapidly changing network states characterized by high throughput potential while decreasing it for slowly changing states. Since CRN estimates the network state, we assume that the sensing rate may depend on $\hat{S}_t$. It is reasonable to assume that due to the physical and the hardware limitations, the transmissions and the observations are mutually exclusive and hence the sensing and the transmission rates are negatively correlated. This condition forms the throughput-sensing tradeoff which was the focus of the study in [10–13] and is an intrinsic part of our model as we point out later.

In the following subsections, we model the environment's dynamics, the cognition cycle and the CRN data transmission process. Then, we unify these models under the entire system framework. Closing the loop makes it possible to analyze the cognition cycle and to evaluate the performance of the CRN. We assume that CRN knows the correct models of both the environment and the transmitter. This assumption reduces the need of updating the models (for example through machine learning methods) and allows us to focus on modeling the perception and decision making phases of the cognition cycle.

2.1 Environment model

In the scenario under consideration, CRN accesses the network channels in an opportunistic manner to create virtual unlicensed bands, i.e., bands that are shared with PU on a non-interfering basis. We consider a general scenario of wireless communication system which consists of M channels. There are M PUs in the system, while every PU has an exclusive access to a single channel. Every PU alternates between transmitting and idle states. The ON (OFF) period of a channel corresponds to the time interval T_{ON} (T_{OFF}) during which a PU is transmitting (idle). We assume that T_{ON} and T_{OFF} intervals are exponentially distributed with parameters α and β, respectively.

CRN uses the channels to form a pool of M spectral bands. In this mode of operation CRN look for "holes" in the spectrum and dynamically adapt its transmissions over unused bands. The holes do not have to be contiguous [20]. Additionally, once CRN detects PU appears in a frequency band all SU leave this band immediately, giving priority to PU and avoiding interference. Since the PU are statistically independent, the number of bands available for SU access S_t ($S_t \in \{0, 1, \ldots, M\}$) at time t is a continuous time Markov chain (CTMC) (see Fig. 1) with transition rates q_{ij} given by.

$$q_{ij} = \begin{cases} (M-i)\alpha, & j = i+1, 0 \le i < M \\ i\beta, & j = i-1, 0 < i \le M \\ 0, & \text{else} \end{cases} \quad (1)$$

2.2 Perception model

Here we model the perception process, which is an aggregation of the observation and the orientation phases of the cognition cycle. In CRN these phases are generally

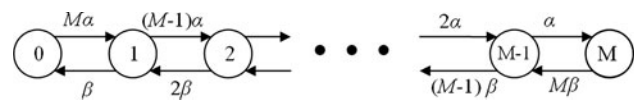

Fig. 1 Aggregate birth–death process of unoccupied bands

comprised of sensing the channels and estimating their availability. There are a lot of works in the literature that focus on the different aspects of sensing and estimation techniques [21, 22]. Instead of going into technical details of these processes, we model them at a system level.

The environment state S_t is unknown and therefore CRN estimates it through sensing. As was assumed before, CRN knows the environment model and its parameters. In our case of structured environment model the parameters are α, β and M. CRN uses the environment model and the data from sensors to obtain the estimation $\hat{S}_t$, which is the output of the unified perception phase of the cognition cycle. The perception process updates $\hat{S}_t$ at random time instants t_n, $n \in \{0, 1, 2, \ldots\}$. We assume that the time it takes to update the estimation is exponentially distributed. We denote the estimations update rate by δ_t. CRN adaptively tunes this update rate according to its current estimate $\hat{S}_t$. The notations of $\hat{S}_n$ and S_n describe the values of $\hat{S}_t$ and S_t at time t_n. At each instant t_n, an accurate sensing is assumed, in which the estimation $\hat{S}_n$ is updated to be the true value of S_n and remains unchanged till the next update instant t_{n+1}.

The compound process $Z_t = \{S_t, \hat{S}_t\}$ describes the mutual evolvement of both the environment and the estimation processes which can be shown to be a CTMC (see Fig. 2). In this CTMC the horizontal transitions describe the changes of environment state S_t. The vertical transitions describe the updates of the estimator $\hat{S}_t$ toward the correct value of S_t. Note that the states for which $S_t = \hat{S}_t$ act as absorbers of the vertical transitions. Once the process enters such a state, the vertical transitions hold off till the moment when the environment state changes.

2.3 Decision making

The decision-making phase of the cognition cycle employs some policy P for both transmission-sensing tradeoff management and for channels allocation. As we already mentioned, at any instant CRN either senses or transmits over a channel. The transmission rate of SU over a single unoccupied channel by PU is μ [bit/sec]. We introduce the tradeoff parameter θ ($0 \le \theta \le 1$) which divides the available bandwidth between the transmissions and sensing, where the portion θ of the channel is assigned for transmission and the remaining part $(1 - \theta)$ is assigned for the sensing process. For a given value of θ, the effective transmission rate over a single channel is therefore

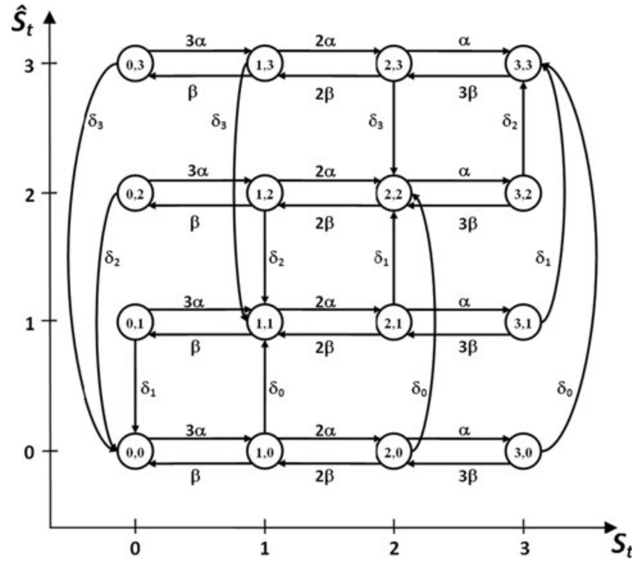

Fig. 2 CTMC of the $Z_t = \{S_t, \hat{S}_t\}$ process for $M = 3$. The horizontal transitions describe the changes of environment state S_t. The vertical transitions describe the updates of the estimator $\hat{S}_t$ to the correct value of S_t

$\theta\mu$ [bit/sec] and the resulting update rate of the estimations is $(1 - \theta)\mu B$ [1/sec]. The constant $1/B$ [bit] is the number of bits required for updating the estimation $\hat{S}_t$ and it is subject to the physical layer issues. The other responsibility of the policy P is the channel allocation C_t, which is the number of channels over which CR tries to transmit at time t.

