This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE Globecom 2010 proceedings.

Optimal Defense Against Jamming Attacks in Cognitive Radio Networks using the Markov Decision Process Approach

Yongle Wu, Beibei Wang, and K. J. Ray Liu
Department of Electrical and Computer Engineering,
University of Maryland, College Park, MD 20742, USA.
{wuyl, bebewang, kjrliu}@umd.edu

***Abstract*—Cognitive radio technology has become a promising approach to increase the efficiency of spectrum utilization. Since cognitive radio users are vulnerable to malicious attacks, security countermeasures are crucial to the successful deployment of cognitive radio networks in the future. In this paper, we focus on defending against the jamming attack, one of the major threats to cognitive radio networks, where several malicious attackers intend to jam the secondary user's communication link by injecting interference. We model this scenario into a jamming game, and derive the optimal strategy through the Markov decision process approach. Furthermore, a learning scheme is proposed for the secondary user to observe the wireless environment and estimate parameters such as primary users' access pattern and the number of attackers. Finally, simulation results are presented to verify the performance.**

I. Introduction

As a revolutionary communication paradigm that enables more efficient and intelligent usage of spectrum resources, cognitive radio technology [1] has been receiving a growing attention in the last decade. In a cognitive radio network, unlicensed users (secondary users) are allowed to access licensed bands on a non-interference basis to legacy spectrum holders (primary users). Since secondary users usually compete for limited spectrum resources and are capable of acting adaptively and intelligently, it is reasonable to assume they are selfish in nature, and hence game theory has been widely applied as a flexible and proper tool to model and analyze their behavior in the network (see [2] and references therein).

Cognitive radio networks are extremely vulnerable to malicious attacks, partly because secondary users do not own the spectrum, and hence their opportunistic access cannot be protected from adversaries. Moreover, malicious attackers are also able to take advantage of technology evolution, such as flexible software/hardware and capabilities of learning and reasoning, which make them even more powerful and dreadful than before. As a result, security countermeasures are crucial to the successful deployment of cognitive radio networks. For instance, in [3], the primary user emulation attack was described and a transmitter verification scheme was proposed to test whether the given signal came from a primary user; [4] discussed one kind of attack where malicious users attempted to mislead the learning process of secondary users; a Hammer model was employed to identify, analyze and assess denial of service attacks in [5]; in [6], a malicious user reporting false sensing results would be found and excluded from the collaborative spectrum sensing when the calculated "suspicious" level was beyond a certain threshold.

In this paper, we mainly focus on the jamming attack, one of the major threats to cognitive radio networks, where several malicious attackers intend to interrupt the communications of secondary users by injecting interference. Considering a situation where a secondary user could hop across multiple bands in order to reduce the probability of being jammed, we derive the optimal defense strategy for the secondary user using the Markov decision process (MDP) approach [7]. The optimal strategy strikes a balance between the cost associated with hopping and the damage caused by attackers.

Moreover, in order to determine the optimal strategy, the secondary user needs to know some information, e.g., the number of attackers, which may not be available directly. The secondary user has to observe and learn from the environment. Therefore, we propose a learning process in this paper that the secondary user estimates the useful parameters based on past observations using the maximum likelihood estimation (MLE).

The rest of this paper is organized as follows. In Section II, some related works are briefly reviewed. In Section III, the system model is described. The optimal defense strategy with perfect information is derived in Section IV, while the learning algorithm is discussed in Section V. Section VI presents simulation results, and Section VII concludes the paper.

II. Related Works

There have been quite a few papers on jamming attacks in wireless ad hoc networks, such as [8] and [9]. A jamming game with transmission costs was formulated in [8], and the blocking probability was analyzed for different kinds of attack strategies and defense strategies in [9]. However, the problem becomes more complicated in a cognitive radio network where primary users' access has to be taken into consideration.

In the context of cognitive radio networks, [10] modeled an attack-and-defense problem as a stochastic game where secondary users reserved several bands to transmit data or control messages. [11] derived the optimal defense strategy when the secondary user equipped with multiple radios could access several bands simultaneously. However, in this paper, we consider the scenario with a single-radio secondary user, and hence the defense strategy is to hop across different bands.

