The Primary User Emulation Attack in Dynamic Spectrum Access Networks: A Game Theoretic Approach

Yi Tan*, Shamik Sengupta† and K.P. Subbalakshmi*

*Department of ECE, Stevens Institute of Technology, Hoboken, NJ

†Department of Math. & Comp. Sci., John Jay College of Criminal Justice,
City University of New York, New York, NY

Email: *{ytan, ksubbala}@stevens.edu, †ssengupta@jjay.cuny.edu

Abstract

Cognitive radio enabled dynamic spectrum access networks are designed to detect and opportunistically utilize the unused or under-utilized spectrum bands. However, due to the open paradigm of cognitive radio networks and lack of proactive security protocols, the dynamic spectrum access networks are vulnerable to various denial-of-service (DoS) threats. In this paper, we propose a game theoretic framework to study the primary user emulation attack (PUEA) on cognitive radio nodes. A non-cooperative dynamic multistage game between the secondary nodes and the adversaries generating the PUEA is formulated. The pure-strategy and mixed-strategy Nash equilibria for the secondary user and malicious attacker are investigated. Moreover, we propose a novel belief updating system for the secondary user to learn the state of the primary user as the game evolves. Simulation results demonstrate that our proposed belief updating system achieves better performance than other models for the secondary user in terms of greater payoff, lower probability of missing primary user and better robustness to the inaccurate estimation of the primary user's state.

Index Terms

Dynamic spectrum access networks, Primary user emulation attack, Non-cooperative dynamic multistage game, Nash equilibrium, Belief updating system

I. Introduction

SEVERAL recent studies [1], [2] on wireless band availability demonstrate that the wireless spectrum suffers from over utilization in some bands and under utilization in others. This observation underscores the suboptimality of the fixed spectrum assignment policies and has led to the recent spectrum policy reforms by the Federal Communication Commission (FCC). This new policy would allow unused, licensed spectrum bands (*white spaces*) to be used by unlicensed secondary users (SUs) under the provision that they would vacate upon the return of the licensed user (*primary user*). The success of this policy depends on the ability of SU to dynamically identify and access unused spectrum bands, detect the return of primary users and switch to a different band promptly upon sensing the primary user. The recently proposed cognitive radio paradigm is anticipated to make dynamic spectrum access (DSA) a reality.

Unlike conventional radios, cognitive radios can intelligently adjust their transmission/reception parameters based on the interaction with the environment and find the best available spectrum bands to use [3]. However, to avoid the interference to the primary transmission, the SU must periodically sense the spectrum bands and vacate it within a period of time (grace period, e.g., 2 seconds for IEEE 802.22) upon the return of the primary user.

To date, a great deal of research work has been done in energy based spectrum sensing [4]–[7]. Unfortunately, none of these sensing mechanisms can ensure perfect 100% accuracy in the detection outcomes due to the inherent unreliable nature of the wireless medium and varying physical separation between the primary and SUs. Such uncertainties in the licensed user detection make the spectrum sensing vulnerable to *denial-of-service* (DoS) threats in the hostile network environment. While other aspects of cognitive radio networks have received significant attention over the past decade, research in the area of DSA network security is still in its nascence [8]–[13]. In this paper, we study a specific class of DoS attacks in cognitive radio networks known as the primary user emulation attack (PUEA). In this type of attack, one or multiple attacking nodes may transmit in forbidden time slots and effectively emulate the primary user to make the protocol compliant SUs erroneously conclude that the primary user is present [10].

In this work, we study a behavioral model for SUs against the PUEA. A non-cooperative dynamic multistage game between the SU and the malicious attacker (MA) is formulated and both pure-strategy and mixed-strategy Nash equilibria are investigated. In this game, both players are rational and have conflicting goals. More specifically, the SU tries to use the free spectrum band without interfering with the primary user, whereas the MA attempts to monopolize all of the available bandwidth by forcing the SU out of it. In this paper, we use the popularly used energy based primary detection to explain this game formulation.

We also assume that the primary user's arrival schedule is unknown to both SU and MA. Thus, both users need to build a probabilistic model to learn the state of the primary user stage by stage. Therefore, we further introduce analytic models for both players, based on which they adjust their strategies dynamically as the game evolves. In particular, for the SU, we propose a new belief updating system, which incorporates weighting factors to balance the difference of cost between false alarm and misdetection; a confidence factor to reflect its confidence level while making a decision and a threshold for the belief update to protect the primary user. In order to evaluate the efficiency of our proposed system, we compare it with other belief updating systems. Simulation results demonstrate that our proposed belief updating system achieves larger payoffs and better robustness for the SU.

The main contributions of this paper are as follows:

- Formulation of the PUEA in DSA networks as a dynamic multistage game between the SU and MA.
- Investigation of both pure-strategy and mixed-strategy Nash equilibria for the game.
- A new belief updating system proposition for the SU to learn the state of the primary user and effectively defend against the MA.

The rest of this paper is organized as follows. Section II reviews the body of prior work that relates to this paper. In Section III, we introduce the system model. In Section IV, we formulate a non-cooperative dynamic multistage game between the SU and MA and derive Nash equilibrium strategies for both players. In Section V, we propose a new belief updating system for the SU to learn the state of the primary user. Numerical and simulation results are presented in Section VI and the conclusions are drawn in the last section.

II. Related Work

Because of the inherent randomness in the propagation characteristic and consequent randomness in the received signal, existing spectrum sensing algorithms are mostly probabilistic in nature. That is, each decision that the SU can take (primary user present/attack in progress/no activity) is associated with a probability. We define three outcomes at the SU's side: (a) correct decision; (b) false alarm and (c) misdetection. Correct decision refers to the event that the SU's decision exactly matches the current situation. False alarm refers to the event that the SU concludes that the primary user has returned, where as in reality an attack is in progress. Finally, misdetection refers to the event that the SU decides that there is no activity when in fact the primary user is transmitting, thereby causing an unintentional spectrum etiquette violation for the SU. A PUEA is defined as a mechanism by which one or multiple malicious nodes masquerade as the primary user causing SU to make wrong decisions as explained above [14].

Several research works have studied PUEA from the perspectives of both SU and MA. The first analytical model to derive the lower bound for the probability of a successful PUEA was proposed in [11]. In order to thwart this attack, a localization based defense method was developed in [10]. However, these localization mechanisms require a dedicated sensor network which may be too expensive to implement in practice. Jin *et al* [15] proposed hypothesis based approaches to mitigate PUEA using an analytical model for the received power at the SUs' side. Their approach does not assume any prior knowledge about the positions of either the MA or SU. A variance based detection method was proposed in [16] to defend against PUEA by exploring different features of communication channels between the primary user and MAs. Nevertheless, the defense mechanisms proposed in [15] and [16] require SUs to collect lots of sensing samples for analysis, which results in much sensing time and overhead.

The interactions between good SUs and MAs can be analyzed from game theoretic perspective. In [17], a one-stage zero-sum game and a multi-stage stochastic game were formulated between the PUEA jammer and SU in a multi-channel scenario. The optimal strategies for game players, a set of probabilities for channel selection, are derived. A Bayesian game was formulated between two selfish CR in [18], in which radios are unsure whether the other CR is a policy-abiding radio or attacker. The measures to control the occurrence of PUEAs were discussed. However, this paper only studied a one-shot game.

