Comput. Sci. Appl.
Volume 1, Number 4, 2014, pp. 205-216
Received: July 31, 2014; Published: October 25, 2014

Computer Science
and Applications
www.ethanpublishing.com

A POMDP Framework for Throughput Optimization MAC Scheme in Presence of Sensing Errors for Cognitive Radio Networks

John Andrew Msumba and Hongjun Xu

College of Agriculture, Engineering & Science, School of Engineering, Discipline of Electrical, Electronic & Computer Engineering, University of KwaZulu-Natal, Durban 4041, South Africa

Corresponding author: Hongjun Xu (xuh@ukzn.ac.za)

Abstract: In this paper, a POMDP (partially observable Markov decision process) framework MAC (media access control) scheme for throughput maximization in cognitive radio networks is presented. Due to hardware limitations and wireless channel dynamics which lead to spectrum sensing errors, the state of the system cannot be observed. Using cross-layer design concept, the design of spectrum sensing at physical layer and access strategy at MAC layer are considered jointly. Spectrum sensor parameter is incorporated into a state space and the problem is formulated as a constrained POMDP. By considering multiple channels with multiple secondary users with random message arrivals, the throughput performance of the proposed scheme is evaluated using greedy sensing approach in presence of spectrum sensing errors. Computer simulation is performed and the results demonstrate that the proposed scheme outperforms the random sensing scheme and approaches that of perfect sensing in terms of the overall throughput.

Key words: POMDP, spectrum sensing, cross-layer design, MAC protocol.

1. Introduction

The MAC (media access control) protocol for CR (cognitive radio) networks is very crucial part due to the fact that it is responsible for providing coexistence between PUs (primary users) and SUs (secondary users) without or with minimum interference. Thus, the MAC protocol in CR networks is responsible for channel access decision which is the key feature of CR. Spectrum sensing is required over a wide range of licensed channels and the decision needs to be made on which channels to sense and in which sequence. Due to the dynamics of wireless channel, spectrum sensing results may contain errors leading to incorrect spectrum access decision which may cause interference to PUs. In order to efficiently implement this concept of CR networks, MAC protocols with optimal decision on spectrum sensing and access have been suggested in Refs. [1-5], aiming at using some optimization techniques for optimal sensing and spectrum access decision. These optimal sensing and access decisions are based on the objectives and constraints. Primary user activities are generally random in nature and therefore stochastic optimization model is usually applied. However, most of the works in the literature consider the design of MAC and physical (PHY) layers jointly in such a way that the binary decision based on the set threshold is made on the presence or absence of the PU while very little attention is given to errors which are due to spectrum sensing.

In this paper, we develop a cross-layer MAC scheme for throughput maximization in CR networks taking into consideration errors due to spectrum sensing. We assume that the CR node has hardware limitations and therefore cannot sense the overall range of radio spectrum. Instead, the node selects a set of channels to sense and access the idle one for transmission. This leads to partial information about the state of the system. Due to dynamics of the wireless channel, spectrum sensing errors are likely to be introduced leading to imperfect spectrum access decisions to be made by a CR node. Thus, we assume that the state of the system is not fully observable due to sensing errors and hardware limitations. We model the system using the theory of POMDP (partially observable Markov decision process) which is a suitable technique in decision making under uncertainty. We extend the concepts presented by previous researchers [3, 5-7], and we hereby wish to acknowledge their valuable contribution in this topic. We analyze the system by considering multiple secondary users with random message arrivals in presence of spectrum sensing errors. This approach was initially presented by Zhao et al. [6] and this forms a bench mark for our formulation. By using cross-layer design concept, the design of spectrum sensing at PHY layer and access policy at the MAC layer are considered jointly in order to maximize the throughput of the CR user. We formulate the problem as a constrained POMDP and use techniques available in the literature to evaluate the optimal policies and hence compute the total reward. We evaluate the throughput performance of the proposed scheme using greedy sensing approach [6, 8] in presence of spectrum sensing errors at PHY layer. The simulation results demonstrate that the proposed system outperforms the random sensing scheme and approaches that of perfect sensing in terms of the overall throughput. The rest of the paper is organized as follows: Section 2 provides an overview of the POMDP framework; Section 3 presents the system model; Section 4 presents the POMDP formulation; Section 5 presents the spectrum sensing error model; Section 6 presents the way to compute the reward of the system; Section 7 presents belief function computation; Section 8 presents the joint optimal policy computation; Section 9 shows the evaluation of the value function by iteration; Section 10 presents simulation results and discussion and Section 11 presents conclusions.

2. POMDP Framework

POMDP provides a natural model for sequential decision making under uncertainty. This model augments a well researched framework of MDP (Markov decision process) to situations where the secondary user cannot reliably identify the underlying environment of spectrum occupancy state. The key characteristic that set POMDP apart from many other probabilistic models is the fact that the state is not directly observable. Instead the agent can only perceive observations which convey incomplete information about the world's state [9]. It is a very general and powerful tool which extends the application of MDP to many realistic problems. An MDP which can be viewed as an extension of Markov chains with a set of decisions (actions) and a state-based reward or cost structure has been most commonly formal model for fully-observable sequential decision process [10]. For each possible state of the process, a decision has to be made regarding which action should be executed in that state. The chosen action affects both the transition probabilities and the costs (or rewards) incurred. The goal is to choose an optimal action in every state to increase some predefined measures of performance. Formally, POMDP is characterized by seven distinct quantities namely states (S), actions (A), observations (Θ), reward (R) and the three probability distributions namely transition probabilities (P), initial belief (b_0), and observation probabilities (θ) [9]. These items together define the probabilistic system model that underlies each POMDP. In this work we do not intend to delve deep into the development and analysis of the POMDP solutions, instead we will make use of available POMDP solutions in the literature in order to formulate and model our optimization problem.

