Journal of Computations & Modelling, vol.3, no.3, 2013, 177-193
ISSN: 1792-7625 (print), 1792-8850 (online)
Scienpress Ltd, 2013

User Scheduling in Cognitive Radio Networks

Jalaj Bhandari[1] **and Nomesh Bolia**[2]

Abstract

We consider spectrum management in Cognitive Radio (CR) networks. We model primary users (PUs) activity on a channel and consider finite number of secondary users (SUs). We study the trade-off between two conflicting objectives of minimizing overall data queue lengths for the system and maintaining the quality of service to users. We use Markov Decision Processes (MDPs) to model this and derive an easy to implement scheduling policy. Using simulations, we demonstrate and compare the performance of our policy.

Mathematics Subject Classification: 90C40
Keywords: Cognitive Radio Networks; User Scheduling; Resource Allocation; Markov Decision Processes

1 Introduction

The popularity and use of wireless networks has increased manifold in the last decade, leading to the emergence of spectrum shortage as one of the major

[1] Indian Institute of Technology, Delhi.
[2] Indian Institute of Technology, Delhi.

Article Info: *Received* : June 2, 2013. *Revised* : July 28, 2013
Published online : September 15, 2013

challenges in the field of wireless communications. Recent studies in this area attribute this shortage to the fixed policy of spectrum allocation (where fixed portions of bandwidth are leased out for long term usage) rather than to the heavy utilization of existing spectrum [1]. In fact, many surveys done reveal a heavy underutilization of allocated spectrum [1], [2]. Spectrum scarcity in wireless networks has had an adverse impact especially on multimedia transmission which is severely bound by delay constraints imposed by the users [3], [4].This has led to the introduction of Cognitive radio (CR) networks [5], [6] which allows unlicensed users (also called secondary users) to utilize available spectrum (unoccupied frequency bands also known as white space or spectrum holes [7]) allocated to licensed users (primary users (PU)) provided no significant interference is generated. This is seen as one of the potential solutions as it allows exploitation of the idle periods between consecutive packet transmission of the primary users especially since many of the wireless communications systems show bursty transmission behaviour [8].

Numerous aspects of cognitive radio networks and its technology have been well researched and documented. In cognitive radio's technological structure, secondary users are required to "periodically sense" and identify the frequency bands that are not occupied by the primary users using a "spectrum agile radio transciever" [9]. This is done by sensing the RF environment, a process known as "spectrum sensing" [10]. The area of spectrum sensing has received significant attention from researchers lately [11] [12]. Efficient sensing methods [13] (including the latest in cooperative sensing [14]), sensing accuracy, interference between primary and secondary users, system throughput and its relationship with sensing frequency, proability of detection and probability of false alarms are some of the well documented research issues [15], [16], [17]. An important consequence of this spectrum sharing mechanism and periodic sensing structure which has caught the eye of researchers is the conflict between primary and secondary users [18]. This leads us to the problem of efficient resource allocation and user scheduling. A similar problem in context of cognitive radio networks has been studied in [19] where a decentralized learning algrorithm was formulated by modelling the system as multi-agent interactioncs.

In this paper, we aim to tackle the problem of user scheduling in a centralized system. We consider a single channel with constant data transfer rate

R, a fixed number of users m and a time slotted framework where exactly one user is served in the n^{th} time slot, $n \in \{0, 1, 2, \ldots, \infty\}$, as long as there is data to be served for at least one user. We assume that data is in the form of packets and data arrival (number of packets per time slot) for each user i is a $Poisson$ process given by:

$$Y_i \sim Poi(\alpha_i). \tag{1}$$

In this framework, we assume that data is queued up for every user in separate queues maintained by a centralized spectrum management agency (CSMA) and there is no bound on system capacity. Let Q_u^n be the number of packets in the queue for user u at the beginning of time slot n and Y_u^n be the number of packets that user u receives during the n^{th} time slot.We assume that in every time slot n, CSMA keeps a track of the queue length vector $Q^n = [Q_1^n, Q_2^n, \ldots, Q_m^n]$ and the data arrival vector $Y^n = [Y_1^n, Y_2^n, \ldots, Y_m^n]$.

In this setting minimizing the total system queue length is a useful objective for resource allocation as the total queue length is a matter of cost for the system. The data actually transferred in the n^{th} time slot if user u is allocated the channel for transmission is $S_u^n = \min[Q_u^n + Y_u^n, R]$. where R is the number of packets that can be transferred in a time slot. Intuitively, one appropriate policy would be to serve the user with the maximum queue length Q_u^n at the beginning of every time slot. We shall refer to this as the "myopic policy" throughout the course of this paper.