The major differences between our work and existing studies of PUEA are as follows. Firstly, we apply a *game theoretical* approach to study the effects of the PUEA, unlike most existing works purely studying defense mechanism [10], [15], [16]. Then, this work investigates a dynamic competition in *multiple stages* between the SU and MA (as opposite to the one-shot game [18]), which is more realistic because in practical DSA networks, they are always competing over a long period. Also, the game model in [17] does not involve the attackers as the active game player in the multi-stage game but incorporates multiple SUs under the presence of the attack. Finally, we assume that the MA is intelligent and can decide whether to launch PUEA or not with some probability and derive the optimal probability for the MA dynamically as game evolves. This is in contrast to existing works where the MA is assumed to always launch attacks [18] or attacks with some fixed probability [17].

Depending on the motivation of the attack, PUE attacks can be classified into two types:

- *Disruptive attacks*: The driving force behind this class of attacks is simply to disrupt transmissions of protocol compliant SUs rather than any need for additional radio resources for the MA itself. Hence, this attack is similar to the jamming attacks.
- *Selfish attacks*: In contrast to the disruptive attacks, the attack intent in this case is to actually occupy the spectrum band when the SUs vacate it. In that sense, this attack is associated with a positive benefit. In this paper, we explore this class of attacks.

III. System Model

We consider the time epochs in the system model to be divided into discrete periods where each period consists of one sensing slot, τ, and one transmission slot, t-τ [19]. For instance, Fig. 1 shows instances of the absence of the primary, the presence of primary and an instance of a PUEA.

In each period, the SU measures the received signal power during the sensing slot. Also, we assume that SU are not naive, i.e., they are aware of the existence of the MA around the network. However, they neither know the locations of MAs nor the slots when the attacks will be launched. Hence, identifying the presence of the primary user is a challenging task for the SU because the received energy might be from the primary user or the MA or both. Finally, neither the MAs nor the SUs know the arrival schedule of the primary user in any given spectrum band.

The intention of MAs is to launch attacks to disrupt the SU's operation and to grab the spectrum band for themselves. On the other hand, the intention of SUs is to resist attacks and make maximum usage of the available spectrum bands while still obeying the spectrum etiquette. Thus we see that there exists an inherent conflict of interest between the MA and SUs: the SU and MA have completely opposite goals and can not win benefits at the same time.

Each user makes their own decision independently as described below:

- As a SU, the decision problem is to choose whether to stay in the spectrum band or switch to another band such that false alarms and misdetections could be minimized.
- As a MA, the decision problem is to choose whether to generate the PUEA or not in order to maximize the benefits with minimum costs.

A. Notations

Before we begin the analysis, we summarize the specific notations that will be used throughout the paper:

- G – Spectrum gain: Denotes the gain, e.g., bandwidth, for the winning side. If the MA succeed in making the evacuate, this gain is applied to the MA. If the SU chooses not to vacate, this gain is applied to the SU.
- c_s – Switching cost incurred by the SU: Refers to the cost due to the energy consumed in switching from one spectrum band to another.
- c_m – Attack cost incurred by the MA: Refers to the cost due to the energy consumed in generating the PUEA.
- C – Penalty for interference: Measures the penalty incurred by a SU when it unintentionally causes interference to the primary user. One way to penalize an erring SU is to prohibit it from using the given spectrum band for a period of time. In this case, the penalty refers to the loss due to the amount of time for which the good SU is prohibited to access.
- $C^{'}$ – Attack gain: Refers to the incentives obtained by the MA for causing the SU to make a wrong decision to interfere with the primary user.
- R – Reputation factor: Refers to the incentives the SU would get for not causing interference to the primary user. The reputation factor goes up each time the SU leaves when the primary activity begins. Essentially, this is a reward to the SU for not missing the primary user.

For mathematical derivation purpose, we assume that all these parameters are expressed in the same units as the scaling factors. The relationship constraints of these parameters for the SU are as follows:

$$C > G > c_s; R > c_s. \tag{1}$$

Similarly, for the MA, we have:

$$G > c_m; C^{'} > c_m. \tag{2}$$

The reasons for these constraints are explained as follows:

- $C > G$: This constraint makes the SU take every measure to avoid interfering with the primary user. If it is not true, the SU would always stay in the spectrum band because the spectrum gain attained from staying (even when the primary user is active) is greater than the penalty for causing interference to the primary user.
- $G > c_s$: This constraint guarantees that the SU has incentives to access unused licensed bands. If it is not true, the SU would be unwilling to use the licensed bands because the switching cost far

outweighs the gains from accessing to spectrum bands. This, in essence, is the driving force behind DSA networks.

- $R > c_s$: This constraint guarantees that the SU can gain some benefits from successfully avoiding interfering with the primary user. If it is not true, the SU would be unwilling to switch to another band because the overall gain for right decisions is negative.
- $G > c_m, C' > c_m$: This constraint reflects the fact that the MA has incentives to launch attacks. If it is not true, the MA would keep silent because the benefit of a successful attack cannot compensate the attacking cost.

IV. The PUEA as a Non-cooperative Game

In this section, we formulate the PUEA in DSA networks as a non-cooperative dynamic multistage game between the SU and the MA based on our system model. The SU attempts to maximize its spectrum utilization without causing interference to the primary user, while the goal of the MA is to launch the PUEA to gain access to the spectrum band in question.

A. Multistage Dynamic Game Formulation

Based on the system model we described before, we consider each time period in Fig. 1 as a stage in the game. The SU (player s) and MA (player m) play repeatedly and make their moves individually without knowing the other's strategy. Each player has two possible moves for each stage: staying or switching for the SU and attacking or not for the MA. Every time some energy is detected during the sensing slot, player s needs to decide whether to switch to another frequency band or to stay in the same band. This decision, in turn, is to some extent guided by its confidence in its own conclusion about whether the energy is from the primary user or MA. On the contrary, if there is no energy detected, player s will definitely stay to use this free spectrum band.

Note that in many cases, if the primary user returns, the SU needs to find another available band to use. On the other hand, the MA is also spectrum-agile and can still launch PUEA in multiple bands. Thus, our game formulation can be extended to multi-frequency model, in which the SUs and MAs switch among multiple spectrum bands and compete with each other continuously. In this paper, we particularly study one spectrum band scenario and the multi-frequency scenario will be investigated in the future work.

Without loss of generality, we analyze k^{th} stage of the game. In each stage of the game, the primary user can be in one of two states. The *ON* state represents that the presence of the primary transmission in the given spectrum band and the *OFF* state represents the absence of the primary transmission in that band. Let p denote the probability that the primary user is in the *ON* state. This information is not known to either player a priori. Later, we will present mechanisms that each player will use to learn the state of the primary user over several stages of the game. Player s chooses its action from its pure strategy space, A_s={*Switch, Stay*} and player m chooses its action from the pure strategy space, A_m={*Attack, No Attack*}. The payoff matrix of one stage can be defined as given in Table I.

TABLE I

PAYOFF MATRIX OF EACH STAGE OF THE GAME

Case 1: Primary user ON (prob. p)

(s, m)	Attack	No Attack
Switch	R-c_s, $-c_m$	R-c_s, 0
Stay	-C, C'-c_m	-C, 0

Case 2: Primary user OFF (prob. $1-p$)

(s, m)	Attack	No Attack
Switch	$-c_s$, G-c_m	$-$
Stay	G, $-c_m$	G, 0

Note: Each pair of values $(.,.)$ denotes the payoffs obtained by the SU and MAs, in that order, respectively.
The blank hyphen in Case 2 means that case does not exist.