978-1-4244-5637-6/10/$26.00 ©2010 IEEE

Hopping as a defense strategy was also considered in [12] which derived the Nash equilibrium in a one shot game and applied this equilibrium strategy to a multi-stage game. This is different from our approaches. In our work, we explicitly model transitions in time as Markov chains, take the cost and damage into account in addition to communication gains, and further develop a learning process to estimate unknown parameters.

III. System Model

Consider the situation where a secondary user (e.g., a secondary base station) opportunistically accesses one of the predefined M licensed bands, and m malicious attackers intend to jam the secondary user's communications.

Assume each licensed band is time-slotted and the access pattern of primary users can be characterized by an ON-OFF model [13]. As shown in Fig. 1, one band can either be busy (ON) or idle (OFF) in one time slot, and the state can be switched from ON to OFF (or from OFF to ON) with a transition probability α (or β). We assume all bands share the same channel model and parameters, but different bands are used by independent primary users.

In order to avoid interference to primary users, the secondary user has to synchronize with the primary network, and detect the presence of the primary user at the beginning of each time slot, as shown in Fig. 2. We assume the secondary user is equipped with a single radio, and hence can only sense and use one of the M candidate bands at any time slot. When the primary user is absent in that band, the secondary user can utilize the spectrum yielding a communication gain R; otherwise, the secondary user has to tune his/her radio to another band and detect the availability of that band at the beginning of the next time slot. The cost associated with this spectrum hopping is denoted by C.

We assume there are $m\,(m \geq 1)$ malicious single-radio attackers attempting to jam the secondary user's communication link. Because primary users' usage of spectrum is enforced by their ownership of bands, attackers do not want to interfere with primary users either. We assume the attackers use energy detectors which cannot distinguish primary users' or secondary users' signals. As illustrated in Fig. 2, an attacker tunes the radio to one of the bands at the beginning of a time slot to sense the presence of the primary user. If the primary user is absent, the attacker continues to detect whether the secondary user is utilizing this band. On finding the secondary user, the attacker will immediately inject jamming power which makes the secondary user fail to decode data packets. When all the attackers coordinate to maximize the damage, they detect m channels in a time slot. We assume that the secondary user suffers from a significant loss L when jammed, since normal communication is interrupted and considerable effort is needed to reestablish the link.

When there are no malicious attackers, considering the hopping cost C, the secondary user should always stay in a fallow licensed band until the primary user reappears. However, in the presence of attackers, the longer the secondary user stays in a band, the higher risk to be exposed to attackers. In other words, sometimes proactive hopping to another band may help to hide from attackers.

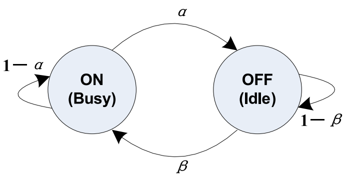


Fig. 1. An ON-OFF model for primary users' spectrum usage.

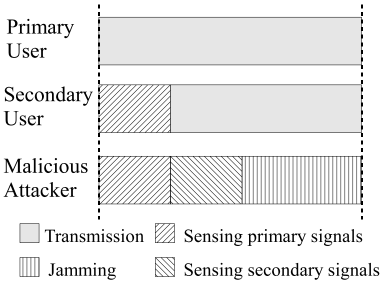


Fig. 2. The time slot structure where the secondary user and the attacker are synchronized to the primary network. The secondary user can access the band if no primary activity is sensed; the attacker senses secondary signals after no primary signals are detected, and jams the band if finding the secondary user.