3. System Model

In this formulation we assume that a radio spectrum with N channels such that these channels are licensed to a slotted primary network. Each channel is assumed to have a bandwidth W_n (n = 1, …, N). The evolution (channel occupancy state) of these channels is assumed to follow a Markov chain model. The reason for choosing Markov process model is the fact that it is the most commonly and widely accepted approach to model a wireless channel evolution [11, 12]. Let the channel state of the n-th channel in slot t be $S_n(t) \in (0, 1)$, where "0" represents a busy channel and "1" represents an idle channel. We consider a finite state space of 2^N and that the transition probability from one state to another is known and remains unchanged for the number of slots in a specific period of time. In this formulation, we adopt the approach presented by Refs. [13, 14] whereby for the total time slot T, the transition probability from one state to another $P(i, j)$ is assumed to be known and does not change. The CR network whose users independently and selfishly searches for and utilizes instantaneous radio spectrum opportunities in these N channels is considered. Fig. 1 shows an example of a discrete Markovian model for one channel with two-states, where P_{00} is the probability of the channel to remain in the busy state, P_{01} is the probability of the channel to change from busy state to idle state, P_{11} is the probability that the channel remain in the idle state and P_{10} is the probability that the channel change from idle to busy state. Fig. 2 shows the network model for the proposed scheme whereby a set of CR users co-exists with primary network users (licensed users). In this paper we assume that the channel fading effect is neglected and that the channel sensing errors are only due to spectrum sensor. Thus, the probability of transition between the busy state and idle state, the probability of the primary user staying in the same state, the stationary distribution and the channel transition matrix can be easily evaluated. For simplicity of presentation, a system with single channel is assumed in the formulation. The transition probability matrix of this channel is $P(i, j)$ which can be represented by a 2 × 2 matrix as shown in Eq. (1) for the case of a single channel.

$$P(i,j) = \begin{bmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{bmatrix} \quad (1)$$

Each CR node is assumed to be equipped with a Neyman-Peason detector [15-17] of which its performance is represented by the ROC (receiver operating characteristic) curves of the detector. At the beginning of each time slot t, CR node with data to transmit chooses a set of channels to sense and the valid spectrum sensor operating point (ε_n, δ_n) on the ROC curve, which simulates the spectrum sensing errors. Based on the sensing outcome, it decides whether or not to access the channel. If it decides to access the channel, it will send RTS (request to send) message to the intended CR receiver. When the receiver receives the RTS message it replies with CTS (clear to send) message provided that the channel is also idle at the receiver side. When the transmitter and receiver are tuned to the same channel after the initial handshake they will hop to the same channel at the beginning of every time slot in order to ensure transceiver synchronization without the need of a

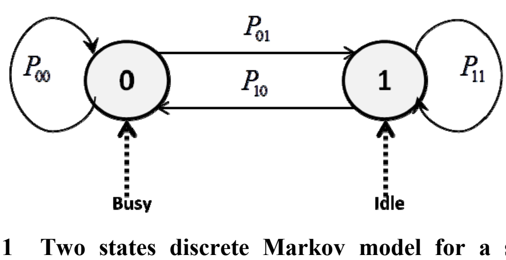

Fig. 1 Two states discrete Markov model for a single channel.

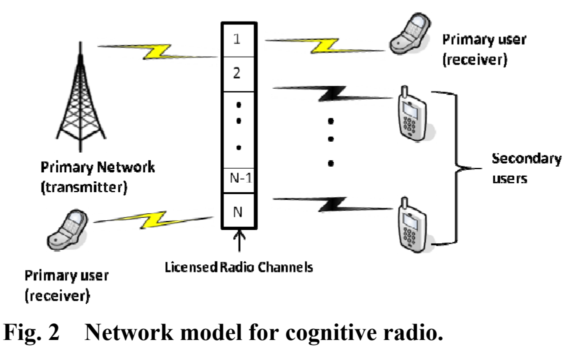

Fig. 2 Network model for cognitive radio.

dedicated channel for synchronization purpose [18]. At the end of the time slot, the CR receiver will acknowledge the successful transmission of the packets. Assumption is made that the packets are discarded if the channel is occupied (busy) by the PU [19], and that a channel only presents opportunity to a pair of CR nodes if and only if it is available to both CR transceivers which necessitate the joint identification of the channel opportunities between the two nodes [6]. The two transceivers can only start communicating over this identified idle channel after the successful exchange of the messages. The simplified sequence of activity in one time slot of N channels and a total of T time slots for the CR transceivers are summarized in Fig. 3 while Fig. 4 shows a simplified flow chart of the protocol[1]. The descriptions of symbols from Fig. 3 are shown in Table 1.

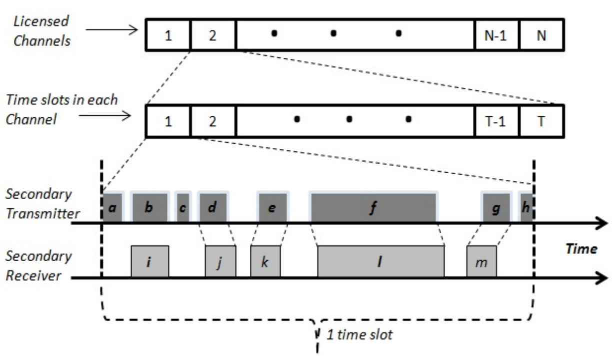

Fig. 3 A sequence of operations in one time slot for proposed MAC scheme.

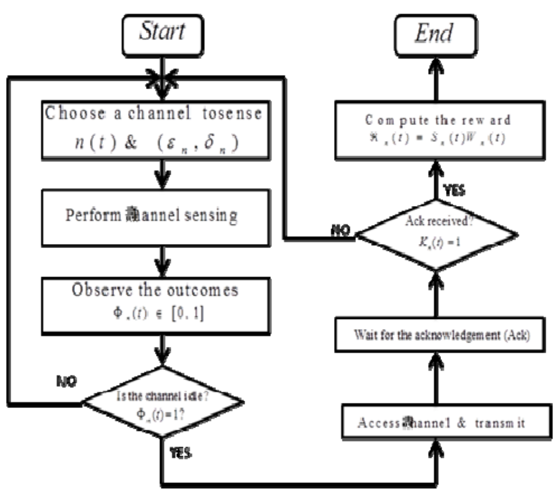

Fig. 4 A simplified one time slot transmission protocol.