Let us suppose that the primary user is present and the spectrum is not available for both players in this stage (corresponds to Case 1 in Table I). The SU, player s, will gain a reputation factor, R, if it chooses *Switch*. However, it will have to incur a switching cost of c_s. Hence, in this case, the total payoff associated with switching is $R - c_s$ for the SU. Conversely, if player s chooses *Stay* and interferes with the primary, it would suffer a penalty resulting in a total payoff of $-C$. We now look at the MA, player m, for this case. Let us suppose that the MA chooses to attack when the primary is ON. This can happen as the MA might fail to detect the presence of the primary and has no a priori knowledge about the arrival of the primary. If player m chooses to attack, then it incurs a cost c_m. The final payoff would depend on whether the attack was successful to confuse player s, $C' - c_m$, or not, $-c_m$. If player m chooses not to attack, the payoff would be 0.

Now we consider the case when the spectrum band is free (corresponds to Case 2 in Table I). Let player m chooses *Attack* now, if player s chooses *Switch*, player m gains the entire spectrum band and the payoff in this case would be $-c_s$ and $G - c_m$ for player s and m respectively. If player s chooses

Stay, player m does not gain anything, but has to incur the cost of attack, c_m. Then, the payoff would be G and $-c_m$ for player s and m respectively. In contrast, if player m chooses *No Attack*, no energy would be detected and player s will definitely choose *Stay*. Hence, the strategy tuple {*Switch*, *No Attack*} does not arise in this case. Using similar reasoning as in the previous cases, the payoff corresponding to the strategy combination {*Stay*, *No Attack*} would be G and 0 for player s and m respectively.

B. Nash Equilibrium Analysis

For every single stage, the common objective of both players is to maximize their own expected payoff. Hence, we now derive the expected payoffs and Nash equilibrium based on the non-cooperative perspective for both players.

(i) Pure-Strategy Nash Equilibrium Analysis

A pure strategy defines a specific move or action that a player will follow in every possible attainable situation [20]. In the strategy space, the strategy $s_i^{'}$ of player i is said to be strictly dominated by his other strategy s_i if

$$u_i(s_i^{'}, s_{-i}) < u_i(s_i, s_{-i}), \forall s_{-i} \in S_{-i}, \tag{3}$$

and weakly dominated by s_i if

$$u_i(s_i^{'}, s_{-i}) \leq u_i(s_i, s_{-i}), \forall s_{-i} \in S_{-i} \tag{4}$$

with strict inequality for at least one $\forall s_{-i} \in S_{-i}$, where $u_i(.,.)$ represents the payoff of player i given a specific strategy profile and S_{-i} represents the strategy spaces for other players except player i [21].

In a two-player game, if each player has a dominant strategy, the game has a unique pure-strategy Nash equilibrium. Thus, we have the following lemma:

Lemma 1: Since the primary user would not maintain one state (*ON* or *OFF*) all the time, there does not exist an unique pure-strategy Nash equilibrium for both players in this game.

Proof: From Table I, we see that in Case 1, there exists a pure-strategy Nash equilibrium as (*Switch*, *No Attack* | *ON*). Similarly, in Case 2, there also exists a pure-strategy Nash equilibrium as (*Stay*, *No Attack* | *OFF*). However, the Nsah equilibrium strategies for player s are different between two cases. Since the primary user would not maintain one state (*ON* or *OFF*) all the time (e.g., present in the spectrum band with probability p), the game will not uniquely fall into either case. Thus, combining both cases, there is no unique pure-strategy Nash equilibrium for both players in this game. Moreover, we notice that player

m has the same Nash equilibrium strategy in both cases, i.e., *No Attack*. However, based on the definition of dominance, player m has no dominant strategy in either case. Hence, this Nash equilibrium strategy is suboptimal for player m. We now investigate the mixed-strategy Nash equilibrium for this game. ■

(ii) Mixed-Strategy Nash Equilibrium Analysis

A mixed strategy is an assignment of a probability to each pure strategy, which corresponds to how frequently each pure strategy is played [20]. We define the mixed strategy space for the SU as $\{$(*Switch*=θ), (*Stay*=$1-\theta$)$\}$ and the mixed strategy space for the MA as $\{$(*Attack*=π), (*No Attack*=$1-\pi$)$\}$. That is, player s decides to switch with probability θ and stay with $(1-\theta)$ and player m decides to attack with π and not attack with $(1-\pi)$.

By definition, the mixed-strategy Nash equilibrium is a probability distribution tuple, (θ^*, π^*), such that no player can increase its payoff by changing the probability unilaterally. Thus, we have the following lemma:

Lemma 2: The mixed-strategy Nash equilibrium exists in this game if

$$\begin{cases} p < \frac{G+c_s}{C+R+G} \\ pC' > c_m > G(1-p) \end{cases} \tag{5}$$

$$\text{or} \begin{cases} p < \frac{G+c_s}{C+R+G} \\ pC' < c_m < G(1-p) \end{cases} \tag{6}$$

is satisfied.

Proof: Fig. 2 illustrates the game tree corresponding to one single stage of the game and shows all possible actions for both players, based on which we can calculate the expected payoffs for both players as follows:

- Expected payoff for player s:

$$\begin{aligned} E(s) &= p(\theta(R-c_s)-(1-\theta)C) \\ &+ (1-p)(\pi(-\theta c_s+(1-\theta)G)+(1-\pi)G) \\ &= p(\theta(C+R-c_s)-C) \\ &+ (1-p)(G-\pi\theta(G+c_s)) \end{aligned} \tag{7}$$

- Expected payoff for player *m*:

$$\begin{aligned} E(m) &= p(\pi(-\theta c_m + (1-\theta)(C^{'} - c_m))) \\ &+ (1-p)(\pi(\theta(G - c_m) - (1-\theta)c_m))) \\ &= p\pi(C^{'} - c_m - \theta C^{'}) + (1-p)\pi(\theta G - c_m) \end{aligned} \tag{8}$$

Based on the definition of Nash equilibrium, we can compute the Nash equilibrium tuple by imposing $\frac{\partial E(s)}{\partial \theta} = 0$ and $\frac{\partial E(m)}{\partial \pi} = 0$, which means one player's strategy is indifferent to its expected payoff if the other player chooses the equilibrium probability. Thus, the π^* can be calculated as:

$$\begin{aligned} \frac{\partial E(s)}{\partial \theta} &= p(C + R - c_s) - \pi(1-p)(G + c_s) = 0 \\ \Rightarrow \pi^* &= \frac{p(C + R - c_s)}{(1-p)(G + c_s)} \end{aligned} \tag{9}$$

Likewise, the θ^* can be calculated as:

$$\begin{aligned} \frac{\partial E(m)}{\partial \pi} &= \theta G(1-p) - c_m + pC(1-\theta) = 0 \\ \Rightarrow \theta^* &= \frac{pC^{'} - c_m}{p(C' + G) - G} \end{aligned} \tag{10}$$

Thus, the mixed-strategy Nash equilibrium for every single stage is (*Switch with* θ^*, *Attack with* π^* | p). It is noted that the mixed strategy for both players exists only when θ^* and π^* are belonging to (0,1). Therefore, the parameters in expressions of θ^* and π^* should satisfy following constraints:

