Determining the Transmission Strategy of Cognitive User in IEEE 802.11 based Networks

Rukhsana Ruby*, Victor C.M. Leung*, John Sydor†
*Department of Electrical and Computer Engineering,
The University of British Columbia,
Vancouver, BC, Canada
†Communication Research Centre (CRC),
Ottawa, Ontario, Canada

***Abstract*—Cognitive radio methodologies have the potential to dramatically increase the throughput of wireless systems. we consider the opportunistic channel access scheme in IEEE 802.11 based networks subject to the interference mitigation strategy where primary and secondary users can superimpose their transmissions on the same time or frequency slots. According to the protocol rule, primary user follows backoff counter based DCF channel access scheme in contrast to traditional retransmission based channel access scheme where secondary user can easily get to know which slot is idle and in which slot a transmission is being scheduled. In this work, we propose an opportunistic channel access procedure where secondary user intelligently picks a backoff counter from a given contention window or stays idle on completion of a transmission. First problem has been derived as a linear program formulation from the markov model of both primary and secondary users. From the insights of problem formulation and structure of the problem, an algorithm has been derived for the strategy of the secondary user. Through numerical calculation, validity of the algorithm has been proven. Later, an online algorithm has been proposed based on reinforcement learning technique. Validity of this algorithm has also been verified through simulation and finally these two algorthm have been compared in terms of throughput given the constraint of primary user's throughput loss and failure probability.**

I. Introduction

Cognitive radio has been the subject of intense research because of its potential to increase the efficiency of wireless networks. Traditional concept of cognitive radio is that the unlicensed secondary users opportunistically access the licensed band while primary user keeps silent. This strategy is called white space approach. There are several works have been conducted based on zero interference rationale [1]–[3]. With this, secondary users sense the channel in order to detect time/frequency slots left unused by the primary users and exploit them for transmission. Main goal of this approach is to not interfere with the primary user at all. But, due to the sensing errors, collision with the primary users is inevitable and thus degrades the throughput of the primary user. The primary user being dumb, after collision, it tries to send packet in the next slot as long as the packet transmission index does not exceed the retry limit. We consider IEEE 802.11 based networks and primary user follows DCF protocol to access the channel. In this work, we have proposed an intelligent strategy for the secondary user who picks a backoff counter for packet transmission or remains idle.

in addition, transmission technique we have considered interference mitigation scheme. This idea has been evolved because of the concept of multiple input/output, where simultaneous transmission of a number of users may enhance the throughput instead of each individual transmission by cancelling the interference somehow. There are some prior literature investigating the coexistence in the same time/frequency band with a focus on physical layer methods for static scenarios [4]–[7]. There is one work similar to our work [8]. They also followed the retransmission based error control scheme as the specification of DCF protocol. Unlike our work, it assumes time is divided into slots and each slot corresponds to one single packet transmission time. Therefore, at the beginning of slot, the secondary user knows whether it is idle or occupied; if occupied, packet's transmission index can also be determined. Their optimal strategy has been based on the state of the slot and thus pretty much straightforward.

However, according to the DCF protocol specification, primary user has to go through DIFS period and then backoff period before flushing the packet into the air. While staying in the DIFS or backoff period, there is no way for the secondary user to know whether the primary user has a packet in a queue or not. Given this situation, secondary user has to determine intelligently a backoff counter or the decision of being idle. In this work, we determine the optimal access control policy for the secondary users in IEEE 802.11 based networks where nodes follow the DCF protocol in order to access the channel. We focus on a network with two mutually interfering links, one primary and one secondary.

We study the interference that the secondary user causes to the primary user and how this interference impacts on the latter. Activity of the primary user is also affected by the secondary user's channel access scheme. Our analysis is based on detailed markov model of both primary and secondary users i.e. secondary user's transmission affects both backoff and retransmission mechanism. Backoff procedure is halted if the secondary user transmits in a particular slot, transmitted packet might be corrupted due to link error or for the collision with the secondary user. Secondary user's backoff procedure and probability of successful transmission are also affected by the primary user's transmission. An accurate stochastic model should capture these behaviors of primary and secondary users nicely. We have developed two markov models in order to detail the activity of both.

In this framework, due to activity of the secondary user affects the steady state distribution of the primary user and thus it affects the achievable throughput. In the similar manner, secondary user is also affected by the primary user. Therefore, constraint on the maximum throughput loss or failure probability controls the achievable throughput of the primary user. The optimization problem can be formalized through a linear program. According to DCF protocol, channel access scheme of the primary user is randomized, optimal strategy of the secondary user is randomized i.e. in the beginning of a slot, with some probability, either it stays idle or picks a backoff counter from the given contentin window. As we do not explicitly know the state of the primary user, the optimal policy is random given the maximum throughput loss and failure probability of the primary user.

The problem though conceptually simple, unveils important issue and general behaviors. As the primary user implemnets a retransmission-based error control mechanism, the activity of the secondary user biases the transmission process via interference. Interference at the primary receiver increases the failure probability of primary user's transmission. Therefore, due to activity of the secondary user the average number of transmissions of primary user's packet gets larger, together with the average time required to return the the primary user's idle state. Interestingly, the increase of the avg number of transmissions of primary user's packets depends on the index of the interfered transmission. For instance, while interference from the secondary user in the first transmission of the of primary user's packets depends on the index of the interfered transmission. For instance, interference from the in the first transmission of the primary user's packets potentially leads to a significant increase of the number of transmission. thus as observed before, the impact of the secondary user's transmissions in the various states critically depend on the state of the primary network.

As this approach assumes that the secondary transmitter has some knowledge of the current state and probabilistic model of the primary transmitter/receiver pair, limiting its applicability. For example, while it is likely that the secondary might read ACKs for the primary system, it is unlikely that the secondary will have knowledge of the pending workload of packets at the primary transmitter or will know the distribution of packet arrivals at the primary transmitter. Therefore, we address this limitation by developing an on line learning approach that uses one feedback bit sent by the primary user and that approximately converges to the optimal secondary control policy. We will show that when the secondary user has access to such tiny knowledge, an online algorithm can obtain performance similar to an offline algorithm with some state information.

Rest of the paper is organized as follows, section II illustrates system model of the network, section III explains the detailed optimization problem and the corresponding strategy picking algorithm. Formulation of the online solution is

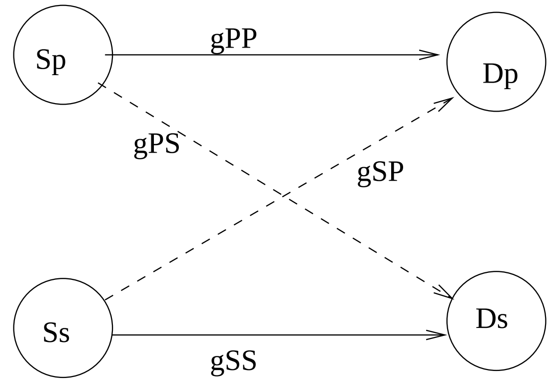

Fig. 1. System Model

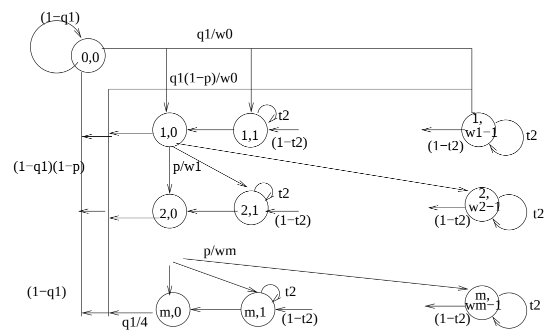

Fig. 2. Markov Model of Primary User

depicted in the section IV. Results obtained from numerical calculation have been shown in section V in order to verify the efficacy of the algorithm. Finally section VI concludes the paper.

