Enhancing Cognitive Radio Dynamic Spectrum Sensing Through Adaptive Learning

Cem Tekin, Steven Hong, Wayne Stark
Department of Electrical Engineering and Computer Science
University of Michigan
Ann Arbor, Michigan 48109-2122

***Abstract*—Cognitive Radio (CR) networks present a difficult set of challenges due to the fluctuating nature of the available spectrum and wide ranging number of applications, each having different Quality of Service (QoS) requirements. This paper studies the key enabling technologies of Cognitive Radio and makes contributions in two key areas: sensing and learning. We shall first present the software testbed which is developed to implement the Cognitive Radio spectrum sensing system. Next, we derive the mathematical relationship between varying parameters and the QoS and test it on our system to verify the overall performance. Novel learning techniques which determine the statistics of primary user (PU) channel usage over time are proposed to enhance the cognitive radio's dynamic spectrum sensing ability. Using our testbed, we shall demonstrate the feasibility of the innovative adaptive learning algorithms and their ability to increase spectrum sensing efficiency and improve performance over time without feedback from the receiver. We will then proceed to the domain where there are multiple non-cooperative cognitive users (secondary users) selfishly applying the learning algorithms to increase their data rate in channels with varying primary user activity. Finally we conclude with discussions about our results and future work.**

I. Introduction

With the explosion in the number of wireless service subscriptions this past decade, spectrum is becoming an increasingly valuable commodity in the wireless community. The traditional method of dividing and licensing spectrum is becoming inadequate as the spectrum becomes more and more crowded and there is no longer any bandwidth left to allocate. While the shrinking amount of licensed spectrum may appear at first glance to be a major problem, on closer inspection, many spectrums are frequently under-utilized with some bands being unoccupied more than 90 % of the time [1]. The problem is that since those bands are already licensed, they cannot be used by another user which would cause interference to the licensed or Primary User (PU). Cognitive Radio (CR) is a technology that seeks to utilize the available frequency spectrum more efficiently through the opportunistic exploitation of unused bands [4]. The main components, from which CR derives its prefix "cognitive", that make it unique are its abilities to sense, learn, and adapt to the environment. For instance, while a frequency band is assigned to a primary wireless system/service at a particular time and location, the same frequency band is unused by this wireless system or service in other times and locations. This results in spectrum holes, which can be exploited by the CR in order to improve spectrum utilization.

In this paper we will look in depth at the key enabling technologies of CR, its sensing and learning capabilities. We will first look at the functions of different spectrum sensing techniques, and the challenges of implementing them in practical settings. More specifically, we focus our attention on an energy detection scheme and analyze how different noise power fluctuations affect our detection performance. While the classic energy detection scheme has been well documented [3], the impacts of Rayleigh fading on the probability of missed detection and false alarm are seldom evaluated in recent literature. Finally, the need for a learning algorithm arises from the fact that secondary users (SUs) do not have a priori knowledge about the statistics of PUs. For a CR, the frequency range of operation is very wide. Sensing all the channels continuously is not efficient in general. We investigate the systems where a SU has limited sensing capability so it cannot monitor all the channels simultaneously. Therefore the sensing strategy focuses on a subset of all the channels that SU can use. In the case where a SU can sense only one channel at a time, the problem reduces to the well known multi-armed bandit problem [7]. In the context of CR, the multi-armed bandit problem corresponds to following: At the end of each transmission period SU chooses one of the channels to sense and receives the available data rate on that channel as a reward. The goal of the SU is to maximize its cumulative reward. [6] considers asymptotically optimal strategy for multi-armed bandit problem with statistical assumptions. However, in a CR network performance may be significantly worse than the asymptotical performance because the channel statistics may change over time. With a different approach, [5] discusses the nonstochastic version of the bandit problem where no statistical assumptions are made about arms (channels in terms of CR) and provides algorithms that satisfies certain regret bounds. Our first learning algorithm is the Exp3 algorithm from [5]. We propose another algorithm based on sum of discounted "weights" with the discount factor depending on the "successes" and "failures" from the sensed channels and compare the performance of it with Exp3 when the PU activity in each channel is modeled as independent from the PU activities on other channels and other time steps or when PU activity is a Markov process. By "success" we mean that the channel that SU sensed was free. By "failure" we mean that the channel that SU sensed was occupied by a PU. Next,

978-1-4244-5239-2/09/$26.00 ©2009 IEEE

we consider the case when there are multiple SUs in which every SU is selfish and neither have information about the total number of SUs in the network nor the strategies and payoffs of other SUs. We formulate the spectrum sensing methods in Section II. In Section III we evaluate the transmission performance of the SU based on the energy detection scheme for different PU SNR. Section IV discusses the proposed learning algorithms and compares their performance in two different channel models, single or multiple SUs. Section V gives the performance results in terms of the average number of transmission failures for different PU activities. Section VI concludes the paper.

