Cognitive MAC Protocols Using Memory for Distributed Spectrum Sharing Under Limited Spectrum Sensing

Jaeok Park and Mihaela van der Schaar, *Fellow, IEEE*

***Abstract*—The main challenges of cognitive radio include spectrum sensing at the physical (PHY) layer to detect the activity of primary users and spectrum sharing at the medium access control (MAC) layer to coordinate access among coexisting secondary users. In this paper, we consider a cognitive radio network in which a primary user shares a channel with secondary users that cannot distinguish the signals of the primary user from those of a secondary user. We propose a class of distributed cognitive MAC protocols to achieve efficient spectrum sharing among the secondary users while protecting the primary user from potential interference by the secondary users. By using a MAC protocol with one-slot memory, we can obtain high channel utilization by the secondary users while limiting interference to the primary user at a low level. The results of this paper suggest the possibility of utilizing MAC protocol design in cognitive radio networks to overcome limitations in spectrum sensing at the PHY layer as well as to achieve spectrum sharing at the MAC layer.**

***Index Terms*—Cognitive medium access control, cognitive radio networks, protocols with memory, spectrum sensing, spectrum sharing.**

I. Introduction

TODAY'S expanding demand for wireless services has necessitated cognitive radio technology in order to overcome the limitations of the conventional static spectrum allocation policy. Cognitive radio technology enables a more efficient use of limited spectrum resources by allowing unlicensed users (or secondary users) to opportunistically utilize licensed spectral bands. The main challenges of cognitive radio include *spectrum sensing* at the physical (PHY) layer to detect the activity of licensed users (or primary users) and *spectrum sharing* at the medium access control (MAC) layer to coordinate access among coexisting secondary users [1]. Spectrum sensing is needed to identify spectrum opportunities or spectrum holes, while spectrum sharing helps secondary users achieve an efficient and fair use of identified spectrum opportunities.

In this paper, we study a MAC protocol design problem for a cognitive radio network in which a primary user shares a spectral band (or a channel) with multiple secondary users. One of the main assumptions of our model is that the secondary users have limited spectrum sensing capability at the PHY layer in the sense that they are unable to distinguish between the activities (i.e., spectrum access) of the primary user and a secondary user. This assumption contrasts with and is weaker than the prevailing assumption, made in previous work on MAC design for cognitive radio, that spectrum sensing is perfect in that secondary users can always detect the presence of primary users (see, for example, [2] and [3]). [4] relaxes the assumption of perfect spectrum sensing and considers sensing errors at the PHY layer. However, [4] requires that the signals of primary users be statistically distinguishable from those of secondary users. On the contrary, our assumption is valid when the signals of primary users are (statistically) indistinguishable from those of secondary users.

Another key assumption we maintain is that explicit coordination messages cannot be communicated between a central controller and a user, or between users. This implies that the primary user cannot broadcast its presence to the secondary users for spectrum sensing and that centralized scheduling schemes such as TDMA cannot be used for spectrum sharing. Again, this assumption contrasts with and is weaker than the assumption made in existing work that there are central controllers or dedicated control channels (see, for example, [2] and [3]). As pointed out in [1], in cognitive radio networks, protocols requiring broadcast messages cause a major problem due to the lack of a reliable control channel as a channel has to be vacated whenever a primary user returns to the channel.

Our protocol design for the secondary users is based on MAC protocols with memory, which are formally presented in [5]. Under a protocol with memory, users adjust their transmission parameters depending on the local histories of their own transmission actions and feedback information. Hence, protocols with memory can be implemented in a distributed way without explicit message passing for any given sensing ability of users. Moreover, by exploiting information embedded in local histories, protocols with memory enable a secondary user to "change its transmitter parameters based on interaction with the environment in which it operates," as demanded by the definition of cognitive radio [6].

In [5], we have focused on the problem of achieving coordinated access among symmetric users by using a protocol with memory. In a cognitive radio network, where a primary user exists, another kind of coordination is needed to ensure that the secondary users do not interfere with the primary user. In this paper, we show that a class of protocols with one-slot memory can achieve high channel utilization by the secondary

Paper approved by S. Roy, the Editor for Cross-Layer Wireless/Cognitive Radio Networks of the IEEE Communications Society. Manuscript received July 2, 2010; revised February 4, 2011.

The authors are with the Electrical Engineering Department, University of California, Los Angeles (UCLA), Los Angeles, CA 90095-1594 USA (e-mail: {jaeok, mihaela}@ee.ucla.edu).

Digital Object Identifier 10.1109/TCOMM.2011.063001.100388

users while protecting the primary user at a desired level. We also show that the system performance can be improved by utilizing longer memory. The results of this paper suggest that a carefully designed MAC protocol can be used in place of an algorithm for primary user detection at the PHY layer. The main contribution of this paper is to illustrate the possibility of utilizing MAC protocol design to overcome limitations in spectrum sensing at the PHY layer as well as to achieve spectrum sharing at the MAC layer.

In recent years, there have been burgeoning research efforts involving cognitive radio networks. Due to space limitations, we review only a few of them, focusing on the most related work, and refer the interested reader to [1] for a comprehensive survey. [2] examines gains from spectrum agility in terms of spectrum utilization. Our model corresponds to the non-agile case of [2] as secondary users in our model stay in the same channel for the considered horizon of time. This is because our model is not equipped with ideal control devices as assumed in [2]. [3] uses a mechanism design approach to determine the allocation of spectrum opportunities to selfish secondary users. [4] analyzes the decision of secondary users to sense and access channels using a partially observable Markov decision process framework. [7] evaluates performance under two spectrum access schemes using different sensing, back-off, and transmission mechanisms. [8] develops a sensing-period optimization mechanism and an optimal channel-sequencing algorithm for efficient discovery of spectrum opportunities. [9] models the interactions between secondary users as a non-cooperative game and derives the price of anarchy. A survey on MAC protocols for cognitive radio networks is presented in [10]–[12].

The rest of this paper is organized as follows. In Section II, we describe the system model and the proposed class of MAC protocols. In Section III, we define performance metrics and provide methods to compute them based on Markov chains. In Section IV, we formulate and analyze a protocol design problem. In Section V, we discuss how to enhance the proposed protocols by utilizing longer memory. In Section VI, we conclude this paper.

II. System Model and Protocol Description

A. System Model

We consider a slotted multiaccess system [13, Ch. 4] where a single primary user and N secondary users share a communication channel. We assume that N is fixed over time. Time is divided into slots of equal length, and the primary and secondary users maintain synchronized time slots. Packets are fragmented into a length that can be transmitted with a time slot. A user can attempt to transmit a packet or wait in a slot when it has a packet to transmit. Due to interference, only one user can have successful transmission in a slot, and simultaneous transmission by more than one user results in a collision.

The traffic of the primary user arrives following a stochastic process. An arrival of traffic generates multiple packets for the primary user. Let T_{pac} be the average number of packets generated by an arrival of traffic. Also, let T_{int} the average time interval (measured in slots) between two consecutive arrivals of traffic. We assume that $T_{int} > T_{pac}$. This assumption will hold when the primary user has bursty traffic. In each slot, the primary user has either a packet to transmit or none depending on traffic arrivals and transmission outcomes. The state of the primary user, denoted by y_p, is *on* if the primary user has a packet to transmit and *off* otherwise. A similar *on-off* model for the primary user can be found in [2] and [8].

Each secondary user always has a packet to transmit. After a user makes a transmission attempt, it learns whether the transmission is successful or not using an acknowledgement (ACK) response. Also, the secondary users perform spectrum sensing in every slot to find out whether the channel is accessed or not. We say that spectrum sensing is *limited* if the secondary users cannot distinguish access by the primary user from that by a secondary user at all, and *perfect* if they can recognize access by the primary user. In this paper, we focus on limited spectrum sensing. In other words, when the channel is sensed busy, the secondary users know that some user accessed the channel, but not whether the primary user or a secondary user did. Using the information obtained from ACK responses and limited spectrum sensing, a secondary user can classify its channel state in a slot into four states, *idle*, *busy*, *success*, and *failure*, as in [14]. The state of secondary user i, denoted by y_i, is *idle* if no user transmits, *busy* if secondary user i does not transmit but at least one other user transmits, *success* if secondary user i transmits and succeeds, and *failure* if secondary user i transmits but fails.