II. System Model

Consider the network in figure 1 with a primary and secondary source, namely S_P and S_S. The primary source S_P and the secondary source S_S transmit packets to their respective destinations, namely D_P and D_S. We consider interference mitigation scenario which assumes primary/secondary user decodes the transmitted packet when it is coincided with other's transmission with some probability.

We assume a quasi static channel, and time is divided into slots. Before initiating a packet transmission, both users first undergo DIFS period and decrements the backoff counter which is as large as each single time slot. While decrementing

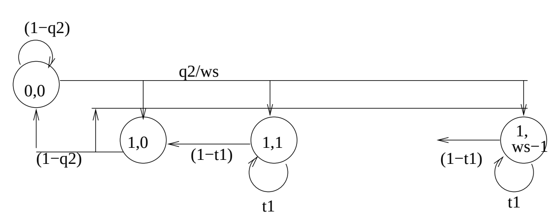

Fig. 3. Markov Model of Secondary User

backoff counter, if the station detects a busy channel, it halts its decrementing process and resumes until it detects idle channel for the time of DIFS period. When the counter reaches to zero, packet is flushed out into the air. Ideally, packet transmission time is variable, but in this work for the sake simplicity, it is multiple of some slots. We denote by g_{PP}, g_{PS}, g_{SS} and g_{SP}, the random variables corresponding to the channel coefficients respectively between S_P and D_P; S_P and D_S; S_S and D_S; S_S and D_P with $\zeta_{PP}(g)$, $\zeta_{PS}(g)$, $\zeta_{SS}(g)$ and $\zeta_{SP}(g)$ their respective probability distribution. The average decoding failure probability at the primary destination D_P associated with a silent secondary source is denoted by $\rho > 0$, while the same probability when the secondary source transmits is $\rho^* > \rho$. Analogously, the average decoding failure probability at the secondary destination D_S when the primary source is silent and transmitting is denoted with $\nu > 0$ and $\nu^* > \nu$ respectively.

Denoting the transmission rate and power of the primary and secondary sources with R_P, P_P, R_S, P_S respectively, we obtain the following failure probabilities for the primary link

$$\rho \ = \ Prob\{R_P \ > \ C(g_{PP}P_P)\}$$

$$\rho^* \ = \ Prob\{R_P \ > \ C(\frac{g_{PP}P_P}{1+g_{SP}P_S})\}$$

where $C(x) \ = \ log(1+x)$
For the secondary link, we obtain

$$\nu \ = \ Prob\{R_S \ > \ C(g_{SS}P_S)\}$$

$$\nu^* \ = \ Prob\{R_S \ > \ C(\frac{g_{SS}P_S}{1+g_{PS}P_P})\}$$

We do not consider a specific physical layer architecture or transmission technique, but rather, we refer to the simple construction based on the average decoding probabilities described before. The primary source S_P accesses the channel in each slot to transmit a fresh packet with fixed probability α, with $0 \ < \ \alpha \ < \ 1$. The secondary source is assumed to be backlogged i.e. it always has a packet to transmit.

The channel access strategy of the secondary source follows a policy κ, where κ is a vector with the length of backoff window size added by one. First element $\kappa(1,0)$ of this vector represents the proportion of time slots the secondary source keeps idle. With $(1-\kappa(1,0))$ probability, the secondary user accesses the channel, however picks a backoff counter with some probability and rest of the elements of vector κ represents this probability. For example, $\kappa(1,j)$ represents the probability of picking backoff counter $j-1$ and according to the fundamental property of probability theory, $\sum_{j=1}^{w_s} \kappa(1,j) \ = \ 1$, where w_s is the contention window size of secondary user.

We have formulated the optimization problem for two types of constraint i.e. throughput loss and failure probability of primary user. Subsection III-A is for the detailed formulation for the first constraint and subsection III-B is for later constraint. The goal of the secondary source is to maximize its own achieved throughput while limiting performance loss to the primary user.

In particular, let us denote as κ_0 the policy by which the secondary source never transmits. The optimization problem can be written as the following infinite horizon constrained markov decision process.

$$\begin{aligned} \hat{\kappa} \ &= \ \arg\max_{\kappa} W_S(\kappa) \\ s.t. \Im_P(\kappa_0) - \Im_P(\kappa) \ &<= \ \gamma \end{aligned}$$

$W_S(\kappa)$ is the secondary source throughput while following policy κ, $\Im_P(\kappa)$ is primary source performance when secondary source follows the policy κ and $\Im_P(\kappa_0)$ is the maximum achievable performance when secondary source does not transmit at all.

III. Markov Chain and Formulation of the Optimization Problem

The state of the network can be modeled as a homogeneous markov process. Two parameters (backoff stage, counter value) referred to as (b, c) describe the state of a user, where c can take any value between 0 and $w_b - 1$. Backoff stage b varies from 1 to maximum backoff stage m. Here, m is the maximum retry limit. Having a transmission failure, each packet is attempted for retransmission at most m numbers. At each backoff stage, if a station reaches state (b, 0) (i.e. backoff counter value becomes 0), the station will send out a packet. If the transmission failure occurs at this point with some probability, the primary user moves to higher backoff stage (b+1, c) with probability $\frac{1}{w_{b+1}}$. If successful packet transmission happens, the primary user goes to idle state (0, 0) (if there is no outstanding packet in queue) or in the initial backoff stage having picked some backoff counter with the probability of $\frac{1}{w_1}$. Markov chain model of primary user has been illustrated in figure 2. Also, packet arrival process in the primary user is distributed as a poisson process with parameter λ_1. Secondary user tries each packet only once, after having transmission, it goes to idle state with some probability or picks a backoff counter j with probability $\kappa(1, j+1)$ for the transmission of new packet from the queue. Note that, secondary user's packet is assumed as backlogged or there is always one packet in the queue. However, in order to meet the performance loss constraint of primary user, secondary user needs to keep silent and therefore we have introduced a fake variable λ_2 i.e. secondary user's packet arrival rate. Markov chain model for the secondary user has been shown in figure 3. First we would like to derive the achieved throughput of primary user. Detailed state transition probabilities are given below:

1) The backoff counter decrements, and the station makes a transition from state (b, c) to state (b, c - 1) when

the medium is idle. Denote transmission probability of secondary user is t_2.

$$Prob\{(b, c-1)|(b, c)\} \; = \; 1 \; - \; t_2$$

2) The backoff counter suspends and the user stays in state (b, c) when the secondary user transmits

$$Prob\{(b, c)|(b, c)\} \; = \; t_2$$

3) The user sends a packet and the packet fails, the user reaches state (b + 1, c).

$$Prob\{(b{+}1, c)|(b, 0)\} \; = \; \frac{t_2^*}{w_{b+1}} \quad b \; = \; 1 \cdots m, \quad c \; = \; 1 \cdots w_{b+1}-1$$