II. Spectrum Sensing

Spectrum sensing is the process of becoming aware of the spectrum usage and existence of PUs in the local geographical area. It enables the SU to adapt to the environment by sensing spectrum holes without damaging the PU transmission [2]. In this section we will consider PU detection method and evaluate the performance of a SU operating in a single channel.

The detection of a signal in the presence of noise requires processing which depends upon what is known of about the noise and signal characteristics. This detection allows the SU to exploit spectral holes and opportunistically use under-utilized frequency bands without causing harmful interference to the PU. There are multiple signal detection techniques and the appropriate processing strategy often depends on the amount of a priori knowledge the SU has about the PU. When we are presented with no a priori information regarding the form of the PU signal, we simply consider the PU signal as a sample function of a random process. The only thing we assume about the PU signal is that it is deterministic and the spectral region to which it is approximately confined is known. The assumption of a deterministic signal allows us to determine that the input to the SU with signal present is Gaussian but not zero mean. In the absence of a priori information concerning the PU signal the best signal detection method is detection by an enery detector. The energy detector measures the energy in the input wave over a specific time interval. Since only the energy of the signal matters, the results apply to any deterministic signal. In order to determine whether or not a specific channel is occupied by a PU, the SU seeks to distinguish between the following two different hypotheses regarding the received signal $r(t)$:

$$r(t) = \begin{cases} n(t), & H_0 \\ s(t) + n(t), & H_1 \end{cases} \tag{1}$$

where $s(t)$ is the PU's transmitted signal and $n(t)$ is the stochastic noise, modelled as AWGN, $n(t) \sim N(0, \sigma_v^2)$. Based on the work done by Urkowitz [3], we know that the distributions of the energy value received by the SU are central and non-central chi-square distributions with test statistic V_T equal to E/N_0, E being the signal energy.

$$V_T = \begin{cases} \sum_{i=1}^{N} a_i^2, & H_0 \\ \sum_{i=1}^{N} (a_i + \rho_i)^2, & H_1 \end{cases} \tag{2}$$

where N is 2 times the duration of the sampling function (T) times the bandwidth of the signal (W) and represents the degrees of freedom. ρ is the noncentrality parameter which is equal to the ratio of signal energy measured over sensing time to noise spectral density (SNR). In a non-fading environment with additive white Gaussian noise, probabilities of detection (P_d), missed detection (P_m) and false alarm (P_f) are given below [3].

$$P_d = P\{E_p > \gamma | H_1\} = Q_u\left(\sqrt{2\rho}, \sqrt{\gamma}\right), \tag{3}$$

$$P_m = P\{E_p \leq \gamma | H_1\} = 1 - Q_u\left(\sqrt{2\rho}, \sqrt{\gamma}\right), \tag{4}$$

$$P_f = P\{E_p > \gamma | H_1\} = \frac{\Gamma\left(\frac{N}{2}, \frac{\gamma}{2}\right)}{\Gamma\left(\frac{N}{2}\right)}. \tag{5}$$

In Rayleigh channels, the P_f expression remains the same due to its independency of the SNR. P_d however, must be averaged over the SNR distribution which is affected by the Rayleigh fading represented with the random variable X.

$$r(t) = \begin{cases} n(t), & H_0 \\ Xs(t) + n(t), & H_1. \end{cases} \tag{6}$$

The probability of missed detections in Rayleigh faded channels is shown below

$$P_m = \epsilon e^{\frac{\frac{B}{2A^2} - \beta^2}{2\alpha^2\sigma^2}} \left[Q\left(-\frac{B}{2A\epsilon}\right) - Q\left(\frac{B}{2A\epsilon}\right)\right] \tag{7}$$

where $\alpha = \sqrt{\frac{2E}{N_0}}$, $\beta = \sqrt{\frac{2\gamma}{N_0}}$, $\epsilon = \frac{\alpha\sigma}{\sqrt{A}}$, $A = \left(1 + \alpha^2\sigma^2\right)$, $B = 2\beta$.

In both the Rayleigh fading and AWGN scenarios, the P_f expression remains the same due to its independene of the SNR. The P_m expression however varies exponentially in the AWGN case and linearly in the Rayleigh fading case with respect to increasing $E/N_0(dB)$. It is also important to note that as the selected threshold γ is increased, P_f improves while P_m worsens. This makes sense intuitively because as the threshold is increased, it will require more signal energy to trigger a false alarm and thus there will be fewer false alarms. However, as the threshold increases, more signal energy will be required to indicate a present PU and this will cause P_m to worsen. Fig. 1 demonstrates the relationship of the power of the PU and the effect of the threshold with misdetections and false alarms for the Rayleigh fading channel case.