In this work, we consider state-dependent policies, meaning that the decisions are based on the estimation of the network state $\hat{S}_t$ and the internal buffer state X_t. The internal buffer state X_t is the number of SU packets waiting for transmission at time t. For the sake of simplicity, in the following modeling we assume that CRN makes decisions based on a greedy policy P_G, $C_t = P_G(\hat{S}_t, X_t)$. The greedy policy aims to increase the throughput by scheduling transmissions over all the channels that are estimated as unoccupied by PU while keeping constant tradeoff parameter:

$$C_t = P_G(\hat{S}_t, X_t) = \begin{cases} \hat{S}_t & X_t > 0 \\ 0 & X_t = 0 \end{cases} \quad (2)$$

This assumption of greedy policy is removed later in Sect. 4 when we search for optimized policies in order to achieve better CRN performance.

2.4 Transmission process

The arrivals generated by SU are modeled as a Poisson process with rate λ [bit/sec] and service time exponentially distributed with rate μ_t [bit/sec], which changes with time dependent on a few factors. These factors are the number of accessed channels C_t, the proportion of the channels bandwidth allocated for transmission θ, the actual state of the environment S_t and the penalty for interfering with PU. The combination of these factors results in.

$$\mu_t = \begin{cases} \theta C_t \mu & C_t \le S_t \\ 0 & C_t > S_t \end{cases} \quad (3)$$

It can be seen from (2) that when the decisions are made according to the greedy policy P_G, we may substitute $\hat{S}_t$ for C_t since transmissions occur only for $X_t > 0$. From (3), our model introduces penalty for CRN when it accesses channels that are in use of PU ($C_t > S_t$). This type of service models the opportunistic spectrum access of CRN giving the highest priority to PUs. For example, during the periods when all the bands are occupied by PUs ($S_t = 0$) no CRN packets are transmitted independently of $\hat{S}_t$.

2.5 System process

Now we aggregate the environment dynamics, the cognition cycle and the transmission process into an unified system model. We define $\{X_t, Z_t\}$ to be the process of the entire system for which at time t there are X_t ($X_t \in \{0, 1, 2, \ldots\}$) queued packets of SU, which is the level of the process, and $Z_t = \{S_t, \hat{S}_t\}$ ($Z_t \in \{0, \ldots, M\} \times \{0, \ldots, M\}$), which is state within the level. This process forms a three dimensional CTMC illustrated in Fig. 3, which is homogeneous, irreducible and stationary.

The exact structure of transitions within the CTMC and its analysis by means of matrix geometric approach are presented in Appendix 1. In the next section, we use the results of the analysis of the CTMC to evaluate the performance of the CRN described by our model.

3 CRN performance evaluation

We first aim at evaluating the performance of the estimator $\hat{S}_t$. The mean square error (MSE) of an estimator is a common way to evaluate its performances. MSE quantifies the difference between an estimator and the true value of the quantity being estimated. In our case

$$MSE = E\left[\left(\hat{S}_t - S_t\right)^2\right] = \sum_{i,j=0}^{M} (i-j)^2 \Pr\left(\hat{S}_t = i, S_t = j\right) \quad (4)$$

The probabilities $\Pr(\hat{S}_t = i,\ S_t = j)$ can be easily obtained by solving a CTMC of the Z_t process, like the one presented in Fig. 2. Based on the probabilities $\Pr(\hat{S}_t = i,\ S_t = j)$, it is possible to calculate various performance measures of interest. For example, $\Pr(\hat{S}_t = S_t)$ is the proportion of time CRN estimates correctly the environment state. Another example is the

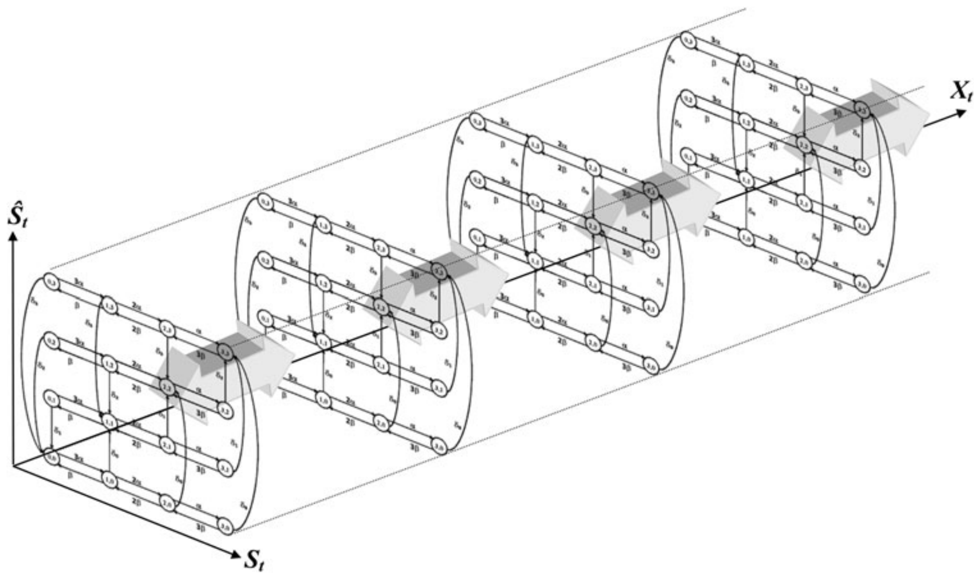

Fig. 3 Illustration of the CTMC of the CRN model. The transitions in the $(S_t, \hat{S}_t)$ plane are identical to those in Fig. 2. The transitions between the levels of the process (along the X_t axis) are ommited here in sake of keeping visuabilty, they are presented in the Appendix 1

probability of interference of SUs with PUs. In the special case when SU access all the channels that are estimated as available for access, this probability is given by $\Pr(\hat{S}_t > S_t)$.