4) At last backoff stage m, when packet transmission fails, user goes to either idle (0, 0) or to the first backoff stage

$$Prob\{(0, 0)|(m, 0)\} \; = \; (1 - q_1)$$

$$Prob\{(1, c)|(m, 0)\} \; = \; \frac{q_1}{w_1}$$

5) Successful transmission happens at the stages other than the last backoff stage, transition probabilities are as follows:

$$Prob\{(1, c)|(b, 0)\} \; = \; \frac{q_1}{w_1}$$

$$Prob\{(0, 0)|(b, 0)\} \; = \; (1 - t_2^*)(1 - q_1)$$

6) User is in idle state. A packet has arrived and it goes to (1,c) state.

$$Prob\{(1, c)|(0, 0)\} \; = \; \frac{\lambda_1}{w_1}$$

Let us denote $\pi_P(b, c)$ is the steady state probability that the user is in state (b, c). From the markov chain model, we see that, when the user is in the state (b,0), the packet is sent out. Under the condition that there is at least one packet to send, the user has probability t_1 to send a packet in any time slot

$$t_1 \; = \; \sum_{b=1}^{m} \pi_P(b, 0)$$

Assuming the transmission probability of secondary user in any particular slot is t_2, transmission failure probability of the primary user is given by

$$t_2^* \; = \; (1 - t_2)\rho + t_2\rho^*$$

Success probability of any particular packet is denoted by

$$Prob_{succ} \; = \; (1 - \pi_P(0, 0))t_1(1 - t_2^*)$$

As for the probability q_1, let us denote the average access delay, the time from when a packet reaches the MAC layer to when it is successfully sent, D_{access}. D_{access} is also the packet service time if we treat each station as a M/M/1/N queue system [9], in which N is the maximum queue length of the link layer queue. For a M/M/1/N queue, the probability that there is a packet in the queue is,

$$q_1 \; = \; max(\lambda_1 D_{access}, 1)$$

For computing D_{access}, as mentioned before, it is the time between when a packet is attempted to transmit until a successful transmission happens. In our analysis, we assume, successful transmission time of the packet is as same as the failure time. If transmission failure happens, the user goes to the next backoff stage, and before another transmission it goes to a number of backoff slots. Backoff counter is picked up uniformly from the backoff window w_b of backoff stage b and therefore, the average number backoff slots passed by the user is $(w_b - 1)/2$. Denoting the transmission time of the user is T_t.

$$D_{access} \; = \; (w_1 - 1)/2 + T_t + \sum_{b=2}^{m}((w_b - 1)/2 + T_t)$$

$Prob_{bkoff}$ is the probability that any slot is backoff slot and given by

$$Prob_{bkoff} \; = \; 1 \; - \; t_1 \; - \; \pi_P(0, 0)$$

As the backoff slot and transmission slot time are different, generic slot time is defined as

$$T_P(slot) \; = \; Prob_{bkoff}\sigma + t_1 T_t + Prob_{bkoff} t_2 T_t$$

Here σ is denoted as backoff time slot and T_t is transmission time slot. Throughput of the primary user then will be given by

$$W_P \; = \; \frac{Prob_{succ}}{T_P(slot)} \qquad (1)$$

Now, we would like to derive the throughput of secondary user. Steady state distribution of state (b,c) is $\pi_S(b, c)$. Transition probabilities of the secondary user are described below:

1) If the secondary user is in backoff state (1,c+1) and the solt is assumed to be idle i.e. the primary user does not transmit.

$$Prob\{(1, c)|(1, c+1)\} \; = \; 1 - t_1$$

2) Backoff slot decrementing procedure is suspended if the primary user transmits.

$$Prob\{(1, c)|(1, c)\} \; = \; t_1$$

3) No matter the packet is successfully transmitted or not, the secondary user either goes to idle state (0,0) or starts transmission of a new packet.

$$Prob\{(0, 0)|(1, 0)\} \; = \; 1 \; - \; q_2$$

$$Prob\{(1, c)|(1, 0)\} \; = \; q_2\kappa_{1,c}$$

4) When the secondary user is in idle state (0,0), it picks the backoff counter for transmission

$$Prob\{(1,c)|(0,0)\} \;=\; q_2\kappa_{1,c}$$

Computation of q_2 is similar to q_1 using λ_2 and has been skipped here.

Transmission probability of secondary user is computed similar to the primary user

$$t_2 \;=\; \pi_S(1,0)$$

Generic slot time of the secondary user can be indicated by

$$T_S(slot) \;=\; t_2T_t + (1-t_2)\sigma + (1-t_2)t_1T_t$$

Throughput of the secondary user can thus be

$$W_S \;=\; \frac{t_2}{T_S(slot)} \tag{2}$$

The optimization problem is equivalent to the following linear program

$$\begin{aligned}\hat{\kappa} \;&=\; \arg\max_{\kappa} W_S\\ s.t. W_P(\kappa_0) - W_P(\kappa) \;&<=\; \gamma_1\\ \sum_{i=1}^{w_s}\kappa(1,i) \;&=\; 1\end{aligned}$$

$W_P(\kappa_0)$ is the achievable throughput of primary user when secondary source is silent. The first constraint bounds the maximum performance loss of the primary user, while the other forces that the probability sum of secondary user's stays idle and not staying idle is 1. Also, the sum probability of choosing the backoff counter is 1.

A. *Solution of Optimization Problem*

In this subsection, we address the structure of the optimal transmission strategy in the particular case i.e. $\nu^* = \nu$, that is, the transmission by the primary source does not affect the successful decoding probability of the packet of the secondary source by the secondary receiver. The transmission strategy maximizing the throughput of the secondary source, given the constraint on the primary source's throughput loss, have some general structure given the optimal portion of time when the secondary user remains idle which can be obtained through exhaustive search. It can be shown that the same structure applies if the constraint is on the failure probability of primary source's packets. The definition for the last case are provided in the subsection III-B.

As discussed before, interference from the secondary source in different states has different effects. In fact, if the secondary source transmits in the backoff slot of primary source, it halts the backoff procedure and therefore incurs delay for the transmission process of later. On the other hand, if secondary source's transmission happens to superimpose with primary source's one, it is more likely that later goes to higher backoff stage and therefore the steady state probability of later backoff stages are increased. Secondary source, in advance does not know which stochastic state currently the primary user is in, only way to control its transmission strategy resides on the probability of being idle and the probability of which backoff counter it will pick. As the DCF protocol itself is not deterministic, interaction between primary and secondary source is not deterministic as well.

Reward of the secondary source, primarily, we consider the throughput which needs to be maximized. Throughput of the secondary source depends on the strategy space of the secondary user. The more portion of time, the secondary source stays idle, the less the throughput it gains. While picking the backoff counter, the smaller the backoff counter it picks, the higher the throughput it obtains. Because of later property, there is some general structure in the strategy finding algorithm of secondary source. However, there is a trade off between picking the probability that the secondary source stays in idle state and the probability of picking of different backoff counters.

As the entire interaction is probabilistic, there is no straightforward algorithm for the secondary source. At primary source's low traffic load, higher throughput is achievable if secondary source stays idle instead of picking higher backoff counter. In this case, it needs to always pick a backoff counter of lower value and otherwise stays in idle state in order to meet the constraint of primary source. On the opposite traffic load, its better to pick a backoff counter of higher value so that transmission of secondary source can be incorporated in the time slot when the primary user is idle. These findings have been observed from our simulation study where there is a primary and a secondary user. Following few theorems help identify the suitable algorithm of secondary source.

Theorem 1: Transmission probability of primary source is inversely proportional to the secondary source's transmission probability.

Theorem 2: Transmission probability of secondary source is inversely proportional to the primary source's transmission probability.

Theorem 3: Backoff counter of lower value incurs more transmission probability for the secondary source than higher backoff counter.

