Throughput Maximization in Uncooperative Spectrum Sharing Networks

Thomas Stahlbuhk*†, Brooke Shrader† and Eytan Modiano*
* Massachusetts Institute of Technology, Cambridge, MA
† MIT Lincoln Laboratory, Lexington, MA

***Abstract*—We consider an opportunistic communication system in which a secondary transmitter communicates over the unused time slots of a primary user. In particular, we consider a system in which the primary user is uncooperative and transmits whenever its buffer is nonempty, and the secondary user relies on feedback from its receiver in order to decide when to transmit. The objective of the secondary user is to maximize its own throughput without degrading the throughput of the primary user. We analyze the maximum achievable throughput of the secondary user by formulating the problem as a partially observable Markov decision process. We derive bounds on the optimal solution and find a channel access policy for the secondary user that is near-optimal when the primary user's exogenous arrival rate is low. These results are then used to characterize the set of arrival rates to the primary and secondary users that may be stably supported by the system.**

I. Introduction

Many currently deployed wireless devices underutilize their allocated spectrum. As a result, the concept of spectrum sharing has emerged as a method for increasing spectral efficiency. In a spectrum sharing system, a secondary user opportunistically communicates over a primary user's channel in order to utilize times when the channel is idle (see Fig. 1). In this paper, we consider an uncooperative primary user that transmits whenever it has packets available to send. Then, the objective of the secondary user is to decide when to transmit in order to maximize its own throughput without degrading the throughput of the primary user.

In many wireless settings, the secondary user may rely upon channel sensing to detect the primary user's activity on the channel in order to avoid collisions. For this setting, the spectrum sharing problem has been well studied in the literature under the assumptions of both perfect and imperfect channel sensing [1]–[7]. Of particular interest are the works of [5] and [6] which consider optimal channel access policies for network models similar to those considered in this paper, with the assumption that the secondary user has access to channel sensing.

In this work, we consider systems in which channel sensing is unavailable. This may arise in settings where the ratio of the channel's propagation delay to packet transmission time is large [8], such as in satellite communications, wide-area wireless networks, or certain spread spectrum systems (e.g., frequency hopping). Our goal is to design a transmission policy for the secondary user that attempts to take advantage of the primary user's idle slots. The key challenge is that the secondary user cannot observe the primary user's queue backlog and must rely on feedback from its receiver in order to make transmission decisions.

In the following, we explore fundamental limits on the maximum throughput that is achievable by the secondary user while still allowing the primary user to meet its throughput requirements. Since the primary user's behavior is driven by whether it has packets available to send, a state which is hidden from the secondary user, the problem can be modeled as a partially observable Markov decision process (POMDP). In general, finding analytical solutions to POMDPs is difficult [9]. Herein, we proceed to derive lower and upper bounds on the optimal solution and relate these results to conditions for queue stability of the primary and secondary users. When the arrival rate to the primary user is low, these results provide a near-optimal secondary user policy. For higher rates, our results provide a meaningful characterization of the spectrum sharing system's maximum throughput.

II. Problem Formulation

We consider a time slotted system with two users communicating over a shared collision channel (see Fig. 1). The first user, which we refer to as the primary user, transmits a packet, in a time slot, whenever its queue is nonempty. Packets arrive to the primary user's queue according to a Bernoulli process with rate $\lambda \in (0, 1)$ and can be transmitted in the slot in which they arrive. Let $A(t) \in \{0, 1\}$ be the number of arrivals to the primary user's queue in time slot t and $Q(t)$ be the number of packets in the queue after arrival $A(t)$.

The objective of the secondary user is to maximize its rate of successful transmissions over the channel. The channel is modeled as a collision channel. Thus, if at time slot t only the primary or secondary user transmits, the transmission is successful, and if both users transmit at time t, both transmissions are unsuccessful. Both users receive ternary feedback immediately after each time slot indicating whether the time slot contained a successful transmission, a collision, or was idle. Primary user packets that are involved in a collision remain in the primary user's queue awaiting a future successful transmission.