Therefore, this situation can be modeled into a multi-stage game in which players are the secondary user and m malicious attackers. At the end of each time slot, the secondary user decides either to *stay* or to *hop* for the next time slot, based on observation of the current and past slots. The secondary user receives an immediate payoff $U(n)$ in the nth time slot, which is the gain minus the cost and damage,

$$\begin{aligned} U(n) =& R \cdot 1(\text{Successful transmission}) - L \cdot 1(\text{Jammed}) \\ & - C \cdot 1(\text{Choosing the action `hop'}), \end{aligned} \quad (1)$$

where $1(\cdot)$ is an indicator function returning 1 when the statement in the parenthesis holds true and 0 otherwise. The average payoff $\overline{U}$, which the secondary user wants to maximize but malicious attackers want to minimize, is a discounted sum of immediate payoffs,

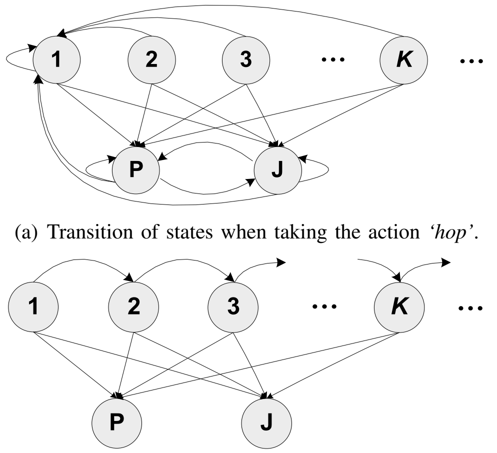
$$\overline{U} = \sum_{n=1}^{\infty} \delta^n U(n), \quad (2)$$

where the discount factor $\delta\,(0 < \delta < 1)$ measures how much the secondary user values a future payoff over the current one.

IV. Optimal Strategy with Perfect Knowledge

In this section, we derive the optimal strategy that the secondary user should adopt when perfect information is available. Learning for unknown parameters will be discussed in the next section.

In order to catch the secondary user as soon as possible, the attackers should coordinately tune their radios randomly to m undetected bands in each time slot, until this process starts over when either all bands have been sensed or the secondary user has been found and jammed. We will derive the optimal



(a) Transition of states when taking the action *'hop'*.

(b) Transition of states when taking the action *'stay'*.

Fig. 3. Markov chains of state transitions when different actions are taken.

defense strategy for the secondary user assuming that attackers stick to this attack strategy.

Under the assumption of the fixed attack strategy, the jamming game can be reduced to a Markov decision process, since only the defense strategy needs to be taken into account. In what follows, we first show how to model the scenario as an MDP, and then solve it using standard approaches.

A. Markov Models

At the end of the nth time slot, the secondary user observes the state of the current time slot $S(n)$, and chooses an action $a(n)$, that is, whether to tune the radio to a new band or not, which takes effect at the beginning of the next time slot. If the primary user occupied the band or the secondary user was jammed in the nth time slot, denoted by $S(n) = P$ and $S(n) = J$, respectively, the secondary user has to hop to a new band, i.e., $a(n) = h$; otherwise, the secondary user has transmitted a packet successfully in the time slot, and possible actions are 'to hop' ($a(n) = h$) and 'to stay' ($a(n) = s$). If this is the Kth consecutive slot with successful transmission in the same band, the state is denoted by $S(n) = K$. For brevity, we drop the time index n wherever there is no room for ambiguity in the rest of the paper. According to (1), the immediate payoff depends on both the state and the action,

$$U(S, a) = \begin{cases} R, & \text{if } S \in \{1, 2, 3, \ldots,\}, a = s; \\ R - C, & \text{if } S \in \{1, 2, 3, \ldots,\}, a = h; \\ -L - C, & \text{if } S = J; \\ -C, & \text{if } S = P. \end{cases} \quad (3)$$

The transition of states can be described by Markov chains, as shown in Fig. 3. The transition probabilities depend on which action has been taken. Hence, we use $p(S'|S, h)$ and $p(S'|S, s)$ to represent the transition probability from an old state S to a new state S' when taking the action h and the action s, respectively.