For the proof of these theorems, readers are encouraged again to check the appendix of this paper. From the equation 1, we observe that throughput of the primary source is proportional to its transmission probability. Moreover, primary source's throughput is inversely proportional to secondary source's transmission probability. Secondary source's throughput is also proportional to its transmission probability. In addition, secondary source's frequent transmission increases the generic slot time of primary user and the throughput of primary user is inversely proportional to its generic slot time. Transmission probability of secondary user is dependent on the portion of time it remains idle. Because of all these observations, there is no any straightforward procedure for finding optimal portion of idle time for the secondary source. Therefore, we have adopted brute-force search in order to find it, and theorem 3 helps find the backoff counter picking strategy and in order to obtain optimal throughput first it gives more emphasize on the backoff counter of lower value and then further. Following 1 is the optimal strategy finding algorithm for the secondary user.

Algorithm 1 Secondary User's Optimal Strategy Finding Algorithm

```
{A}ssume maximum performance loss ℑ_p^max
κ_0 := 0
while κ_0 <= 1 do
  κ_1 := 0
  while κ_1 <= 1 do
    ...
    κ_{w_s} := 0
    while κ_{w_s} <= 1 do
      if Σ_{i=1}^{w_s} κ_i > 1 then
        continue
      end if
      {C}ompute performance loss ℑ_p
      if ℑ_p <= ℑ_p^max then
        {R}ecord the secondary if it is larger than the
        currently recorded throughput
        {R}ecord other parameters
      end if
      κ_1 := κ_1 + ε
    end while
    ...
    κ_{w_s} := κ_{w_s} + ε
  end while
  κ_0 := κ_0 + ε
end while
```

B. Constraint on the Average Failure probability

As shown in the previous subsection, if the constraint on the average performance loss at the primary transmitter is defined for the throughput loss, then the algorithm proposed described above both brute-force and some insights of the problem formulation helps cutting down the number of search iterations. Remarkably, the same structure applies to an analogous optimization problem in which the constraint is defined for primary source's packet failure probability constraint. The failure of a particular packet happens when it fails to be transmitted successfully upto the retry limit times. Thus the failure probability is denoted by:

$$\Im_P^{fp}(\hat{\kappa}) \;=\; \prod_{i=1}^{m} t_2^*$$

IV. Formulation of Online Solution

A. Optimization Problem

Let us define the cost functions $X_i\left(\phi, u\right) : \chi \times \bigcup \rightarrow R$ as the average cost incurred by the markov process in state $\phi \in \chi$ if action $u \in \bigcup$ is chosen. Note that, $u = 0$ represents the secondary source keeps silent and $u = 1$ represents the picking of a backoff counter from secondary backoff counter window i.e. $[0, 1, \cdots, w_s - 1]$. And, average generic cost function yields to

$$X_i(u) \;=\; \lim_{n->\infty} \frac{1}{n} \sum_{t=1}^{n} E\left[X_i(\phi_t, u_t, \epsilon_t(\phi_t, u_t))\right]$$

where $\bigcup = \{u_1, u_2, \cdots\}$ is the sequence of actions of the secondary source and $\epsilon_t(\phi_t, u_t)$ is an exogenous random variable which is not instantaneously obtained due to protocol specific behavior. For example, if secondary user picks a backoff counter j, it has to go through first DIFS and j times backoff slots before having transmission. While passing through the backoff slots, it might be halted by the transmission of primary user and reduces the overall throughput than the case of not being halted. This incidence is also true for the primary user as well. Moreover, state variable ϕ_t is not explicit to the secondary user, because secondary user does not know if the primary user is in backoff stage or in idle slot. However, secondary user can sense the primary user's presence if the primary user transmits in a slot. Considering all these issues, our high level cost functions have been derived below.

$$X_0(\phi, u, \epsilon) \;=\; \begin{cases} \theta_S(\nu) & if\ u = 1\ \&\ \phi \;\neq\; (b, 0) \\ \theta_S(\nu^*) & if\ u = 1\ \&\ \phi \;=\; (b, 0) \\ \theta_S & if\ u = 0\ \&\ \forall \phi \in \chi \end{cases}$$

$$X_1(\phi, u, \epsilon) \;=\; \begin{cases} \theta_P(\rho) & if\ u = 0\ \&\ \phi \;\neq\; (0, 0) \\ \theta_P(\rho^*) & if\ u = 1\ \&\ \phi \;\neq\; (0, 0) \end{cases}$$

$\theta_S(\nu)$ and $\theta_S(\nu^*)$ are the instantaneous calculated secondary user's throughput assuming the failure probability of transmitted packet is ν and ν^* respectively. As discussed previously, ν^* and ν are the failure probability of secondary user's transmitted packet when primary user transmits and does not transmit respectively. Besides these two cases, X_0 is just the throughput of secondary user considering the current time slot as we know secondary user's queue is backlogged. Sitting idle in other's transmission time and backoff slots are taken account into the calculation of throughput.

$$X_2(\phi, u, \epsilon) \;=\; \begin{cases} 1 & if\ tx\ fails\ \&\ \phi \;=\; (m, 0) \\ 0 & otherwise \end{cases} \qquad X_3(\phi, u, \epsilon) \;=\; \begin{cases} 1 & if\ \phi \;=\; (1, 0) \\ 0 & otherwise \end{cases}$$

And again, $X_2(u)$ can be interpreted as the fraction of time slots in which the primary source fails the last allowed transmission and the packet would not be delivered and $X_3(u)$ is the fraction of time slots where the primary begins the service of a new packet. In this paper we define the failure probability as the average ratio of dropped packets after m retransmission, to the total number of new packets sent, one can see that $\frac{X_2(u)}{X_3(u)}$ is equivalent to the failure probability of the primary sources packets. The optimization problem is then given by

$$\arg\min_{\kappa} X_0(u) \quad s.t.\ \theta_P^{max} - \theta_P \le \gamma_1 \; OR \; \frac{X_2(u)}{X_3(u)} \le \gamma_2 \quad (3)$$

It is shown in the section III that the optimization problem in equation 3 is solved by formulating the problem as a linear program. Parameters in the formulation have been derived from the steady state distribution of the markov chain. Finally, the obtained optimal strategy has been denoted by a vector κ. Elements of this vector holds the proportion of time secondary user keeps silent or in which probability should it picks the backoff counter from the given contention window. Solution needs a little brute force search with some standard policy that have been proven analytically.

B. State Knowledge

The offline solution of the optimization problem requires full knowledge of state ϕ_t, which corresponds to the transmission index and queue state of the primary source, as well as knowledge of the transition probabilities and cost functions. However, the full knowledge of ϕ_t requires an explicit exchange of information.

We address this limitation in two steps. First, by assuming that the secondary only has information about what can be directly observed about the primary, and second, by using an on line learning technique that learns the necessary parameters without requiring knowledge of the transition probabilities.

By sensing the channel, primary user cannot instantaneously detect the channel condition as primary user follows DCF protocol. Therefore, there is no way to get the information about the state of the primary user when it is in backoff state or in idle state i.e. primary user's queue is empty. Secondary user can get to know if primary user transmits in a certain slot by sensing the channel. In some cases, secondary user can get knowledge if primary user's transmitted packet is new or old. The header includes the sequence number of the packet, which increases if the transmitted packet is a new one and remains the same if it is a retransmission. However, when the retransmitted packet reaches to its maximum limit, there is no way for the secondary user to know, in the next slot whether it will go through another backoff stage with fresh packet, since the buffer state is completely unknown to the secondary user.