III. Cognitive Radio Transmission Performance

Sensing errors potentially have two effects. False alarm errors could potentially reduce the data rate and bandwidth efficiency because the available channel will not be used for transmission when a false alarm occurs. However, if the SU can find another channel to transmit on, then this is inconsequential. Missed detection causes collisions so PU's signal will distort the SU's transmission. However, if the SU

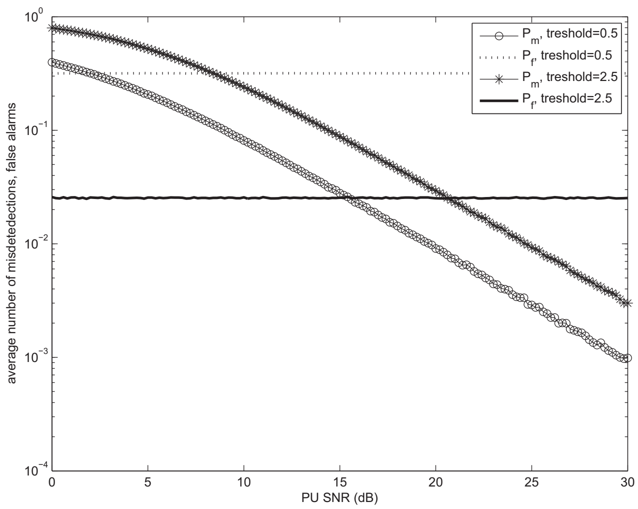

Fig. 1. PU detection with Rayleigh fading.

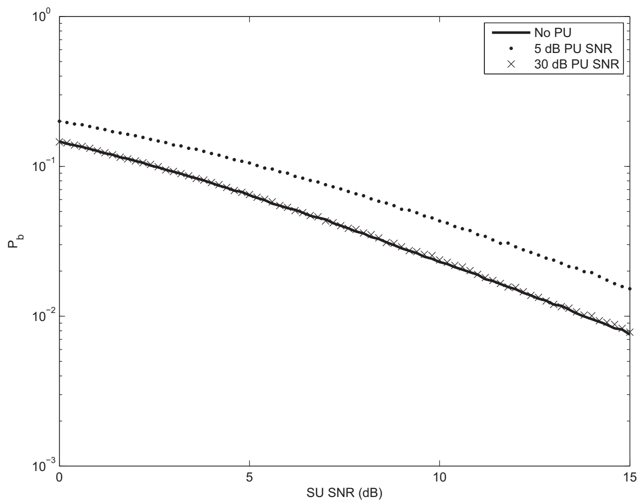

Fig. 2. Bit error probability with Rayleigh fading.

transmits with enough power, the receiver may overcome this data collision and decode the data correctly. We should also note that missed detection errors can result in data distortion for the PU, if the SU SNR is high. There are two main performance metrics to be considered when evaluating the performance of the SU: data rate and probability of bit error, which are directly affected by the threshold set by the energy detector and the power of the PU. The threshold affects the number of false alarms and missed detections. An increasing threshold decreases the number of false alarms and increases the number of missed detections and vice versa. An increase in the number of missed detections leads to an increase in the probability of bit error and an increase in the number of false alarms leads to lower data rates for the SU. We consider a SU as a receiver-transmitter pair and a bit error can occur only when the SU is transmitting, i.e., it concluded that there is no active PU. Fig. 2 demonstrates the bit error probability for the SU in Rayleigh fading channel. At PU interference of $30dB$, there are very few misdetection errors. Therefore the bit error performance of SU is closer to the performance in the absence of PU.

Knowing the PU SNR relationship to a given P_f and P_d will allow SU to determine its sensing strategy based on available PU SNR information. In this way, SU takes more samples at high PU SNR values because it knows results are more likely to be accurate whereas it takes fewer samples or selects a different channel when the PU SNR is low. Varying the sensing duration at each time interval allows SU to meet the specified P_d and P_f values in a faster and more reliable way. Therefore, SUs utilize the spectrum more efficiently without interfering with PUs.