B. Protocol Description

1) Protocol for the Primary User: The decision rule for the primary user is to transmit whenever it has a packet to transmit. Note that the primary user does not need to modify its decision rule for coexistence with the secondary users, which is consistent with the requirements of cognitive radio networks.

2) Protocol for the Secondary Users: The decision rule for the secondary users is prescribed by a protocol with one-slot memory [5]. A protocol with one-slot memory specifies a transmission probability for each possible state of the previous slot, and thus it can be formally represented by a function $f : \mathcal{Y}_s \to [0,1]$, where $\mathcal{Y}_s$ is the set of the states of a secondary user, i.e., $\mathcal{Y}_s = \{\textit{idle}, \textit{busy}, \textit{success}, \textit{failure}\}$. A secondary user whose state is $y \in \mathcal{Y}_s$ in the previous slot transmits with probability $f(y)$ in the current slot.

Definition 1: A protocol f with one-slot memory is *non-intrusive* if $f(\textit{busy}) = 0$.

When the secondary users follow a non-intrusive protocol, they wait in a slot following a busy slot. Thus, a non-intrusive protocol allows the primary user not to be interrupted by the secondary users once it obtains a successful transmission.

Definition 2: A protocol f with one-slot memory has the *fairness level* $\theta \in (0, 1]$ if the average number of consecutive successes by a secondary user while the primary user does not transmit is $1/\theta$, or

$$1 - f(\textit{success})(1 - f(\textit{busy}))^{N-1} = \theta. \quad (1)$$

Suppose that there is no transmission by the primary user. Once a secondary user succeeds, it has a successful transmission in the next slot with probability $f(\textit{success})(1 - f(\textit{busy}))^{N-1}$, and thus the average number of consecutive successes by a secondary user is given by $1/[1 - f(\textit{success})(1 - f(\textit{busy}))^{N-1}]$. As the fairness level is smaller, a secondary user keeps using the channel for a longer period once it succeeds, which makes other secondary users wait longer until they have a successful transmission. In [15], a protocol with fairness level θ is said to be M-short-term fair if $1/\theta \leq M$.

In this paper, we restrict attention to non-intrusive protocols with one-slot memory. We focus on protocols with one-slot memory because they are simple to design and implement [5]. Also, we impose non-intrusiveness in order to give priority to the primary user. Given a protocol, its fairness level can be computed by (1). In what follows, we consider protocols having a particular fairness level $\theta \in (0, 1]$, interpreting θ as the fairness level most preferred by the protocol designer. Once the fairness level is specified as θ, (1) together with non-intrusiveness implies that $f(\textit{success}) = 1 - \theta$. We denote the other elements of a protocol, $f(\textit{idle})$ and $f(\textit{failure})$, by q and r, respectively. For simplicity, we call hereafter a non-intrusive protocol with one-slot memory having fairness level θ a θ-fair non-intrusive protocol. To sum up, a θ-fair non-intrusive protocol can be expressed as

$$f(\textit{idle}) = q,\ f(\textit{busy}) = 0,$$
$$f(\textit{success}) = 1 - \theta,\ f(\textit{failure}) = r,$$

for some $q, r \in [0, 1]$.

Note that in this paper all the secondary users adopt the same protocol, following which each secondary user chooses a transmission probability depending on its channel state in the previous slot. We can consider a more general scenario where the secondary users can have different channel conditions, which change dynamically, and the probability of successful transmission by a user in a slot is determined by the channel conditions of all users transmitting in that slot. In such a scenario, a protocol for the secondary users can be extended to prescribe a transmission probability for a user depending on its channel conditions as well as on its channel states. Moreover, the secondary users may be given different protocols, for example, if there are several classes of users to be treated differently. In this paper, we simplify the protocol design problem by assuming a channel model with perfect reception or collision as in [13, Ch. 4] and specifying the same protocol for all the secondary users, while leaving the extension for future research.

III. Performance Metrics

A. Definition

1) Collision Probability of the Primary User: In overlay spectrum sharing, it is important to protect the primary user from interruption by the secondary users. We measure interference experienced by the primary user by the collision probability of the primary user, defined as

$$P_c = \frac{\text{No. of collisions experienced by PU}}{\text{No. of transmission attempts by PU}},$$

where PU represents "primary user." That is, the collision probability of the primary user is the probability that it experiences a collision when it attempts to transmit a packet.

2) Success Probability and the Channel Utilization Rate of the Secondary Users: We measure the utilization of spectrum opportunities by the success probability of the secondary users, defined as

$$P_s = \frac{\text{No. of successes by SUs}}{\text{No. of slots in which PU is } \textit{off}},$$

where SU represents "secondary user." In other words, the success probability of the secondary users is the probability that a secondary user has a successful transmission when the primary user has no packet to transmit. The channel utilization rate (or throughput) of the secondary users is defined as the proportion of time slots in which a secondary user has a successful transmission, i.e.,

$$C_s = \frac{\text{No. of successes by SUs}}{\text{No. of slots}}.$$

3) Channel Utilization Rate of the System: The channel utilization rate of the system is defined as the proportion of time slots in which a successful transmission (either by the primary user or by a secondary user) occurs, i.e.,

$$C = \frac{\text{No. of successes}}{\text{No. of slots}}.$$

B. Computation

1) Success Probability of the Secondary Users: We define an *on* period and an *off* period as a period in which the state of the primary user is *on* and *off*, respectively, between two consecutive arrivals of traffic. We first study the operation of the system in an *off* period, in which the primary user is inactive. To analyze the performance of a protocol in an *off* period, we construct a Markov chain whose state space is $\{0, 1, \ldots, N\}$, where state k represents transmission outcomes in which exactly k secondary users transmit. The transition probability from state k to state k' in an *off* period, denoted by $P_{\textit{off}}(k'|k)$, under a θ-fair non-intrusive protocol is given by

$$P_{\textit{off}}(k'|0) = \binom{N}{k'} q^{k'} (1-q)^{N-k'} \quad \text{for } k' = 0, \ldots, N, \quad (2)$$

$$P_{\textit{off}}(k'|1) = \begin{cases} \theta & \text{for } k' = 0 \\ 1-\theta & \text{for } k' = 1 \\ 0 & \text{for } k' = 2, \ldots, N, \end{cases}$$

$$P_{\textit{off}}(k'|k) = \begin{cases} \binom{k}{k'} r^{k'} (1-r)^{k-k'} & \text{for } k' = 0, \ldots, k \\ 0 & \text{for } k' = k+1, \ldots, N, \end{cases}$$
$$\text{for } k = 2, \ldots, N. \quad (3)$$

The transition matrix of the Markov chain can be written in the form of

$$\mathbf{P}_{\textit{off}} = \begin{array}{c} \\ 0 \\ 2 \\ \vdots \\ N-1 \\ N \\ 1 \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 2 & \cdots & N-1 & N & \quad 1 \end{array} \\ \left(\begin{array}{ccccc|c} \diamond & \diamond & \cdots & \diamond & \diamond & \diamond \\ * & * & \cdots & 0 & 0 & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ * & * & \cdots & * & 0 & * \\ * & * & \cdots & * & * & * \\ \hline \theta & 0 & \cdots & 0 & 0 & 1-\theta \end{array} \right) \end{array}$$

where the entries marked with a diamond and an asterisk can be found in (2) and (3), respectively.

Consider a slot t in which the state of the primary user has changed from *on* to *off*, i.e., $y_p^{t-1} = on$ and $y_p^t = \mathit{off}$, where y_p^τ is the state of the primary user in slot τ. Since such a transition can occur only if the primary user transmitted a packet successfully in slot $t-1$, it must be the case that $y_i^{t-1} = busy$ for every secondary user i, where y_i^τ is the state of secondary user i in slot τ. By non-intrusiveness, no secondary user transmits in slot t, and thus an *off* period always begins with an idle slot (state 0). Starting from an idle slot, the secondary users contend with each other until one of them obtains a success, i.e., until state 1 is reached. When a secondary user obtains a success, it transmits with probability $1-\theta$ in the next slot while all the other secondary users wait. A period of consecutive successes by a secondary user ends with an idle slot, when the successful user waits. In short, an *off* period can be considered as the alternation of a contention period and a success period, which is continued until traffic arrives to the primary user. A success period consists of slots with consecutive successes by a secondary user, whereas a contention period begins with an idle slot and lasts until a secondary user succeeds. Since all the secondary users transmit with the same transmission probability following an idle slot, they have an equal chance of becoming the successful user for the following success period at the point when a contention period starts.