An interesting observation, which characterizes the performance of CRN, is its dependence not only on the fraction of time available for SU to access the channel, but also on the pattern of spectrum usage of PU. Let $k > 0$ be a scaling factor that multiplies both the transition rates α and β. It is obvious that the fraction of time that the channel is available for the SU remains constant for all $k > 0$. The difference k causes is in the pattern of spectrum usage of PU. For low values of k the rates are slow and PU are characterized by a persistent behavior in which they remain in transmitting or idle states for long periods. When k is high, PU behave in an oscillatory manner alternating quickly between the transmitting and idle states. In particular, the average length of the OFF period becomes $1/(k\beta)$, and the average length of the ON period becomes $1/(k\alpha)$. Figure 4 shows that MSE of the estimator $\hat{S}_t$ improves for increasing values of δ_t. As the PU oscillate more frequently (increased k) the update rate δ should be significantly increased in order to keep the same MSE value.

Next we assess the communication performance of the SU. Since there is no loss of packets in our model, as long as the system is stable, the throughput of the SU equals to the rate of the arrivals λ. Next, we focus on the delay of SU. In Appendix 1, a way to obtain π_0 and R is presented. Using these quantities we calculate N the number of SU's packets in the system. N consists of the packets queued in the buffer and the packet in the transmitter:

$$N = \sum_{j=1}^{\infty} j\pi_j e = \pi_1 \sum_{j=1}^{\infty} jR^{j-1} e = \pi_0 R(I-R)^{-2} e \qquad (5)$$

where e is a column vector of 1's of length $(M+1)^2$, $\pi_i \equiv (\pi_{i,0}, \pi_{i,1}, \ldots, \pi_{i,M^2})$ and $\pi_{i,j}$ are the stationary

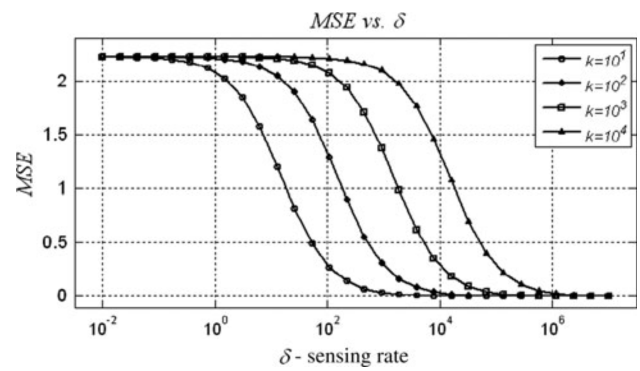

Fig. 4 MSE of $\hat{S}_t$ for parameters $M = 5$, $\alpha = 0.5$, $\beta = 1$, $\mu = 1$, $\lambda = 1$, $k = \{10, 100, 1{,}000, 10{,}000\}$. The MSE improves for growing values of δ_t

probabilities of the process $\{X_t, Z_t\}$ to be at level i and state j within that level. Using N_q and Little's law, we can obtain the average delay (waiting time and service time) of SU:

$$W = N/\lambda = \left(\pi_0 R(I-R)^{-2} e\right)/\lambda \qquad (6)$$

We examine the case when the estimation is perfect, i.e. $\hat{S}_t = S_t$ for all t. This situation is achieved for $\delta_t \to \infty, \forall \hat{S}_t$ The resulting performance of SU dependent on k is presented in Fig. 5. It is noticeable that for persistent behavior of the PU, the CRN performance weaken and SU have to wait longer periods on average although the channel is available for secondary access the same fraction of time.

Next we examine the behavior of the average delay for finite update rates, see Fig. 6. When the update rate $\delta_t\,(\forall \hat{S}_t)$ is significantly higher than the transition rates (α, β), the estimator $\hat{S}_t$ is characterized by a small MSE (see Fig. 7). As a result the curves coincide in the corresponding interval of k values. In this case the average delay behaves in the same

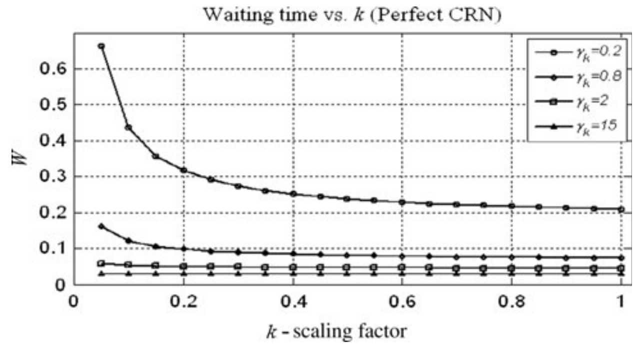

Fig. 5 Average delay versus k for different values of γ. Parameter values $\alpha = 1$, $\beta = 2$, $\mu = 1$, λ varies in the stable region of the system $\rho = \lambda/\mu < M\alpha/(\alpha + \beta)$

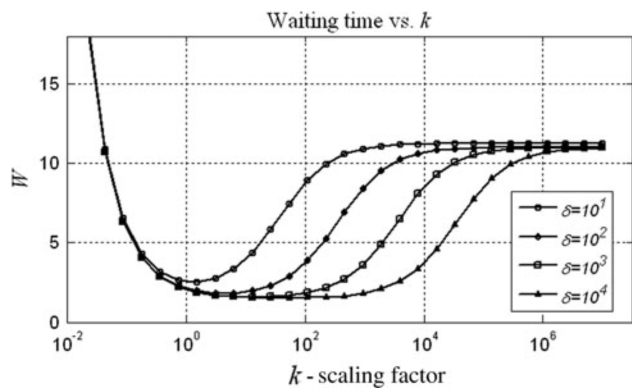

Fig. 6 Average delay of CTMC model for parameters $M = 5$, $\alpha = 0.5$, $\beta = 1$, $\mu = 1$, $\lambda = 1$, $\delta = \{10, 100, 1{,}000, 10{,}000\}$