IV. Learning Algorithms

A. Single Secondary User

A defining property of CR networks is their ability to use the spectrum in an intelligent way. Generally, the SUs may not be able to sense all the channels due to limited sensing and hardware capability. In order to increase spectral sensing efficiency, SUs should focus sensing channels with low PU activity. This can be accomplished by using an online learning algorithm. However there is a tradeoff between “exploration” and “exploitation” in online learning algorithms. “Exploration” is sensing different channels to search for a newly emerged good channel while “exploitation” is sensing a known good channel repeatedly because it offers high data rates. A good online learning algorithm should balance this tradeoff. We first define the CR network.The set of SUs is $\mathcal{N} = \{1, ..., n\}$, the set of channels is $\mathcal{M} = \{1, ..., m\}$ and $\mathcal{B} = \{B_1, ..., B_m\}$ where B_j is the bandwidth of channel j. Time is divided into slots, each slot is denoted by $t \in \{1, ..., T\}$. Each SU senses one channel from $\mathcal{M}$ at the beginning of each time slot, i.e, SU i senses $\alpha_i(t) \in \mathcal{M}$. We denote the sensing result of SU i at time t as $R_{\alpha_i(t)}(t)$. After sensing at the beginning of time slot t if SU i finds the channel to be free, i.e., $R_{\alpha_i(t)}(t) = 1$, it transmits on that channel. If the channel is not free, i.e., $R_{\alpha_i(t)}(t) = 0$, the SU waits for the beginning of the next time slot without transmitting. For each successful transmission the reward the SU receives is the bandwidth of the channel that it had sensed. Therefore, the reward at time t is $R_{\alpha_i(t)}(t) * B_{\alpha_i(t)}(t)$. To simplify the problem let $B_j = 1$ for all $j \in \mathcal{M}$. In the independent model, PU activity on each channel at time t is independent from the PU activities on other channels and other time steps. We define $Ppu_j(t)$ as the probability that there exists a primary user in channel j at time t. In the Markov model $P0_j(t)$ is the probability that there is no primary user in channel i at time t given that it was empty at time $t-1$ and $P1_j(t)$ is the probability that there is a primary user in channel j at time t given that it was occupied by a primary user at time $t-1$. In both learning algorithms SUs are using randomized sensing strategies. We denote by $\Delta(\mathcal{M})$ the set of probability distributions over $\mathcal{M}$. $\sigma_i(t) = [P_{i1}(t), ..., P_{im}(t)] \in \Delta(\mathcal{M})$ is the strategy profile of the SU i at time t. Let $X(\sigma_i(t))$ be a random variable which takes values in the set $\mathcal{M}$ with distribution $\sigma_i(t)$, i.e, $P_r\{X(\sigma_i(t))=j\}$=$P_{ij}(t)$. Then, $O(X(\sigma_i(t)))$ which is the outcome of the random variable $X(\sigma_i(t))$ is $\alpha_i(t)\in\mathcal{M}$.

First, we consider the learning algorithm Exp3 [5] whose code is shown in Fig. 4. When SU i uses Exp3, it starts with a uniform strategy to select the channel to sense from $\mathcal{M}$. The weight assigned to channel j by SU i at time step t is $\omega_{ij}(t)$. When $R_{\alpha_i(t)}(t)=1$, the weight of channel $\alpha_i(t)$ is updated by a term exponential in the inverse of the probability assigned that channel. Thus, the impact of the successes on rarely selected channels are more than the impact of the successes on the frequently selected channels. When $R_{\alpha_i(t)}(t)=0$ the weight of the channel is not updated. The $\gamma/|\mathcal{M}|$ in step 1 of algorithm Exp3 provides uniform term to the distribution $\sigma_i(t)$. Thus, every channel has a positive probability to be sensed at every time step. To evaluate the performance of SU i using Exp3 with respect to the channel with highest sum of rewards let

$$G_{i,\text{Exp3}}(T)=\sum_{t=1}^{T}R_{\alpha_i(t)}(t)$$

be the cumulative reward of SU i using Exp3. The cumulative reward of the best channel is

$$G_{\max}(T)=\max_j\sum_{t=1}^{T}R_j(t).$$

Corollary 3.2 of [5] states that for any $T>0$, $g\geq G_{\max}$, $\gamma=\min\left\{1,\sqrt{\frac{|\mathcal{M}|\ln|\mathcal{M}|}{(e-1)g}}\right\}$

$$G_{\max}(T)-\mathbf{E}\left[G_{i,\text{Exp3}}(T)\right]\leq 2\sqrt{(e-1)}\sqrt{g\,|\mathcal{M}|\ln|\mathcal{M}|}. \quad (8)$$

This gives a bound on the performance of Exp3 that will hold regardless of the PUs and other SUs in the system.