(1) For the expression of π^* to be valid, we must have

$$\begin{aligned} &0 < \tfrac{p(C+R-c_s)}{(1-p)(G+c_s)} < 1 \\ \Rightarrow\ &0 < p(C + R - c_s) < (1-p)(G + c_s). \end{aligned} \tag{11}$$

According to the relationship of utility parameters in Eqn (1), it is obvious that $p(C+R-c_s) > 0$. Hence, we can derive the constraints for the expression of π^* as follows:

$$p < \frac{G + c_s}{C + R + G}. \tag{12}$$

(2) Similarly, for the expression of θ^* to be valid, we must have:

$$\begin{aligned} & 0 < \tfrac{pC^{'} - c_m}{p(C^{'}+G)-G} < 1 \\ \Rightarrow \quad & 0 < pC^{'} - c_m < p(C^{'} + G) - G \\ \text{or} \quad & p(C^{'} + G) - G < pC^{'} - c_m < 0. \end{aligned} \tag{13}$$

Using the same logic, we derive the constraints for the expression of θ^* as follows:

$$\begin{aligned} & pC^{'} > c_m > G(1-p) \\ \text{or} \quad & pC^{'} < c_m < G(1-p). \end{aligned} \tag{14}$$

■

It is important to emphasize that, in practice, the probability of the primary user being *ON*, p, is not known a priori to either player. Without knowing p, both players cannot analytically figure out their strategies. Hence, in this multistage game, it is necessary for each player to build *an analytic model to update the state of the primary user stage by stage*, which facilitates them to adjust their strategies dynamically.

V. Proposed Belief Updating Mechanism

We denote t_k as the time period corresponding to the k^{th} stage. Let n_e be the total number of stages in which the energy has been detected by the SU in the past, α be the total number of stages in which the primary user was present in the spectrum band. For example, in Fig. 1, at the end of stage 3, the MA has attacked once (stage 3) and the primary has been present in the spectrum band once (stage 2), giving $n_e = 2$ and $\alpha = 1$. Furthermore, for the SU, player s, we denote $\mu_k^s(ON)$ as the belief that the primary user is *ON* at t_k.

However, in a practical cognitive radio environment, the SU does not have accurate information about the source of the received energy (i.e., whether it is from the primary transmission or PUEA). Hence, to establish the belief about the primary user's activity, the SU will try to *estimate* the number of times the primary user has been active in the past stages, denoted by α_e. Thus, instead of the real value of α, the SU will use the estimated value, α_e, to build its belief updating system. The methods described in [22] can be applied in estimating the value of α.

There are several ways for the SU to make the estimation. For example, it potentially possible to incorporate some policing nodes which will periodically monitor spectrum activity across spectrum bands. This information can be used to cross check if there is any policy violation or PUEA. For example, suppose at some time slot t_k, a particular spectrum band was declared to be used by the primary user, but some secondary activity was visible in this $\langle$spectrum, time$\rangle$ tuple, then it is possible to conclude that a PUEA was launched while the good SU vacated the spectrum band and the MAs occupied it. Note however, that this information cannot be available real-time. The good SU can request the past spectrum information from these policy nodes and cross-check this information with results concluded by itself. However, policy nodes cannot monitor every time slot for every single spectrum band because of the sheer number of frequencies and much overhead in monitoring them. Hence, the spectrum information obtained from policy nodes is incomplete and the SU will still have to estimate the value of α based on these information for its real-time use.

In this paper, we are not discussing about the specific estimation measures and the details of estimation mechanisms are beyond the scope of this paper. It is also noted that the estimated value, α_e, would not be equal to the real one, α, and sometime deviates a lot. Thus, in the next section, we will investigate the results under different estimation results, including both under-estimation and over-estimation scenarios.

A. Basic Bayesian Update Function

Let $q(ON|E)$ be the ratio of the total number of stages in which primary user has been *ON* to the total stages when some energy has been detected in the past. Then, $q(ON|E) = \frac{\alpha_e}{n_e}$. Similarly, $q(OFF|E) = 1 - q(ON|E)$. In addition, $q(E|ON)$ is an estimation of the conditional probability that some energy is detected given that the primary user is *ON*. Intuitively, if the primary user is *ON*, the SU will detect some energy. Thus, $q(E|ON) = 1$. Similarly, $q(E|OFF)$ denotes that conditional probability that some energy is detected given that the primary user is not present at the spectrum band. Based on our assumptions, at k^{th} stage, there were $k - 1 - \alpha$ stages where the primary user is in *OFF* state and $n_e - \alpha$ stages where some energy is detected due to PUEAs in the past. Thus, $q(E|OFF)$ is given by:

$$q(E|OFF) = \frac{n_e - \alpha_e}{k - 1 - \alpha_e}. \tag{15}$$

It is noted that player s needs to update the belief only in stages where some energy is detected. If no energy is detected, the SU will definitely stay in the current spectrum band and the belief remains the same as the previous stage as:

$$\mu_k^s(ON) = \mu_{k-1}^s(ON). \tag{16}$$

Now, we look into the stages where some energy is detected. Traditionally, the belief updating rules in game theory are based on the classic Bayesian function [23], [24]. Using the logic of Bayesian function to update the belief of player s receiving some energy at t_k gives:

$$\begin{aligned} \mu_k^s(ON) &= \frac{q(E|ON)\mu_{k-1}^s(ON)}{q(E|ON)\mu_{k-1}^s(ON) + q(E|OFF)\mu_{k-1}^s(OFF)} \\ &= \frac{\mu_{k-1}^s(ON)}{\mu_{k-1}^s(ON) + q(E|OFF)\mu_{k-1}^s(OFF)}, \end{aligned} \tag{17}$$

where $\mu_{k-1}^s(OFF) = 1 - \mu_{k-1}^s(ON)$.

B. Weighted Coefficients and Confidence Factor

A major drawback of using the Bayesian function is that the traditional Bayesian function assumes equal impact for both the *ON* and *OFF* states. This would work well if the effect of making both types of wrong decisions (false alarm and misdetection) were equal. However, in our case, the costs are remarkably different. It is much more important to not miss the primary in order to maintain spectrum etiquette.

For instance, if player s decides that the primary user is *ON* when in fact, the MA was transmitting, then it will not obtain the spectrum gain and pay a switching cost, c_s. Thus, the SU will lose a total of $G + c_s$ in the event of false alarm. However, if it misses the primary user, then it will also not receive a spectrum gain but interfere with the primary user, which incurs a penalty, C. Thus, the SU will lose a total of $G+C$ in the event of misdetection. *This asymmetry in the loss is important to assure that the spectrum etiquette is given the highest importance*. Therefore, we define two normalized weighting coefficients, w & $w^{'}$ and incorporate them in our system to emphasize the priority of primary transmission:

$$w = \frac{C + G}{C + G + c_s + G}, \tag{18}$$

$$w^{'} = \frac{c_s + G}{C + G + c_s + G}, \tag{19}$$

where w represents the factor for the *ON* state and $w^{'}$ represents the factor for the *OFF* state. Since the penalty for interference, C, is much greater than the switching cost, c_s, it is easy to know that $w > w^{'}$.