If the secondary user hops to a new band, transition probabilities do not depend on the old state, and furthermore, the only possible new states are P (the new band is occupied by the primary user), J (transmission in the new band is detected by an attacker), and 1 (successful transmission begins in the new band). When the total number of bands M is large, i.e., $M \gg 1$, we can assume that the probability of primary user's presence in the new band equals the steady-state probability of the ON-OFF model in Fig. 1, neglecting the case that the secondary user hops back to some band in very short time,

$$p(P|S, h) = \frac{\beta}{\alpha + \beta} \triangleq \gamma, \ \forall S \in \{P, J, 1, 2, 3, \ldots,\}. \quad (4)$$

Provided that the new band is available, the secondary user will be jammed with the probability m/M, since each attacker detects one band without overlapping. As a result, transition probabilities are

$$\begin{aligned} p(J|S, h) &= (1 - \gamma)\frac{m}{M}, \ \forall S \in \{P, J, 1, 2, 3, \ldots,\}; \\ p(1|S, h) &= (1 - \gamma)\frac{M - m}{M}, \ \forall S \in \{P, J, 1, 2, 3, \ldots,\}. \end{aligned} \quad (5)$$

On the other hand, if the secondary user stays in the same band, the primary user may reclaim the band with probability β given by the ON-OFF model. With the primary user absent, the state will go to J if transmission is jammed, and will increase by 1 otherwise. Note that s is not a feasible action when the state is in J or P. At state K, only $\max(M - Km, 0)$ bands have not been detected by attackers, but another m bands will be detected in the upcoming time slot; therefore, the probability of jamming conditioned on the absence of primary user is given by

$$f_J(K) = \begin{cases} \frac{m}{M - Km}, & \text{if } K < \frac{M}{m} - 1; \\ 1, & \text{otherwise}. \end{cases} \quad (6)$$

To sum up, transition probabilities associated with the action s are as follows: $\forall K \in \{1, 2, 3, \ldots\}$,

$$\begin{aligned} &p(P|K, s) = \beta, \\ &p(J|K, s) = (1 - \beta) f_J(K), \\ &p(K + 1|K, s) = (1 - \beta)(1 - f_J(K)). \end{aligned} \quad (7)$$

B. Markov Decision Process

If the secondary user stays in the same band for too long, he/she will eventually be found by an attacker, as it can be seen from (6) and (7) that $p(K + 1|K, s) = 0$ if $K > M/m - 1$. Therefore, we can limit the state S to a finite set $\{P, J, 1, 2, 3, \ldots, \bar{K}\}$, where $\bar{K} = \lfloor M/m - 1 \rfloor$ and the floor function $\lfloor x \rfloor$ returns the largest integer not greater than x.

An MDP consists of four important components, namely, a finite set of states, a finite set of actions, transition probabilities, and immediate payoffs. As we have already specified all of them, the defense problem is modeled by an MDP, and the optimal defense strategy can be obtained by solving the MDP.

For an MDP, a *policy* is defined as a mapping from a state to an action, i.e., $\pi : S(n) \to a(n)$. In other words, a policy π specifies an action $\pi(S)$ to take whenever the user is in state S. Among all possible policies, the optimal policy is the one that maximizes the expected discounted payoff. The value of a state S is defined as the highest expected payoff given the MDP starts from state S, i.e.,

$$V^*(S) = \max_{\pi} E\left(\sum_{n=1}^{\infty} \delta^n U(n) \middle| \text{the initial state is } S\right), \quad (8)$$

where the optimizer is the optimal policy. The optimal policy is the optimal defense strategy that the secondary user should adopt since it maximizes the expected payoff.

An important but straightforward idea is that after a first move the remaining part of an optimal policy should still be optimal. Hence, the first move should maximize the sum of immediate payoff and expected payoff conditioned on the current action. This is the well-known Bellman equation [7],

$$V^*(S) = \max_{a\in\{h,s\}} \left(U(S,a) + \delta \sum_{S'} p(S'|S,a) V^*(S') \right). \quad (9)$$

The values of states can be calculated from a standard procedure called *value iteration* [7]. With all values known, the optimal policy $\pi^*(S)$ is the maximizer to the Bellman equation.