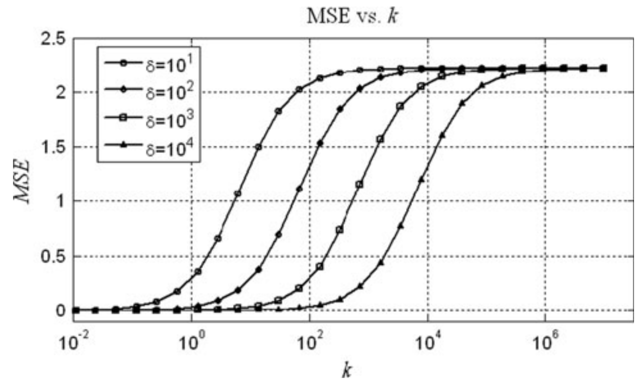

Fig. 7 MSE of $\hat{S}_t$ for parameters $M = 5$, $\alpha = 0.5$, $\beta = 1$, $\mu = 1$, $\lambda = 1$, $\delta = \{10, 100, 1{,}000, 10{,}000\}$

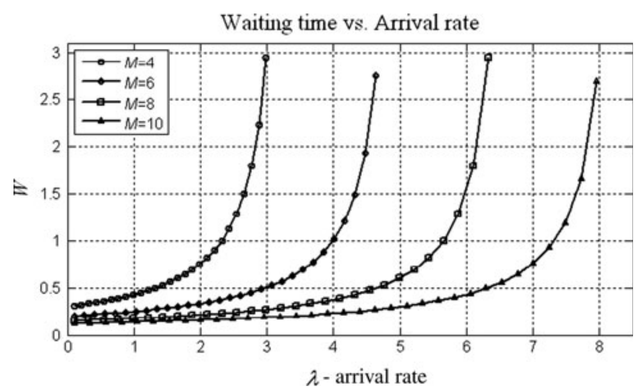

Fig. 8 Average delay of CTMC model for $M = \{4, 6, 8, 10\}$, $\alpha = 10$, $\beta = 2$, $\mu = 1$

manner as if the estimation process had small MSE. However, when the transition rates grow, the performance of CRN becomes sensitive to the estimation process. As it can be seen from Fig. 6, for $k > 1$, the accuracy of the estimation process affects the performance of SU significantly. When the updates of $\hat{S}_t$ occur too slowly compared to the environment dynamics, the average delay increases. Each of the curves describes longer average delay dependent on δ. Further, it can be seen from the graphs that the average delay saturates in a rapidly changing environment, however the system remains stable. This can be explained by the fact that when the environment state fluctuates quickly, the probability $\Pr(S_t = i, \hat{S}_t = j)$ remains positive and independent of k or δ, which can be seen from Fig. 7.

As a summary of the analysis we plot in Fig. 8 the performance curves of CRN for different values of M assuming perfect estimation. It is clear that the performance for different systems (different M values) saturate when λ approaches high values. It is interesting to notice that from this plot one can learn about system trade-offs. For example, one can answer the question whether better performance could be reached by splitting the SU in two groups generating half the original traffic rate (0.5λ) and with separated spectrum pools of $M/2$ channels. Comparing the average delay at the point $\lambda = 6$ on curve $M = 8$ to the point $\lambda = 3$ on curve $M = 4$ shows that using a larger spectrum pool improves the performance.

4 CRN policy optimization

In previous sections, we modeled the cognition cycle, analyzed it and evaluated its impact on the performance of the CRN. The evaluation was carried out for CRN that makes decisions based on some arbitrarily chosen greedy policy P_G. In this section, we aim to improve the performance of the cognition cycle and the CRN by optimizing the decision making process, i.e., by optimizing the policy.

4.1 Problem formulation

In our framework, a policy P governs the decision-making phase of the cognition cycle. This policy is responsible for managing the sensing-transmission tradeoff by tuning the continuous parameter θ_t, and for allocation of channels, C_t. The values of C_t and θ_t are determined dependently on current estimation of the networks state $\hat{S}_t$, current CRN buffer state X_t, entire system model and its parameters, which we denote by Ω:

$$(C_t, \theta_t) = P(\hat{S}_t, X_t; \Omega) \tag{7}$$

We aim to optimize CRN performance by minimizing the average average delay W of SU. In the previous section, we calculated W by applying the matrix geometric analysis to the 3-D CTMC and the Little's law. The 3-D CTMC structure embeds the policy P as follows: the levels transitions (X_t) are affected by the service rate μ_t (3) and the state transitions (Z_t) within the level are affected by the estimation update rate δ_t given by $\delta_t = (1 - \theta_t)\mu B$. Therefore, given the system structure and its parameters Ω, we regard the average average delay W of SU as a function of the policy P, $W = W(P; \Omega)$. The resulting optimization problem is given by:

$$P^* = \arg\min_{P \in \Pi} W(P; \Omega) \tag{8}$$

where Π is the set of all the feasible policies, i.e., policies which for valid inputs $\hat{S}_t \in \{0, \ldots, M\}$ and $X_t \in \{0, 1, 2, \ldots\}$ decide on valid values of $\theta \in [0,1]$ and $C_t \in \{0, 1, 2, \ldots, M\}$. Our optimization problem (8) is a complicated one. First, it can be shown that the problem is not convex, and the gradient-based techniques are not applicable since it is difficult to obtain a gradient for W. Next, the set Π consists of policies comprising both continuous (θ) and discrete (C_t) action spaces, which requires special approach for optimization. Additionally, the problem exhibits a high computational complexity, due to the rapidly growing (with M) set of feasible policies Π.

We solve this problem by applying the cross-entropy (CE) method of stochastic optimization. CE method is a state-of-the-art method for solving combinatorial and multi-extremal optimization problems. In the following subsection, we review briefly the CE method and demonstrate its application for our optimization problem. The readers interested in further details are referred to [25].

4.2 Cross-entropy based stochastic optimization

The main idea behind the CE method is to define for the original optimization problem an associated stochastic problem and then to solve efficiently the associated problem by an adaptive scheme. The described below procedure sequentially generates random solutions which converge stochastically to the optimal or near-optimal one.

We define a stochastic policy $P((C_t, \theta_t)|\sigma(\hat{S}_t, X_t))$ as the associated stochastic problem for (8). $P((C_t, \theta_t)|\sigma(\hat{S}_t, X_t))$ is the probability of choosing action (C_t, θ_t) when CRN's state is $(\hat{S}_t, X_t)$ according to the parameter $\sigma(\hat{S}_t, X_t)$. In the following we use shorthand notation of σ for $\sigma(\hat{S}_t, X_t)$. For the defined associated stochastic problem, the CE method iteratively draws sample policies $P^{(j)}$ ($j = 1, 2, \ldots, J$) from the defined above probability and calculates the average delay $W(P^{(j)}; \Omega)$ for each sample. Then N ($N < J$) best samples graded by their related average delay are used to update the parameters σ, in order to produce better samples in the next iteration. The algorithm stops when the score of the worst selected sample no longer improves significantly. The exact CE algorithm is presented in Appendix 2.