This work was sponsored by NSF Grants CNS-1524317, CNS-1217048 and AST-1547331 and the Department of the Air Force under Air Force Contract FA8721-05-C-0002. Opinions, interpretations, conclusions and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

978-1-5090-1806-2/16/$31.00 ©2016 IEEE

The secondary user does not have access to channel sensing and must rely only on its history of actions and ternary feedback up to time slot $t-1$ in order to decide whether to attempt transmission at time slot t. We assume the secondary user knows that $A(t)$ is Bernoulli and the value of λ but cannot directly observe either $A(t)$ or $Q(t)$. This leads to a partially observable Markov decision process with the objective of maximizing the time average expected reward rate (i.e., the expected rate of successful secondary user transmissions). Our goal is to devise a policy that maps the history of actions and ternary feedback to a decision on whether or not to transmit at the given time slot such that the rate of expected successful transmissions by the secondary user is maximized. Note that, from the above, the secondary user can only obtain successful transmissions during time slots when the primary user's queue is empty. Therefore, in order to have a nonzero throughput, the secondary user must allow the primary user to service its arrival process.

III. POMDP Model

We proceed to give a concrete formulation of the problem as a POMDP following the notation of [9]. At each discrete time t, the state of the system is the queue size of the primary user, $Q(t) \in \{0, 1, 2, \dots\}$, and the secondary user may take an action $u(t) \in \mathcal{U} = \{TR, NT\}$ where TR denotes transmission and NT no transmission. The state $Q(t)$ is unobservable to the secondary user. Instead, ternary feedback provides an observation $z(t) \in \mathcal{Z} = \{S, C, I\}$ where S denotes a successful transmission, C a collision, and I an idle slot. The observed ternary feedback is given by the following function of $Q(t)$ and $u(t)$

$$z(t) = \begin{cases} S, & \text{if } Q(t) = 0 \text{ and } u(t) = TR \\ S, & \text{if } Q(t) > 0 \text{ and } u(t) = NT \\ C, & \text{if } Q(t) > 0 \text{ and } u(t) = TR \\ I, & \text{if } Q(t) = 0 \text{ and } u(t) = NT \end{cases}.$$

To inform its decisions, the secondary user has at its disposal the history $H(t)$ of all past actions and observations (i.e., $H(t) = \{u(0), \dots, u(t-1), z(0), \dots, z(t-1)\}$). Without loss of generality, we assume $Q(-1) = 0$ and this fact is known to the secondary user.

A solution to the POMDP is a policy $\pi = \{f_0, f_1, f_2, \dots\}$ that provides for each time t a mapping f_t such that $f_t(H(t)) \in \mathcal{U}$. Define by $B(t)$ the reward process of the POMDP where

$$B(t) = \begin{cases} 1, & \text{if } Q(t) = 0,\ u(t) = TR \\ 0, & \text{otherwise} \end{cases}.$$

(i.e., a reward of 1 is accrued for each successful secondary user transmission.) An optimal solution to the POMDP is a policy that achieves

$$\mu^* \triangleq \max_{\pi} \limsup_{T\to\infty} E\left[\frac{1}{T}\sum_{t=0}^{T-1} B(t)\right]. \tag{1}$$

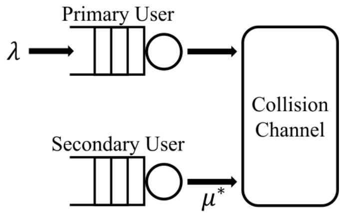

Fig. 1. Primary and secondary users transmitting over a collision channel. The secondary user's objective is to maximize its throughput.

This objective is equivalent to maximizing the throughput of the secondary user given that at each time t the secondary user always has an available packet to send.

Given $H(t)$, a probability mass function on the distribution of $Q(t)$ may be computed. In general, evaluating the optimal solution to a POMDP is difficult [9]. In the following sections we derive lower and upper bounds on (1) as a function of λ.