Since the probability of being jammed will be larger when the secondary user stays in the same band for a longer time, we can expect that there is a critical state $K^*(K^* \leq \bar{K})$ beyond which the damage overwhelms the hopping cost. If the secondary user stays in the same band for a short period ($\leq K^*$ time slots), he/she should stay to exploit more; otherwise, he/she should proactively hop to another band since the risk of being jammed becomes significant. K^* can be obtained from solving the MDP, and the optimal strategy becomes

$$a^* = \pi^*(S) = \begin{cases} s, & \text{if } 1 \leq S \leq K^*; \\ h, & \text{otherwise.} \end{cases} \quad (10)$$

V. Learning the Parameters

In the previous section, we have shown that the secondary user has an optimal strategy with perfect knowledge. Although one may argue that sometimes it is a reasonable assumption to know primary users' parameters β and γ as *a priori*, in general it is quite difficult to know the exact number of attackers m beforehand, as the secondary user cannot expect reliable information from adversaries. Both overestimating and underestimating the threat may result in inappropriate degrees of protection. Therefore, in this section, we propose a learning scheme in which the secondary user learns the parameters of the environment using the maximum likelihood estimation.

The secondary user simply sets a value $\hat{K}^*$ as an initial guess of the optimal critical state K^*, and follows the strategy (10) with the estimate $\hat{K}^*$ during the whole learning period. This guess needs not to be accurate, as the goal is merely to observe transitions occurred during the learning period that can be used for parameter estimation. After the learning period, the secondary user gains knowledge of the environment, and updates the critical state K^* accordingly.

With full history available including states and actions, the secondary user is able to count the occurrences of transitions given either action. For example, the notation $N^{(h)}_{S,S'}$ gives the total number of transitions from S to S' with the action h taken, whereas $N^{(s)}_{S,S'}$ is the total number of transitions with the action s taken. We define $K_L = \max\{K : N^{(s)}_{K,K+1} > 0\}$, $\mathbb{H} = \{P, J, K_L + 1\}$, and $\mathbb{S} = \{1, 2, \ldots, K_L\}$. Given the sequence of transitions in history, the likelihood that such a sequence has occurred can be written as a product over all feasible transition tuples $(S, a, S') \in \{P, J, 1, 2, 3, \ldots, K_L + 1\} \times \{s, h\} \times \{P, J, 1, 2, 3, \ldots, K_L + 1\}$,

$$\Lambda = \prod_{(S,a,S') \,:\, p(S'|S,a)>0} \left(p(S'|S,a)\right)^{N^{(a)}_{S,S'}}. \quad (11)$$

Moreover, if we define $\rho \triangleq m/M$ and relax it to any real number, the following proposition gives the MLE of the parameters β, γ, and ρ.

Proposition 1: Given $N^{(h)}_{S,S'}$, $S \in \mathbb{H}$ and $N^{(s)}_{S,S'}$, $S \in \mathbb{S}$ counted from history of transitions, the MLE of primary users' parameters are

$$\beta_{\text{ML}} = \frac{\sum_{K\in\mathbb{S}} N^{(s)}_{K,P}}{\sum_{K\in\mathbb{S}} \left(N^{(s)}_{K,P} + N^{(s)}_{K,J} + N^{(s)}_{K,K+1}\right)}, \quad (12)$$

$$\gamma_{\text{ML}} = \frac{\sum_{S\in\mathbb{H}} N^{(h)}_{S,P}}{\sum_{S\in\mathbb{H}} \left(N^{(h)}_{S,P} + N^{(h)}_{S,J} + N^{(h)}_{S,1}\right)}, \quad (13)$$

and the MLE of attackers' parameters ρ_{ML} is the unique root within an interval $(0, 1/(K_L + 1))$ of the following $(K_L + 1)$-order polynomial,

$$\frac{1}{\rho}\left(\sum_{S\in\mathbb{H}} N^{(h)}_{S,J} + \sum_{K\in\mathbb{S}} N^{(s)}_{K,J}\right) = \sum_{K\in\mathbb{S}} \frac{N^{(s)}_{K,P}}{\frac{1}{K} - \rho} + \frac{N^{(s)}_{K_L,K_L+1}}{\frac{1}{K_L+1} - \rho}. \quad (14)$$