Let T_s and T_{ns} be the average duration (measured in slots) of a success period and a contention period, respectively. T_s is determined by the fairness level θ, where the relationship is given by $T_s = 1/\theta$. Let $\mathbf{Q}_{\mathit{off}}$ be the N-by-N matrix in the upper-left corner of $\mathbf{P}_{\mathit{off}}$. Also, let $\mathcal{A} = \{(q,r) \in [0,1]^2 : q = 0, r = 1, \text{or } (q,r) = (1,0)\}$. If $(q,r) \in \mathcal{A}$, then the average duration of a contention period is infinitely long (i.e., $T_{ns} = +\infty$), yielding $P_s = 0$. Suppose that $(q,r) \notin \mathcal{A}$. Then $(\mathbf{I} - \mathbf{Q}_{\mathit{off}})^{-1}$ exists and is called the fundamental matrix for $\mathbf{P}_{\mathit{off}}$, when state 1 is absorbing (i.e., $\theta = 0$) [16]. The average number of slots in state $k \neq 1$ starting from state 0 (an idle slot) is given by the $(1,k)$-entry of $(\mathbf{I} - \mathbf{Q}_{\mathit{off}})^{-1}$. Hence, the average number of slots to hit state 1 (a success slot) for the first time starting from an idle slot is given by the first entry of $(\mathbf{I} - \mathbf{Q}_{\mathit{off}})^{-1}\mathbf{e}$, where $\mathbf{e}$ is a column vector of length N all of whose entries are 1. Hence, we obtain $T_{ns} = [(\mathbf{I} - \mathbf{Q}_{\mathit{off}})^{-1}\mathbf{e}]_1$, where $[\mathbf{v}]_k$ denotes the k-th entry of vector $\mathbf{v}$. Note that T_{ns} is independent of θ. That is, the average duration of a contention period is not affected by the average duration of a success period. The success probability of the secondary users can be computed by

$$P_s = \frac{T_s}{T_{ns} + T_s} = \frac{1}{\theta[(\mathbf{I} - \mathbf{Q}_{\mathit{off}})^{-1}\mathbf{e}]_1 + 1}, \tag{4}$$

for $(q,r) \notin \mathcal{A}$.

An alternative method to compute the success probability of the secondary users when $(q,r) \notin \mathcal{A}$ is to use a stationary distribution. Since $\theta \in (0,1]$, all states communicate with each other under the transition matrix $\mathbf{P}_{\mathit{off}}$. Hence, the Markov chain is irreducible, and there exists a unique stationary distribution $\mathbf{w}_{\mathit{off}}$, which satisfies

$$\mathbf{w}_{\mathit{off}} = \mathbf{w}_{\mathit{off}}\mathbf{P}_{\mathit{off}} \text{ and } \mathbf{w}_{\mathit{off}}\mathbf{e} = 1. \tag{5}$$

Let $w_{\mathit{off}}(k)$ be the entry of $\mathbf{w}_{\mathit{off}}$ corresponding to state k, for $k = 0, 1, \ldots, N$. Then $w_{\mathit{off}}(k)$ gives the probability of state k during an *off* period. In particular, the success probability of the secondary users is given by $w_{\mathit{off}}(1)$. Since contention and success periods alternate from the beginning of an *off* period, the stationary distribution yields the probabilities of states for any duration of an *off* period (assuming that T_{off} is sufficiently larger than $T_{ns} + T_s$), not just the limiting probabilities as an *off* period lasts infinitely long. By manipulating (5), we can derive that $w_{\mathit{off}}(1) = P_s$, whose expression is given in (4). To sum up, the success probability of the secondary users under a θ-fair non-intrusive protocol can be computed as

$$P_s(q,r) = \begin{cases} \frac{1}{\theta[(\mathbf{I} - \mathbf{Q}_{\mathit{off}})^{-1}\mathbf{e}]_1 + 1}, & \text{if } (q,r) \notin \mathcal{A}, \\ 0, & \text{if } (q,r) \in \mathcal{A}. \end{cases} \tag{6}$$

2) Collision Probability of the Primary User: We next study the operation of the system in an *on* period, in which the primary user always transmits. To analyze the performance of a protocol in an *on* period, we construct another Markov chain with the same state space $\{0, 1, \ldots, N\}$ as before. Again, state k corresponds to transmission outcomes in which exactly k secondary users transmit. The transition probability from state k to state k' in an *on* period, denoted by $P_{on}(k'|k)$, under a θ-fair non-intrusive protocol is given by

$$P_{on}(k'|k) = \begin{cases} \binom{k}{k'} r^{k'}(1-r)^{k-k'} & \text{for } k' = 0, \ldots, k \\ 0 & \text{for } k' = k+1, \ldots, N, \end{cases}$$
$$\text{for } k = 0, \ldots, N. \tag{7}$$

The transition matrix of the Markov chain can be written in the form of

$$\mathbf{P}_{on} = \begin{array}{c} \\ 1 \\ 2 \\ \vdots \\ N-1 \\ N \\ 0 \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & 2 & \cdots & N-1 & N & 0 \end{array} \\ \left(\begin{array}{ccccc|c} \star & 0 & \cdots & 0 & 0 & \star \\ \star & \star & \cdots & 0 & 0 & \star \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \star & \star & \cdots & \star & 0 & \star \\ \star & \star & \cdots & \star & \star & \star \\ \hline 0 & 0 & \cdots & 0 & 0 & 1 \end{array} \right) \end{array}$$

where the entries marked with a star can be found in (7). Note that state 0, which corresponds to a success by the primary user, is absorbing because once the primary user has a successful transmission, its transmissions in the following slots are not interrupted by the secondary users. Hence, collisions in an *on* period occur only before the primary user obtains a successful transmission. Also, the average number of collisions experienced by the primary user in an *on* period, denoted by T_{col}, is independent of the length of traffic, T_{pac}. If $q = 0$, then only idle slots arise in an *off* period, which leads to $T_{col} = 0$. If $q > 0$ and $r = 1$, then colliding secondary users do not back off and collisions are absorbing states in an *off* period, which leads to $T_{col} = +\infty$. If $(q,r) = (1,0)$, then an idle slot and a collision of all the secondary users alternate in an *off* period, which leads to $T_{col} = 1/2$. Suppose that $(q,r) \notin \mathcal{A}$. Let $\mathbf{Q}_{on}$ be the N-by-N matrix in the upper-left corner of $\mathbf{P}_{on}$. Since $r \neq 1$, the matrix $\mathbf{I} - \mathbf{Q}_{on}$ is invertible,

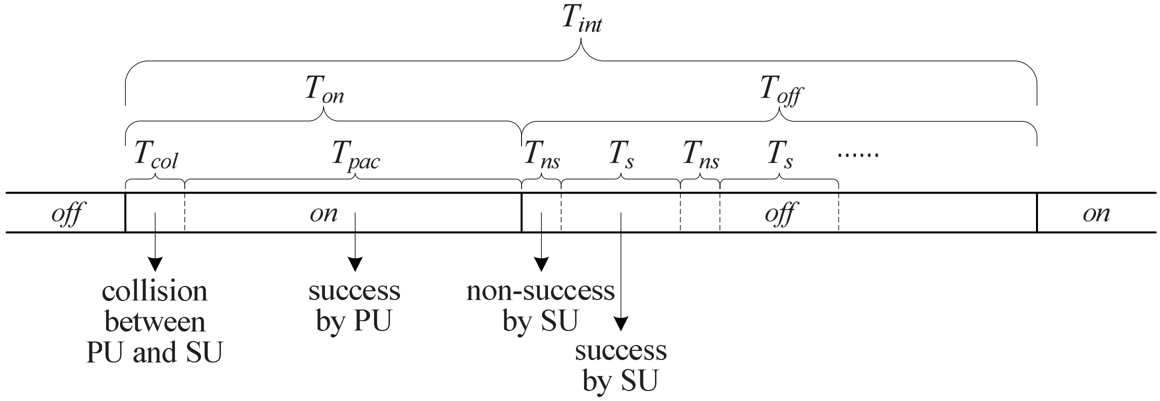


Fig. 1. Operation of the system under a θ-fair non-intrusive protocol.

and the average number of slots until the first success by the primary user starting from state k is given by the k-th entry of $(\mathbf{I}-\mathbf{Q}_{on})^{-1}\mathbf{e}$, for $k=1,\dots,N$.