IV. Lower Bound

We derive a lower bound on (1) by evaluating the performance of a class of simple, suboptimal randomized stationary policies, π'. Policy π' is defined as follows: at each time slot t, π' does not transmit if the ternary feedback indicates a collision at time slot $t-1$ and otherwise transmits with probability p. For a given value of λ, p can be optimized to obtain the maximum secondary user throughput attainable by this class of policies. This gives the following result.

Theorem 1. *There exists a value of p, denoted p^*, such that the randomized stationary policy π' achieves a secondary user throughput*

$$\mu_{lb}^* \triangleq \lim_{T\to\infty} E\left[\frac{1}{T}\sum_{t=0}^{T-1} B(t)\right] = \begin{cases} 1-2\lambda, & \textit{for } \lambda \le \frac{1}{3} \\ \frac{(1-\lambda)^2}{4\lambda}, & \textit{for } \lambda > \frac{1}{3} \end{cases} \tag{2}$$

We plot (2) as a function of λ in Fig. 2. The proof is based on a Markov chain analysis of the policy and is omitted for brevity. Note that under policy π', the primary user's queue satisfies the following notion of stability.

Corollary 1. *Under policy π' with $p = p^*$, the primary user's queue is recurrent.*

This result is intuitive, since the secondary user may only obtain successful transmissions during time slots when the primary user's queue is empty.

V. Upper Bound

We next provide an upper bound on (1). We begin with the following two lemmas.

Lemma 1. *A policy that minimizes the expected time between successful secondary user transmissions, maximizes the secondary user's throughput.*

Proof. Recall that we assume that at time $t = -1$, $Q(-1) = 0$, and this fact is known by the secondary user. Now, assume at time $t-1$, the secondary user makes a successful transmission. Then, $H(t) = \{u(0), \ldots, u(t-1) = TR, z(0), \ldots, z(t-1) = S\}$ and the secondary user knows $Q(t-1) = 0$. However, given $Q(t-1) = 0$, the actions $u(0), \ldots, u(t-1)$ and observations $z(0), \ldots, z(t-1)$ are independent of all future events. Additionally, the future starting at time t is statistically the same as at time 0. Therefore, every successful transmission by the secondary user renews the system and there exists an optimal policy that, after each successful transmission, ignores all actions and observations that preceded the successful transmission.

Under this optimal policy, the intervals between successful transmissions are independent and identically distributed and thus form a renewal process. By the elementary renewal theorem [10, Theorem 5.6.2] we have

$$\limsup_{T\to\infty} E\left[\frac{1}{T}\sum_{t=0}^{T-1} B(t)\right] = \frac{1}{E\left[D\right]}. \qquad (3)$$

where $E\left[D\right]$ is the mean time between successful transmissions. Thus, a policy that minimizes $E\left[D\right]$ must maximize throughput (cf. (1)). □

Next, we show that there exists an optimal policy under which the secondary user chooses to not transmit after collisions.

Lemma 2. *There exists a policy that maximizes the secondary user's throughput and adopts the following rule: if $z(t-1) = C$, then action $u(t) = NT$ is selected.*

The intuition behind Lemma 2 is simple. Since after a collision we know that the primary user's queue is nonempty, a transmission by the secondary user will result in another collision. We omit the complete proof for brevity.

A. Augmented System Model

Define $N(t)$ to be the maximum number of packets that could potentially be in the primary user's queue at time slot t (i.e., the maximum possible value of $Q(t)$ given $H(t)$). Note that every time we observe $z(t-1) = I$ or obtain a successful secondary user transmission, we know $N(t) = 1$ since either event implies $Q(t-1) = 0$.

Now, to derive an upper bound on the secondary user's throughput, we augment the observation space $\mathcal{Z}$ to include additional information beyond ternary feedback. Any policy that ignores this additional feedback is admissible under the original system. Thus, the optimal solution to the augmented system is an upper bound on the original system.