Even though primary user can gather such little information, in the proposed solution, secondary user does not rely on these information. Rather, the proposed solution depends on a simple bit which indicates whether the performance constraint of the primary user is satisfied in the current time slot or not. This information is sent by the primary user as piggy backed form in either ACK or the actual packet's header. Having this information, secondary user regulates its transmission strategy.

It is shown in the following section via numerical results that this such partial knowledge is sufficient to implement a learning algorithm operating close to the limit provided by full state knowledge. Note that, the state of the primary user is overlooked here, it does not help in the decision making process of the secondary user. Rather the cost functions are most important driving factor of the proposed online algorithm.

C. Learning Algorithm

Most approaches to optimal control require knowledge of an underlying probabilistic model of the system dynamics which requires certain assumptions to be made, and this entails a separate estimation step to estimate the parameters of the model. In particular, in our optimization paradigm III, the optimal randomized stationary policy can be found if the failure probabilities ρ, ρ^*, ν, ν^* are known to the secondary user, together with some knowledge of state ϕ. In this section we describe how we can use an adaptive learning algorithm called Q-learning [10], [11] to find the optimal policy without a priori knowledge about our probabilistic model.

The Q-learning algorithm is a long-term average reward reinforcement learning technique. It works by learning an action-value function $R_t(\phi, u)$ that gives the expected utility of taking a given action u in a given state ϕ and following a fixed policy thereafter. Intuitively, the Q-function captures the relative cost of the choice of a particular allocation for the next time-step at a given state, assuming that an optimal policy is used for all future time steps. Q-learning is based on the adaptive iterative learning of Q factors. However, as discussed previously, it is almost impossible to get to know about the information of primary user's current state and thus it ignores the current state ϕ_t while learning the system and behavioral parameters of primary user. Since, secondary user overlooks current state ϕ_t while taking any action, we can call it as the variant of markov decision process(MDP). The original MDP means, the agent takes action based on the current state of the environment.

No matter, secondary user follows MDP or variant of MDP, it needs to fix a cost function which is typically named as reward. Ultimate reward of the secondary user is its own throughput i.e. $X_0(\phi, u, \epsilon(\phi, u))$ which it wants to maximize. However, in order to maximize throughput, we have adopted some indirect approach to get the maximized value of $X_0(\phi, u, \epsilon(\phi, u))$. Cost function is associated with the action of secondary user. Proability of each action of secondary user is resided in the vector $\kappa = [0, 1, \ldots, w_s]$. Length of this vector is $w_s + 1$ (w_s is backoff window size of secondary user). Index 0 denotes the proportion of time secondary user keeps itself silent, subsequent indexes i denote the portion of time backoff counter $(i-1)$ is chosen by the secondary user. As discussed previously, outcome of secondary user's action is not obtained instantaneously until the secondary user has its transmission. Due to the interaction of secondary and primary user, the obtained throughput from each action vary and our cost function is the obtained average throughput (added to the long term average throughput) resultant from the taken action. Let $X_0^p(\phi, u, \epsilon(\phi, u))$ is the average throughput of secondary user while taking the action u and $X_0^n(\phi, u, \epsilon(\phi, u))$ is the average throughput when the secondary user really completes its packet transmission. Then, the cost function at time t is defined as follows:

$$c(\phi, u, \epsilon(\phi, u)) \;=\; X_0^n(\phi, u, \epsilon(\phi, u)) \;-\; X_0^p(\phi, u, \epsilon(\phi, u))$$

And our optimization problem thus stands to

$$\arg\max_{\kappa} \lim_{n->\infty} \frac{1}{n} \sum_{t=1}^{n} E\left[c_t(\phi_t, u_t, \epsilon_t(\phi_t, u_t))\right] \qquad (4)$$

And the Q-Learning algorithm for solving equation 4 is illustrated as follows.

- **Step 1:** Let the time step $t = 0$. Initialize each element of reward vector $R_t(\kappa)$ as some small number, such as 0.
- **Step 2:** Check if the constraint of the primary user is satisfied. If not, choose the action of $u = 0$. Otherwise, choose the action u with index $j = [0, 1, \cdots, w_s]$ that has the highest $R_t(j)$ value with some probability say $1 - \tau_t$, else let u be a random exploratory action. In other words,

$$u_t \;=\; \arg\max_{\kappa} R_t(\kappa)$$

- **Step 3:** Carry out action u_t. Wait until secondary user completes its transmission if it picks any backoff counter. Or secondary user may choose the option of being silent. In either case, Calculate the cost function $c_t(u_t)$ and update the reward variable for the corresponding action. If the current state is ϕ and the resultant state is ϕ' after taking the action u_t, reward is updated as follows.

$$R_t(\phi, u_t) \;=\; (1 - \alpha_t)R_t(\phi, u_t) + \alpha_t(c_t + \gamma \max_{u'} R_t(\phi', u_t')) \qquad (5)$$

- **Step 4:** Set the current state as ϕ' and repeat step 2. When convergence is achieved, set $\tau_t = 0$.

This is the typical Q-learning algorithm. In our case, we don't know the primary user's exact current state and also don't know what the next state will be. Therefore the equation 5 reduces to

$$R_t(u_t) \;=\; (1 \;-\; \alpha_t)R_t(u_t) + \alpha_t c_t$$

In order to obtain the optimal value of α_t, we have found the following theorem.

Theorem 1: Step size parameter $\alpha_t = \frac{1}{t}$ gives the convergence to the algorithm.

Proof: The choice $\alpha_t \;=\; \frac{1}{t}$ results in the sample-average method, which is guaranteed to converge to the true action values by the law of large numbers. A well-known result in stochastic approximation theory gives us the conditions required to assure convergence with probability 1:

$$\sum_{t=1}^{\infty} \alpha_t \;=\; \infty \quad and \quad \sum_{t=1}^{\infty} \alpha_t^2 \;<\; \infty$$

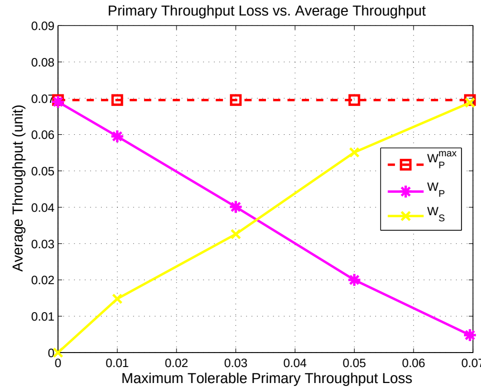

Fig. 4. Primary User's Tolerable Throughput Loss vs. Average Throughput for $\lambda_1 \;=\; 0.05$

The first condition is required to guarantee that the steps are large enough to eventually overcome any initial conditions or random fluctuations. The second condition guarantees that eventually the steps become small enough to assure convergence. Note that, both convergence conditions are met for the sample-average case, $\alpha_t = \frac{1}{t}$.

V. Performance Evaluation

In this section, numerical results validating the findings and observations throughout the paper are presented. Throughout the study, we assume that the buffer size of the primary source is $B \;=\; 4$ and the maximum retransmission time is $m = 4$. Backoff window size in each stage are 4, 6, 8, and 10 respectively. Secondary user's backoff window size is as $w_s = 3$. We set the failure probabilities for the transmission of the primary source $\rho \;=\; 0.2$, $\rho^* \;=\; 0.5$, depending on the fact that secondary is silent or not, respectively. Similarly, the failure probabilities of the secondary source are set to be $\nu \;=\; \nu^* \;= 0.3$. As, both users follow DCF protocol, each backoff slot is considered as one time unit, and without losing generality, transmission slot combined with DIFS period and ACK time is considered as 10 time units. Each packet has unit size in terms of bit. Once again, the goal of the algorithm is to maximize the throughput of the secondary source.