Consider a slot t in which the state of the primary user has changed from *off* to *on*, i.e., $y_p^{t-1}=\mathit{off}$ and $y_p^t=\mathit{on}$. Then an *on* period begins from slot t. The number of collisions that the primary user expects to experience in the *on* period depends on the transmission outcome in slot $t-1$, the last slot of the preceding *off* period. Suppose that there was a collision among $k\geq 2$ secondary users in slot $t-1$. Then the Markov chain starts from state k in slot $t-1$. Since the *on* period starts in slot t, the number of collisions in the *on* period does not include the collision in slot $t-1$. Hence, the average number of collisions until the first success in an *on* period when the preceding *off* period ended with k transmissions is given by

$$d(k)=[(\mathbf{I}-\mathbf{Q}_{on})^{-1}\mathbf{e}]_k-1,$$

for $k=2,\dots,N$.

Suppose that there was a success in slot $t-1$. Then the successful secondary user transmits with probability $1-\theta$ while all the other secondary users wait in slot t. Thus, with probability θ, the primary user succeeds in slot t, and with probability $1-\theta$, state 1 occurs in slot t, from which it takes $[(\mathbf{I}-\mathbf{Q}_{on})^{-1}\mathbf{e}]_1$ collisions on average to reach a success by the primary user. Therefore, the average number of collisions until the first success in an *on* period when the preceding *off* period ended with a success is given by

$$\begin{aligned}d(1)&=\theta\cdot 0+(1-\theta)[(\mathbf{I}-\mathbf{Q}_{on})^{-1}\mathbf{e}]_1\\&=(1-\theta)[(\mathbf{I}-\mathbf{Q}_{on})^{-1}\mathbf{e}]_1.\end{aligned}\tag{8}$$

Suppose that slot $t-1$ was idle. Then with probability $\binom{N}{k}q^k(1-q)^{N-k}$, slot t contains transmission by k secondary users, for $k=0,\dots,N$. With probability $(1-q)^N$ the primary user experiences no collision, while with probability $\binom{N}{k}q^k(1-q)^{N-k}$ the *on* period begins with state k, for $k=1,\dots,N$. Therefore, the average number of collisions until the first success in an *on* period when the preceding *off* period ended with an idle slot is given by

$$\begin{aligned}&d(0)\\&=(1-q)^N\cdot 0+\sum_{k=1}^{N}\binom{N}{k}q^k(1-q)^{N-k}[(\mathbf{I}-\mathbf{Q}_{on})^{-1}\mathbf{e}]_k\\&=\sum_{k=1}^{N}\binom{N}{k}q^k(1-q)^{N-k}[(\mathbf{I}-\mathbf{Q}_{on})^{-1}\mathbf{e}]_k.\end{aligned}$$

The probability that the last slot of an *off* period has k transmissions is given by $w_{off}(k)$, for $k=0,1,\dots,N$. Hence, the average number of collisions that the primary user experiences before its first success in an *on* period is given by

$$T_{col}=\sum_{k=0}^{N}w_{off}(k)d(k).$$

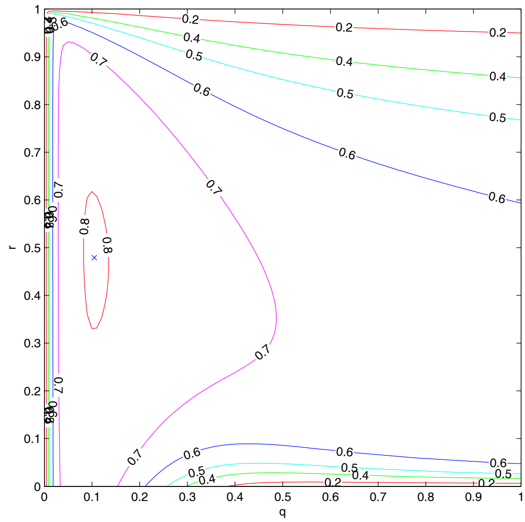
To sum up, the average number of collisions experienced by the primary user in an *on* period under a θ-fair non-intrusive protocol can be computed as

$$T_{col}(q,r)=\begin{cases}\sum_{k=0}^{N}w_{off}(k)d(k) & \text{if } (q,r)\notin\mathcal{A}\\ 0 & \text{if } q=0\\ +\infty & \text{if } q>0 \text{ and } r=1\\ \frac{1}{2} & \text{if } (q,r)=(1,0).\end{cases}\tag{9}$$

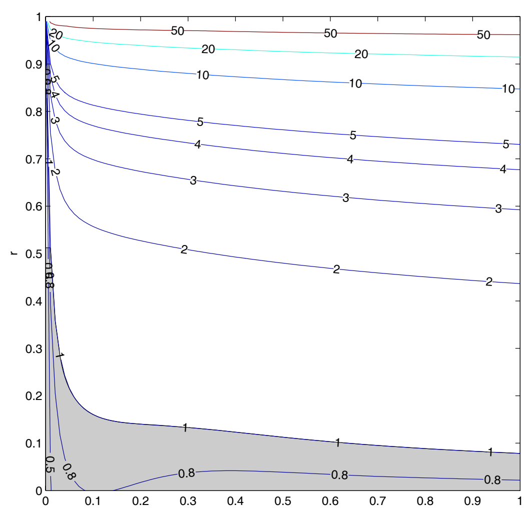
The collision probability of the primary user can be computed using the relationship $P_c=T_{col}/(T_{pac}+T_{col})$.

3) Channel Utilization Rate of the System: Let T_{on} and T_{off} be the average lengths (measured in slots) of an *on* period and an *off* period, respectively. Then the average time interval between two consecutive arrivals of traffic can be decomposed as $T_{int}=T_{on}+T_{off}$. A protocol determines the value of T_{col}, as shown in (9). We assume that the adopted protocol satisfies $T_{col}<T_{int}-T_{pac}$ in order to assure the stability of the system.[1] Since the primary user has either a successful transmission or a collision in an *on* period, T_{on} can be decomposed as $T_{on}=T_{pac}+T_{col}$, which leads to the relationship $T_{off}=T_{int}-T_{pac}-T_{col}$. The channel utilization rate of the primary user is given by $C_p=T_{pac}/T_{int}$, while that of the secondary users is $C_s=P_s\cdot T_{off}/T_{int}=P_s\cdot(T_{int}-T_{pac}-T_{col})/T_{int}$. Hence, the channel utilization rate of the system can be computed as $C=C_p+C_s=(T_{pac}+P_s\cdot T_{off})/T_{int}$. The operation of the system under a θ-fair non-intrusive protocol is summarized in Fig. 1.

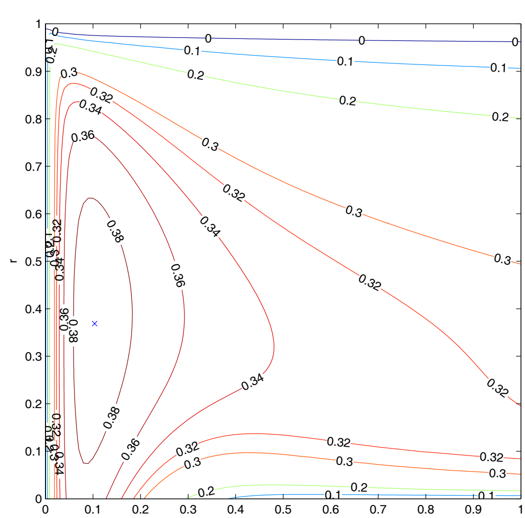
[1]This inequality holds for optimal protocols studied in Section IV.



(a) P_s



(b) T_{col}



(c) C_s

Fig. 2. Contour curves of P_s, T_{col}, and C_s as functions of (q, r) when $\theta = 0.1$, $N = 10$, $T_{int} = 100$, and $T_{pac} = 50$.