However, in doing this we risk being unfair to those users for which Q_u^n is not the highest. Thus some users could be starved badly in myopic policy. In the extreme case, a user could remain unserved forever. Therefore, the CSMA has two conflicting objectives: maintaining fairness and reducing the total system queue length. A useful measure for fairness is "starvation age"; which is the time interval between successive servings for a user. A lower starvation age (or simply age) implies a better quality of service and greater fairness. Intuitively, the most fair of all policies is simply the "round robin rule" where users are served in cyclic order. It can be easily shown that the average age for this policy is minimum ($m*(m-1)/2$ where m is the number of users).The objective of this work is to find a scheduling algorithm that resolves this trade-off between system queue length and fairness to users.

We use Markov Decision Processe (MDPs) to model this problem. However, we observe that it is infeasible to solve this problem optimally. We use a single step policy improvement algorithm (which is well documented in literature and has been previously used for scheduling problems) to come up with an Index Policy for our framework. Its noteworthy that the procedure described in this paper is more general in nature and can be applied outside the domain of cognitive radio networks as well using suitable assumptions and system models.The rest of the paper is organized as follows: we formulate the problem as an MDP in section 2, in section 3 we apply the one step policy improvement approach and derive our policy. We provide simulation results in section 4, conclude in section 5.

2 MDP Formulation

We formulate this scheduling problem as an MDP. We start with a model for data arrival process $\{Y^n, n \geq 0\}$, the queue length process $\{Q^n, n \geq 0\}$, channel availability $\{X^n, n \geq 0\}$ and age process $\{A^n, n \geq 0\}$.

2.1 Queue length process

We assume that the data arrival processes are independent for all the users i.e. $\{Y_k^n, n \geq 0\}$ is independent of $\{Y_l^n, n \geq 0\}$ for $k \neq l$. Recall that Q_u^n denotes the total data in the queue for user u, at the beginning of the n^{th} time slot. Let $v(n)$ denote the user served in the n^{th} time slot. Therefore $\{Q_u^n, n \geq 0\}$ changes according to the following:

$$Q_u^{n+1} = \begin{cases} Q_u^n + Y_u^n & \text{if } u \neq v(n) \\ \max[Q_u^n + Y_u^n - R, 0] & \text{if } u = v(n), \end{cases} \tag{2}$$

We define Q^n as the vector of queue lengths for all users, i.e. $Q^n = \{Q_1^n, Q_2^n, \ldots, Q_m^n\}$.

2.2 Channel Availability

We model the arrival of a PU on the channel as a Poisson process with rate λ and the sojourn time distribution as exponential with rate μ. Thus once a PU arrives in the channel, it stays there for an exponentially distributed time period.Both of these are reasonable, since we expect both inter arrival and sojourn times to be completely random (i.e. without any memory). Let X^n be the channel availability defined as: $X^n = 0$ if the PU doesn't occupy the channel (making it available for secondary user) and 1 otherwise. Then, $\{X^n; n \geq 0\}$ can be shown to be a Discrete Time Markov Chain. Let p_{ij} for $i, j \in 0, 1$ be the transition probability from i to j.Then if t denotes the length of the time slot,

$$p_{00} = 1 - \int_0^t \lambda * e^{-\lambda * s} ds; \;\; p_{01} = (1 - p_{00}) \tag{3}$$

$$p_{10} = \int_0^t \mu * e^{-\mu * s} ds; \;\; p_{11} = (1 - p_{10}).$$

Here p_{01} denotes the probability that a PU arrives in the channel in the current time slot even though the channel was available at the beginning of the slot. When this happens, we assume that no data is transmitted in the entire slot. Let π be the long run probability that the system is in state 0. Then using standard DTMC theory [20], it can be shown that

$$\pi = p_{01}/(p_{01} + p_{10}) \tag{4}$$

2.3 Starvation Age

Recall that starvation age is defined as the time interval between two slots in which the user was successfully able to transmit data. Let A_u^n denote the age of user u at the beginning of the n^{th} time slot. Let $v(n)$ be the user served in the nth time slot. the age process is given by:

$$A_u^{n+1} = \begin{cases} 0 & \text{if } u = v(n) \text{ (successful data transmission)} \\ A_u^n + 1 & \text{otherwise} \end{cases} \tag{5}$$