Using Alg1 in Fig. 3, SU i starts with a uniform distribution to sense a channel from $\mathcal{M}$. We can view $\phi_{ij}(t)$ as a discount factor of SU i for channel j at time step t which depends on the results of past and present sensing actions. If a channel is sensed free often, then $\phi_{ij}(t)$ for that channel goes to 1, in which case the past rewards from that channel has more effect on the weight of that channel. In other words, the algorithm considers a channel that is empty for most of the subsequent sensing results to be a good channel. On the contrary, channels that are occupied by a primary user in subsequent sensing results are treated as bad channels. For a bad channel, the past sensing results has little impact on the weight of the channel. These properties of Alg1 suits the dynamic nature of the problem well, since Alg1 responds to the changes in channel statistics quickly. First, when a good channel becomes bad, the discount factor for that channel decreases to compensate the effects of good results in the past by the bad results that occurred recently. Thus, the weight of the channel drops quickly resulting in a rapid adjustment of probability distribution of the sensing strategy. Second, when a bad channel becomes good, the discount factor for that channel goes to 1 and subsequent good results from that channel increases the weight of that channel so the probability of sensing that channel is increased. Moreover, the uniform distribution existing in Exp3 is not included in Alg1, so it is possible that the performance of Alg1 approaches to the performance of the algorithm that uses the best fixed available channel asymptotically when the PU distributions are stationary.

Alg1
$\forall i\in\mathcal{N}$
Parameters: $\beta_1,\beta_2,\eta\in(0,1)\,,k\in\mathbb{N}_+$
Initialization: $\omega_{ij}(1),\phi_{ij}(1)=1,\ \forall j\in\mathcal{M}$
for $t\in\{1,...,T\}$
1.

$$P_{ij}(t)=\frac{\omega_{ij}(t)}{\sum_{k=1}^{m}\omega_{ik}(t)},\forall j\in\mathcal{M}$$

2. $\alpha_i(t)=O(X(\sigma_i(t)))$
3.

$$\phi_{i\alpha_i(t)}(t+1)=\begin{cases}\min\left(1,\phi_{i\alpha_i(t)}(t)+\beta_1\right), & R_{\alpha_i(t)}(t)=1\\ \max\left(\eta,\phi_{i\alpha_i(t)}(t)-\beta_2\right), & R_{\alpha_i(t)}(t)=0\end{cases}$$

$$\omega_{i\alpha_i(t)}(t+1)=\begin{cases}\omega_{i\alpha_i(t)}(t)\phi_{i\alpha_i(t)}(t+1)+k, & R_{\alpha_i(t)}(t)=1\\ \omega_{i\alpha_i(t)}(t)\phi_{i\alpha_i(t)}(t+1), & R_{\alpha_i(t)}(t)=0\end{cases}$$

Fig. 3. Algorithm 1.

Exp3
$\forall i\in\mathcal{N}$
Parameters: $\gamma\in(0,1]$
Initialization: $\omega_{ij}(1)=1,\forall j\in\mathcal{M}$
for $t\in\{1,...,T\}$
1.

$$P_{ij}(t)=(1-\gamma)\frac{\omega_{ij}(t)}{\sum_{k=1}^{m}\omega_{ik}(t)}+\frac{\gamma}{|\mathcal{M}|},\forall j\in\mathcal{M}$$

2.$\alpha_i(t)=O(X(\sigma_i(t)))$
3.$\omega_{i\alpha_i(t)}(t+1)=\omega_{i\alpha_i(t)}(t)\exp\left(\frac{\gamma R_{\alpha_i(t)}(t)}{P_{i\alpha_i(t)}(t)|\mathcal{M}|}\right)$

Fig. 4. Algorithm Exp3.

B. Multiple Secondary Users

In this section we extend the use of algortihms to the multiple SU case. Note that the algortihms do not require external knowledge about PU activity. This motivates us to use these algortihms in multiple SU setting where a SU does not require external knowledge about other SUs to adjust its transmission. Therefore, SUs operate in a non-cooperative fashion, each SU selfishly wants to exploit the best available channels. Now, the reward of a channel does not only depend on the PU activity on that channel but also depends on the activity of other SUs using that channel. In this setting all SUs are runnning either Alg1 or Exp3, using same parameters.

Each SU decides which channel to sense at time slot t at the beginning of that time slot. Then it draws a waiting time τ' from a uniform distribution in $(0 \quad \tau]$ where τ is much smaller than the length of the time slot t. After the waiting time SU senses the channel and transmits on that channel if the channel is found to be free. Otherwise it waits until the beginning of the next time slot. By this scheme, SU with the smallest τ' gains access to the channel. Therefore, the expected rate of a channel is inversely proportional to the expected number of SUs sensing it.