Proof: With transition probabilities specified in (4) – (7) and the fact that the number of transitions into a state equals the number of transitions out of that state[1], the likelihood of observed transitions (11) can be decoupled into a product of three terms $\Lambda = \Lambda_\beta \Lambda_\gamma \Lambda_\rho$, where

$$\begin{aligned}
\Lambda_\beta &= \beta^{\sum_{K\in\mathbb{S}} N^{(s)}_{K,P}} (1-\beta)^{\sum_{K\in\mathbb{S}}\left(N^{(s)}_{K,J} + N^{(s)}_{K,K+1}\right)}, \\
\Lambda_\gamma &= \gamma^{\sum_{S\in\mathbb{H}} N^{(h)}_{S,P}} (1-\gamma)^{\sum_{S\in\mathbb{H}}\left(N^{(h)}_{S,J} + N^{(h)}_{S,1}\right)}, \\
\Lambda_\rho &= \rho^{\sum_{S\in\mathbb{H}} N^{(h)}_{S,J} + \sum_{K\in\mathbb{S}} N^{(s)}_{K,J}} \cdot (1 - (K_L+1)\rho)^{N^{(s)}_{K_L,K_L+1}} \\
&\quad \cdot \prod_{K\in\mathbb{S}} (1 - K\rho)^{N^{(s)}_{K,P}}.
\end{aligned} \quad (15)$$

Then, by differentiating $\ln \Lambda_\beta$, $\ln \Lambda_\gamma$, and $\ln \Lambda_\rho$ and equating them to 0, we obtain the MLE (12) (13) and (14).

To ensure that the likelihood is positive, ρ has to lie in the interval $(0, 1/(K_L + 1))$. Within this interval, the left-hand side of equation (14) decreases monotonically and approaches positive infinity as ρ goes to 0, whereas the right-hand side increases monotonically and approaches positive infinity as ρ goes to $1/(K_L + 1)$. Therefore, there must be a unique value of $\rho \in (0, 1/(K_L + 1))$ which is the root of the equation, and meanwhile, is the MLE ρ_{ML}. ■

After the learning period, the secondary user rounds $M \cdot \rho_{\text{ML}}$ to the nearest integer as an estimation of m, and calculate the optimal strategy using the MDP approach described in the previous section.

[1] It is completely true if the beginning state and the ending state are identical; otherwise, there will be a difference of one transition associated with the beginning and ending states, but the impact could be negligible when the learning period is long enough.

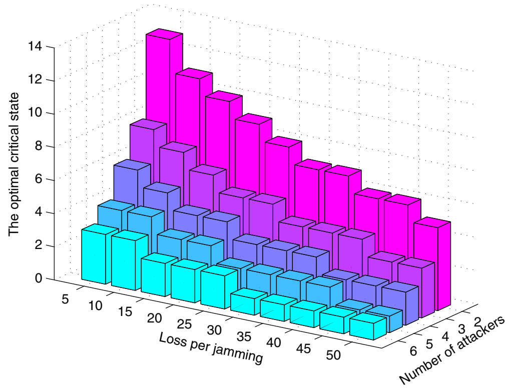


Fig. 4. The critical state K^* with different attack strengths and damages.

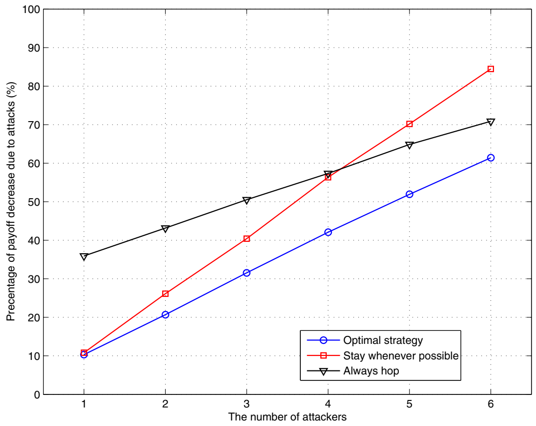


Fig. 5. The percentage of payoff decrease due to jamming attacks with different numbers of attackers.

VI. Simulation Results

In this section, we present some simulation results to evaluate the proposed defense strategy against jamming attacks. In the simulation, we fix a set of parameters to gain some insight of the defense strategy. The parameters are as follows: the communication gain $R = 5$, the hopping cost $C = 1$, the total number of bands $M = 60$, the discount factor $\delta = 0.95$, and the primary users' access pattern $\beta = 0.01, \gamma = 0.1$.