4.3 Cross-entropy optimized policies

We present here policies obtained from CE optimization and examine them in order to get insights concerning the optimal decision-making process in CRN. As in the previous sections we are interested to reveal the impact of the cognition cycle and the dynamics of the environment on the optimal policy. We set the parameters of the environment (Ω): the number of PU channels is $M = 6$, and the transmission rate over every channel is $\mu = 1$, the constant B is set to unity, the parameters responsible for the environment dynamics are set to $\alpha = \beta = k$—as before we will check the performance for different values of $k = \{0.001, 1, 1{,}000\}$, the arrival rate of CRN traffic is $\lambda = 4$. This set of parameters Ω initializes the algorithm for CE based policy search described in Appendix 2, the additional parameters controlling the algorithm are: population size $N = 1{,}000$, number of best samples $J = 10$, maximum iterations $T = 100$, threshold values $d = 5$ and $\varepsilon = 1e - 4$.

In our associated stochastic problem the policy chooses action (C_t, θ_t) when CRN is in state $(\hat{S}_t, X_t)$. We assume that, C_t is a discrete random variable that takes integer values $\{0, 1, \ldots, M\}$, while the tradeoff parameter θ_t is normally distributed according to a truncated normal distribution in the range [0,1]. Note that our policy is state dependent. We distinguish between the cases $X_t = 0$ and $X_t > 0$. Obviously, for $X_t = 0$ CRN has no packets to transmit and in this case it is reasonable to allocate the bandwidth resources to the sensing process ($\theta_t = 1$). The CE algorithm optimizes the policy for $X_t > 0$.

The resulting CE optimized policies are presented in Fig. 9. For the case $k = 1{,}000$, CRN fails to keep track of the rapidly changing network state. This can be seen through the channel allocation $C = (3, 3, 3, 3, 3, 3, 4)$, which is insensitive to the estimation $\hat{S}_t$, and the number of accessed channels is approximately the average number of unoccupied channels $E[S_t]$. Nevertheless, the tradeoff parameter $\theta = (0.57, 0.91, 0.93, 0.99, 0.72, 0.59, 0.37)$ shows that CRN tries to avoid collisions with PU; a simple analysis of the CTMC (in Fig. 2) shows that for $\alpha = \beta$, S_t resides only a small portion of time in the states 0 and M while it spends more time in the inner states. This fact is reflected in the low values of θ when $\hat{S}_t$ is 0 or M. In order to better react to the fast network changes, CRN accelerates the sensing rate $\delta = (1 - \theta)\ \mu B$ in these states.

For the case $k = 1$, the resulting policy is more sensitive to the estimation of the environment state $\hat{S}_t$, and the number of accessed channels $C = (1, 1, 2, 3, 4, 5, 5)$ is approximately $\hat{S}_t$ except for the rapidly switching states 0 and M. As in the previous case, the tradeoff parameter $\theta = (0.3, 0.7, 0.8, 0.92, 0.92, 0.87, 0.7)$, allocates more bandwidth for transmissions when $\hat{S}_t$ indicates that the network state is a persistent one. When the environment changes occur in a significantly slower manner compared to the rate of the perception process $k = 0.001$, the tradeoff parameter $\theta = (0.99, 0.99, 0.99, 0.99, 0.99, 0.98, 0.98)$ takes very high values independently of the estimation $\hat{S}_t$. The allocation of the channels $C = (1, 1, 2, 3, 4, 5, 6)$ is equal to $\hat{S}_t$ even for the rapidly switching state M.

Note that the optimized policies allow the sensing rate and the number of accessed channels to be a function of the current state estimation. This is crucial, because when some states are very likely to persist for longer periods, the cognition cycle may choose a more efficient course of action.

In Fig. 10, we compare the performance of the CRN under greedy and CE optimized policies. Under the greedy policy, the average delay W decreases when the network transitions accelerate ($k < 0.1$). This happens since CRN tracks well the channels and efficiently utilizes the vacant ones. For $k > 0.1$, W grows since CRN fails to track the fluctuating state of the network. When comparing the two policies, it can be seen that the average delay, under CE optimized policy, does better in orders of magnitude for the entire range of network dynamics ($k \in [10^{-3}, 10^{4}]$).

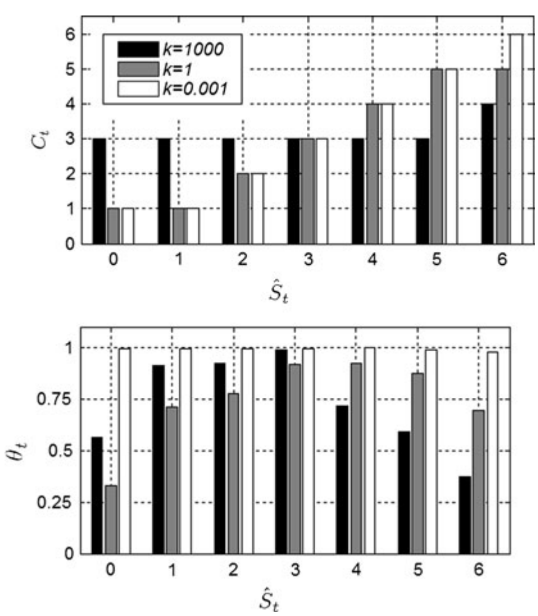

Fig. 9 The CE optimized policy for parameters $M = 6$, $\mu = 1$, $\lambda = 4$, $\alpha = \beta = k = \{1{,}000, 1, 0.001\}$

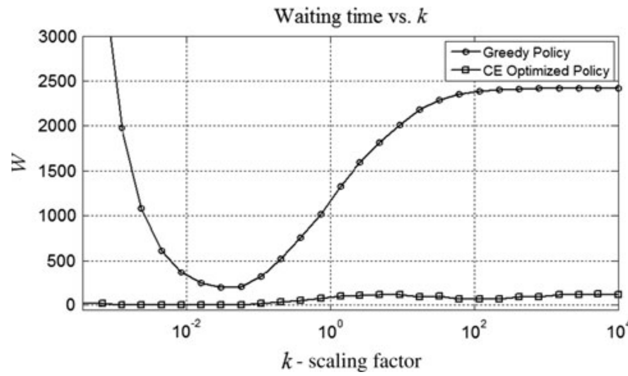

Fig. 10 CRN average delay under greedy and CE optimized policies for parameters $M = 4, B = 1, \mu = 1, \lambda = 0.5, \alpha = \beta = k \in [10^{-3}, 10^{4}]$

4.4 Protection of PUs' transmissions

Generally, the PUs legacy communication systems lack the mechanisms for coordination of their medium access with SUs. Therefore, it is CRN responsibility to maintain hierarchical operation of DSA scheme, i.e., giving priority and reducing interference to PUs. Here, we study the important objective of CRN to protect PUs' transmissions.