Let the vector of starvation age A^n for all users be given by:

$$A^n = \{A_1^n, A_2^n, \ldots, A_m^n\} \tag{6}$$

The system state at the beginning of n^{th} time slot is given by $[X^n, Q^n, A^n]$. Clearly, when $X^n = 1$, the channel is not available for transmission, there is no user allocation decision to be made. When $X^n = 0$ CSMA observes the $2N$ components vector $\{Q^n, A^n\}$ and decides which user to serve in slot n. We assume that there is a mechanism to sense whether the channel is available for transmission, and the time between consecutive sensing is our time slot. The problem of scheduling a user in a given time slot can be formulated as an MDP after imposing a cost structure. The decision epochs are $\{1, 2, \ldots, \infty\}$. The state at the beginning of n^{th} slot is given by $\{Q^n, A^n\}$ with Markovian evolution as described above. Recall that our objective is to minimize the queue length while ensuring a degree of fairness. To make the allocation decision optimally, we describe a cost structure below. We define the one step cost (the cost associated with serving user u in n^{th} time slot), $C_u^n(Q^n, A^n)$, as follows

$$C_u^n(Q^n, A^n) = \sum_{l \neq u} Q_l^n (1 + k_u * A_l^n) \tag{7}$$

where k_u is a constant which is a measure of the relative priority for the CSMA between the queue lengths and the starvation age. The first term of $C_u^n(Q^n, P^n)$ accounts for the objective of minimizing queue lengths while the second term accounts for fairness. The constant k_u gives us a handle to model the relative priority between fairness and queue length. For instance, $k_u = 0$ for all users implies "myopic policy" while increasing k_u for all users makes the system fairer. It is noteworthy that in our model we incur the one step cost only when the channel is available for transmission, when it is not, no allocation decision is made and hence no cost is incurred.

We next describe the value function for this MDP. Let $q^n = \{q_1^1, q_2^1, \ldots, q_m^1\}$ denote the queue length vector, $t^n = \{t_1^1, t_2^1, \ldots, t_m^1\}$ denote the starvation age vector and $y^n = \{y_1^1, y_2^1, \ldots, y_m^1\}$ denote the vector of data arrival for all users in slot n. Let $q^{u,n+1,k}$ and $t^{u,n+1,k}$ be the queue length and age vectors respectively in the $(n+1)th$ time slot given $v(n) = u$ and $X^{n+1} = k$. Let $V_T(0, q^n, t^n)$ denote the optimal reward starting from state $\{X^n, Q^n, A^n\} = \{0, q^n, t^n\}$ at time n over the next T time slots. If user u is served at time n and is able to successfully transmit data (no PU arrives in that time slot, i.e. $X^{n+1} = 0$), then $t^{u,n+1,0} = (t_1^{n+1}, t_2^{n+1}, \ldots, t_{u-1}^{n+1}, 0, t_{u+1}^n + 1, \ldots, t_m^{n+1})$. Using standard dynamic programming arguments [21], the Bellman equation is given by:

$$\begin{aligned} V_T(0, q^n, t^n) = \max_{u=1,2,\ldots,m} [-C_u^n(q^n, t^n) + \\ \sum_{k=0,1} p_{0k} V_{T-1}(k, q^{u,n+1,k}, t^{u,n+1,k}) \end{aligned} \tag{8}$$

$$V_T(1, q^n, t^n) = \sum_{k=0,1} p_{1k} V_{T-1}(k, q', t') \tag{9}$$

where $q^{u,n+1,k}$ and $t^{u,n+1,k}$ are given by:

$$q^{u,n+1,1} = q^n + y^n = q', t^{u,n+1,1} = t^n + 1 = t' \tag{10}$$

$$q^{u,n+1,0} = \begin{cases} q_l^n + y_l^n & \text{for all } l \neq u \\ \max[q_l^n + y_l^n - R, 0] & \text{for } l = u \end{cases} \tag{11}$$

$$t^{u,n+1,0} = \begin{cases} t_l^n + 1 & \text{for all } l \neq u \\ 0 & \text{for } l = u \end{cases} \tag{12}$$