[1]RTS and CTS messages are omitted in order to simplify the presentation.

4. POMDP Formulation

Let us denote the instantaneous state of the system by s, whereby the finite set of all states denoted by $S = \{s_1, s_2, \ldots\}$ as well as the state of the n-th channel at time slot t as $s_n(t)$. It should be noted here that under POMDP framework, the state of the system is not directly observable and therefore the CR node can only compute a belief over the state space. In order to infer a belief regarding the state of the system, the CR node takes sensor measurements (spectrum sensing results). A set of all sensing results (observations) is denoted by $\Theta = \{\Phi_1, \Phi_2, \ldots\}$. The observation of n-th channel at time t is denoted by $\Theta_n(t)$. Thus the observation $\Theta_n(t)$ is usually an incomplete projection of the system state $s_n(t)$ due to spectrum sensing errors. The observation is very much depending on the sensor capability and channel estimation technology used [13].

In order to define a POMDP framework precisely, probabilistic laws that describe the state transitions and observations have to be specified. These include the initial belief probability distribution (b_0), which is the probability that the system is in state s at time $t = 0$. This distribution is defined over all states in S.

Table 1 The descriptions of symbols in Fig. 3.

Symbol	Description
a	Transmitter (TX) chooses a channel to sense
b	TX performs sensing action
c	TX observes the sensing results
d	TX sends RTS signals
e	TX waits for CTS signal
f	TX transmits
g	TX receives acknowledgement
h	TX compute rewards
i	Receiver (RX) performs sensing
j	RX receives RTS signal
k	RX sends CTS signal
l	RX receives traffic
m	RX sends acknowledgement

Mathematically it can be represented as

$$b_o := \Pr(s_o = s) \quad (2)$$

Another quantity is the state transition probability $P(i, a, j)$, which is defined as the probability of transition from state i to state j given that the CR node was initially in state i and chooses the action a for any (i, a, j). This quantity can be expressed mathematically as

$$P(i,a,j) = \Pr(s(t) = j \mid S(t-1) = i,\ a(t-1) = a) \quad (3)$$

Since P(•) is a conditional probability distribution, it follows that

$$\sum_{j \in S} P(i,j) = 1 \quad \forall (i,a) \quad (4)$$

which also suggests that $P(i, j)$ is time-invariant, and thus the stochastic matrix $P(i, j)$ does not change over time. We also need to specify the observation probability distribution, $\Theta(s, a, \Phi)$ which is defined as the probability that CR node will perceive observation Φ upon executing the action a in state s. This can be expressed mathematically as

$$\Theta(s,a,\Phi) := \Pr\left(\Phi(t) = \Phi \mid \Phi(t-1) = \Phi, a(t-1) = a\right) \quad (5)$$

The conditional probability is defined over all (s, a, Φ) triplets of which

$$\sum_{\Phi \in \Theta} \Theta\left(s,a,\Phi\right) = 1 \quad \forall \left(s,a\right) \quad (6)$$

5. Spectrum Sensing Error Model

In the literature, some commonly spectrum sensing techniques which briefly discuss several spectrum sensing algorithms in the context of cognitive radio [20-26] are presented. However, in this paper we do not intend to develop any spectrum sensing scheme, instead we make use of the ROC for the energy detector that was developed in our previous work. The ROC curve presents the measure of false alarm P_F and the probability of miss detection P_{MD} which are the performance measure for the spectrum sensor. Thus, we assume that the CR nodes are equipped with a Ney-man Pearson ED (energy detector) which is a preferred approach for spectrum sensing in CR due to its simplicity and applicability as well as its low computational and implementation costs. We adopt our previous work [27] where the spectrum sensor of the CR node performs a binary hypothesis tests. We have H_0 indicates that the channel is idle and only AWGN (additive white Gaussian noise) is present, while H_1 indicates that the channel is busy. We can express this binary hypothesis in terms of channel sensing observations as

$$\begin{aligned} H_0 &: \Phi_n(t) = 1 \\ H_1 &: \Phi_n(t) = 0 \end{aligned} \quad (7)$$

Thus, the results of the binary hypothesis test $\Phi_n \in \{0, 1\}$ means that "If the sensor mistakes H_0 for H_1 in a channel, a false alarm occurs which lead to spectrum opportunity overlooked by the sensor." This generally means that the sensing outcome indicates that the channel is busy while in reality the channel is idle. On the other hand, if the sensor mistakes H_1 for H_0, we have a miss detection which leads to collision with the PU. We denote the probability of false alarm and the probability of miss detection respectively, by ε and δ which can be defined mathematically as

$$\begin{aligned} \varepsilon_n(t) &\triangleq \Pr\left\{\Phi_n(t) = 0 \mid \Phi_n(t) = 1\right\} \\ \delta_n(t) &\triangleq \Pr\left\{\Phi_n(t) = 1 \mid \Phi_n(t) = 0\right\} \end{aligned} \quad (8)$$

Usually, the performance of the spectrum sensor is characterized by ε and δ on the ROC curve. Fig. 5 shows the ROC curve achieved by the ED which is the theoretical and simulation results based on Eqs. (9) and (10):

$$\varepsilon = \frac{1}{2} erfc\left(\frac{\left(\frac{r - N\sigma^2}{\sqrt{2N\sigma_w^4}}\right)}{\sqrt{2}}\right) \quad (9)$$

$$\delta = 1 - Q\left[\frac{Q^{-1}(P_F) - SNR\sqrt{\frac{N}{2}}}{1 + SNR}\right] \quad (10)$$

These results are incorporate into the proposed MAC scheme focusing on the trade-off between ε and δ. The idea is to find in which point on the ROC curve should spectrum sensor operate for the optimal performance of the proposed MAC scheme. Fig. 6 shows how the valid sensor operating point $(\varepsilon_n, \delta_n)$ lies below