Now, the belief updating function by incorporating the weighted factors is given by:

$$\mu_k^s(ON) = \frac{w\mu_{k-1}^s(ON)}{w\mu_{k-1}^s(ON) + w' q(E|OFF)\mu_{k-1}^s(OFF)}. \tag{20}$$

It is noted that in this weighted Bayesian function, the belief at t_k mainly depends on that at t_{k-1}. However, the confidence of the player s in terms of the number of stages it has experienced would also be an important factor in the belief updating system. Hence, it is necessary to take into account how confident player s would be while updating the belief. In other words, the more evidence the player has, the more confidence it has, e.g., in the multistage game, the player at stage 100 has gathered much more experience than at stage 10, and thus the belief at stage 100 is more fine-tuned. Therefore, in order to evaluate the confidence level of player s in the course of updating the belief, we introduce a confidence factor at stage k, indexed by $\phi(k)$.

The confidence factor in our situation refers to the degree to which player s can confidently learn the state of the primary user. It also reflects whether player s possesses sufficient experience from past stages. Consequently, the confidence factor is directly influenced by the number of times in the past when some energy has been detected, n_e. The larger this value is, the more experience the SU has gained resulting in a higher confidence factor. It is also to be noted that the initial value or $\mu_{k-1}^s(ON) = \mu_{k-1}^s(OFF) = 0.5$ represents that the belief update system does not have a priori knowledge about the behavior of the primary. Under this situation, player s is most uncertain about the primary user's state, which would result in the lowest confidence factor. Thus, we define the confidence factor of player s at t_k as follows:

$$\phi(k) = 1 - \frac{\mu_{k-1}^s(ON)(1-\mu_{k-1}^s(ON))}{n_e}. \tag{21}$$

Combine the confidence factor with weighted Bayesian function in Eqn (20), we propose a new belief updating system for the SU as follows:

$$\mu_k^s(ON) = \frac{\phi(k)\cdot w\mu_{k-1}^s(ON)}{w\mu_{k-1}^s(ON) + w' q(E|OFF)\mu_{k-1}^s(OFF)}. \tag{22}$$

On the other hand, for the MA, player m, we denote $\varphi_k^m(ON)$ as its prediction of the probability of the primary user being *ON* at t_k, which is computed by summarizing its observations from all past $k-1$ stages as:

$$\varphi_k^m(ON) = \frac{\alpha}{k-1}. \tag{23}$$

Note that if there is no primary user and only a MA exists in the spectrum band, the system becomes a SU-MA pair. However, the SU has no sure way of knowing about the absence of the primary user because there is no information exchange between them. In this scenario, the belief of primary user being *ON* for the SU, $\mu_k^s(ON)$, will gradually converge to 0. Consequently, the SU will eventually realize that the received energy is most likely from the MA rather than the primary user. Hence, in the case without the primary user, the estimated value α_e would get smaller and if the game is infinite, α_e will eventually converge to 0.

C. The Strategies for Both Players at t_k

Since both the SU and MA are independent of each other and cannot know the strategy the opponent will take at any given time, the safest way for them to choose their actions is to follow the Nash equilibrium strategy based on their own belief about the state of the primary user. Moreover, in order to protect the primary transmission, it is necessary for the SU to set a threshold of the belief. That is, if the belief that the primary user is present exceeds the threshold, the SU will evacuate to reduce the risk of interfering with the primary user. From Eqn (10), we can see that θ^* has a pole at the point $p = \frac{G}{G+C'}$. Thus, the threshold can be set as $\frac{G}{G+C'}$.

Based on analytic models for both players, the strategic tuple at t_k can then be extended from Eqns (9) and (10) and presented as follows:

- For the SU, player s:

  (i) If no energy is detected at stage t_k,

$$\theta^*(t_k) = 0. \tag{24}$$

  (ii) If some energy is detected at stage t_k,

$$\theta^*(t_k) = \begin{cases} \frac{\mu_k^s(ON)C' - c_m}{\mu_k^s(ON)(C'+G)-G}, & \mu_k^s(ON) < \frac{G}{G+C'} \\ 1, & \mu_k^s(ON) \geq \frac{G}{G+C'} \end{cases} \tag{25}$$

- For the MA, player m:

$$\pi^*(t_k) = \frac{\varphi_k^m(ON)(C+R-c_s)}{(1-\varphi_k^m(ON))(G+c_s)}, \tag{26}$$

where $\theta^*(t_k)$ and $\pi^*(t_k)$ represent the switching probability and attack probability for the SU and MA at stage t_k respectively.

Therefore, the expected payoffs at t_k for both players are as follows:

$$\begin{aligned} E_s(t_k) &= p(\theta^*(t_k)(C+R-c_s)-C) \qquad (27)\\ &+ (1-p)(G-\pi^*(t_k)\theta^*(t_k)(G+c_s)), \\ E_m(t_k) &= p\pi^*(t_k)(C^{'}-c_m-\theta^*(t_k)C^{'}) \qquad (28)\\ &+ (1-p)\pi^*(t_k)(\theta^*(t_k)G-c_m), \end{aligned}$$

from which we can calculate the average per-stage payoffs over N stages for both players as follows:

$$\bar{E}_s = \frac{1}{N}\sum_{k=1}^{N} E_s(t_k), \qquad (29)$$

$$\bar{E}_m = \frac{1}{N}\sum_{k=1}^{N} E_m(t_k). \qquad (30)$$

D. The Definition of Probability of Missing Primary user

Another important task for the SU is to minimize the interference to the primary user because the protection of the primary transmission is the first priority in the DSA network protocols. Hence, we define the probability of missing primary user, P_m, to evaluate the interference impact to the primary user for the proposed belief updating system, which is given by:

$$P_m = \frac{\text{Number of times interfering with the primary user}}{\text{Number of times the primary user is in } \textit{ON} \text{ state}}. \qquad (31)$$

In the next section, we will conduct simulations to evaluate the performance of the proposed belief updating system in terms of P_m.

E. Discussions

The proposed belief updating system is a probabilistic updating method and based on the Bayesian function and the estimation of previous results in the past stages. It estimates the posterior probability for the state of the primary user. This is essentially different from traditional machine learning algorithms (e.g., reinforcement learning, neural network and SVM) which are built on training data rather than using learn-on-the-air mechanism like our proposed method.

Moreover, the one SU-MA game studied in this paper can also be extended to the case of multiple users model where multiple SUs and MAs switch among multiple free spectrum bands. Under this situation, there exist two attack mechanisms, coordinated attacks where all MAs will launch PUEA cooperatively to maximize overall utilities, and distributed attacks where each MA launches PUEAs without coordination to maximize its individual utility. The extension to multiple users will be studied in the future work.

VI. Numerical and Simulation Results

In this section, we first present the numerical results of mixed-strategy Nash equilibrium for a single stage and then conduct the simulations for the multistage case. The simulations are conducted in MATLAB environment and the results are averaged over 100,000 Monte Carlo simulations.

The parameters used for the simulations are: $G = 50$, $C = 60$, $R = 20$, $C^{'} = 40$, $c_s = 15$, $c_m = 25$. We use three different values values of p: $p = 0.05$, $p = 0.2$ and $p = 0.35$ in the simulations. Note that all parameters satisfy the constraints defined in Eqns (1) and (2).

A. Simulations for the Single Stage

We first conduct the numerical analysis for the single stage with $p = 0.2$. From the Eqns (7) and (8), we calculate the theoretical expected payoffs for both players at Nash equilibrium point, which are *E(s)*=28 and *E(m)*=0 for the SU and MA respectively. Next, we obtain experimental values of the expected payoff for each player with the strategy of the other player fixed. Fig. 3 shows the expected payoff of the SU with the probability of *Switch* varying between 0 and 1 while the MA adheres to its Nash equilibrium strategy. Similarly, Fig. 4 shows the expected payoff of the MA with the probability of *Attack* varying between 0 and 1 while the SU adheres to its Nash equilibrium strategy. As expected, the payoffs of both players obtained through numerical analysis perfectly match the theoretically calculated payoffs at Nash equilibrium point.