Clearly, $k = 1$ denotes the situation where a PU arrives in the time slot which was empty at the beginning, and so user u allocated to this slot will not be able to transmit any data. Thus $q^{u,n+1,1}$ and $t^{u,n+1,1}$ in 11 and 12 are not dependent on u. Hence we can drop index u from this notation to yield:

$$q^{n+1,1} = Q^{n+1} = q', t^{n+1,1} = t^n + 1 = t' \tag{13}$$

Equation 8 can be expanded as:

$$\begin{aligned} V_T(0, q^n, t^n) = \max_{u=1,2,\ldots,m} [&-C_u^n(q^n, t^n) + \\ &p_{00} V_{T-1}(0, q^{u,n+1,0}, t^{u,n+1,0}) + \\ &p_{01} V_{T-1}(1, q^{n+1,1}, t^{n+1,1}) \end{aligned} \tag{14}$$

Since $q^{n+1,1}$ and $t^{n+1,1}$ are independent of the action u, $V_{T-1}(1, q^{n+1,1}, t^{n+1,1})$ is constant w.r.t. u.

Let $\gamma = p_{01} V_{T-1}(1, q^{n+1,1}, t^{n+1,1})$. Equation 8 can now be rewritten as:

$$\begin{aligned} V_T(0, q^n, t^n) = \gamma + \max_{u=1,2,\ldots,m} [&-C_u^n(q^n, t^n) + \\ &p_{00} V_{T-1}(0, q^{u,n+1,0}, t^{u,n+1,0}) \end{aligned} \tag{15}$$

Clearly, 9 describes the state when the channel is occupied by the primary user and is not available for transmission. In this case, no action for allocating any user needs to be taken. In the next section, we develop a method to determine action $u(0, q^n, t^n)$ that maximizes the long run average reward $\lim_T \to \infty$ $[V_T(q^n, t^n)/T]$ From equation 15, using standard MDP theory, it is well known that such a policy exists if there is a constant g and a bias function $w(0, q^n, t^n)$ satisfying:

$$\begin{aligned} g + w(0, q^n, t^n) = \max_u [&-C_u^n(q^n, t^n) + \\ &p_{00} w(0, q^{u,n+1,0}, t^{u,n+1,0}) \end{aligned} \tag{16}$$

We re-emphasize that the term corresponding to p_{01} is an additive constant and hence can be ignored while computing the action u that maximizes the long run average reward.

3 Index Policy

Since the MDP we developed cannot be solved optimally in higher dimensions, we formulate our scheduling algorithm based on a single step of the Policy Improvement Algorithm [21], [22]. Given the state of the system $(0, q^n, t^n)$ in all time slots, our policy will yield an index for each user $u \in \{1, 2, \ldots, N\}$ solely based on its state $(0, q_u^n, t_u^n)$. Our policy is to serve the user whose index is maximized.We first define an initial policy, δ^0 and then derive δ^1, the recommended (one step improved) policy.

3.1 Initial Policy

Consider a stationary state independent policy δ^0 (we also refer to this as the randomized policy) which serves user u with probability β_u. Thus, $\beta_u \geq 0$ and $\sum_u \beta_u = 1$. Let $\beta = [\beta_1, \beta_2, \ldots, \beta_m]$ be the initial policy (we also refer to this as the randomized policy) which we use to derive the final index policy through the one step policy improvement approach. Let g_β be the long run average reward and $w_\beta(0, q^n, t^n)$ be the bias vector for this policy. Then:

$$\begin{aligned} g_\beta + w_\beta(0, q^n, t^n) = \sum_{u=1}^{N} \beta_u * [&-C_u^n(q^n, t^n) + \\ &p_{00} V_{T-1}(0, q^{u,n+1,0}, t^{u,n+1,0})] \end{aligned} \tag{17}$$

Recall that $y_u^n \sim Poi(\alpha_u)$ is the data arrival process for user u and α_u is the long run average number of packets that arrive for user u in one time slot.

From standard stochastic theory [23], it can be shown that the randomized policy is stable if every user's queue is stable i.e., if $\alpha_u < [\beta_u * R * \pi_1]$.

For the purposes of the derivation of our index policy and simulations, we assume that the data queues for all users are stable.