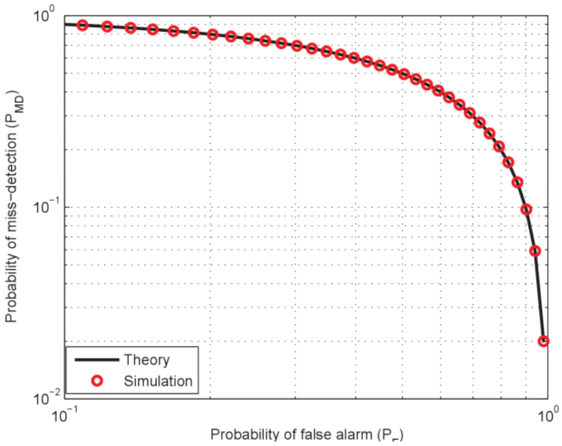

Fig. 5 ROC (receiver operating characteristic) curves for energy detector.

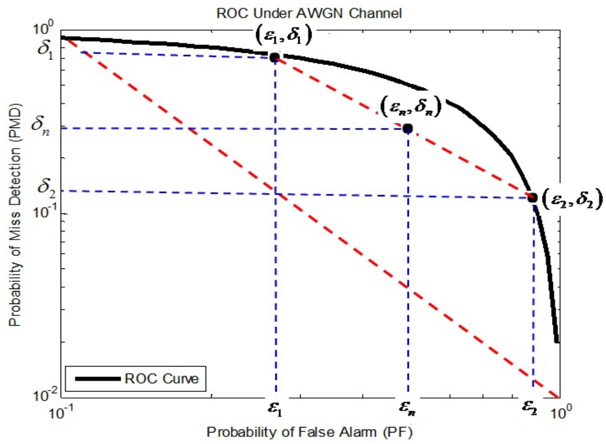

Fig. 6 Receiver ROC curve for ED showing the trade-offs between ε and δ.

the ROC curve, and thus the operating points can be achieved by randomizing between the two operating Ney-man Pearson detectors with properly chosen conditions (constraints) on the probability of false alarm (ε) [16, 17]. Thus the sensor operating point (ε_n, δ_n) in Fig. 6 can be achieved by applying the optimal Ney-man Pearson detector under the condition that $P_F \leq \varepsilon_n$, with probability p, such that $p = \frac{\varepsilon_n - \varepsilon_2}{\varepsilon_1 - \varepsilon_2}$ and Ney-man Pearson detector under the constraint that $P_F \leq \varepsilon_n$ with $1 - p$ which makes the design of spectrum sensor to reduce to the choice of the desired sensor operating point (ε_n, δ_n) on the ROC curve. The main objective is to find the optimal sensor operating point $\left(\varepsilon_n^*, \delta_n^*\right)$ on the ROC curve in order to achieve the best trade-off between false alarm and miss-detection.

6. The Reward Computation

The objective of the POMDP framework is to optimize action selection, so that the CR node is given a reward function describing its performance. The reward function assigns a numerical value quantifying the utility of performing action a while in state s. The goal of CR node is to maximize the sum of its reward over time. The reward gained by CR node can be defined in many ways depending on the design objective such as minimizing the BER (bit error rate), minimizing the time delay, maximizing the data rate, and so on. However, in this scheme we define the reward as the amount of information transmitted by the CR node under the constraint that the collision with primary user is avoided. We adopt the assumptions made by Refs. [2, 5, 6, 28] whereby the number of bits delivered by CR node over a channel is assumed to be proportional to the channel bandwidth. Let $\Re_n^k(t)$ denote the reward gained by a CR node after transmitting in channel n having the bandwidth of W which is in state S in time slot t and receiving acknowledgement k from the CR receiving node, then the reward can be expressed as

$$\Re_n^k(t) = S_n^t(t) W_n(t) \tag{11}$$

7. Belief Computation

Similar to our problem formulation, the system cannot be fully observable because of partial sensing of the radio spectrum (due to hardware limitation) as well as the spectrum sensing errors. The CR node can instead maintain a complete trace of all observations and all actions it ever executed, and use this to select its actions. The action/observation trace is known as history [10], and this can be very long as time goes on. A well-known fact is that this history does not need to be represented explicitly, but can instead be summarized via a belief distribution, which is the posterior distribution over the states [9]. Smallwood et al. in Ref. [29] suggested that the belief vector which is defined as the probability distribution over the state space can summarize their knowledge of all past actions and observations. Let us denote the belief vector space as $B_i(t)$. This can be defined as the probability that the current state in time slot t is i, and that at the end of each time slot, the belief vector is updated using Bayes' rule. Thus,

$$B_i(t) \triangleq \{b_0(t), b_1, ..., b_S(t)\} \tag{12}$$

and

$$B_i(t+1) = \frac{\sum_i B_i(t)P(i,j)\Theta^a_{j,\Phi}}{\sum_{i,j} B_i(t)P(i,j)\Theta^a_{j,\Phi}} \tag{13}$$

Given belief vector $B_i(t)$, the distribution of the system state in the time slot t after the state transition is then given by

$$\Pr\{S(t)=i\} = \sum_{j\in S} B_i(t)P(i,j) \quad \forall j \in S \tag{14}$$

This belief vector $B_i(t)$ is a sufficient statistic for the history as it suffices to condition for selection of the action on $B_i(t)$, instead of ever-growing sequence of past observations and actions. The belief $B_i(t)$ is calculated recursively, thus using only the belief one time step earlier $B_i(t-1)$ along with the most recent action $a(t-1)$ and the observation $\Phi(t)$. After taking the action a which is based on the spectrum sensing and the observation from the acknowledgement, the belief vector is updated by

$$B_i(t+1) = \frac{\sum_i B_i(t)P(j,i)\Theta_{i,k}a_{\{s,\delta,c\}}}{\sum_i\sum_j B_i(t)P(j,i)\Theta_{i,k}a_{\{s,\delta,c\}}} \tag{15}$$