A. Numerical Results of Linear Program Formulation

First, we want to show results relevant to the constraint, i.e. maximum fraction of throughput loss by the primary source. In figures 4 and 5, the throughput and secondary source's transmission strategy vector are depicted as a function of throughput loss constraint. In the figure, $W_{p^{max}}$ corresponds to the maximum throughput achieved by the primary source when the secondary source is silent. The throughput of the secondary source increases as the constraint is increased. A larger throughput loss constraint allows the secondary source to interfere more with the primary source. The throughput actually achieved by the primary source decreases according to the increased maximum performance loss allowed, and it can be observed that the policy of the secondary source lowers

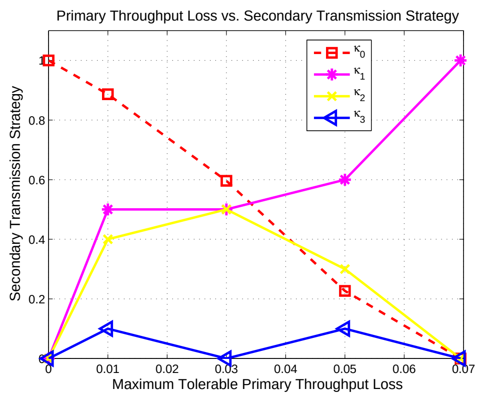

Fig. 5. Primary User's Tolerable Throughput Loss vs. Secondary Transmission Strategy for $\lambda_1 = 0.05$

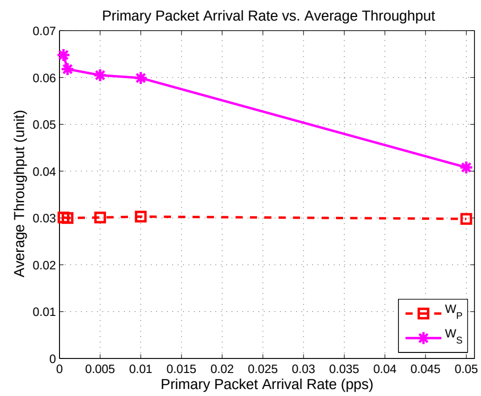

Fig. 6. Primary User's Packet Arrival Rate (pps) vs. Average Throughput for $\gamma_1 = 0.04$

the throughput of the primary one as much as possible in order to maximize the secondary throughput. When the secondary source's policy turns out that it will always transmit with zero backoff, throughput achieved by it reach to the maximum whereas the primary source stucks and stays at the backoff stage unless it chooses counter with zero value. From the current simulation, it chooses backoff counter valued zero and thus it's throughput turns into zero at this point. When $\gamma_1 = 0$ or the primary source cannot tolerate any throughput loss, ideally secondary user should only use the slots which are idle. However, in every slot, if it finds current throughput loss is higher than tolerable throughput, it keeps itself silent. Thus, in this case, we observe secondary source's throughput is zero.

From the figure 5, primarily, we observe that as throughput loss constraint is increasing, with the increased secondary throughput (resultant from increased transmission probability), fraction of slots being idle is reduced. As secondary source is more allowed to interfere with the primary source, it tends to always transmit rather than staying idle. As described in the previous algorithm, given the proportion of idle slots, it tries to find the strategy of selecting backoff counter by putting more probability first to the lower value and then further. Except than this property, there is no any other insights that we can draw from the figure.

A larger λ_1 means the primary source is accessing the channel more often. Therefore, the number of slots in which the secondary source can transmit while meeting the constraint on throughput loss of the primary source decreases. However, there is another effect of a large λ_1 that needs to be considered besides the scarcity of empty slots. With larger λ_1, with increased secondary source's transmission probability, its more likely that primary source's transmission is lingered or coincided by the former one. When both users' transmissions superimpose, primary user more likely moves to larger backoff stage. Thus primary user is prompt to retransmit its corrupted packets instead of transmitting fresh packets and this way, transmission behavior of primary source is affected by the value of λ_1 and achieved throughput as well depicted in

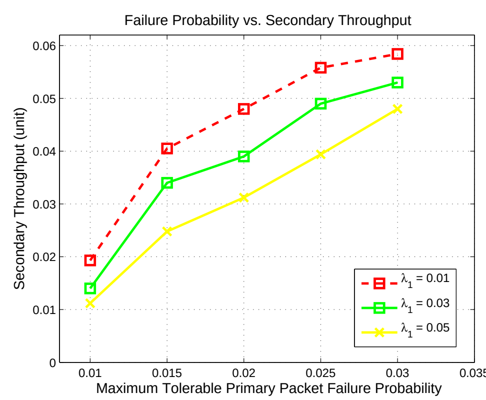

Fig. 7. Primary User's Tolerable Failure Probability vs. Secondary User's Throughput

figure 6. Similar to the previous figures, we also have figure for the transmission strategy of secondary user for different arrival rate λ_1 (For the space limit, it has been skipped here). In general, we can say from the figure that at high traffic load, optimal strategy is picking backoff counter of larger value rather than staying idle, while for the low traffic load secondary user more likely picks backoff counter of lower value or stays idle whenever it exceeds the performance loss constraint.

Now, we show the results for the optimization problem with a constraint defined on the primary source's packet failure probability. The activity of the secondary source increases the failure probability. Figure 7 shows the throughput of the secondary source as a function of failure probability constraint. Intuitively, the throughput as well as the overall fraction of slots in which the secondary source transmits, increase as the maximum failure probability of the primary source's packets increase. Higher arrival rate results lower throughput for the secondary user which has been depicted in this figure. The transmission strategy of the secondary source follows the structure discussed throughout this paper.

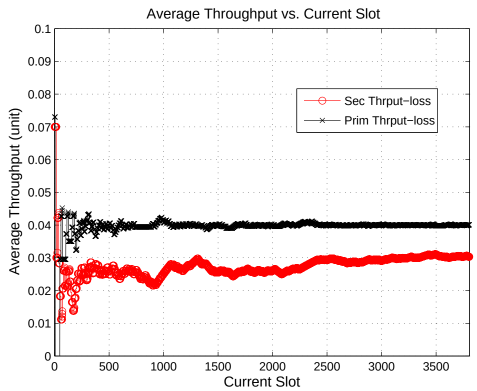

Fig. 8. Average Throughput (unit) vs. Current Slot for $\lambda_1 = 0.05$ and $\gamma_1 = 0.04$

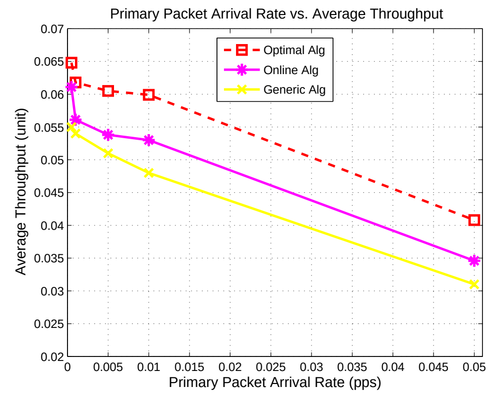

Fig. 10. Primary User's Packet Arrival Rate (pps) vs. Average Throughput for $\gamma_1 = 0.04$