3.2 Policy Improvement Step

For the given state $(0, q^n, t^n)$ the policy improvement step seeks to maximize $-C_u(q,t)+p_{00}w_\beta(q^{u,n+1,0}, t^{u,n+1,0})$ over all $u \in \{1, 2, \ldots, m\}$. This is equivalent to maximizing

$$\begin{aligned} I_u(0, q^n, t^n) = -C_u^n(q^n, t^n) + p_{00}[w_\beta(0, q^{u,n+1,0}, t^{u,n+1,0}) - \\ w_\beta(0, q^{u,n+1,1}, t^{n+1,1})] + \sum_{l=1,2,\ldots,m} q_l^n(1 + k_l * t_l^n) \end{aligned} \tag{18}$$

over all $u \in \{1, 2, \ldots, m\}$, since for a given state $(0, q^n, t^n)$ the additional term given below is independent of the action taken

$$-p_{00}w_\beta(0, q^{u,n+1,1}, t^{n+1,1}) + \sum_{l=1,2,\ldots,m} q_l^n(1 + k_l * t_l^n).$$

Recall that while formulating equation 15 we had argued that $q^{u,n+1,1}$ is independent of the action u, hence adding this terms doesn't affect derivation of policy. Using 7

$$\sum_{l=1,2,\ldots,m} q_l^n(1 + k_l * t_l^n) - C_u^n(q^n, t^n) = q_u^n(1 + k_u * t_u^n).$$

The index can be further simplified to the following form:

$$\begin{aligned} I_u(0, q^n, t^n) = q_u^n(1 + k_u * t_u^n) + p_{00}[w_\beta(0, q^{u,n+1,0}, t^{u,n+1,0}) \\ - w_\beta(0, q^*, t^{n+1,1})] \end{aligned} \tag{19}$$

To compute $I_u(0, q^n, t^n)$, we need an expression for

$$w_\beta(0, q^{u,n+1,0}, t^{u,n+1,0}) - w_\beta(0, q^{u,N=1,1}, t^{n+1,1}). \tag{20}$$

We derive a linear index in Theorem 1 by imposing certain additional constraints on the system.The proof of this is provided in the appendix.

Theorem 1. *Given the initial policy β and the reward structure as given in section 2, $I_u(0, q^n, t^n)$ is given by:*

$$I_u(0, q^n, t^n) = q_u^n(1 + k_u * t_u^n) + \frac{k_U * (q_u^n) * (t_u^n + 1)}{(\beta_u * Pr(Y_u \leq R)} + Z_u \tag{21}$$

For simplicity let $\beta_u = 1/m$; $u \in \{1, 2, \ldots, m\}$. Then, the index can be written as:

$$I_u(0, q^n, t^n) = q_u^n(1 + k_u * t_u^n) + \frac{k_u * m * (q_u^n) * (t_u^n + 1)}{Pr(Y_u \leq R)} + Z_u \tag{22}$$

4 Simulation Results

The objective of this paper is to optimize between two objectives: minimize the system ququе length while maintaining a degree of fairness among users. In this section, we use simulations that our Index policy gives a robust decision making tool for scheduling users. We evaluate the performance of our resource allocation algorithm by studying the following performance measures: Q' = Long run expected queue length of the system T' = Long run expcted age starvation age of a user. While minimizing Q' is important for the service provider (CSMA) to ensure full utilization of the existing infrastructure and minimizing costs; T' is a measure of fairness to the users and is directly related to customer satidfaction (and hence is a measure of Quality of Service (QoS) to individual users).

In this section we compare the performance of three algorithms; namely the myopic policy, the round robin rule and the Index policy (that we derive and test); by plotting the two performance measures (Q' vs T'). As explained earlier, k_u gives us a handle to model the relative proirity between queue lengths and starvation age.It can also be used to differentiate between users (depending on their type and priority, k_u values can be assigned to different users). For simplicity, we assume $k_u = k$ for all users u.For our Index policy; we plot the two performance measues for different values of k. We also assume that the number of users is a constant. The code for simulation has been written in Matlab. We begin by formulating the estimators for these parameters below.

4.1 The Estimators

We use $\hat{Q}'$ to denote the estimator of Q'. Let L be the number of independent sample paths simulated and D be the number of time slots in each path. Let $Q^{n,l}$ be the queue length in the n^{th} time slot of the l^{th} sample path, $l = \{1, 2, \ldots, L\}$ and $n = \{1, 2, \ldots, D\}$.The estimator $\hat{Q}'$ of Q' is given by:

$$\hat{Q}' = \frac{1}{L} \sum_{l=1,L} \left(\frac{1}{D} \sum_{n=1,T} Q^{n,l} \right) \tag{23}$$

Similarly, we use $\hat{T}'$ to denote the estimator of T'. Let $t_u^{n,l}$ denote the age of the u^{th} user in the n^{th} time slot along the l^{th} sample path; $l = \{1, 2, \ldots, L\}$, $n = \{1, 2, \ldots, D\}$ and $u = \{1, 2, \ldots, N\}$. The estimator $\hat{T}'$ of T' is given by:

$$\hat{T}' = \frac{1}{L} \sum_{l=1,L} \left[\frac{1}{D} \sum_{n=1,T} \left(\frac{1}{N} \sum_{u=1,N} t_u^{n,l} \right) \right] \tag{24}$$

These estimators are for long run performance measures and so while compiling the results, we collect samples only from the stationary region of the markov chain $\{X^n, n \geq 0\}$.