V. Simulation Results

We simulated algorithms Exp3 and Alg1 for different PU distributions for both single SU and multiple SU cases. A transmission failure occurs whenever SU senses a channel and finds it to be occupied. We simulated the average number of transmission failures for each time step $t \in \{0, ..., T\}$ with $|\mathcal{M}|$=10.

A. Simulation 1

For simulation 1 we use the following parameters: $|\mathcal{N}|$=1, T=3500, β_1=0.05, β_2=0.4, η=0.01, k=1. Activity of the PU on each channel at time t is independent from the PU activities on other channels and other time steps. We set $g = T$ so γ=0.062 and Exp3 satisfies (8). For $t < 1000$, **Ppu**=$[Ppu_1, ...Ppu_m]$=[0.1, 0.9, 0.9, 0.9, 0.9, 0.1, 0.9, 0.9, 0.9, 0.9]. For $t \geq$1000, **Ppu**=[0.9, 0.9, 0.9, 0.9, 0.9, 0.9, 0.9, 0.9, 0.1, 0.9]. In this setting there are two types of channels: one with low PU activity and one with high PU activity. Results in Fig. 5 show that Alg1 outperforms Exp3.

B. Simulation 2

The parameters for simulation 2 are same as the parameters of simulation 1 except T=2000 and for t <400, **Ppu** = $[0.01, 0.9, 0.9, 0.9, 0.01, 0.9, 0.9, 0.9, 0.01, 0.9]$. For $t \geq 400$, **Ppu** = $[0.5, 0.999, 0.2, 0.99, 0.5, 0.5, 0.2, 0.99, 0.01, 0.99]$. For $t \geq 1000$, **Ppu**=[0.01, 0.999, 0.2, 0.01, 0.5, 0.5, 0.2, 0.01, 0.8, 0.99]. The channel conditions in this setting are more heterogenous. The results shown in Fig. 6 show that Alg1 outperforms Exp3 in this setting.

C. Simulation 3

In this simulation PU activity on each channel is modeled as a Markov process. T=6000, γ=0.048, **P0**=$[P0_1, ..., P0_m]$=[0.99, 0.99, 0.99, 0.99, 0.99, 0.99, 0.99, 0.99, 0.99, 0.99], **P1**=$[P1_1, ..., P1_m]$=[0.99, 0.99, 0.99, 0.99, 0.99, 0.99, 0.99, 0.99, 0.99, 0.99]. Other parameters are identical to the ones in Simulation 1 and Simulation 2. Results are given in Fig. 7. In this model, if a channel is empty it will stay empty with high probability or if it is occupied it will remain occupied with high probability. Alg1 perfoms better than Exp3 since it can adapt its probability distribution rapidly to the changing channel conditions. Exp3's performace is close to the performance of the best channel. However, concentrating on the same channel is not a good strategy since it will be occupied by a PU half of the time.

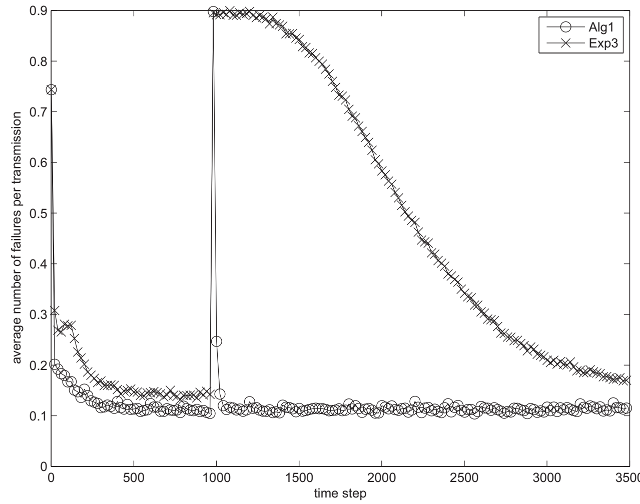

Fig. 5. Transmission failures for a single SU in independent model.

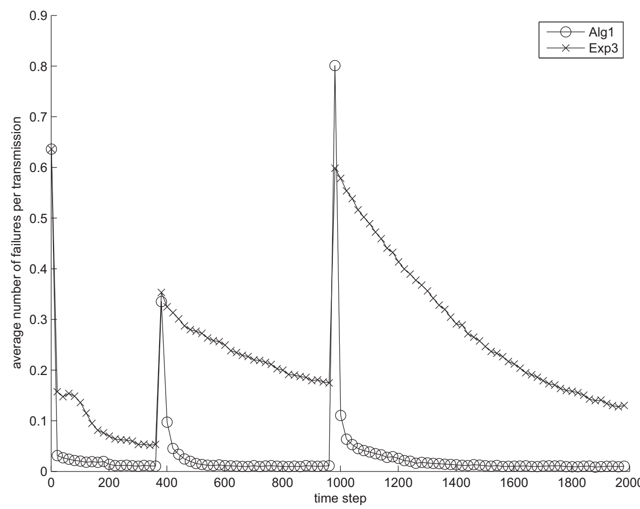

Fig. 6. Transmission failures for a single SU in independent model.