We define I to be the proportion of PUs that are interfered on average by CRN transmissions. This performance measure is given by:

$$I = \frac{1}{M} \sum_{i,j=0}^{M} [C(i) - j]^{+} \Pr\left(\hat{S}_t = i, S_t = j\right) \quad (9)$$

where $[x]^{+} = \max(x, 0)$ is a positive part function and $C(i)$ is the number of channel accessed by CRN when the estimation is $\hat{S}_t = i$. It is obvious that optimizing for I would result in an undesirable policy for which $C(i) = 0$, $i \in \{0, 1, 2, \ldots, M\}$. Instead, this metric could serve as additional criteria for evaluation of the CE optimized policies. We calculate the interference for the three cases presented in the previous section. Calculations show that for the values $k = 1{,}000$, 1 and 0.001, we obtain $I = 0.08$, 0.12 and 0.03, respectively. The explanation for this non-monotonic behavior comes from looking closely at the policies. It turns out that for $k = 1$, the CE optimized policy tries to access the channels in a greedy manner while it can be seen from the corresponding values of θ, that the sensing process requires significant rates. This means that in some cases the optimization of CRN performance could result in high interference to PU's. In these situations a special care should be taken. For example, the CRN could be designed to deliberately reducing the number of the channels accessed till the interference decreases bellow some threshold value.

5 Summary and future work

In this paper, a three-dimensional CTMC process has been introduced to model the operation of CRN where PU form a birth–death process and SU can queue. The analytical framework combines the environment dynamics, perception and decision making components of the cognition cycle and the spectrum access processes. The model was analyzed using matrix geometric approach. The analysis results give insights about the behavior of CRN in general and the impact of sensing rate and the system dynamics on the average delay of secondary users in particular.

The cognition cycle is treated as an integral part of the system's overall behavior, and we optimize policies controlling simultaneously the interdependent perception and transmission processes. In this way, the resources are allocated according to the needs of the overall task. The CE optimized policies demonstrate adaptive behavior in which the resources are intelligently allocated to the perception and the transmission processes in a task-relevant manner. In the future work we plan to further enhance the proposed perception model by introducing inaccurate sensing and prediction of the environment state changes.

Appendix 1: Analysis of the 3-D CTMC

We present here the analysis of the 3-D CTMC in Fig. 4.

CTMC structure

In order to make the analysis of the system easier we numerate the states of Z_t lexicographically, i.e. (0, 0), (0, 1), …, (0, M), (1, 0), (1, 1), …, (M, M) and index them 1 to $(M+1)^2$. This new order of states turns our CTMC to two dimensional since now $Z_t \in \{1, 2, \ldots, (M+1)^2\}$. Then again we order the states lexicographically, i.e. (0, 1), (0, 2), …, (0, $M+1$), (1, 1), (1, 2), … and construct the generator matrix Q of this CTMC which is given by:

$$Q = \begin{pmatrix} B_{00} & B_{01} & 0 & 0 & 0 & \\ B_{10} & B_{11} & A_0 & 0 & 0 & \\ 0 & A_2 & A_1 & A_0 & 0 & \cdots \\ 0 & 0 & A_2 & A_1 & A_0 & \\ 0 & 0 & 0 & A_2 & A_1 & \\ & & \vdots & & & \ddots \end{pmatrix}$$

where $B_{00} = \{B_{00}(i, j)\}$, $B_{01} = \{B_{01}(i, j)\}$, $B_{10} = \{B_{10}(i, j)\}$, $B_{11} = \{B_{11}(i, j)\}$, $A_0 = \{A_0(i, j)\}$, $A_1 = \{A_1(i, j)\}$ and $A_2 = \{A_2(i, j)\}$ are $(M+1)^2 \times (M+1)^2$ matrices. A 0 entry in Q (and in other matrices) is a matrix of all zeros of the appropriate dimension. It can be seen that in our model $B_{01} = A_0 = \mathrm{diag}\{\lambda, \lambda, \ldots, \lambda\}$. For each value $z_{i,j} = (i, j)$ the process Z_t can take, the service rate is $\mu_{i,j} = \mu \min\{i, j\}$. We order the elements $\mu_{i,j}$ in the same way as we did for Z_t and obtain a vector of service rates $\underline{\mu}$. It can be seen that $B_{10} = A_2 = \mathrm{diag}\{\underline{\mu}\}$, while the matrices B_{00} and $B_{11} = A_1$ are more complicated:

$$B_{00}(i,j) = \begin{cases} -(\lambda + (M - \lfloor i/(M+1) \rfloor)\alpha + \lceil i/(M+1) \rceil \beta) & j = i \\ (M - \lfloor i/(M+1) \rfloor)\alpha & j = i + M \\ \lceil i/(M+1) \rceil \beta & j = i - M \\ \delta & j = \lfloor i/(M+1) \rfloor (M+2) \cap i \neq j \\ 0 & \text{else} \end{cases}$$

and

$$A_1(i,j) = \begin{cases} -(\lambda + i\mu + (M - \lfloor i/(M+1) \rfloor)\alpha + \lceil i/(M+1) \rceil \beta) & j = i \\ (M - \lfloor i/(M+1) \rfloor)\alpha & j = i + M \\ \lceil i/(M+1) \rceil \beta & j = i - M \\ \delta & j = \lfloor i/(M+1) \rfloor (M+2) \cap i \neq j \\ 0 & \text{else} \end{cases}$$

Stationary probabilities

We define the stationary probabilities $\pi_{i,j}$ of the process to be at level i and state j within that level. Calculating the stationary probabilities will allow evaluating interesting quantities, mainly the average delay of SU. The calculations here follow [23] and are adopted for our model.