4.2 Results

We provide simulation results for 2 combinations of parameters (R, N as indicated on the plots) in Figures 1 and 2. PU's activity parameters in both cases are taken to be the same: $\lambda = 5$ and $\mu = 0.1$. For the first set of $N = 25$ users, the mean data arrival rate is taken to be $\alpha \in \{2, 4, \ldots, 50\}$, while for the second set of $N = 50$ users, it is taken as $\alpha \in \{2, 4, \ldots, 100\}$. We use $D = 5000$ time slots and collect samples from $L = 50$ sample paths.

We plot the queue length vs starvation age for our ploicy as well as both the myopic and round robin policy. Each point on the curve for our policy corresponds to a different k value. The value of k increases as we move from right to left on the curves. As expected, for $k = 0$ the results coincide with that of the myopic policy; this can be easily verified from the derivation and the final expression of our policy.

On increasing the k value, there is a significant reduction in the average age without a corresponding increase in queue lengths which makes our Index

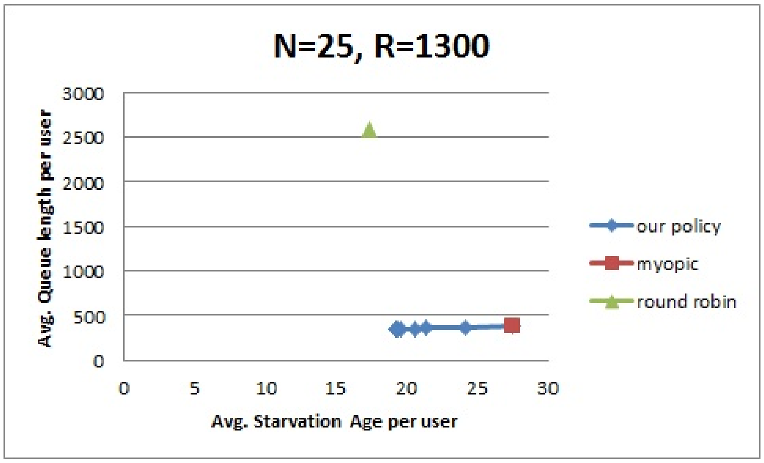

Figure 1: [N = 25]

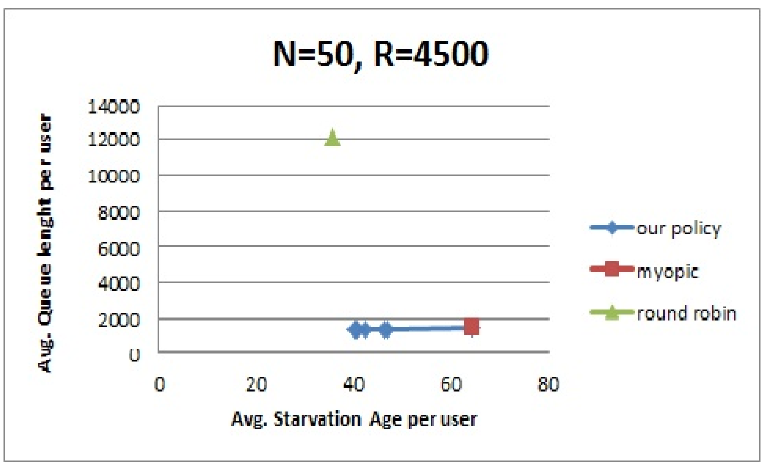

Figure 2: [N = 50]

policy robust. We observe similar behavior for various other parameter sets. Clearly, the recommended policy is the one with a very high k value. At this point, we achieve best of both the worlds; i.e. the mean starvation age of our policy is very close to that of the round robin rule while the increase in queue lengths is almost insignificant. However that does not make all the intermediate points insignificant. It is quite possible that the variance in age and queue length shows an increasing trend in k. This is something that we are exploring as part of our future work.