D. Simulation 4

When multiple SUs are present in the system, simulation results show that in both algorithms each SU has the same average number of unsuccesfull transmisions for each time step. By using these algortihms SUs settle to an equilibrium point in terms of the average number of unsuccesfull transmission where all players experience the same average number of unsuccessful transmission at each time step. Fig. 8 and Fig. 9 shows the performance of algorithms when $|\mathcal{N}|$=3 for independent and Markov models respectively. Results in Fig. 8 are obtained using same parameters as Simulation 2 while Results in Fig. 9 are obtained by using same parameters as Simulation 3 for each SU.

E. Simulation 5

Simulation 5 shows the performance of the energy detector scheme combined with Alg1 or Exp3 for different levels of PU SNR. Here the number of unsuccessful transmissions

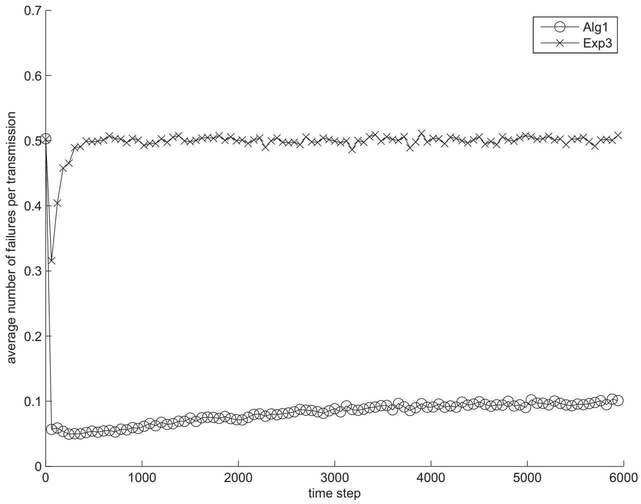

Fig. 7. Transmission failures for a single SU in Markov model.

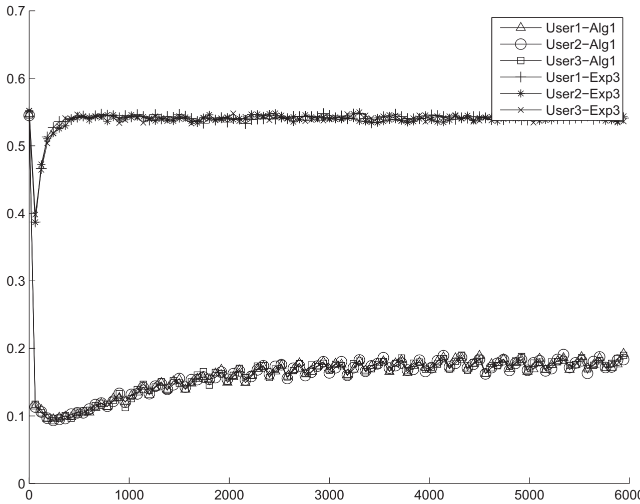

Fig. 9. Transmission failures for 3 SUs in Markov model.

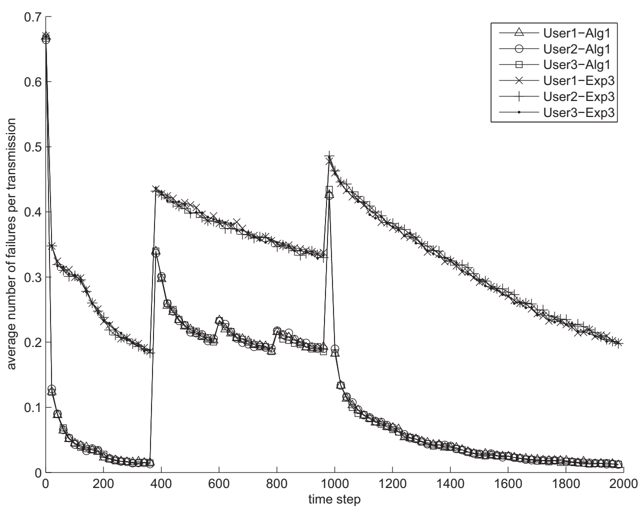

Fig. 8. Transmission failures for 3 SUs in independent model.