IV. Protocol Design Problem and Optimal Protocols

A. Formulation of the Protocol Design Problem

We formulate a problem solved by the protocol designer to determine a protocol. The protocol designer aims to maximize the channel utilization rate of the system while keeping the collision probability of the primary user below a certain threshold level specified as $\eta \in (0, 1)$. The protection level η can be considered as a requirement imposed by the primary user or by spectrum regulators. The protocol design problem can be formally expressed as

$$\max_{(q,r)\in[0,1]^2} C \text{ subject to } P_c \leq \eta.$$

Since T_{pac} and T_{int} are exogenous parameters independent of the prescribed protocol, the protocol design problem can be rewritten as

$$\max_{(q,r)\in[0,1]^2} C_s = P_s \frac{T_{int} - T_{pac} - T_{col}}{T_{int}} \text{ subject to } T_{col} \leq \gamma, \tag{10}$$

where $\gamma = (\eta/(1-\eta))T_{pac} > 0$ is the threshold level for T_{col}, derived from the relationship $P_c = T_{col}/(T_{pac}+T_{col})$ and the requirement $P_c \leq \eta$. We say that a protocol is optimal if it is an optimal solution to the protocol design problem (10). Note that T_{col} appears both in the objective function and in the constraint. The protocol designer prefers small T_{col} for two reasons. Smaller T_{col} implies less interference to the primary user and at the same time longer *off* periods that the secondary users can utilize.

B. Investigation into the Protocol Design Problem and Optimal Protocols

Using the expressions in (6) and (9), we can check that P_s is continuous and T_{col} is lower semi-continuous in (q, r) on $[0, 1]^2$. This implies that the objective function of the protocol design problem (10) is upper semi-continuous while the constraint set is compact. Therefore, there always exists an optimal protocol. Let C_s^o be the optimal value of the problem (10). Note that we can always find a protocol (q', r') that satisfies $T_{col} \leq \gamma$ and $T_{col} < T_{int} - T_{pac}$ by choosing q' and r' positive but sufficiently small. Also, since $(q', r') \in (0, 1)^2$, we have $P_s > 0$. This shows the existence of a protocol that yields $C_s > 0$ while satisfying the constraint $T_{col} \leq \gamma$. Therefore, we have $C_s^o > 0$. As γ approaches zero, q needs to be close to zero in order to satisfy the constraint, which yields P_s near zero. Thus, C_s^o converges to zero as γ approaches zero.

In Fig. 2, we show graphically the dependence of the performance metrics, P_s, T_{col}, and C_s, on the protocol (q, r). To obtain the results, we use parameters $\theta = 0.1$, $N = 10$, $T_{int} = 100$, and $T_{pac} = 50$. The maximum possible value of C_s is thus 0.5, while $T_s = 10$. Fig. 2(a) plots the contour curves of P_s. The success probability of the secondary users P_s is maximized at $(q, r) = (0.11, 0.48)$, and the maximum value of P_s is 0.804, which corresponds to the minimum value of T_{ns} as 2.44. The value of (q, r) that maximizes P_s can be explained as follows. Following an idle slot in an *off* period, every secondary user transmits with probability q, and thus the probability of success is maximized when

$q = 1/N$ [17]. During an *off* period, a collision cannot follow a success, and following an idle slot, a collision involving two transmissions is most likely among all kinds of collisions when $q \approx 1/N$. Since non-colliding users do not transmit following a collision under a non-intrusive protocol, the probability of success between two contending users is maximized when $r = 1/2$. r is chosen slightly smaller than $1/2$ because collisions involving more than two transmissions occur with small probability.

Fig. 2(b) plots the contour curves of T_{col}. As q and r are large, secondary users transmit aggressively in a contention period, intensifying interference to the primary user when it starts transmitting. Thus, we can see that T_{col} tends to increase with q and r. The set of protocols that satisfy the constraint $T_{col} \leq \gamma$ can be represented by the region below the contour curve of T_{col} at level γ. For example, the shaded area in Fig. 2(b) represents the constraint set corresponding to $T_{col} \leq 1$. Since $P_c = T_{col}/(T_{pac} + T_{col})$, P_c is monotonically increasing in T_{col}, and thus the contour curves of P_c have the same shape as those of T_{col}. Fig. 2(c) plots the contour curves of C_s. Let $(q^*, r^*) \in \arg\max_{(q,r)\in[0,1]^2} C_s$. That is, (q^*, r^*) is an optimal protocol when the constraint $T_{col} \leq \gamma$ is nonbinding.[2] The maximum of C_s is attained at $(q^*, r^*) = (0.10, 0.37)$, while the maximum value of C_s is 0.390. Since C_s is decreasing in T_{col}, which tends to increase with q and r, both transmission probabilities q^* and r^* are smaller than the corresponding ones in the protocol that maximizes P_s.

We can see from Fig. 2(b) that the constraint set is not convex. Also, Fig. 2(c) shows that there are upper contour sets of C_s that are not convex (e.g., $\{(q, r) \in [0, 1]^2 : C_s(q, r) \geq 0.34\}$), which implies that the objective function is not quasiconcave. Therefore, the protocol design problem (10) is not a convex optimization problem. This makes it difficult to establish analytically the uniqueness of an optimal solution and the convergence of an algorithm. Moreover, due to matrix inversion involved in computing P_s and T_{col}, it is difficult to obtain a closed-form expression for an optimal solution. As a result, we rely on a numerical method using MATLAB to obtain optimal protocols.[3] The structure of the problem is qualitatively the same for most values of parameters, and we can locate a unique global maximizer by using the graphs of the objective function and the constraint set, as explained below. This allows us to verify whether the protocol obtained by the numerical method is indeed the global maximizer, not just a local maximizer.

Fig. 3 shows the contour curves of C_s and T_{col} in the same graph as well as the locus of optimal protocols as γ varies. Depending on the value of γ, we can classify the protocol design problem into three cases. First, when γ is sufficiently large (more precisely, $\gamma \geq \gamma^*$, where γ^* is the value of T_{col} at (q^*, r^*)), the constraint is nonbinding and the optimal protocol, denoted by (q^o, r^o), is given by (q^*, r^*), independently of γ. Second, when γ is sufficiently small (i.e., $\gamma \leq \tilde{\gamma}$ for some $\tilde{\gamma} > 0$), the optimal protocol is at a corner

[2]A constraint is binding if its removal results in a strict improvement in the objective value and nonbinding otherwise.

[3]We use the function *fmincon*, which uses one of three constrained nonlinear optimization algorithms: active-set, interior-point, or trust-region-reflective.

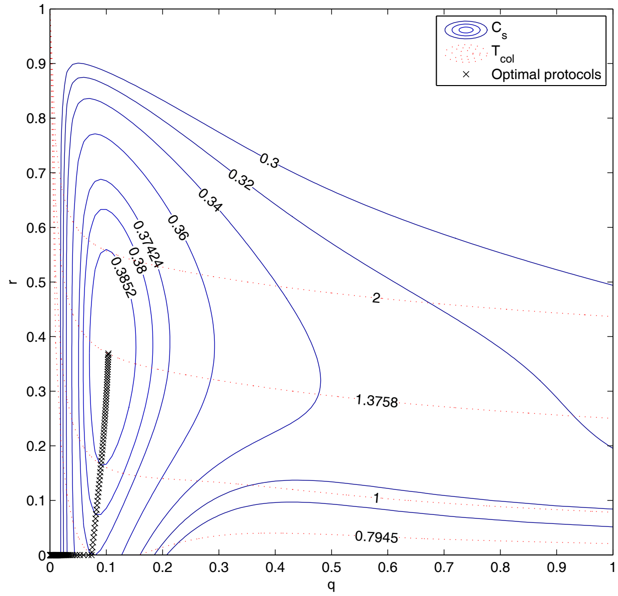


Fig. 3. Locus of optimal protocols as γ varies.

with $q^o > 0$ and $r^o = 0$, while q^o converges to zero as γ decreases to zero. Finally, when γ takes a medium value (i.e., $\tilde{\gamma} < \gamma < \gamma^*$), the optimal protocol is interior with $0 < q^o < q^*$ and $0 < r^o < r^*$. Both q^o and r^o decrease with γ, but due to the shape of the contour curves of C_s, r^o changes more rapidly than q^o does as γ varies in this region. In short, as a more stringent constraint is imposed on T_{col}, the protocol designer first reduces r^o mainly and then q^o when r^o cannot be reduced further.