5 Conclusion

We develop an MDP based framework for dynamic scheduling of users on a channel of a CR network. We use the one step policy improvement approach to develop our Index policy which is intuitive and has a closed form expression. We demonstrate the robust performance of our policy on the chosen performance measures; Average Queue Length and Quality of Service (QoS) criteria (Starvation Age). Our policy performs most optimally since the queue length is reasonably close to the best possible (i.e., the queue length of myopic policy) and the QoS(starvation age) also is fairly close to the best possible (i.e.,starvation age of the round robin policy). Future work includes analyzing the performance of our policy for intermediate values of k as indicated above, and extending the framework to multiple channels.

Appendix

Proof of Theorem 1:

Proof. To find an expression for the index, we only need an expression for the differences in the biases as given by 20. Consider two sample paths followed under the randomized policy β of the process $\{X^n, Q^n, A^n; n \geq 0\}$ denoted by $\{X^{n,\nu}, Q^{n,\nu}, A^{n,\nu}\}$ for $\nu = 1, 2$. We model the system as follows:

At $n = 0$, $X^{0,1} = X^{0,2} = 0$; $Q^{0,1} = q^{u,n+1,0}, Q^{0,2} = q^{u,n+1,1}$ and $A^{0,1} = t^{u,n+1,0}, A^{0,2} = t^{n+1,1}$.

Let $l^{n,\nu}$ denote the user that is chosen to transmit in slot n along the sample path $\nu = 1, 2$. Recall that $Y^n = \{Y_1^n, Y_2^n, \ldots, Y_m^n\}$ is the vector of data arrival process for all users $u \in \{1, 2, \ldots, m\}$; where $Y_u^n \sim Poi(\alpha_u)$. Let $Y^{n,\nu}; \nu = 1, 2$ denote the data arrival process across the two sample paths. For $n > 0$, the two sample paths are coupled in such a way that:

$$X^{n,1} = X^{n,2}; \quad Y^{n,1} = Y^{n,2}; \quad l^{n,1} = l^{n,2}$$

The conditions described above clearly imply that $Q_u^{n,2} - Q_u^{n,1} = Q_u^{0,2} - Q_u^{0,1} = R$; $Q_l^{n,1} = Q_l^{n,2} \forall l \neq u$. Also $A_u^{0,1} = 0; A_u^{0,2} = t_u^n + 1$ which implies that $A_u^{n,1} = n; A_u^{n,2} = t_u^n + 1 + r$ and $A_l^{n,1} = A_l^{n,2} \forall l \neq u$. Let S_u^ν denote the first time when $Q_u^{n+1,\nu} \leq Q_u^{n,\nu}$. Clearly due to the coupling of both the sample paths, $S_u^1 = S_u^2 = S_u(say)$. Then S_u can be shown to be geometrically distributed with parameter p_u where p_u is defined as:

$$p_u = p_{00} * \beta_u * Pr(Y_u \leq R) \tag{25}$$

$$Pr(S_u = a) = (1 - p_u)^{a-1} * p_u \tag{26}$$

Let $C_\nu(S_u)$ be the total cost incurred along the sample path ν. From standard MDP theory, it can be easily argued that the difference in the bias functions is equal to the difference in the expected costs along both the sample paths. While computing the costs, we assume that the probability $P(Q_u^n > R)$ is negligible. This is a good approximation when data transmission rates are high. Thus $Q_u^n = 0$ for $n > 0$. Clearly

$$E[C_1(S_u)] = \sum_{a=1}^{\infty} \sum_{r=1}^{a} Q_u^{r,1} * (1 + k * A_u^{r,1}) * Pr(S_u = a) \tag{27}$$

$$E[C_2(S_u)] = \sum_{a=1}^{\infty} \sum_{r=1}^{a} Q_u^{r,2} * (1 + k * A_u^{r,2}) * Pr(S_u = a) \tag{28}$$

But $Q_u^{n,2} - Q_u^{n,1} = R$ for $n \leq S_u$, and 0 for $n > S_u$. On simplifying and neglecting the constant terms, we get:

$$E[C_2(S_u) - C_1(S_u)] = k * (q^{u,n+1,0}) * \frac{t_u^{n+1}}{p_u} + Z_u \tag{29}$$

where $Z_u = (0.5)*(\alpha_u)*k*(t_u+1)*p_u \sum_{a=1}^{\infty} a*(a+1)*(1-p_u)^{a-1}$. Substituting the value for p_u, the final index can be written as:

$$I_u(0, q^n, t^n) = q_u^n(1 + k_u * t_u^n) + \frac{k * (q_u^n) * (t_u^n + 1)}{(\beta_u * Pr(Y_u \leq R)} + Z_u \tag{30}$$

□

References

[1] I. Akyildiz, W. Lee, M. Vurun and S. Mohanty, Next generation/dynamic spectrum access/cognitive wireless networks: A survey, *Computer Networks*, **13**, (2006), 2127-2159.