B. Simulations of Multistage Game with Belief Update

We now consider the multistage game. Since both players have no prior knowledge about the state of the primary user, both players use 0.5 for the initial value of the estimate of the state of the primary user. That is $\mu_1^s(ON) = \varphi_1^m(ON) = 0.5$. As the game evolves, each player updates its belief based on its own analytic model. In order to evaluate the efficiency of our proposed belief updating system for the SU, we compare it with two other systems: traditional Bayesian belief update [25] and uncertainty-related belief update [26].

In terms of the average per-stage payoff for the SU, we first assume that the SU makes the accurate estimation of the value of α, i.e., $\alpha_e = \alpha$. Later, we will relax this assumption and look into cases where the SU does not have an accurate estimation of the value of α.

First of all, we study the relationship between the primary user's arrival probability and player s's average per-stage payoff. Fig. 5 shows the average per-stage payoff for player s calculated over 100 stages. As evident from this figure, the simulation results are very *close* to the theoretical results. The average per-stage payoff for player s decreases monotonically with the increase in the probability of the primary user being *ON*. This is because higher probability for the primary user being *ON* implies less opportunity for the SU to use the spectrum band.

Fig. 6 shows the average per-stage payoffs of the SU, player s, with three different belief updating systems for different values of p. As illustrated in this figure, our proposed system clearly outperforms the other existing systems in the sense of achieving larger average per-stage payoff for the player s. It is also noted that our proposed belief updating mechanism performs much better than the traditional Bayesian belief update and the uncertainty-related belief update system even when the number of stages played is small. This is because, our proposed belief updating system which incorporates the weighting and confidence factors, allows the SU to obtain better belief at the initial stages.

Next, we compare the probability of missing the primary user for these three systems based on Eqn (31). Fig. 7 shows the probability of missing primary user calculated over 100 game stages with different probabilities of the primary user being *ON* for three systems. As evident, our proposed belief updating system results in much smaller miss probability than the other two systems. This advantage is achieved by incorporating the normalized factors w and $w^{'}$ and setting a threshold for the SU in our system to reduce the risk of interference to the primary user. Another observation from Fig. 7 is that since the primary user is transmitting in the spectrum band more often, the chances of the secondaries interfering with the primary user also increase, thereby increasing this missing probability.

TABLE II
THE PROBABILITY OF MISSING PRIMARY USER FOR THE SU IN CASE OF UNDER-ESTIMATION OF α

Belief updating system	α	0.9α	0.8α	0.7α	0.6α
Proposed system	0.0445	0.0456	0.0460	0.0465	0.0471
Uncertainty	0.1128	0.1201	0.1259	0.1311	0.1373
Bayesian	0.1643	0.1673	0.1708	0.1745	0.1779

Now, we look at the cases of over-estimation and under-estimation of α, i.e., $\alpha_e \neq \alpha$, from the perspective of player s. Fig. 8 plots the average per-stage payoff for the player s with $p = 0.2$. As shown in this figure, with the increase in the over-estimation rate, the SU's average per-stage payoff will decrease. This is because, the over-estimation of α results in greater $\mu_k^s(ON)$ for player s as well as its switching probability. As a consequence, the SU will lose some spectrum opportunities because of its "over cautiousness". On the other hand, Fig. 9 shows the average per-stage payoffs for player s in the under-estimation scenario. As shown in Fig. 9, as the under-estimation rate becomes smaller, the average per-stage payoffs for the other systems gradually approach our proposed system. However, the corresponding probabilities of missing primary user (shown in Table II) is significantly larger for the other systems, implying that the proposed system has a much better chance of adhering to the spectrum etiquette than the other systems. Moreover, it is observed from Fig. 8, Fig. 9 and Table II that the proposed belief updating system is demonstrated to be the most robust to over-estimation and under-estimation cases for the SU in terms of the magnitude of difference in both the payoff and the probability of missing the primary user.

VII. Conclusion

In this paper, we formulated the PUEA in DSA networks as a non-cooperative dynamic multistage game between the SU and the MA. Assuming that the arrival schedule of the primary user in the spectrum band is unknown to both players, we derived the mixed strategy for each player that would achieve the Nash equilibrium. We also proposed a novel belief updating system for the SU to defend against the MA in the multistage version of the game, based on which the SU can learn the state of the primary user and intelligently adjust its strategy stage by stage. Numerical and simulation results demonstrated that by using the proposed belief updating system, the SU can obtain larger payoffs and less probability of missing primary user compared to traditional Bayesian and uncertainty-related updating models. We also showed that the proposed system achieves better robustness to errors in estimating the state of the primary user.

Acknowledgement

This research is partially funded by NIJ # 2009-92667-NJIJ, NSF # 0916180 and PSC-CUNY Award # 60079-40 41.

References

[1] F. C. C., "Spectrum policy task force report," *IEEE Trans. Information Forensics and Security*, pp. 02–155, Nov 2002.

[2] F. C. C., “In the matter of unlicensed operation in the TV broadcast bands,” *Second Report and Order and Memorandum Opinion and Order*, no. FCC-08-260A1, Nov. 2008.

[3] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, “Next generation/dynamic spectrum access/cognitive radio wireless networks: a survey,” *Comput. Netw.*, vol. 50, no. 13, pp. 2127–2159, 2006.

[4] G. Ganesan and Y. Li, “Cooperative spectrum sensing in cognitive radio networks, Part I: Two user networks and Part II: Multiuser networks,” *IEEE Trans. Wireless Communication*, vol. 6, no. 6, June 2007.

[5] P. De and Y.-C. Liang, “Blind spectrum sensing algorithms for cognitive radio networks,” *IEEE Trans. Vehicular Technology*, vol. 57, no. 5, pp. 2834–2842, Sept. 2008.

[6] A. Ghasemi and E. Sousa, “Interference aggregation in spectrum-sensing cognitive wireless networks,” *Selected Topics in Signal Processing, IEEE Journal of*, vol. 2, no. 1, pp. 41–56, Feb. 2008.

[7] Z. Quan, S. Cui, A. H. Sayed, and H. V. Poor, “Optimal multiband joint detection for spectrum sensing in cognitive radio networks,” *IEEE Trans. Signal Processing*, vol. 57, no. 3, pp. 1128–1140, March 2009.

[8] Q. H. Mahamound, *Cognitive Networks: Towards Self-Aware Networks, Chapter 11, pp. 271-289*. Wiley Press, Sep. 2007.

[9] G. Jakimoski and K. P. Subbalakshmi, “Denial-of-service attacks on dynamic spectrum access networks,” *Communications Workshops, 2008. ICC Workshops ’08. IEEE International Conference on*, pp. 524–528, May 2008.

[10] R. Chen, J.-M. Park, and J. Reed, “Defense against primary user emulation attacks in cognitive radio networks,” *Selected Areas in Communications, IEEE Journal on*, vol. 26, no. 1, pp. 25–37, Jan. 2008.

[11] S. Anand, Z. Jin, and K. P. Subbalakshmi, “An analytical model for primary user emulation attacks in cognitive radio networks,” *Proceedings, IEEE DySPAN 2008*, Oct. 2008.