Let $\pi_i \equiv (\pi_{i,1}, \pi_{i,2}, \ldots, \pi_{i,M^2})$ and $\pi \equiv (\pi_0, \pi_1, \pi_2, \ldots)$. The stationary distribution is the unique set of $\pi_i \geq 0$, $i \geq 0$, that solves

$$\begin{cases} \pi Q = \underline{0} \\ \pi \underline{e} \; = \; 1 \end{cases} \tag{10}$$

where $\underline{e}$ ($\underline{0}$) denotes an appropriately dimensioned column (row) vector of 1's (0's). From the first equation in (10) we may write down for the repeating portion of the process ($j \geq 1$):

$$\pi_{j-1}A_0 + \pi_j A_1 + \pi_{j+1}A_2 = \underline{0} \quad (j \geq 1) \tag{11}$$

For this type of CTMC characterized by a boundary conditions in the first column of Q followed by a repetitive portion of columns containing matrices A_0, A_1 and A_2, there exist some constant matrix R such that

$$\pi_j = \pi_{j-1}R, \qquad (j \geq 1) \tag{12}$$

and that the values of $\pi_j, j \geq 1$, have a matrix geometric form:

$$\pi_j = \pi_0 R^j, \quad (j \geq 1) \tag{13}$$

substituting (13) into (11) yields

$$A_0 + RA_1 + R^2 A_2 = 0 \tag{14}$$

This quadratic equation in R is typically solved numerically. There is more than one R that solves (14). When the CTMC is ergodic, there is an unique stationary distribution π that satisfies (10). Analogous to the scalar case where the utilization factor should be less than unity, in our case all eigenvalues of R must be less then unity for the normalization constraint in (10) to hold [24].

After solving for R, in order to determine the stationary probabilities, we continue with the boundary conditions:

$$\pi_0 B_{00} + \pi_1 A_2 = \pi_0(B_{00} + RA_2) = 0 \tag{15}$$

Equation (15) alone is not enough to solve for π_0 since it is not of full rank and we must use the normalization constraint in (10):

$$\pi e \; = \; \left(\sum_{j=0}^{\infty} \pi_j\right) e = \pi_0 (I - R)^{-1} e = 1 \tag{16}$$

Combining (15) and (16) we have

$$\pi_0 \left[(I - R)^{-1} \underline{e}, \; (B_{00} + RA_2)^* \right] = [1, \underline{0}] \tag{17}$$

where $(B_{00} + RA_2)^*$ is the result from removal of the first column from the matrix $(B_{00} + RA_2)$, and $[1, \underline{0}]$ is a row vector consisting of a 1 followed by $(M + 1)^2 - 1$ zeros. Equation (17) is solved by appropriate numerical methods.

Appendix 2: Cross-entropy algorithm for CRN policy optimization

In this appendix we present the CE algorithm for CRN policy optimization.

Input:

- function $W(P; \Omega)$
- system parameters $\Omega = \{\alpha, \beta, M, \lambda, \mu\}$
- probability density families $\{p_C(\cdot; \sigma_C)\}$ and $\{p_\theta(\cdot; \sigma_\theta)\}$,
- initial parameters $\sigma_{C,0}$ and $\sigma_{\theta,0}$
- parameters N, J, T, d, ε
- $t \leftarrow 0$

Repeat

1: Generate samples $C^{(j)}$ ($j = 1, 2, \ldots, J$) from $p_C(\cdot; \sigma_{C,t-1})$
2: Generate samples $\theta^{(j)}$ ($k = 1, 2, \ldots, J$) from $p_\theta(\cdot; \sigma_{\theta,t-1})$
3: Compose policy samples $P^{(j)} = (C^{(j)}, \theta^{(j)})$ ($j = 1, 2, \ldots, J$)
4: Calculate $W^{(j)} = W(P^{(j)}; \Omega)$ for each sample ($j = 1, 2, \ldots, J$)
5: Keep N ($N < J$) best samples graded by their $W^{(j)}$ value and discard the other samples
6: $V_t = \min_j(W^{(j)})$ (minimize over the saved N best samples)
7: Using the N best samples update the parameters

$$7.1\text{: } \sigma_{C,t} \leftarrow \arg\max_{\sigma_C} \sum_{n=1}^{N} \ln\left(p_{\sigma_C}\left(C^{(n)}; \sigma_C\right)\right)$$

$$7.2\text{: } \sigma_{\theta,t} \leftarrow \arg\max_{\sigma_\theta} \sum_{n=1}^{N} \ln\left(p_{\sigma_\theta}\left(\theta^{(n)}; \sigma_\theta\right)\right)$$

8: $t \leftarrow t + 1$

Until ($t > T$ or $|V_t - V_{t-\tau}| < \varepsilon$, $\tau = 1, 2, \ldots, d$)
Output:
$P^* = (C^*, \theta^*)$—best sample, $W^* = W(P^*; \Omega)$—best value

References

1. Zhao, Y., Mao, S., Neel, J., & Reed, J. (2009). Performance evaluation of cognitive radios: Metrics, utility functions and methodologies. *Proceedings of the IEEE, 97*(4), 642–659.
2. Mitola, J., et al. (1999). Cognitive radio: Making software radios more personal. *IEEE Personal Communications, 6*(4), 13–18.
3. Mitola, J. (2000). Cognitive radio—an integrated agent architecture for software defined radio, Ph.D. dissertation, Royal Institute of Technology, Kista, Sweden, May 8 2000.
4. Haykin, S. (2005). Cognitive radio: Brain-empowered wireless communications. *IEEE Journal on Selected Areas in Communications, 23*, 201–220.
5. Federal Communications Commission (2003). ET Docket No 03-222 Notice of proposed rule making and order, December 2003.