Our augmentation is as follows. Assume the primary user serves packets in first come first served (FCFS) order and each packet is timestamped *with the time slot in which it arrived to the primary user's queue*. The new observation space consists of $\mathcal{Z}$ and the timestamps on the primary user's successfully transmitted packets. Upon observing the timestamp the secondary user knows that all packets that may have arrived to the primary user's queue prior to the timestamp have been successfully transmitted. Define $\tau(t)$ to be the most recently observed timestamp by the secondary user by the start of time slot t. Then, uncertainty only remains over those packets that may have arrived between time $\tau(t)$ and current time t. Clearly, following the observation of a new timestamp, $N(t) = t - \tau(t)$ is equal to the number of time slots in which a packet may have arrived to the primary user's queue since the most recent successfully transmitted packet and is the maximum number of packets that could potentially be in the primary user's queue at time t.

B. Augmented System as a Total Cost Problem

We now formulate a stochastic shortest path problem with the goal of minimizing the expected time until the next successful transmission by the secondary user. By Lemma 1, this is equivalent to maximizing the throughput. Furthermore, by Lemma 2, there exists an optimal policy that, given a collision at time $t-1$, chooses not to transmit at time t; thereby allowing the previously collided primary user packet to be successfully transmitted. We therefore restrict our attention to policies that take action $u(t) = NT$ whenever $z(t-1) = C$. In the following, we analyze the augmented system model in order to bound the optimal solution to the original system.

Given the above, we formulate an infinite horizon Markov decision process (MDP) over stage index $k \in \{0, 1, \ldots\}$. The index k defines a subsequence t_k of the sequence t, where the subsequence corresponds to the time slots t not following a collision (i.e., $z(t-1) \neq C$). Given that we force the policy to choose $u(t) = NT$ whenever $z(t-1) = C$, one can see that t_k corresponds to those time slots where decisions about which action to take must be made. Note that $N(t_k)$ completely characterizes the secondary user's knowledge about $Q(t_k)$ at time t_k. Moreover,

$$P\left(Q(t_k) = q | N(t_k) = n\right) = \binom{n}{q} \lambda^q (1-\lambda)^{n-q}.$$

Therefore, an optimal policy may be a function of $N(t_k)$ instead of $H(t_k)$ [9].

We define the states of our MDP as $N(t_k) \in \mathcal{N} = \{0, 1, 2, \ldots\}$ and actions as $u(t_k) \in \mathcal{U}$. Our state space $\mathcal{N}$ is composed of two parts. Prior to obtaining a successful transmission, the system is in states $\{1, 2, \ldots\}$ which corresponds to the maximum possible size of $Q(t_k)$, as defined above. However, when the secondary user obtains a successful transmission, the system enters a trapping state, 0, and remains there for all future indices at no further cost. We now give the state transition probabilities and (negative) reward function of the MDP. They are subsequently explained.

For $n \geq 1$, the transition probabilities are given by

$$P\left(N(t_{k+1}) = m | N(t_k) = n, u(t_k)\right) = \begin{cases} (1-\lambda)^n, & \text{for } u(t_k) = TR, \quad m = 0 \\ \lambda(1-\lambda)^{n+1-m}, & \text{for } u(t_k) = TR, \quad 2 \leq m \leq n+1 \\ (1-\lambda)^{n-1}, & \text{for } u(t_k) = NT, \quad m = 1 \\ \lambda(1-\lambda)^{n-m}, & \text{for } u(t_k) = NT, \quad 2 \leq m \leq n \end{cases} \qquad (4)$$

For $n \geq 1$, the reward function, which counts the number of time slots between successful transmissions by the secondary user, is given by

$$g\left(N(t_k) = n, u(t_k), N(t_{k+1}) = m\right) = \begin{cases} 0, & \text{for } u(t_k) = TR, \quad m = 0 \\ -2, & \text{for } u(t_k) = TR, \quad m \geq 2 \\ -1, & \text{for } u(t_k) = NT, \quad m \geq 1 \end{cases} \tag{5}$$