[2] F.C. Commission, Spectrum policy task force report, f.c.c 02-155, (2002).

[3] R.-T. Sheu and J.-L. Wu, Performance analysis of rate control with scaling qos parameters for multimedia transmissions, *Proceedings os IEEE Conference - Communications*, **150**(5), (2003), 361-366.

[4] H. Kushwaha, Y. Xing, R. Chandramouli and Heffes, Reliable multimedia transmission over cgnitive radio networks using fountain codes, *Proceedings of IEEE*, **96**(1), (2008), 155-165.

[5] S. Haykin,Cognitive radio: Brain-empowered wireless communications, *IEEE Journal on Selected Areas of Communication*, **23**, (2005), 201-220.

[6] J. Mitola, *Cognitive radio: An integrated agent architecture for software defned radio*, PhD thesis, Royal Institute of Technology, Stockholm, 2000.

[7] I. Akyildiz, W.Y. Lee and K. Chowdhury, Crahns: Cognitive radio ad hoc networks, *Ad hoc networks 7*, **5**, (2009), 810-836.

[8] P. Si, Y. Zhang, R. Yang and J. Zhang, Resource allocation policy of primary users in proactively-optimization cognitive radio networks, *Third International Conference on Computing Intelligence, Communication Systems and Networks*, (2011), 343-347.

[9] G. Zhu, I.F. Akyildiz and G. Kuo, Stod:a spectrum-tree based on-demand routing protocol for multi-hop cognitive radio networks, *Proceedings of IEEE Globecom*, (2008), 1-5.

[10] J. Ma., G. Li and B. Juang, Signal processing in cognitive radio, *Proceedings of IEEE 97*, **5**, (2009), 805-823.

[11] Y. Zeng, Y.-C. Liang, A.T. Haong and R. Zhang, A review on spectrum sensing for cognitive radio: Challenges and solutions, *EURASIP Journal on Advances in Signal Processing*, (2010).

[12] E. Axell, G. Leus, E.G. Larsson and H.V. Poor, State-of-the-art and recent dadvances in spectrum sensing for cognitive radio, *IEEE Signal Processing Magazine (print)*, 101-116.

[13] T. Ycek and H. Arslan, A survey of spectrum sensing algorithms for cognitive radio applications, *IEEE Communications - Survey & Tutorials*, **11**.

[14] I. Akyiliz, B.F. Lo and R. Balakrishnan, Cooperative spectrum sensing in cognitive radio networks: A survey, *Physical communications*, **4**, (2011), 40-62.

[15] Y.C. Liang, Y. Zeng, E.C. Peh and A.T. Hoang, Sensing-throughput tradeoff for cognitive radio networks, *IEEE Transactions on Wireless Communications*, **7**, (2008), 1326-1337.

[16] W. Lee and I.F. Akyildiz, Optimal spectrum sensing framework for cognitive radio networks, *IEEE Transactions on Wireless Communications*, (2008), 3845-3857.

[17] X. Zhou, J. Ma, G.Y. Li, Y.H. Kwon, and A.C.K. Soong, Probability based optimization of inter-sensing duration and power control in cognitive radio, *IEEE Transactions in Wireless Communications*, **7**.

[18] X.-L. Huang, G. Wang, F. Hu and S. Kumar, The impact of spectrum sensing frequency and packet loading scheme on multimedia transmission over cognitive radio networks, *IEEE Transactions on Multimedia*, **13**, (2011), 748-761.

[19] H.-P. Shiang and M.V.D. Schaar, Queuing based dynamic channel selection for hetrofenous multimedia applications over cognitive radio networks, *IEEE Transactions on Multimedia*, **10**.

[20] V. Kulkarni, Modeling and Analysis of Stochastic Systems, Chapman & Hall, Inc, 1995.

[21] M. Puterman, *Markov Decision Processes - Discrete Stochastic Dynamic Programming*, John Wiley & Sons, Inc, 1994.

[22] N. Bolia and V. Kulkarni, Policies for resource allocation in wireless networks, *IEEE Transactions on Vehicular Technology*, **58**(4), (2009), 1823-1835.

[23] D. Gross and C. Harris, *Fundamentals of Queueing Theory*, John Wiley & Sons, Inc, 1985.