Eq. (15) specifies the probability of observing $K_n = k$ when the belief vector is given by $B_i(t)$ and the action $a_{\{s,\delta,c\}}$ is taken as $p\{k|a_{\{s,\delta,c\}}, B_i(t)\}$ which is obtained by averaging the conditional observation probability over the current spectrum occupancy state [30]. Thus, Eq. (15) becomes

$$\begin{aligned} &p\left\{k \mid a_{\{s,\delta,c\}}, B_i(t)\right\} \\ &= \sum_i\sum_j B_i(t)P(j,i)\Theta_{i,k}a_{\{s,\delta,c\}} \end{aligned} \tag{16}$$

8. Joint Optimal Policy Computation

The key objective of POMDP perspective is to compute a joint policy for choosing actions such that the expected cumulative reward is maximized. In this scheme, we need three different policies namely: optimal spectrum sensing policy π_s^*, optimal sensor operating point policy π_δ^* and optimal transmission policy π_c^*. Let $(\pi_s^*, \pi_\delta^*, \pi_c^*)$ be the joint optimal policy strategy to maximize the total number of information bits that can be delivered by the CR node in the finite period of time. Mathematically this policy can be expressed as

$$\begin{aligned} \{\pi_s^*, \pi_\delta^*, \pi_c^*\} &= \arg\max_{\pi_\delta,\pi_s,\pi_c} \sum_{t=1}^{T} \mathrm{E}_{\{\pi_\delta,\pi_s,\pi_c\}} \\ &\quad \times\left[\Re_{K_n}^{(a,\Phi_n)}(t) \mid B(1)\right] \\ &= \arg\max_{\pi_\delta,\pi_s,\pi_c} \sum_{t=1}^{T} \mathrm{E}_{\{\pi_\delta,\pi_s,\pi_c\}} \\ &\quad \times\left[S_{K_n}^{(\Phi_n)}(t)W_n(t) \mid B(1)\right] \\ S.t \quad & \Pr\{\Phi_n = 1 \mid S_n = 0\} \le \tau \end{aligned} \tag{17}$$

where $\mathrm{E}_{\{\pi_\delta,\pi_s,\pi_c\}}$ is the mathematical expectation for given policies $\{\pi_s, \pi_\delta, \pi_c\}$ and $B(1)$ is the initial belief state and τ is the maximum allowable probability of miss detection. The optimal policy strategy which is the design objective in Eq. (17) is a constrained POMDP which usually requires randomized policies to achieve optimality. Chen et al. [31] established a separation principle for the optimal joint design. This separation principle reveals the existence of deterministic optimal sensing and access policies, leading to significant complexity reduction. The proof of this separation principle can be found in Ref. [30]. With a separation principle theorem, the optimal access policy is time-invariant and belief independent. Thus the optimal transmission probabilities are solely determined by the sensor operating point δ and the maximum allowed probability of miss detection τ. For any chosen action a, the information state (belief) $B(1)$ and time slot t, the transmission probability can be evaluated as

$$\left(f_a^1(B(t),t), f_a^0(B(t),t)\right) = \begin{cases} \left(1, \dfrac{\tau-\delta}{1-\delta}\right), & for\ \delta < \tau \\ (1,0), & for\ \delta = \tau \\ \left(\dfrac{\tau}{\delta}, 0\right), & for\ \delta > \tau \end{cases} \tag{18}$$

where f_a^1 is the transmission probability after taking action a when the channel is detected idle and f_a^0 is the transmission probability after taking action a when the channel is detected busy. Given the information state $B(1)$ at the beginning of time slot t, the constraint $\Pr\{\Phi_n = 1|S_n = 0\}\leq \tau$ can be rewritten as

$$\begin{aligned}&\Pr\{\Phi_n = 1 \mid S_n = 0\}\\ &= \delta f_a^1(B(t),t) + (1-\delta) f_a^0(B(t),t)\end{aligned} \tag{19}$$

Algorithm 1 in the Appendix section shows how the joint optimal policy is computed iteratively.

9. Value Function Iteration

Computing the optimal policy is very challenging mainly due to two reasons namely the curse dimensionality and the curse history [9, 10, 32]. These two problems are related in such a way that the higher the dimension of a belief space, the more room it has for a distinct histories. But they often act independently. The complexity can grow exponentially with time horizon even in problem with only few states, and problems with a large number of physical states may still only have a small number of relevant histories [10]. Sondik et al. [33] proposed value-iteration approach as the most straight forward approach to find optimal policies where iterations of dynamic programming are applied to compute increasingly more accurate values for each belief state. We need to evaluate a value function (maximum expected remaining reward) that can be accumulated starting from slot 1 to slot T, $(1 \leq t \leq T)$ for a given initial belief vector $B_i(t)$. Let a value function be denoted by $V_t(B_i(t))$. Similar to approach in Ref. [6], this value function has two parts namely, immediate reward obtained in slot t and the maximum expected remaining reward which is obtained starting from slot $t + 1$ given the belief vector $B_i(t + 1)$. Given the optimal sensor operating point policy δ^* and the optimal access policy c^*, the value function can be obtained recursively by

$$\begin{aligned}V_t(B_i(t)) &= \max_{a_{\{s,\delta,c\}}\in A} \sum_i \sum_j B_j(t) P_{i,j}\\ &\times \sum_{K_a=0}^{1} (\Pr\{K_a = 1 \mid S(t) = i\})\\ &\times [K_a W_a + V_{t+1}(B_i(t) \mid a, K_a)]\end{aligned} \tag{20}$$

For $1 \leq t \leq T$,

$$V_T(B_i(t)) = \max_{a_{\{s,\delta,c\}}\in A} \sum_i \sum_j B_j(t) Q_i(1) W_a \tag{21}$$

where $Q_i(1) \triangleq \Pr\{K_a = 1 \mid S(t) = i\}$ which is the probability of successful transmission under the current spectrum state. Likewise, belief vector can be updated using Baye's rule as

$$B_i(t+1) = \frac{\sum_j B_i(t) P_{j,i} Q_i(K_a)}{\sum_i \sum_j B_j(t) P_{j,j+1} Q_{j+1}(K_a)} \tag{22}$$

The value function in Eq. (20) need to be evaluated recursively in order to obtain the maximum total reward accumulated in T slots. Smallwood et al. [29] proved that, the value function $V_t(B_i(t))$ is piecewise linear and convex. Therefore, the joint optimal policy is evaluated by linear programming technique which is an iterative solution. Algorithm 2 in the Appendix section shows how the value function is evaluated iteratively.