We now explain (4) and (5). Suppose at stage k, $N(t_k) = n \geq 1$. If the policy selects $u(t_k) = TR$, with probability $(1-\lambda)^n$ the primary user's queue will be empty and the secondary user will obtain a successful transmission. Then, $N(t_{k+1}) = 0$ and no cost is incurred. However, with probability $1-(1-\lambda)^n$ the primary user's queue will be nonempty. When this happens, a collision occurs at time t_k, the secondary user elects to not transmit at the next time slot to allow the primary user to successful transmit the previously collided packet, and the next subsequence decision point $t_{k+1} = t_k + 2$. Thus, a total of two time slots are lost by this event (cf., (5) case 2) and a new value $N(t_{k+1})$ is obtained according to the distribution of (4) case 2.

If, on the other hand, the policy selects $u(t_k) = NT$ the secondary user cannot obtain a successful transmission but also cannot incur a collision. Then, the next subsequent decision point is $t_{k+1} = t_k + 1$, one time slot is lost (cf., (5) case 3), and the next value of $N(t_{k+1})$ can be shown to be given by (4) cases 3 and 4.

Given the transition probabilities (4) and reward process (5), our objective is to obtain a policy $\pi = \{f_0, f_1, \dots\}$ that maximizes

$$\sigma \triangleq \lim_{K\to\infty} E\left[\sum_{k=0}^{K-1} g\left(N(t_k), u(t_k), N(t_{k+1})\right) \middle| N(t_0) = 1\right]. \tag{6}$$

We define σ^* to be the maximum, attainable value of (6) over the set of all policies.

Note that following a successful secondary user transmission, $N(t) = 1$. Then, (6) can be seen to be the (negative) expected number of time slots until the beginning of the next successful transmission by the secondary user starting from a time slot immediately following a successful transmission. Using (3), an upper bound on μ^* is

$$\mu^*_{ub} \triangleq \frac{1}{1-\sigma^*}.$$

The specified Markov decision process is a negative expected total-reward problem with a countable state space and finite action space. By [11, Theorem 7.3.6] such a problem has an optimal deterministic, stationary policy (i.e., there exists an optimal policy $\pi^* = \{f^*, f^*, \dots\}$ such that f^* deterministically maps $f^*(N(t_k)) \in \mathcal{U}$).

Now, consider a deterministic stationary policy that chooses to transmit in states $N(t_k) \in \{1, \dots, M-1\}$ and not transmit in state $N(t_k) = M$ (i.e., the smallest numbered state in which the policy decides to not transmit is M). It follows that states $\{M+1, M+2, \dots\}$ are unreachable from states $\{0, 1, 2, \dots, M\}$. Thus, states $\{0, 1, 2, \dots, M\}$ form a finite state Markov chain under this policy, and Bellman's equations for this policy have a solution over these states. Denote by $V(n)$ the cost-to-go of state $n \in \{0, 1, \dots, M\}$. From equations (4) and (5) we see that Bellman's equations for the policy are given by the following. For states $1, 2, \dots, M-1$ in which the policy transmits

$$V(n) = \sum_{m=1}^{n} \lambda(1-\lambda)^{n-m}\left(V(m+1) - 2\right), \quad \text{for } n = 1, 2, \dots, M-1. \tag{7}$$

For state M in which the policy does not transmit

$$V(M) = (1-\lambda)^M\left(V(1) - 1\right) + \sum_{m=1}^{M} \lambda(1-\lambda)^{M-m}\left(V(m) - 1\right). \tag{8}$$

Furthermore, $V(0) = 0$, since once entering state 0 no more negative reward is accrued.

By the definition of cost-to-go, $V(1)$ defined by (7) is the same as (6), (i.e., $V(1) = \sigma$). For a given value of λ we now optimize $V(1)$ over integer value M in order to characterize σ^*. Note that if $M = 1$, the expected time until a successful transmission is unbounded. Thus, we restrict our attention to $M \geq 2$. The optimization is over the set of all policies that do not transmit in at least one state. However, it may be shown that as M goes to infinity, $V(1)$ approaches the value of σ obtained by the policy that transmits in all states.