Fig. 4(a) plots optimal protocols as γ varies between 0.01 and 2. With the considered parameters, the threshold values for γ are given by $\tilde{\gamma} = 0.80$ and $\gamma^* = 1.38$. Let T_{col}^o be the value of T_{col} at (q^o, r^o). Fig. 4(b) shows the values of T_{col}^o and C_s^o as γ varies. As expected, T_{col}^o is equal to γ when the constraint is binding (i.e., $\gamma < \gamma^*$) and remains at γ^* when the constraint is nonbinding (i.e., $\gamma \geq \gamma^*$). C_s^o is increasing with γ when the constraint is binding and remains the same when the constraint is nonbinding. The rate of change in C_s^o with respect to γ is smaller when $\tilde{\gamma} < \gamma < \gamma^*$ than when $\gamma < \tilde{\gamma}$. This implies that the optimal dual variable on the constraint $T_{col} \leq \gamma$ is small when $\tilde{\gamma} < \gamma < \gamma^*$ and is zero when $\gamma \geq \gamma^*$.

Fig. 4 also shows a comparison with the benchmark case of perfect spectrum sensing. When spectrum sensing is perfect, it is possible to require the secondary users to wait following a slot in which they detect the activity of the primary user. When this requirement is imposed, the primary user experiences at most one collision in an *on* period, and the average number of collisions T_{col} is reduced to $w_{off}(0)[1-(1-q)^N]+w_{off}(1)(1-\theta) + \sum_{k=2}^{N} w_{off}(k)[1-(1-r)^k]$ if $(q, r) \notin \mathcal{A}$ and 1 if $q > 0$ and $r = 1$ while remaining the same otherwise. The reduction in the number of collisions experienced by the primary user (in both average and maximum senses) can be considered as the benefit of perfect spectrum sensing. We can solve the protocol design problem in the case of perfect spectrum sensing by using the modified expression of T_{col}. We can see from Fig. 4 that the threshold value γ^* for a nonbinding

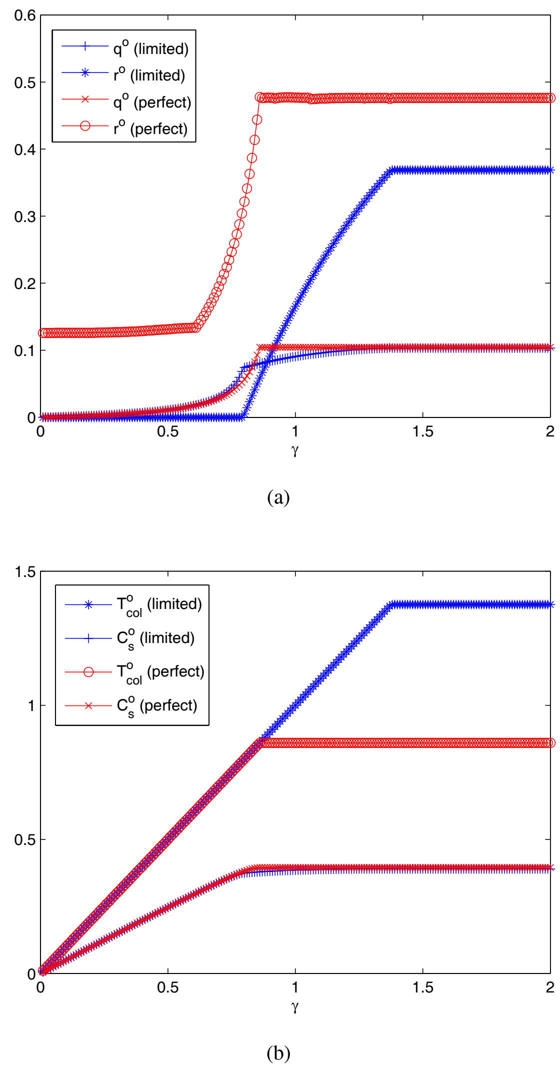


Fig. 4. Solution to the protocol design problem for γ between 0.01 and 2 when $\theta = 0.1$, $N = 10$, $T_{int} = 100$, and $T_{pac} = 50$ in the cases of limited and perfect spectrum sensing: (a) optimal protocols, and (b) the values of T_{col} and C_s at the optimal protocols.

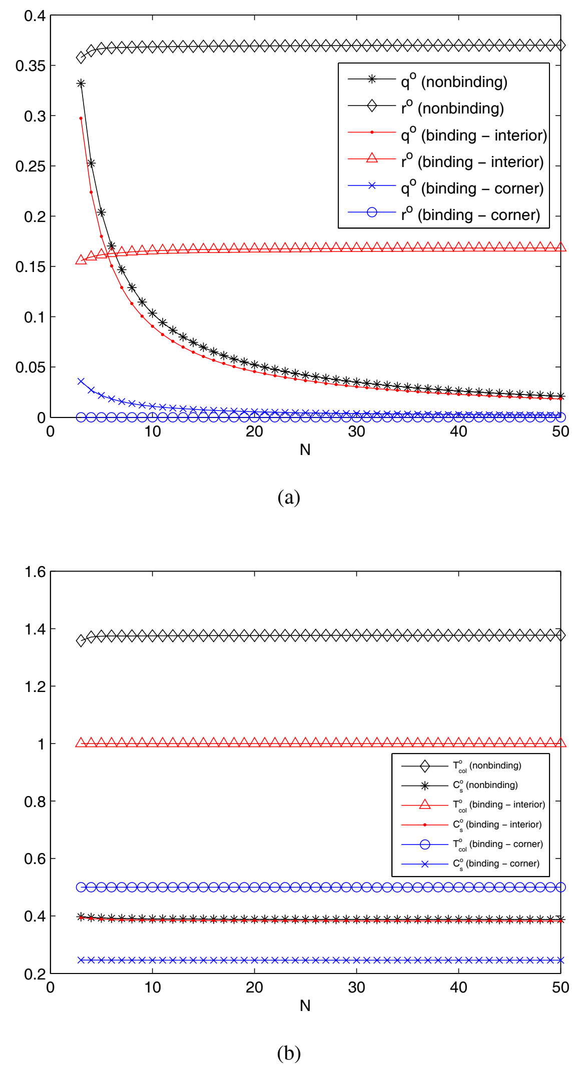


Fig. 5. Solution to the protocol design problem for N between 3 and 50 when $\theta = 0.1$, $T_{int} = 100$, and $T_{pac} = 50$ ($\gamma = 1$ for a binding constraint with an interior solution, and $\gamma = 0.5$ for a binding constraint with a corner solution): (a) optimal protocols, and (b) the values of T_{col} and C_s at the optimal protocols.

constraint is reduced to 0.86, compared to 1.38 in the case of limited spectrum sensing. Fig. 4(a) shows that, as γ decreases in the region $(0, \gamma^*)$, q^o approaches zero while r^o stabilizes to a positive value. Also, Fig. 4(b) shows that the optimal value C_s^o is not much reduced by perfect spectrum sensing. This result suggests that, in our formulation, limited spectrum sensing causes little performance loss in terms of channel utilization although perfect spectrum sensing can achieve less interference to the primary user.