[12] T. Clancy and N. Goergen, “Security in cognitive radio networks: Threats and mitigation,” *Cognitive Radio Oriented Wireless Networks and Communications, 2008. CrownCom 2008. 3rd International Conference on*, pp. 1–8, May 2008.

[13] Y. Tan, S. Sengupta, and K. P. Subbalakshmi, “Coordinated Denial-of-Service attacks in IEEE 802.22 networks,” *IEEE International Conference on Communications, ICC 2010*, May 2010.

[14] R. Chen and J.-M. Park, “Ensuring trustworthy spectrum sensing in cognitive radio networks,” *Networking Technologies for Software Defined Radio Networks, 2006. SDR ’06.1st IEEE Workshop on*, pp. 110–119, Sept. 2006.

[15] Z. Jin, S. Anand, and K. P. Subbalakshmi, “Detecting primary user emulation attacks in dynamic spectrum access networks,” *IEEE International Conference on Communications, ICC 2009*, pp. 1–5, June 2009.

[16] Z. Chen, T. Cooklev, C. Chen, and C. Pomalaza-Raez, “Modeling primary user emulation attacks and defenses in cognitive radio networks,” in *IEEE 28th International Performance Computing and Communications Conference (IPCCC), 2009*, Dec. 2009, pp. 208–215.

[17] H. Li and Z. Han, “Dogfight in spectrum: Jamming and anti-jamming in multichannel cognitive radio systems,” pp. 1–6, Dec. 2009.

[18] R. Thomas, R. Komali, B. Borghetti, and P. Mahonen, “A bayesian game analysis of emulation attacks in dynamic spectrum access networks,” in *IEEE Symposium on New Frontiers in Dynamic Spectrum, 2010*, Apr. 2010, pp. 1–11.

[19] Y.-C. Liang, Y. Zeng, E. Peh, and A. T. Hoang, “Sensing-throughput tradeoff for cognitive radio networks,” *IEEE Trans. Wireless Communications*, vol. 7, no. 4, pp. 1326–1337, April 2008.

[20] D. Fudenberg and J. Tirole, *Game Theory*. MIT press, 1991.

[21] M. Felegyhazi and J.-P. Hubaux, “Game Theory in Wireless Networks: A Tutorial,” Tech. Rep., 2006.

[22] T. Schonhoff and A. Giordano, *Detection and Estimation Theory*. Prentice Hall, 2006.

[23] Y. Liu, C. Comaniciu, and H. Man, “A bayesian game approach for intrusion detection in wireless ad hoc networks,” in *GameNets ’06: Proceeding from the 2006 workshop on Game theory for communications and networks*, 2006, p. 4.

[24] H. Otrok, N. Mohammed, L. Wang, M. Debbabi, and P. Bhattacharya, “A moderate to robust game theoretical model for intrusion detection in manets,” *Networking and Communications, 2008. WIMOB ’08.*, pp. 608–612, Oct. 2008.

[25] R. B. Myerson, *Game Theory: Analysis of Conflict*. Harvard University, 1997.

[26] F. Li and J. Wu, “Hit and run: A bayesian game between malicious and regular nodes in manets,” *SECON ’08. 5th Annual IEEE Communications Society Conference on*, pp. 432–440, June 2008.

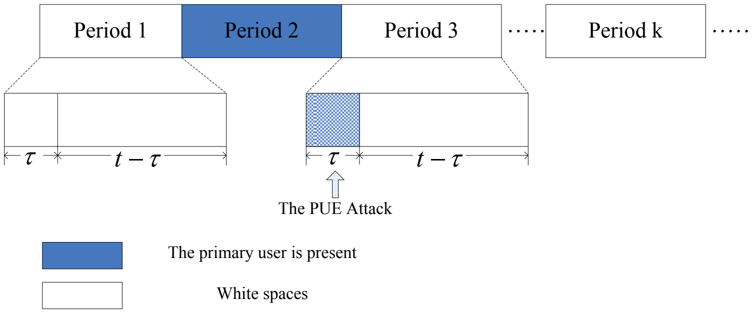

Fig. 1. Discrete sensing/transmission periods. In each period, τ and $t-\tau$ represent the sensing slot and transmission slot respectively. Note that the shaded period indicates that the primary user is using the spectrum band and the grid filled sensing slot indicates that the attack has been launched.

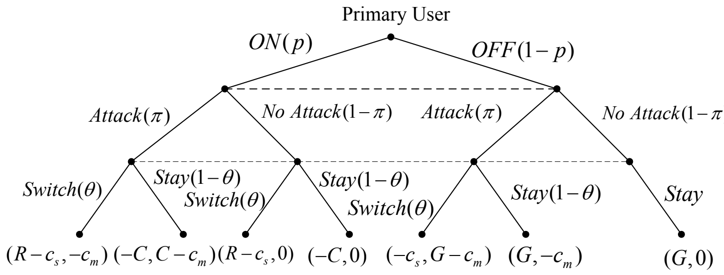

Fig. 2. The game tree corresponding to one single stage of the game, showing all possible actions for both players

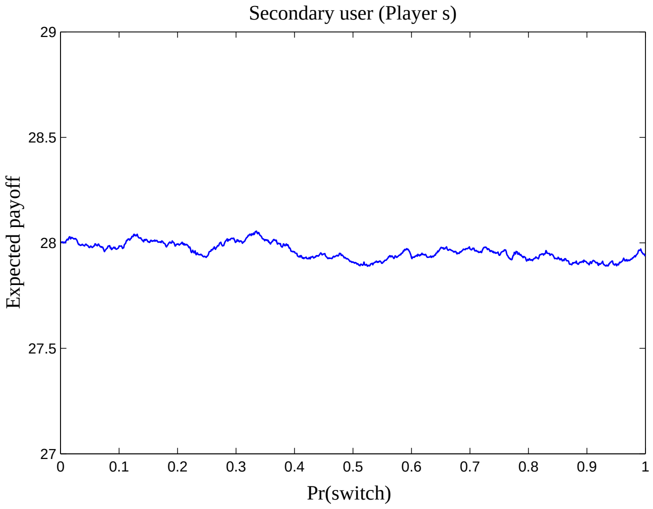

Fig. 3. The expected payoff in the single stage game for player s with varying probability of *Switch* , Pr(switch), while player m adheres to its Nash equilibrium strategy. Note that p=0.2.

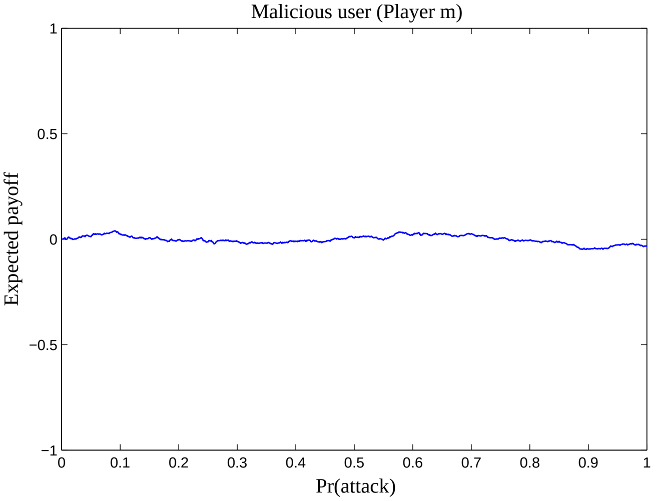

Fig. 4. The expected payoff in the single stage game for player m with varying probability of *Attack*, Pr(attack), while player s adheres to its Nash equilibrium strategy. Note that p=0.2.