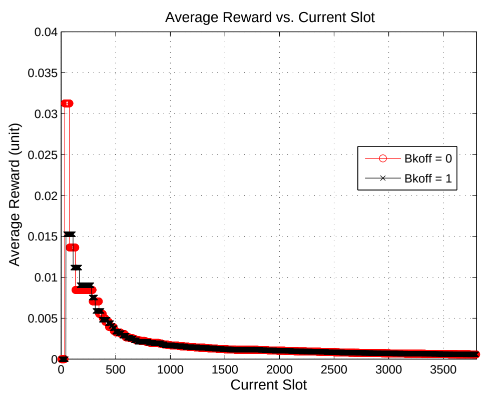

Fig. 9. Secondary User's Average Reward (unit) vs. Current Slot for $\lambda_1 = 0.05$ and $\gamma_1 = 0.04$

B. Results of Online Algorithm

Figure 8 depicts the convergence of secondary and primary user's throughput from 0'th iteration to some number of iterations. Throughput loss is defined as the difference between maximum achievable throughput and instantaneous throughput at a particular slot. From the given parameters, maximum achievable throughput is calculated considering only a single user (primary or secondary) is acting on the channel. We see the convergence of throughput loss happens after a few iterations.

In order to extrapolate the cost functions of our algorithm, we also have shown convergence process of two actions picked up by the secondary user, i.e. probability of picking backoff counter 0 and 1 respectively in the figure 9. We have initialzed cost of all actions at time slot zero. As the algorithm moves along with time, it updates its average reward according the formula presented in the algorithm. The algorithm is more prone to pick backoff counter with lower value that will be shown in the subsequent figures. However, in terms of general rule, algorithm does not pick the same action repeatedly. This is because, due to the interaction between primary and secondary users, the repeated action may cause to the degradation of the primary user's performance or it may degrade its own average reward value than the other actions. Consequently, the algorithm moves to the other action and the average reward value over the time for different actions look similar.

Figure 10 shows the throughput of primary and secondary source with the increased packet arrival rate λ_1 for a fixed tolerable primary source's throughput loss. As expected, throughput of the secondary source decreases as λ_1 is increased gradually. A larger λ_1 means that the primary source is accessing the channel more often. Therefore, the number of slots in which the secondary source can transmit while meeting the constraint on the throughput loss of the primary source decreases. In addition, in this figure, we have projected the result obtained by our optimal algorithm III. Optimal algorithm though due to the protocol behavior is not fully aware of state of the system, has some better information than our proposed online algorithm. Therefore, it incurs better performance in terms of achievable throughput for different λ_1 value. Whereas, our online algorithm though does not look like have similar performance, but gains better one than other blind generic algorithm. Generic algorithm means, here secondary user picks its backoff counter uniformly. With this strategy, we see the performance for the secondary user is the worst. Even worst news is that, this algorithm is completely blind about the performance constraint of primary user.

C. Comparison of Both Approach

Figure 11 compares the obtained secondary user's strategy for both our optimal and online algorithms. We have presented the proportion of idle slots and probability of picking backoff counter 0. For the sake of page limit, we have skipped other results here. In this result apparently, we don't see any match between two algorithms. However, we can explain the difference. In fact, online algorithm is mostly dependent on the primary user' performance loss violation indicator and its own reward value for different actions. It tries to pick the action with maximum value, which is usually the backoff counter with lower value. Otherwise, upon the signal of constraint violation, it keeps silent. Therefore, we see that

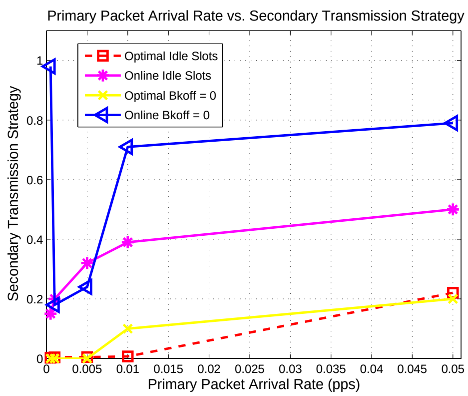

Fig. 11. Primary User's Packet Arrival Rate (pps) vs. Secondary Transmission Strategy for $\gamma_1 = 0.04$

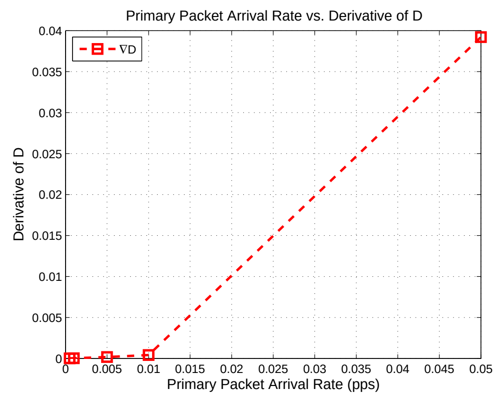

Fig. 12. Primary User's Packet Arrival Rate (pps) vs. Derivative of D

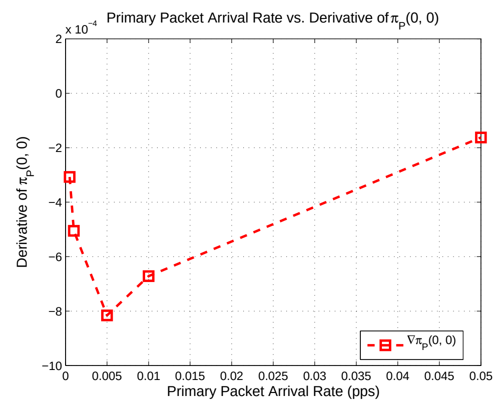

Fig. 13. Primary User's Packet Arrival Rate (pps) vs. Derivative of $\pi_P(0,0)$

online algorithm puts more weights to the backoff of lower value and again backoff counter of lower value breaks the constraint more often and thus it keeps more silent than offline algorithm. Whereas, optimal algorithm knows the arrival rate of primary user, it runs a near brute-force algorithm in order to find the optimal strategy of secondary user.

VI. Conclusion

In this paper, we have provided an opportunistic channel access scheme for the secondary user in IEEE 802.11 based networks. For maintaining the simplicity of the problem, we have assumed there is one primary and one secondary user in the network where both of them adopt interference mitigation scheme in order to extract a transmitted packet successfully. In the fisrt half of this paper, problem has been formulated as a linear program and then an algorithm has been proposed which has been proven to be optimal. Solution of this problem require that secondary user has some knowledge of the network, however in practice, it might not be possible. Given this fact, in the second half of the paper, an online algorithm has been proposed which does not require to know all information of the network or the secondary user can adaptively learn the state of the network while making a transmission decision. This approach relies only on the little performance violation feedback of the primary transmitter and uses Q-learning to converge to nearly optimal secondary transmitter control policies. Numerical results have been provided in order to justify the efficacy of these algorithms and it has also been shown that both algorithms achieve comparable performance.