10. Simulation Results and Discussions

In order to study the performance of the proposed scheme, we considered multiple channels and multiple CR nodes co-exist with primary network users. Random message arrivals were considered with the aim to investigate the performance of greed sensing approach in multiple CR users in presence of spectrum sensing errors.

We considered four simulation scenarios which enabled us to study the throughput performance using simulations. The basic simulation parameters are shown in Table 2. In the first scenario we consider 10 independent channels (N = 10) with equal bandwidth (W = 2) and equal transition probability (P_{01}, P_{11}) = (0.2, 0.8) for each channel. It was assumed that the message arrives at the CR node form a Poisson process with arrival rate of λ. It was further assumed that the message length is geometrically distributed with average message length of 50 packets. The transmission time of one packet was assumed to be one time slot.

Upon the arrival of the message, the whole message will be randomly assigned to a secondary user. In each slot, those CR nodes that do not have packets to transmit will turn to sleep in order to conserve energy.

Table 2 Simulation setup for a proposed MAC scheme.

Parameter	Assigned value
Number of independent channels	10
Bandwidth per channel	2
Transition probability for each channel $P(i, j)$	(0.2, 0.8) equal for each node
Error model (ε, δ)	ROC curve (10 cooperative sensing users)
Message arrival rate (λ)	Form a Poisson process
Message length	50 packets geometrically distributed
Packet transmission time (per packet)	Assumed to be one time slot

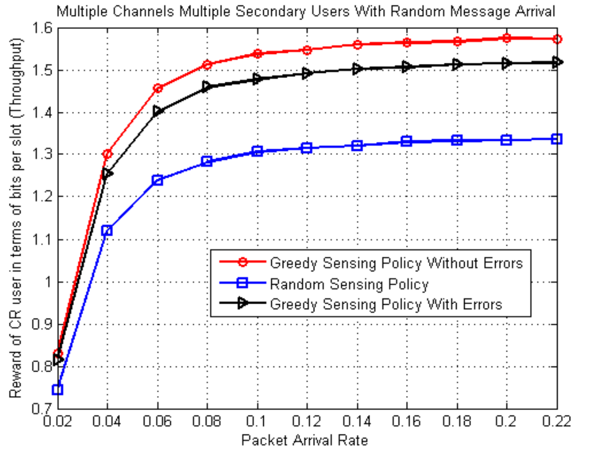

Fig. 7 Performance of multiple SUs multiple channels scheme in presence of sensing errors.

They do not participate in channel selection and sensing, and their belief states are updated according to the Markovian model of spectrum occupancy. On the other hand, those CR nodes with data to transmit will choose a channel according to greedy approach, and then update their belief states according to the sensing outcomes. When an available channel is chosen by multiple users, the assumption was made that one of these users will succeed. For the purpose of comparisons of the schemes, we implemented the scheme using greedy sensing algorithm as was previously presented by Zhao et al. [6]. Fig. 7 shows the overall throughput performance as a function of message arrival rate (λ) for multiple SUs accessing multiple channels in presence of spectrum sensing errors. We implemented the greed sensing algorithm for 10 independent channels with equal bandwidth and equal transition probabilities. It can be seen from Fig. 7 that the proposed greedy sensing scheme with sensing errors outperforms that of random sensing scheme in terms of throughput. In the second scenario we aimed at studying the throughput performance of the proposed multi channel multi user scheme as a function of prescribed collision probability (τ) (which indicates the level of sensing errors). During simulation, we considered 10 independent channels (N = 10), with equal bandwidth (W = 1) and equal transition probabilities (P_{01}, P_{11}) = (0.5, 0.5). It can be seen from Fig. 8 that the average throughput of the proposed scheme (with sensing errors) is outperformed by that without errors. This is due to the fact that by introducing errors for the spectrum sensor in the system, the overall throughput of the system is affected. However this scheme reflects the practical scenario since it is almost impossible to have an error free spectrum sensor in wireless communication. Some more research is needed to improve the throughput performance under dynamic wireless environment where spectrum sensing errors are unavoidable.

In the third scenario, we aimed at studying the performance of the proposed MAC scheme by considering 3 independent channels (N = 3) with equal bandwidth (W = 1) and equal transition probability (P_{01}, P_{11}) = (0.2, 0.8). It can be seen in Fig. 9 that under such condition, the performance of a proposed scheme under greedy sensing algorithm approaches that of optimal sensing and greedy sensing that was proposed by Zhao et al. [6].

In the fourth scenario we assumed the same number of channels (N = 3), but different bandwidth and different transition probabilities. We used the same values of bandwidth and transition probabilities as used in Ref. [6] for comparison purposes. Thus P_{01} = [0.8, 0.6, 0.4], P_{11} = [0.6, 0.4, 0.2], and W = [0.75, 1, 1.5]. It can be seen from Fig. 10 that the average throughput of the proposed scheme outperforms the greedy and the random sensing schemes.