We show the critical state K^* obtained from the value iteration of the MDP, when we change the value of damage L and the number of attackers m. We assume that the secondary user has perfect knowledge of the environment. As shown in Fig. 4, if the damage from each jamming L is fixed, say $L = 10$ for example, the critical state K^* decreases from 11 to 3 when the number of attackers m increases from 2 to 6. Similarly, when the number of attackers m is fixed, the critical state K^* also decreases as the value of L increases. The reason is that the secondary user should proactively hop more frequently (i.e., K^* is smaller) to avoid being jammed when the threat from attackers are more stronger (more attackers and/or more severe damage if jammed).

In Fig. 5, we present the damage caused by attackers when the number of attackers varies, in terms of percentages of payoff loss compared with a network without malicious attackers. The damage L is set to 20 in this simulation. Besides the optimal strategy (10), another two naive strategies are simulated and compared. If the "always hopping" strategy is employed, the secondary user will hop every time slot; if the "staying whenever possible" strategy is adopted, the secondary user will always stay in the band unless the primary user reclaims the band or the band is jammed by attackers. When the number of attackers is small, it is better to stay than to hop, but when the number of attackers is large, hopping outperforms staying. The optimal strategy, however, beats both naive strategies in the entire range, as shown by the smaller decrease in payoffs in the figure. For all strategies, more damage is caused when there are more attackers.

VII. Conclusions

In this paper, we have investigated the proactive hopping as a defense strategy against jamming attacks in a cognitive radio network with multiple available bands. Since the attackers want to find the secondary user as soon as possible, they should adopt the strategy that randomly scans all the bands, and the attack-and-defense problem can be reduced to a Markov decision process, in which the optimal defense can be obtained from the value iteration of the MDP. Because not all the information may be available, a learning scheme has been proposed to estimate the parameters through the maximum likelihood estimation. Simulation results have been shown to verify the performance.

References

[1] J. Mitola III, "Cognitive radio: an integrated agent architecture for software defined radio," Ph.D. Thesis, KTH Royal Institute of Technology, Stockholm, Sweden, 2000.

[2] B. Wang, Y. Wu, and K. J. R. Liu, "Game theory for cognitive radio networks: an overview," *Computer Networks*, to appear.

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

[4] T. C. Clancy and N. Goergen, "Security in cognitive radio networks: threats and mitigation," *International Conference on Cognitive Radio Oriented Wireless Networks and Communications*, Singapore, May 2008.

[5] A. Sethi and T. X. Brown, "Hammer model threat assessment of cognitive radio denial of service attacks," *IEEE DySPAN*, Oct. 2008.

[6] W. Wang, H. Li, Y. Sun, and Z. Han, "CatchIt: detect malicious nodes in collaborative spectrum sensing," *IEEE Globecom*, Hawaii, Dec. 2009.

[7] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, John Wiley & Sons, 1994.

[8] E. Altman, K. Avrachenkov, and A. Garnaev, "A jamming game in wireless networks with transmission cost," *NET-COOP 2007. Lecture Notes in Computer Science*, vol. 4465, pp. 1–12, 2007.

[9] S. Khattab, D. Mosse, and R. Melhem, "Jamming mitigation in multiradio wireless networks: reactive or proactive?," *International Conference on Security and Privacy in Communication Netowrks*, Istanbul, Turkey, Sept. 2008.

[10] B. Wang, Y. Wu, and K. J. R. Liu, "An anti-jamming stochastic game for cognitive radio networks," *IEEE Journal on Selected Areas in Communications*, submitted.

[11] Y. Wu, B. Wang, and K. J. R. Liu, "Optimal power allocation strategy against jamming attacks using the Colonel Blotto game," *IEEE Globecom*, Hawaii, Dec. 2009.

[12] H. Li and Z. Han, "Dogfight in spectrum: jamming and anti-jamming in multichannel cognitive radio systems," *IEEE Globecom*, Dec. 2009.

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