C. Solving Bellman's Equations

In this section we show that μ^*_{ub} can be found using a simple search over integer values M. We omit the proofs for brevity. We begin by giving the solution to $V(1)$ from Bellman's equations, (7) and (8).

Proposition 1.

$$V(1) = \begin{cases} \frac{(2-4\lambda)(\lambda(1-\lambda))^{M-1} + 2\lambda(1-\lambda)^{M-1} - \lambda^{M-1}}{(1-2\lambda)(\lambda^{M-1}-1)(1-\lambda)^{M-1}}, & \text{for } \lambda \neq \frac{1}{2} \\ \frac{\left(\frac{1}{2}\right)^{M-1} + M - \frac{3}{2}}{\left(\frac{1}{2}\right)^{M} - \frac{1}{2}}, & \text{for } \lambda = \frac{1}{2} \end{cases} \tag{9}$$

We now give the following proposition for (9).

Proposition 2. *Over integers $M > 1$, there exists an M^* such that $V(1)$ is monotonically increasing for $1 < M \leq M^*$ and monotonically decreasing for $M \geq M^*$.*

Proposition 2 implies that we may find μ^*_{ub} from (9) by a simple linear search over the integers $M \geq 2$. In Fig. 2, we plot μ^*_{ub} which was computationally found in this way. We observe, μ^*_{ub} implies that policy π' with $p = p^*$ performs very well when the rate λ is low.

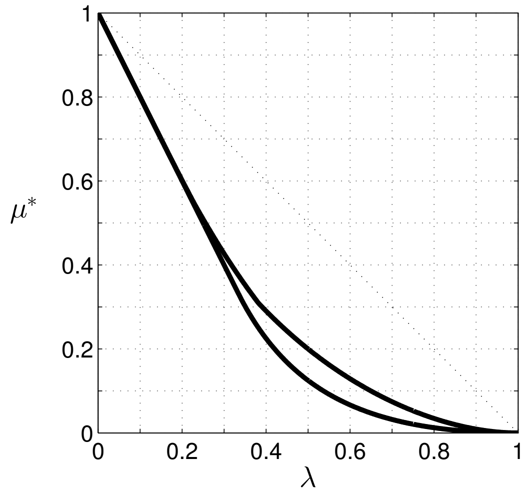

Fig. 2. The upper and lower bounds on μ^*. The line $1 - \lambda$ is shown for reference.

VI. Secondary User Queue Stability

The previous sections gave lower and upper bounds on the maximum secondary user throughput, μ^*, assuming that the secondary user always had an available packet to send at every time slot. In this section, we remove this assumption and instead assume that packets arrive to the secondary user as a Bernoulli process $A_s(t)$ with rate λ_s. Packets awaiting transmission by the secondary user are kept in a FCFS queue. The queue backlog at time t is denoted by $Q_s(t)$. We now relate μ_{lb}^* and μ_{ub}^* to conditions on λ_s for stability of the secondary user's queue.

A. Finite Queueing Delay for $\lambda_s < \mu_{lb}^*$

By definition, the arrival process $A_s(t)$ has i.i.d. interarrival times. Additionally, assume we apply policy π' with the modification that transmissions are only attempted when $Q_s(t) > 0$. Then, from the definition of π', it may be seen that the service time of each packet is i.i.d. with rate μ_{lb}^*. Define, W_r as the queueing delay of the r^{th} packet to arrive to the queue. Then a well known result for GI/GI/1 queues, in general, is that

$$P(W_r < \infty) = 1$$

for all r if $\lambda_s < \mu_{lb}^*$ [12]. Thus, every packet that arrives to the secondary user's queue, exits in finite time with probability 1 if $\lambda_s < \mu_{lb}^*$.