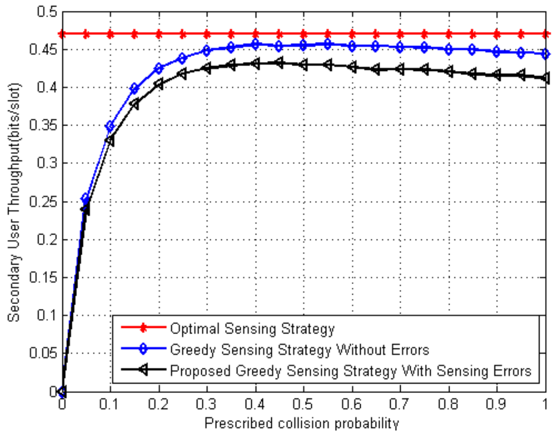

Fig. 8 Performance of the proposed scheme in presence of sensing errors for equal bandwidth and equal transition probabilities.

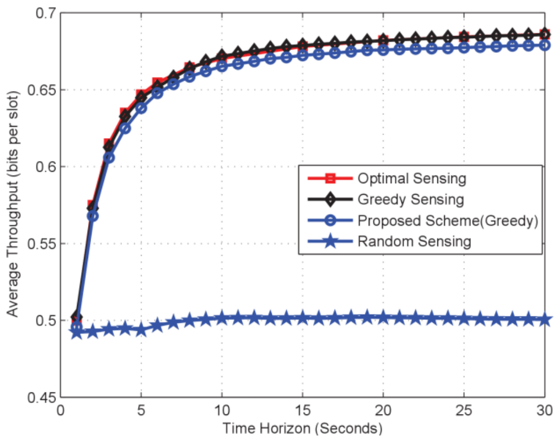

Fig. 9 Performance of the proposed scheme under greedy sensing algorithm: 3 independent channels with equal bandwidth and equal transition probabilities.

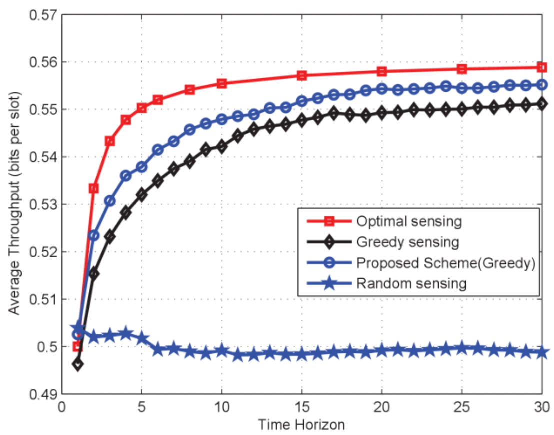

Fig. 10 Performance of the proposed scheme using greedy sensing algorithm: 3 independent channels with equal bandwidth and equal transition probability.

10. Conclusions

Spectrum sensing errors in CR are inevitable. In the event of a false alarm, a spectrum opportunity is overlooked by the sensor, and eventually wasted if the access strategy trusts the sensing outcome. On the other hand, miss-detections may lead to collisions with primary users. The trade-off between false alarm and miss-detection is captured by the receiver operating characteristic of the spectrum sensor, which relates the probability of detection and the probability of false alarm. The design of the spectrum sensor and the choice of the sensor operating point are thus important issues and should be addressed by considering the impact of sensing errors on the MAC layer performance in terms of throughput and collision probability. A decision-theoretic framework for the optimal joint design MAC scheme based on the theory of partially observable Markov decision processes can be established to capture the fundamental design trade-offs between the probability of detection and the probability of false alarm for the design of a CR network. The POMDP framework is very useful for modeling problems in which the state of the system cannot be directly observed due noise and fading in a wireless channel. Within this framework, the optimal MAC strategy is given by the optimal policy of a constrained POMDP. However, while powerful in problem modeling for decision making under uncertainty, POMDP suffers from the curse of dimensionality and does not easily lend itself to tractable solutions. Constraints on a POMDP further complicate the problem, often demanding randomized policies to achieve optimality. There is no known a closed form solution for a constrained POMDP framework. However, due to the fact that the belief vector is piecewise linear and convex, linear programming can be used to compute the optimal policy iteratively. The proposed POMDP framework has proved to be a convenient and realistic way for modeling a decision process under uncertainty.

Acknowledgments

The authors would like to thank the Centre of Excellence (CoE) in the School of engineering at the University of KwaZulu-Natal (UKZN) and the Dar Es salaam Institute of Technology (DIT), for their financial support.

References

[1] H. Kim, K.G. Shin, Efficient discovery of spectrum opportunities with MAC-layer sensing in cognitive radio networks, IEEE Transactions on Mobile Computing 7 (5) (2008) 533-545.

[2] Q. Zhao, L. Tong, A. Swami, Decentralized cognitive MAC for dynamic spectrum access, in: Proceedings of 1st IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN 2005), 2005, pp. 224-232.

[3] J. Jia, Q. Zhang, X. Shen, HC-MAC: A hardware constrained cognitive MAC for efficient spectrum management, Selected Areas in Communications 26 (1) (2008) 106-117.

[4] A. Motamedi, A. Bahai, MAC protocol design for spectrum-agile wireless networks: Stochastic control approach, in: Proceedings of 2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN 2007), 2007, pp. 448-451.

[5] Y. Chen, Q. Zhao, A. Swami, Distributed cognitive MAC for energy-constrained opportunistic spectrum access, in: IEEE Military Communications Conference (MILCOM 2006), Washington, DC, 2006.

[6] Q. Zhao, L. Tong, A. Swami, Y. Chen, Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework, Selected Areas in Communications 25 (3) (2007) 589-600.

[7] Q. Zhao, S. Geirhofer, L. Tong, B.M. Sadler, Optimal dynamic spectrum access via periodic channel sensing, in: Proceedings of the 2007 Wireless Communications and Networking Conference, 2007, pp. 33-37.

[8] J. Unnikrishnan, V.V. Veeravalli, Algorithms for dynamic spectrum access with learning for cognitive radio, IEEE Transactions on Signal Processing 58 (2) (2010) 750-760.

[9] J. Pineau, Tractable planning under uncertainty: Exploiting structure, Ph.D. Dissertation, Rutgers University, 2004.

[10] D. Braziunas, Pomdp Solution Methods, Technical Report, Department of Computer Science, University of Toronto, 2003.