Appendix

Proof of Theorem 1

Due to the steady state distribution of the markov model, we obtain

$$\sum_{b=1}^{m}\sum_{j=0}^{w_b-1}\pi_P(b,i)+\pi_P(0,0) \;=\; 1 \qquad (6)$$

We consider $m = 4$ and contention window size in each backoff stage are described in the section V. Solving the equation 6 and equation due to the state transition probabilities of all states, we obtain

$$\pi_P(0,0) \;=\; \frac{1-q_1}{1-q_1+Dt_2^{*3}R_4+\cdots\cdots+q_1R_4+Dq_1R_1(1-t_2^{*4})}$$

$$\pi_P(1,0) \;=\; \pi_P(0,0)D$$

where

$$D \;=\; \frac{q_1}{1-q_1(1-t_2^{*4})}$$

$$R_4 \;=\; \sum_{j=1}^{w_4-1} j \quad and \quad R_1 \;=\; \sum_{j=1}^{w_1-1} j$$

However, we know $q_1 = min(1,\lambda_1 D_{access})$ and D_{access} is also the function of secondary user's transmission probability t_2. Therefore, its hard to find closed form term when we differentiate the above terms with respect to t_2. In order to prove this theorem, we have plotted the numerical differenttion as a function of the packet arrival rate of primary

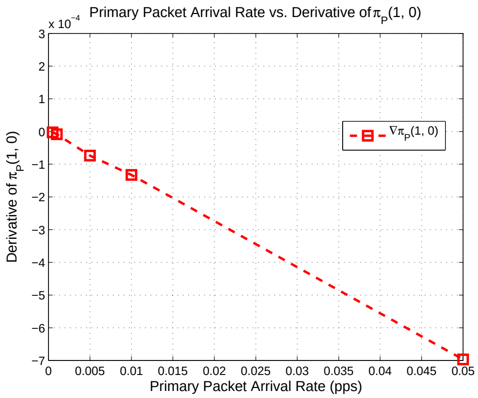

Fig. 14. Primary User's Packet Arrival Rate (pps) vs. Derivative of $\pi_P(1,0)$

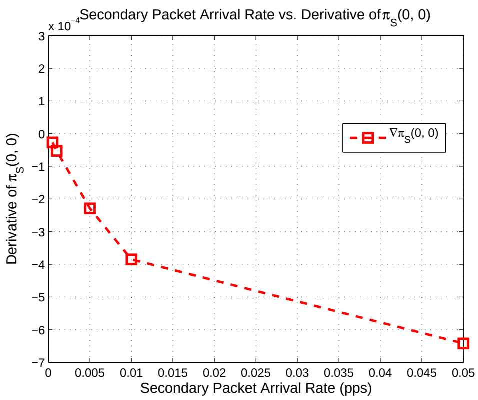

Fig. 15. Secondary User's Packet Arrival Rate (pps) vs. Derivative of $\pi_S(0,0)$

user. Values of packet arrival rate are as same as in the section V. Figures 12, 13 and 14 depict the derivatives of D, $\pi_P(0,0)$ and $\pi_P(1,0)$ with respect to t_2. From the definition of differentiation we know

$$\nabla f = \frac{f(x+\Delta)-f(x)}{\Delta}$$

We observer in the figures that no matter the packet arrival rate, diffrention of $\pi_P(0,0)$ and $\pi_P(1,0)$ is negative with respect to t_2. It means, as we increase the value of t_2, value of $\pi_P(0,0)$ and $\pi_P(1,0)$ i.e. idle slots and transmission probability of fresh packet transmission rate are decreased. Thus, it proves theorem I.

Proof of Theorem 2

Considering the secondary user's contention window size $w_s = 4$ and solving the steady state distribution equation and equations due to secondary user's state transition, we obtain

$$\pi_S(0,0) = \frac{(1-t_1)(1-q_2)}{(1-t_1)(1-q_2)+q_2(1-t_1)+(1-q_2)q_2P+q_2^2P}$$

$$\pi_S(1,0) = \frac{\pi_S(0,0)q_2}{1-q_2}$$

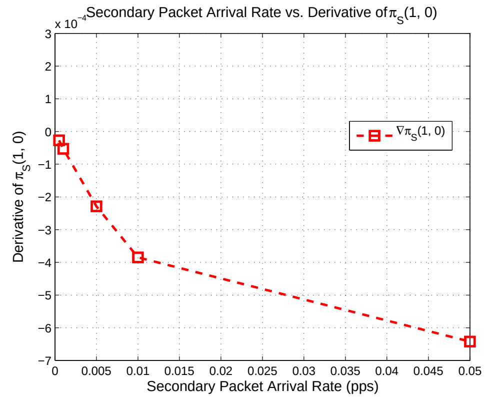

Fig. 16. Secondary User's Packet Arrival Rate (pps) vs. Derivative of $\pi_S(1,0)$

where $P = \sum_{j=0}^{w_s-1} j\kappa(j)$. For different secondary user's fake packet arrival rate, we have plotted the differentiation of $\pi_S(0,0)$ and $\pi_S(1,0)$ in the figures 15 and 16. Negative values of these differentiations prove the theorem II.

Proof of Theorem 3

Higher probability in the backoff counter of lower value means lower probability in higher valued backoff counter which results in lower number of backoff slots before having a transmission by the secondary user. Keeping primary and secondary users' transmission probability constant, increased probability of picking lower valued backoff counter incurs on average lower number of backoff slots and thus higher transmission probability. This statement proves the theorem III.

References

[1] O. Simeone, Y. Bar-Ness, and U. Spagnolini, "Stable throughput of cognitive radios with and without relaying capability," *Communications, IEEE Transactions on*, vol. 55, no. 12, pp. 2351 –2360, dec. 2007.

[2] S. Geirhofer, L. Tong, and B. Sadler, "Cognitive radios for dynamic spectrum access - dynamic spectrum access in the time domain: Modeling and exploiting white space," *Communications Magazine, IEEE*, vol. 45, no. 5, pp. 66 –72, may 2007.

[3] R. Urgaonkar and M. Neely, "Opportunistic scheduling with reliability guarantees in cognitive radio networks," in *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, april 2008, pp. 1301 –1309.

[4] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *Selected Areas in Communications, IEEE Journal on*, vol. 23, no. 2, pp. 201 – 220, feb. 2005.

[5] W. Zhang and U. Mitra, "Spectrum shaping: a new perspective on cognitive radio-part i: coexistence with coded legacy transmission," *Communications, IEEE Transactions on*, vol. 58, no. 6, pp. 1857 –1867, june 2010.

[6] L. Cao and H. Zheng, "Distributed rule-regulated spectrum sharing," *Selected Areas in Communications, IEEE Journal on*, vol. 26, no. 1, pp. 130 –145, jan. 2008.

[7] H. Su and X. Zhang, "Cross-layer based opportunistic mac protocols for qos provisionings over cognitive radio wireless networks," *Selected Areas in Communications, IEEE Journal on*, vol. 26, no. 1, pp. 118 –129, jan. 2008.

[8] M. Levorato, U. Mitra, and M. Zorzi, “Cognitive interference management in retransmission-based wireless networks,” in *Communication, Control, and Computing, 2009. Allerton 2009. 47th Annual Allerton Conference on*, 30 2009-oct. 2 2009, pp. 94 –101.

[9] D. Gross and C. M. Harris, *Fundamentals of queueing theory (2nd ed.)*. New York, NY, USA: John Wiley & Sons, Inc., 1985.

[10] R. S. Sutton and A. G. Barto, *Reinforcement Learning*. New York, NY, USA: The MIT Press, 1998.

[11] S. Mahadevan, “Average reward reinforcement learning: foundations, algorithms, and empirical results,” *Mach. Learn.*, vol. 22, pp. 159–195, January 1996. [Online]. Available: http://dl.acm.org/citation.cfm?id=225667.225681