6. Federal Communications Commission (2002). Spectrum Policy Task Force Report. ET Docket No. 02-135, November 2002.
7. FCC (2003). Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies, notice of proposed rule making and order, Docket No. 03-108, Dec 30, 2003.
8. Mitola, I. J. (1993). Software radios: Survey, critical evaluation and future directions. *IEEE Aerospace and Electronic Systems Magazine, 8*, 25–36.
9. Akyildiz, I. F., Lee, W.-Y., Vuran, M. C., & Mohanty, S. (2006). NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey. *Computer Networks, 50*, 2127–2159.
10. Hoang, A. T., Liang, Y. C. (2007). Adaptive scheduling of spectrum sensing periods in cognitive radio networks, In *Proceedings of IEEE GlobeCommunications*, Washington.
11. Liang, Y. -C., Zeng, Y., Peh, E., Hoang, A. T. (2006). Sensing-throughput tradeoff for cognitive radio networks, In *Proceedings IEEE International Conference Communications (ICC)*, June 2006, pp. 5330–5335.
12. Pei, Y., Hoang, A. T., Liang, Y. -C. (2007). Sensing-throughput tradeoff in cognitive radio networks: How frequently should spectrum sensing be carried out? In *Proceedings of the 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'07)*, Athens, Greece, September 2007.
13. Ghasemi, A., Sousa, E. S. (2007). Optimization of spectrum sensing for opportunistic spectrum access in cognitive radio networks, In *Proceedings of 4th IEEE Consumer Communications. Networking Conference (CCNC),* Jan. 2007, pp. 1022–1026.
14. Zhao, Q., Tong, L., Swami, A., & Chen, Y. (2007). Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework. *IEEE Journal on Selected Areas in Communications, 25*(3), 589–600.
15. Cormio, C., & Chowdhury, K. R. (2009). A survey on MAC protocols for cognitive radio networks. *Ad Hoc Networks, 7*(7), 1315–1329.
16. Maldonado, D., Le, B., Hugine, A., Rondeau, T. W., & Bostian, C. W. (2005). *Cognitive radio applications to dynamic spectrum allocation: A discussion and an illustrative example*. Nov: DySPAN.
17. Shu, T., Krunz, M. (2009). Throughput-efficient sequential channel sensing and probing in cognitive radio networks under sensing errors. In *Proceedings of the 15th Annual international Conference on Mobile Computing and Networking* (Beijing, China, September 20–25, 2009). MobiCom '09. ACM, New York, NY, pp 37–48.
18. Rashid, M. M., Hossain, M. J., Hossain, E., & Bhargava, V. K. (2009). Opportunistic spectrum scheduling for multiuser cognitive radio: A queueing analysis. *Transaction Wireless Communication, 8*(10), 5259–5269.
19. Oklander, B., & Sidi, M. (2010). *Modeling and analysis of system dynamics and state estimation in cognitive radio networks, PIMRC'10 (CogCloud Workshop)*. Turkey: Istanbul.
20. Weiss, T. A., & Jondral, F. K. (2004). Spectrum pooling: An innovative strategy for the enhancement of spectrum efficiency. *IEEE Communications Magazine, 42*(3), 8–14.
21. Yucek, D. T., Arslan, H. (2009). A survey of spectrum sensing algorithms for cognitive radio applications. IEEE Communications Surveys & Tutorials, 11(1). First Quarter 2009.
22. Zeng, Y., Liang, Y., Hoang, A. T., Zhang, R. (2010). A review on spectrum sensing for cognitive radio: Challenges and solutions. EURASIP Journal on Advances in Signal Processing Volume 2010.
23. Nelson, R. (1991). Matrix geometric solutions in markov models: A mathematical tutorial, IBM Research Report RC 16777, 1991.
24. Horn, R. A., Johnson, C. R. (1987). Matrix analysis, Cambridge, 1987.
25. Rubinstein, R. Y., & Kroese, D. P. (2004). The cross entropy method. A unified approach to combinatorial optimization, Monte-Carlo simulation, and machine learning. In M. Jordan, J. Kleinberg, B. Scholkopf, F. Kelly, & I. Witten (Eds.), *Information science and statistics*. Heidelberg: Springer.
26. Mannor, S., Rubinstein, R. Y., Gat, Y. (2003). The cross-entropy method for fast policy search, In *Proceedings 20th International Conference on Machine Learning (ICML-03)*, Washington, US, 21–24 August 2003.

Author Biographies

Boris Oklander received the B.Sc., M.Sc. and Ph.D. degrees in Electrical Engineering from the Technion-Israel Institute of Technology, Haifa, Israel in 2001, 2008 and 2013, respectively. Currently he is a research associate at Intelligent Systems and Networks Group, EEE Department, Imperial College London. His current research interests are in the modeling and the performance evaluation of cognitive communications networks.

Moshe Sidi received the B.Sc., M.Sc. and the D.Sc. degrees from the Technion-Israel Institute of Technology, Haifa, Israel, in 1975, 1979 and 1982, respectively, all in electrical engineering. In 1982 he joined the faculty of Electrical Engineering Department at the Technion where he is currently a Professor holding the Louis and Miriam Benjamin Chair in Computer Communication Networks. During 2003–2004 he served as the Dean of the Electrical Engineering Department and during 2006–2012 he served as the Technion Executive Vice President for Academic Affairs. During the academic year 1983–1984 he was a Post-Doctoral Associate at the Electrical Engineering and Computer Science Department at the Massachusetts Institute of Technology, Cambridge, MA. During 1986–1987 he was a visiting scientist at IBM, Thomas J. Watson Research Center, Yorktown Heights, NY. He coauthors the book "Multiple Access Protocols: Performance and Analysis," Springer Verlag 1990. He served as the Editor for Communication Networks in the IEEE Transactions on Communications from 1989 until 1993, as the Associate Editor for Communication Networks and Computer Networks in the IEEE Transactions on Information Theory from 1991 until 1994, as an Editor in the IEEE/ACM Transactions on Networking from 1993 until 1997, and as an Editor in the Wireless Journal 1993–2001. He also served as the General Chair for Infocom 2000. He was awarded the 1983–1984 Rothschild Fellowship for postdoctoral research, the 1983–1984 Fulbright Award for postdoctoral research, the 1985–1986 Bat-Sheva De Rothschild Fund for Young Distinguished Researchers, the 1989–1990 New England

Academic Award and the 1997 Muriel and David Jacknow Award for Excellence in Teaching. His research interests are in wireless networks and multiple access protocols, traffic characterization and guaranteed grade of service in high-speed networks, queueing modeling and performance evaluation of computer communication networks.