[11] A. Konrad, B.Y. Zhao, A.D. Joseph, R. Ludwig, A Markov-based channel model algorithm for wireless networks, Wireless Networks 9 (3) (2003) 189-199.

[12] M. Hassan, M.M. Krunz, I. Matta, Markov based channel characterization for tractable performance analysis in wireless packet networks, IEEE Transactions on Wireless Communications 3 (3) (2004) 821-831.

[13] Q. Zhang, S.A. Kassam, Finite-state Markov model for Rayleigh fading channels, IEEE Transactions on Communications 47 (11) (1999) 1688-1692.

[14] Q. Zhao, B. Krishnamachari, Structure and optimality of myopic sensing for opportunistic spectrum access, in: Proceedings of the IEEE International Conference on Communications (ICC'07), 2007, pp. 6476-6481.

[15] H.V. Poor, An Introduction to Signal Detection and Estimation, Springer-Verlag, New York, 1988, p. 559.

[16] H.V. van Trees, Detection, Estimation, and Modulation Theory, Part I, John Wiley and Sons, 1968.

[17] H.L. van Trees, Detection, Estimation, and Modulation Theory, Part IV: Optimum Array Processing, John Wiley & Sons, 2004.

[18] Y. Chen, Q. Zhao, A. Swami, Joint PHY-MAC design for opporunistic spectrum access with multichannel sensing, in: IEEE 8th Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2007), pp. 1-5.

[19] F. R. Yu, B. Sun, V. Krishnamurthy, S. Ali, Application layer QoS optimization for multimedia transmission over cognitive radio networks, Wireless Networks 17 (2) (2011) 371-383.

[20] S. Haykin, D.J. Thomson, J.H. Reed, Spectrum sensing for cognitive radio, Proceedings of the IEEE 97 (5) (2009) 849-877.

[21] W. Wang, Spectrum sensing for cognitive radio, in: Proceedings of the 3rd International Symposium on Intelligent Information Technology Application Workshops (IITAW'09), 2009, pp. 410-412.

[22] Z. Quan, S. Cui, A.H. Sayed, Optimal linear cooperation for spectrum sensing in cognitive radio networks, Selected Topics in Signal Processing 2 (1) (2008) 28-40.

[23] G. Ganesan, Y. Li, Cooperative spectrum sensing in cognitive radio: Part I. Two user networks, IEEE Transactions on Wireless Communications 6 (6) (2007) 2204-2213.

[24] A. Ghasemi, E.S. Sousa, Spectrum sensing in cognitive radio networks: Requirements, challenges and design trade-offs, IEEE Communications Magazine 46 (4) (2008) 32-39.

[25] T. Yucek, H. Arslan, A survey of spectrum sensing algorithms for cognitive radio applications, IEEE Communications Surveys & Tutorials 11 (1) (2009) 116-130.

[26] C. Sun, W. Zhang, K. Ben, Cluster-based cooperative spectrum sensing in cognitive radio systems, in:

Proceedings of IEEE International Conference on Communications (ICC'07), 2007, pp. 2511-2515.

[27] J. Msumba, H. Xu, Spectrum sensing for cognitive radio networks: The need for cross-layer design approach for multimedia applications, in: AFRICON 2011, Livingstone, 2011.

[28] Y. Chen, Q. Zhao, A. Swami, Distributed spectrum sensing and access in cognitive radio networks with energy constraint, IEEE Transactions on Signal Processing 57 (2) (2009) 783-797.

[29] R.D. Smallwood, E.J. Sondik, The optimal control of partially observable Markov processes over a finite horizon, Operations Research 21 (5) (1973) 1071-1088.

[30] Y. Chen, Q. Zhao, A. Swami, Proof of the Separation Principle for Opportunistic Spectrum Access, Tech. Rep., University of California, Davis, 2006.

[31] Y. Chen, Z. Qing, S. Ananthram, Joint design and separation principle for opportunistic spectrum access, in: Proceedings of IEEE 40th Asilomar Conference on Signals, Systems and Computers (ACSSC'06), 2006, pp. 696-700.

[32] G.E. Monahan, State of the art—A survey of partially observable Markov decision processes: Theory, models, and algorithms, Management Science 28 (1) (1982) 1-16.

[33] E.J. Sondik, The Optimal Control of Partially Observable Markov Processes, DTIC Document, Tech. Rep., 1971.

Appendix

Algorithm 1 Optimal policy computation by iteration.

Input: Policy π to be computed
Output: Optimal policy (π^*) for action taken

$\mathcal{B}_i(1) = 0, \forall_i \in \mathbb{S}$ $\{initialize\}$
$\pi^* \leftarrow \{a, \delta, c\}$
Repeat
$\hat{b} = 0$
for *all* $i \in \mathbb{S}$ **do**
$b = \mathcal{B}_i(1)$
$\mathcal{B}_i(1) = \sum_a \pi(s,a) \sum_j P(j|i,a) \left(\mathcal{R}(j|i,a) + \tau \mathcal{B}_i(j)\right)$
if $a_i > max$ **then**
  $max \leftarrow a_i$
return $Optimal\ policy\ (\pi^*)$

Algorithm 2 Value iteration for optimal policy computation.

$Assumption:\ policy\ \pi\ is\ proper$
$Initialize\ \mathcal{V}_0^\pi\ arbitrally\ for\ each\ state$
$t \leftarrow \emptyset$
repeat
$t \leftarrow t + 1$
for $i \in \mathbb{S}$ **do**
  $Compute\ \ \mathcal{V}_t^\pi \leftarrow \sum_{j \in \mathbb{S}} P\{i, \pi(i), j\} \left[C(i, \pi(i), j) + \mathcal{V}_{t-1}^\pi(j)\right]$
  $Compute\ \ resdual_t(i) \leftarrow |\mathcal{V}_t^\pi(i) - \mathcal{V}_{t-1}^\pi(i)|$
end
until $max_{i \in \mathbb{S}} residual_t(i) < Threshold$
return $\mathcal{V}_t^\pi$