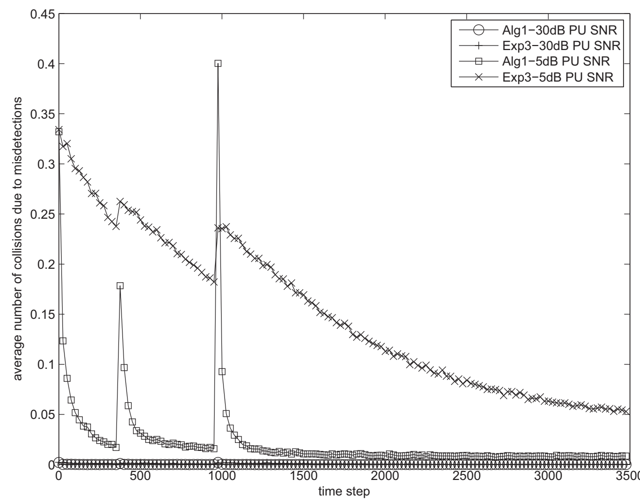

Fig. 10. Collisions due to misdetections for different PU SNR values.

is the sum of the number of misdetections, false alarms and corect sensing results which indicates that the sensed channel is used by a PU. As the PU SNR increases the number of misdetections decreases. For low PU SNR value the performance degradation is due to the high number of misdetections. Fig. 10 shows collision due to misdetections, Fig. 11 shows unsuccessful transmissions resulting from active PU on the sensed channel and false alarms. For E_p/N_0=30dB, average number of misdetections is close to 0. Number of unsuccessful transmissions resulting from PU activity and false alarms does not depend on PU SNR in Alg1 while dependence is observed in Exp3. The treshold value of the energy detector is 2.5 N_0^2.

VI. Conclusion

In this paper, we have proposed innovative adaptive learning algorithms which improves cognitive radio dynamic spectrum sensing and combined them with an energy detection scheme.

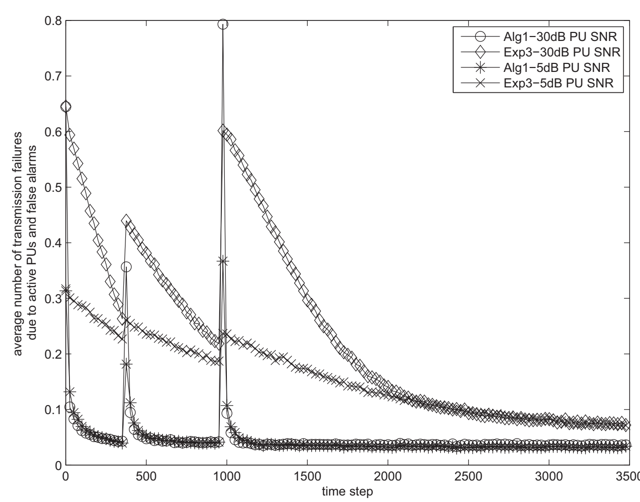

Fig. 11. Transmission failures not including collisions for different PU SNR.

We evaluated the performance of learning algorithms and demonstrated their ability to increase spectrum sensing efficiency by determining the channel to sense over a set of channels. The feasibility of the learning algorithms in dynamic environments where the distribution of PU activity is changing over time and where multiple SUs are present, is shown. Future work involves developing more sophisticated spectrum sensing techniques such as feature detection of the PU signal and optimizing the learning algorithms in scenarios where the SUs cooperate.

REFERENCES

[1] N. Devroye, M. Vu, V. Tarokh, "CR Networks," *IEEE Signal Processing Magazine*, pp. 12-23, Nov. 2008.

[2] I. Akyildiz, W. Lee, M. Vuran, S, Mohanty, "A Survey on Spectrum Management in Cognitive Networks," *IEEE Communications Magazine*, pp. 40-48, Apr. 2008.

[3] H. Urkowitz, "Energy Detection of an Unknown Deterministic Signal," *Proceedings of the IEEE*, vol. 55, no. 4, pp. 523-531, Apr. 1967.

[4] S. Haykin, "Cognitive Radio : Brain Empowered Wireless Communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, Feb. 2005.

[5] P. Auer, N. Cesa-Bianchi, Y. Freund, R. Schapire, "The non-Stochastic Multi-Armed Bandit Problem," *SIAM Journal on Computing*, vol. 32, pp. 48-77, 2003

[6] T. L. Lai, H. Robbins, "Asymptotically Effi cient Adaptive Allocation Rules," *Adv. in Appl.Math*, 6:11, pp. 4-22, 1985 vol. 52(4), pp. 1545-1561, Apr. 2006.

[7] H. Robbins, "Some aspects of the sequential design of experiments," *Bull. Amer. Math. Soc.*, 55, pp. 527-535 1952