C. Varying the Number of Secondary Users

We study how the solution to the protocol design problem changes with the number of secondary users. We vary N between 3 and 50 while fixing other parameters as before. We consider $\gamma = 0.5$ and 1 to illustrate the cases of a binding constraint with a corner solution and that with an interior solution, respectively. Fig. 5(a) shows optimal protocols (q^o, r^o), while Fig. 5(b) plots the values of T_{col}^o and C_s^o. First, consider the case of a nonbinding constraint. In this case, as N increases, q^o decreases at a diminishing rate while r^o increases slightly but remains almost constant. This is because q that maximizes P_s is close to $1/N$ while r that maximizes P_s is close to $1/2$ regardless of N. As N increases, T_{col}^o increases slightly and C_s^o decreases slightly, but both remain almost constant. This shows that the degree of contention increases with the number of secondary users but only slightly. Almost constant T_{col}^o shows that, even without a constraint on T_{col}, interruption to the primary user can be kept below a certain level. This is because under an optimal protocol the primary user is likely to contend with at most two secondary users when it starts transmitting, regardless of the total number of secondary users. Also, almost constant C_s can be interpreted as that optimal protocols are capable of resolving contention among the secondary users efficiently even when there are many secondary users sharing the channel. Next, consider the case of a binding constraint with an interior solution. As mentioned in Section IV-B, imposing such a constraint reduces

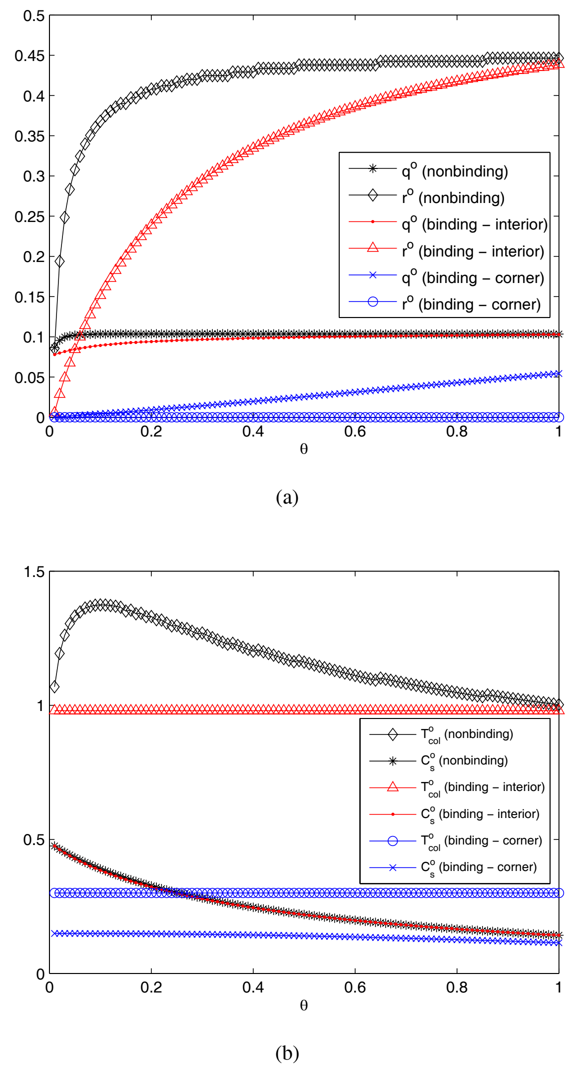


Fig. 6. Solution to the protocol design problem for θ between 0.01 and 1 when $N = 10$, $T_{int} = 100$, and $T_{pac} = 50$ ($\gamma = 0.98$ for a binding constraint with an interior solution, and $\gamma = 0.3$ for a binding constraint with a corner solution): (a) optimal protocols, and (b) the values of T_{col} and C_s at the optimal protocols.

the values of q^o and r^o, but it impacts r^o more than q^o. Also, it reduces C_s^o very slightly, thus preserving the property of C_s^o being almost constant. Lastly, consider the case of a binding constraint with a corner solution. In this case, both q^o and C_s^o are reduced significantly compared to the other two cases in order to meet the stringent constraint. However, C_s^o is almost constant, as in the other two cases.

D. Varying the Fairness Level

We investigate the impact of the fairness level on optimal protocols and their performance. We vary θ from 0.01 to 1 while fixing other parameters at $N = 10$, $T_{int} = 100$, and $T_{pac} = 50$. Here we use $\gamma = 0.3$ and 0.98 to illustrate the cases of a binding constraint with a corner solution and that with an interior solution, respectively. Fig. 6(a) shows optimal protocols (q^o, r^o), while Fig. 6(b) plots the values of T_{col}^o and C_s^o. First, consider the case of a nonbinding constraint. As θ increases, q^o stabilizes around 0.10 quickly while r^o increases at a diminishing rate. When θ is small, a success period lasts long, and thus r can be chosen small to limit T_{col} without affecting C_s much. On the contrary, when θ is large, contention periods occur frequently, and thus r is chosen close to $1/2$ in order to resolve contention among the secondary users quickly. As θ increases, T_{col}^o increases, reaches a peak at $\theta = 0.1$, and then decreases, whereas C_s^o decreases monotonically. The negative relationship between C_s and θ can be interpreted as a trade-off between channel utilization and short-term fairness. Next, consider the case of a binding constraint with an interior solution. As before, imposing such a constraint affects r^o more than q^o while it reduces C_s^o only slightly. Lastly, consider the case of a binding constraint with a corner solution. As θ increases, q^o increases while r^o is fixed at zero. Also, C_s^o decreases with θ but is almost constant. When θ is close to zero, $d(1)$ is close to one. In order to achieve a small value of T_{col}, we need small q to keep $w_{\mathit{off}}(1)$ small. In this case, the trade-off between channel utilization and short-term fairness is not as severe as in the other two cases, because the stringent constraint limits the frequency of success slots in an *off* period.

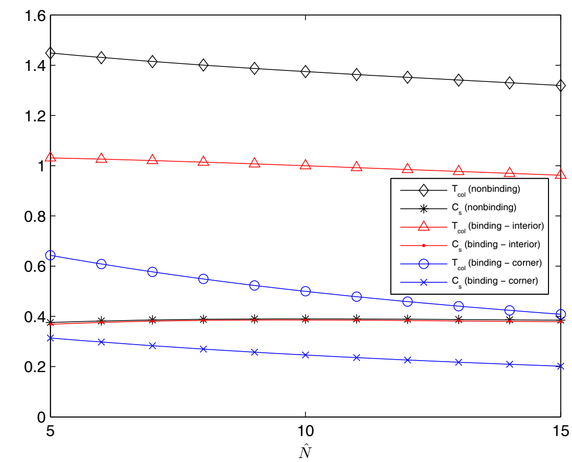


Fig. 7. Values of T_{col} and C_s for $\hat{N}$ between 5 and 15 when $\theta = 0.1$, $N = 10$, $T_{int} = 100$, and $T_{pac} = 50$ ($\gamma = 1$ for a binding constraint with an interior solution, and $\gamma = 0.5$ for a binding constraint with a corner solution).

E. Estimated Number of Secondary Users

Suppose that the protocol designer solves the protocol design problem for each possible N and prescribes the obtained protocols for the secondary users as a function of N. So far we have assumed that the secondary users know the exact number of secondary users sharing the channel so that they can adopt the correct optimal protocol. Here we consider an alternative scenario where the secondary users choose an optimal protocol based on their (possibly incorrect) estimates of the number of secondary users. For simplicity, we assume that all the secondary users have the same estimate. We consider $N = 10$ and the estimated number of secondary users, denoted by $\hat{N}$, between 5 and 15 while fixing other parameters at $\theta = 0.1$, $T_{int} = 100$, and $T_{pac} = 50$. In Fig. 7, we plot the values of T_{col} and C_s when the N secondary users follow the optimal protocol computed assuming $\hat{N}$ secondary users. As in Section IV-C, we consider $\gamma = 0.5$ and 1 to illustrate the cases of a binding constraint with a corner solution and that with an interior solution, respectively. We have seen in Fig. 5(a) that q^o decreases with the number of secondary users while r^o is almost constant. As a result, the overall interference level from the secondary users reduces as $\hat{N}$ increases, and thus T_{col} decreases with $\hat{N}$. In the cases of a nonbinding constraint or a binding constraint with an interior solution, C_s is not affected much by $\hat{N}$, reaching a peak at $\hat{N} = N$. In contrast, in the case of a binding constraint with a corner solution, C_s decreases with $\hat{N}$. These results suggest that the channel utilization rate of the secondary users is robust to errors in the estimation of the number of secondary users as long as the optimal solution remains interior. Note that, in the case of a binding constraint, the constraint is violated when an underestimation occurs, i.e., $\hat{N} < N$. In order to avoid this, the protocol designer may prescribe an estimation procedure that is biased toward overestimation, or specify a smaller γ than the required threshold. An estimation procedure is also needed in a scenario where the number of secondary users that have packets to transmit varies over time. For example, we can consider an estimation procedure under which a secondary user updates its estimate by comparing its expected throughput and its actual throughput over a certain period.