B. Necessary Condition for Mean Rate Stability

A discrete-time queue is defined to be mean rate stable if [13]

$$\lim_{T\to\infty} \frac{E\left[Q_s(T)\right]}{T} = 0.$$

Mean rate stability is a weak form of queue stability in that many other definitions of queue stability imply it under general assumptions [13]. Now, at every time T,

$$Q_s(T) \geq \sum_{t=0}^{T-1} A_s(t) - \sum_{t=0}^{T-1} B(t).$$

This implies,

$$\begin{aligned}
&\liminf_{T\to\infty} E\left[\frac{Q_s(T)}{T}\right] \\
&\geq \lim_{T\to\infty} E\left[\frac{1}{T}\sum_{t=0}^{T-1} A_s(t)\right] - \limsup_{T\to\infty} E\left[\frac{1}{T}\sum_{t=0}^{T-1} B(t)\right] \\
&\geq \lambda_s - \mu^*.
\end{aligned}$$

Thus, a necessary condition for mean rate stability is $\lambda_s \leq \mu^* \leq \mu_{ub}^*$.

VII. Conclusion

In this paper, we studied the throughput of a secondary user opportunistically communicating over the unused time slots of an uncooperative primary user. In particular, we considered the setting where channel sensing is unavailable and the secondary user must rely upon feedback in order to make transmission decisions. We characterized the achievable throughput of the secondary user by deriving upper and lower bounds and established conditions on the arrival rate for the system to be stable.

Possible extensions to our work include the consideration of correlated arrivals to the primary user, as well as multiple primary channels. We expect that correlated arrivals would make it easier for the secondary user to predict the primary user's behavior, leading to higher throughputs. The case of multiple primary channels would give rise to the additional problem of channel selection. That is, in addition to deciding whether or not to transmit, the secondary user must decide which channel to transmit on.

References

[1] Y. Chen, Q. Zhao, and A. Swami, "Joint Design and Separation Principle for Opportunistic Spectrum Access in the Presence of Sensing Errors," *IEEE Trans. Inf. Theory*, Vol. 54, No. 5, 2008.

[2] S. H. A. Ahmad, et al., "Optimality of Myopic Sensing in Multichannel Opportunistic Access," *IEEE Trans. Inf. Theory*, Vol. 55, No. 9, 2009.

[3] K. Liu and Q. Zhao, "Indexability of Restless Bandit Problems and Optimality of Whittle's Index for Dynamic Multichannel Access," *IEEE Trans. Inf. Theory*, Vol. 56, No. 11, 2010.

[4] M. Levorato, U. Mitra, and M. Zorzi, "Cognitive Interference Management in Retransmission-Based Wireless Networks," *IEEE Trans. Inf. Theory*, Vol. 58, No. 5, 2012.

[5] R. Urgaonkar and M. J. Neely, "Opportunistic Cooperation in Cognitive Femtocell Networks," *IEEE J. Selected Areas in Communications*, Vol. 30, No. 3, 2012.

[6] S. Kompella, et al., "Cooperation in Cognitive Underlay Networks: Stable Throughput Tradeoffs," *IEEE/ACM Trans. Networking*, Vol. 22, No. 6, 2014.

[7] J. Jeon, et al. "The Stability Property of Cognitive Radio Systems with Imperfect Sensing," *IEEE J. Selected Areas in Communications*, Vol. 32, No. 3, 2014.

[8] D. P. Bertsekas and R. G. Gallager, *Data Networks*, Second Edition, Prentice-Hall, 1992.

[9] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, Vol. I, Third Edition, Athena Scientific, 2005.

[10] R. G. Gallager, *Stochastic Processes: Theory for Applications*, Cambridge University Press, 2013.

[11] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, John Wiley and Sons, 2005.

[12] R. Srikant and L. Ying, *Communication Networks: An Optimization, Control, and Stochastic Networks Perspective*, Cambridge University Press, 2014.

[13] M. J. Neely, "Stability and Capacity Regions for Discrete Time Queueing Networks," arXiv preprint arXiv:1003.3396, 2010.