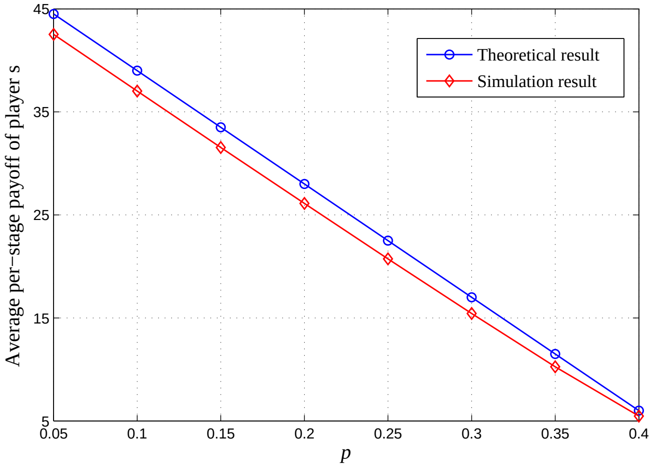

Fig. 5. The average per-stage payoffs calculated over 100 stages, for player s corresponding to different values of the probability of the primary user being *ON*, p.

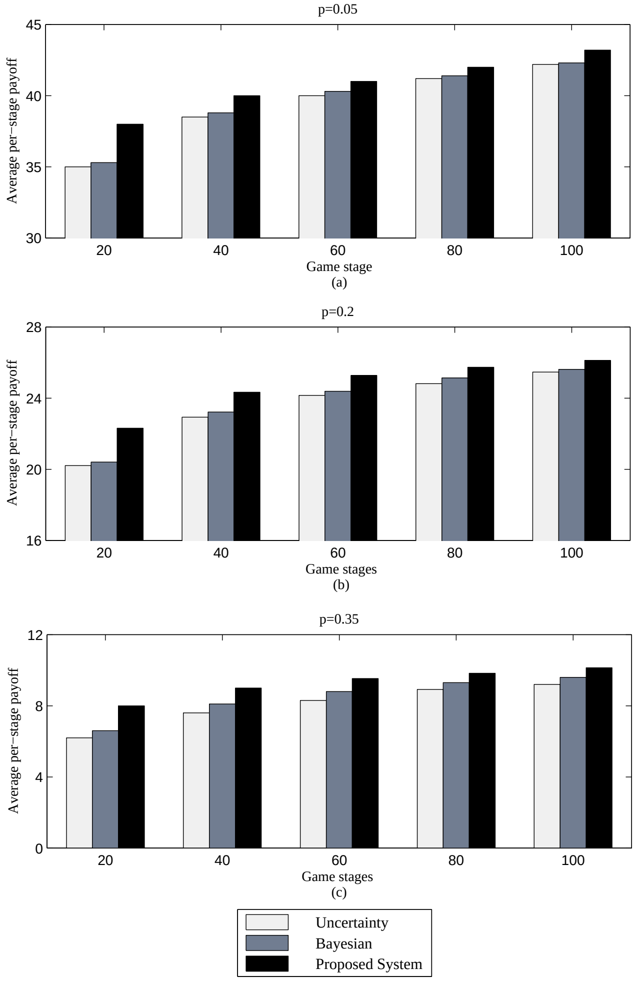

Fig. 6. Comparison of experimentally obtained average per-stage payoffs for player s based on three belief updating systems for three different values of p. (a) $p = 0.05$, $E(s)$=44.5; (b) $p = 0.05$, $E(s)$=28; (c) $p = 0.05$, $E(s)$=11.5. $E(s)$ represents the theoretical average per-stage payoff for player s.

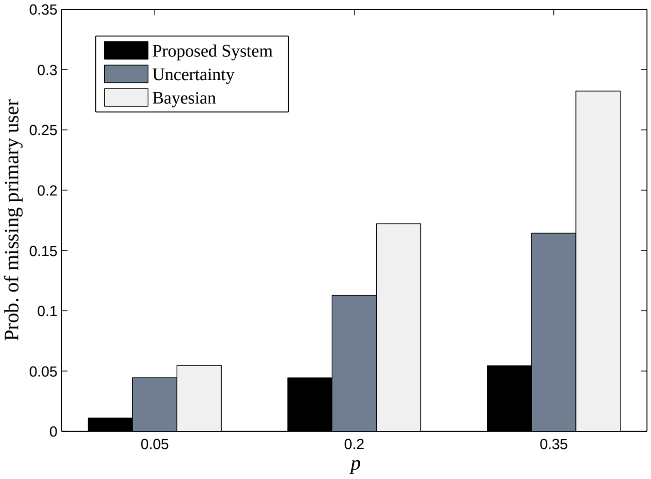

Fig. 7. The probability of missing primary user for different probabilities of the primary user being *ON*, p, over 100 game stages.

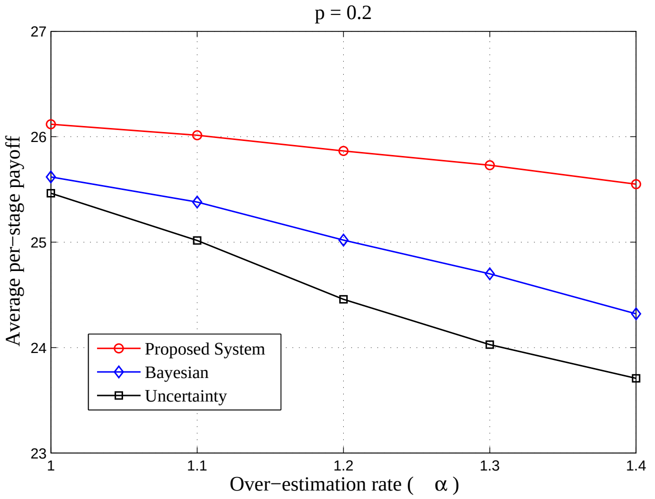

Fig. 8. Comparison of the the average per-stage payoffs of three belief updating systems for player s in case of over-estimation of α. Note that $p = 0.2$ and the results are calculated over 100 stages.

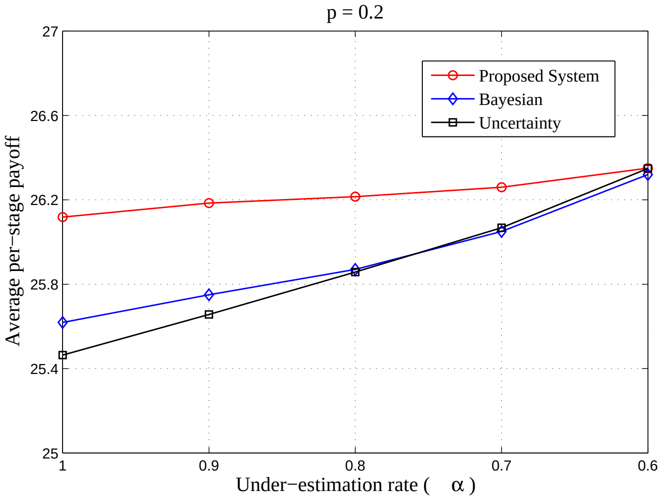

Fig. 9. Comparison of the the average per-stage payoffs of three belief updating systems for player s in case of under-estimation of α. Note that $p = 0.2$ and the results are calculated over 100 stages.