V. Enhancement Using Longer Memory

So far we have focused on protocols with one-slot memory for their simplicity in terms of implementation and analysis. In this section, we explain how longer memory can enhance protocols by reducing the average number of collisions and bounding the maximum number of collisions experienced by the primary user in an *on* period. Let p_i^τ be the transmission probability of secondary user i in slot τ. A protocol with B-slot memory that enhances a θ-fair non-intrusive protocol f is described below.

(P1) If $y_i^{t-2} = success$ and $y_i^{t-1} = failure$, then $p_i^t = 0$.
(P2) If $y_i^{t-B} = \cdots = y_i^{t-1} = failure$, then $p_i^t = 0$.
(P3) Otherwise, $p_i^t = f(y_i^{t-1})$.

(P1) requires that a secondary user that experiences a collision following a success back off. Note that a collision cannot follow a success in an *off* period due to non-intrusiveness, and thus (P1) does not affect the performance in an *off* period. The only possible occasion in which a collision follows a success is when the primary user starts transmitting. Therefore, if a secondary user experiences a collision following a success, it can infer that an *on* period has started. According to a θ-fair non-intrusive protocol, a secondary user transmits with probability r after a collision, which yields $d(1) = (1-\theta)[(\mathbf{I} - \mathbf{Q}_{on})^{-1}\mathbf{e}]_1$ as shown in (8). By imposing (P1), we can reduce $d(1)$ to $1-\theta$, which in turn reduces T_{col}. For example, with $\theta = 0.1$, $N = 10$, $T_{int} = 100$, $T_{pac} = 50$, and $(q, r) = (q^*, r^*) = (0.10, 0.37)$, (P1) reduces $d(1)$ from 1.426 to 0.9 and T_{col} from 1.376 to 0.954.

In the range of parameter values considered in Section IV, the average number of collisions experienced by the primary user in an *on* period is reasonably small, not exceeding 1.5 slots, even without a constraint imposed on it. However, as colliding secondary users transmit with probability $r > 0$, the realized number of collisions in an *on* period can be arbitrarily large with positive probability. That is, the worst-case number of collisions in an *on* period is unbounded under a θ-fair non-intrusive protocol. We can bound the maximum number of collisions in an *on* period by imposing (P2), which requires a secondary user that experiences B consecutive collisions to back off. Since non-colliding secondary users wait after a collision, colliding secondary users must have the same number of consecutive collisions in any slot. Thus, secondary users experiencing B consecutive collisions back off simultaneously, yielding a slot that can be utilized by the primary user. Therefore, the primary user cannot experience more than B collisions in an *on* period. When B is chosen moderately large, B consecutive collisions rarely occur in an *off* period, and thus (P2) has a negligible impact on the success probability of the secondary users P_s while it reduces T_{col}. (P2) can be considered as a safety device to limit the number of collisions that the primary user can experience during an *on* period.

VI. Conclusion

In this paper, we have considered a scenario in which a primary user shares a channel with secondary users that cannot distinguish the signals of the primary user from those of a secondary user. We have shown that a class of distributed MAC protocols can coordinate access among the secondary users while restricting interference to the primary user, thereby overcoming the limited sensing ability of the secondary users at the PHY layer. The basic ideas underlying the proposed protocols can be exploited to build more sophisticated protocols. For example, in a CSMA/CA network, protocols with memory can be used to adjust the back-off parameters of secondary users based on their own transmission results and obtained channel information. Also, we can provide quality-of-service differentiation to secondary users by specifying different protocols for secondary users. Lastly, in a multipacket reception scenario with dynamic channel conditions, the proposed protocols can be extended so that secondary users adjust their transmission probabilities to their channel conditions as well.

References

[1] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: a survey," *Computer Networks*, vol. 50, no. 13, pp. 2127–2159, Sep. 2006.

[2] C.-T. Chou, S. Shankar, H. Kim, and K. G. Shin, "What and how much to gain by spectrum agility?" *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 576–588, Apr. 2007.

[3] A. R. Fattahi, F. Fu, M. van der Schaar, and F. Paganini, "Mechanism-based resource allocation for multimedia transmission over spectrum agile wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 601–612, Apr. 2007.

[4] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: a POMDP framework," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 589–600, Apr. 2007.

[5] J. Park and M. van der Schaar, "Medium access control protocols with memory," *IEEE/ACM Trans. Netw.*, vol. 18, no. 6, pp. 1921–1934, Dec. 2010.

[6] FCC, ET Docket No 03-222, "Notice of proposed rule making and order," Dec. 2003.

[7] S. Huang, X. Liu, and Z. Ding, "Opportunistic spectrum access in cognitive radio networks," in *Proc. IEEE INFOCOM*, 2008, pp. 1427–1435.
[8] H. Kim and K. G. Shin, "Efficient discovery of spectrum opportunities with MAC-layer sensing in cognitive radio networks," *IEEE Trans. Mobile Computing*, vol. 7, no. 5, pp. 533–545, May 2008.
[9] L. Law, J. Huang, M. Liu, and S. Li, "Price of anarchy for cognitive MAC games," in *Proc. IEEE GLOBECOM*, 2009.
[10] H. Wang, H. Qin, and L. Zhu, "A survey on MAC protocols for opportunistic spectrum access in cognitive radio networks," in *Proc. Int. Conf. Comput. Sci. Software Eng.*, 2008, pp. 214–218.
[11] T. V. Krishna and A. Dasa, "A survey on MAC protocols in OSA networks," *Computer Networks*, vol. 53, no. 9, pp. 1377–1394, June 2009.
[12] C. Cormio and K. R. Chowdhury, "A survey on MAC protocols for cognitive radio networks," *Ad Hoc Networks*, vol. 7, no. 7, pp. 1315–1329, Sep. 2009.
[13] D. Bertsekas and R. Gallager, *Data Networks*, 2nd edition. Prentice Hall, 1992.
[14] A. H. Mohsenian-Rad, J. Huang, M. Chiang, and V. W. S. Wong, "Utility-optimal random access without message passing," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1073–1079, Mar. 2009.
[15] R. T. Ma, V. Misra, and D. Rubenstein, "An analysis of generalized slotted-ALOHA protocols," *IEEE/ACM Trans. Netw.*, vol. 17, no. 3, pp. 936–949, June 2009.
[16] C. M. Grinstead and J. L. Snell, *Introduction to Probability*. AMS, 1997.
[17] J. L. Massey and P. Mathys, "The collision channel without feedback," *IEEE Trans. Inf. Theory*, vol. 31, no. 2, pp. 192–204, Mar. 1985.

Jaeok Park received the B.A. degree in economics from Yonsei University, Seoul, Korea, in 2003, and the M.A. and Ph.D. degrees in economics from the University of California, Los Angeles, in 2005 and 2009, respectively. He is currently a Postdoctoral Scholar in the Electrical Engineering Department at the University of California, Los Angeles. From 2006 to 2008, he served in the Republic of Korea Army. His primary research interests include game theory, mechanism design, network economics, and wireless communication.

Mihaela van der Schaar (F'10) is a Professor in the Electrical Engineering Department at the University of California, Los Angeles. Her research interests include multimedia networking, communication, processing, and systems, dynamic multi-user networks and system designs, online learning, energy-efficient designs, network economics, and game theory.

Prof. van der Schaar is an IEEE Fellow, a Distinguished Lecturer of the Communications Society for 2011–2012, the Editor-in-Chief of the IEEE TRANSACTIONS ON MULTIMEDIA and a member of the editorial board of the IEEE JOURNAL ON SELECTED TOPICS IN SIGNAL PROCESSING. She received an NSF CAREER Award (2004), the Best Paper Award from the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY (2005), the Okawa Foundation Award (2006), the IBM Faculty Award (2005, 2007, 2008), the Most Cited Paper Award from the *EURASIP: Image Communications Journal* (2006), the GameNets Conference Best Paper Award (2011), and the 2011 IEEE Circuits and Systems Society Darlington Best Paper Award. She received three ISO awards for her contributions to MPEG video compression and streaming international standardization activities, and holds 33 granted US patents.