# Game Theoretic Anti-jamming Dynamic Frequency Hopping and Rate Adaptation in Wireless Systems

Manjesh K. Hanawal, Mohammad J. Abdel-Rahman, and Marwan Krunz
Dept. of Electrical and Computer Engineering, University of Arizona, Tucson, AZ 85721, USA
{mhanawal, mjabdelrahman, krunz}@email.arizona.edu

***Abstract*—Wireless transmissions are inherently broadcast and are vulnerable to jamming attacks. Frequency hopping (FH) and transmission rate adaptation (RA) have been used to mitigate jamming. However, recent works have shown that using either FH or RA (but not both) is inefficient against smart jamming. In this paper, we propose mitigating jamming by jointly optimizing the FH and RA techniques. We consider a power constrained "reactive-sweep" jammer who aims at degrading the goodput of a wireless link. We model the interaction between the legitimate transmitter and jammer as a zero-sum Markov game, and derive the optimal defense strategy. Numerical investigations show that the new scheme improves the average goodput and provides better jamming resiliency.**

***Keywords—Dynamic frequency hopping, jamming, Markov decision processes, Markov games, rate adaptation.***

## I. Introduction

Due to their broadcast nature, wireless networks are vulnerable to various security threats, including jamming attacks. Adversaries can use readily available off-the-shelf commercial products to launch stealth jamming attacks [6], [8], [18]. In a jamming attack, an adversary injects interfering power into the wireless medium that can hinder legitimate communication in one of two ways: (i) the jamming power can degrade the signal-to-interference-plus-noise ratio (SINR) at a legitimate receiver, and (ii) in carrier sensing networks, continuous jamming may prevent the legitimate transmitter from accessing the medium, hence, causing a denial-of-service attack. In this paper, we consider the attack of the first type.

Common jamming models in the literature include random, constant, proactive, and reactive jammers [3], [8]. In this paper, we consider a multi-channel "reactive-sweep" jammer. The jammer sweeps through channels, jamming $m$ channels at a time according to a random pattern. The jammer has a listening capability. It reactively adjusts its actions based on the listening outcome. Although transmitting continuously at the maximum power will enable the jammer to cause the maximum harm, this happens at the cost of high energy consumption and, more importantly, a high likelihood of being detected. In this work, we assume a power-constrained jammer.

Frequency hopping (FH) [10], [19] and rate adaptation (RA) [13], [20] are commonly used techniques to mitigate jamming. However, these techniques are shown to be ineffective when applied separately. In the case of RA with no FH, it is shown that by merely randomizing its power levels the jammer can force the transmitter to *always* operate at the lowest rate, if the average jamming power reaches a given threshold [5]. Experiments on IEEE 802.11 networks with different RA schemes (e.g., SampleRate [9], AMRR [12], Onoe [11]) also confirm this observation. On the other hand, FH is shown to be largely inadequate in coping with jamming attacks in current 802.11 networks [14]. In particular, when the number of channels is small and channels are not perfectly orthogonal, experimental studies in [14] show that the jammer can degrade the link goodput significantly. Our aim in this paper is to study the effectiveness of a jointly optimized RA and FH technique in mitigating jamming.

In a multi-channel system, the transmitter can run away from the jammer by hopping from one channel to another. However, hopping results in a throughput loss, as the transmitter will not be able to start transmitting on the new channel instantaneously. Therefore, the hopping rate needs to be set carefully. If the hopping rate of the transmitter is too high, a significant throughput loss will be incurred. On the other hand, the transmitter cannot reside on the same channel for long time as the sweep jammer may reach that channel. In adopting the transmission rates, the transmitter faces a similar dilemma. Using high rates increases the chances of getting jammed. On the other hand, using low rates will achieve low throughput. We seek to derive a jointly optimal FH and RA policy for the transmitter against a reactive-sweep jammer. This policy informs the transmitter when to hop to another channel and when to stay on the current channel. Furthermore, it gives the best rate to use in both cases (hop and stay).

**Main Contributions**–We model the interaction between a legitimate transmitter and a power-constrained reactive-sweep jammer as a zero-sum Markov game. The transmitter dynamically decides when to switch the operating channel and what transmission rate to use. The optimal defense strategy of the transmitter is derived using Markov decision processes (MDPs), and the structure of the optimal policy is shown to be threshold type. We analyze the "constrained Nash equilibrium (NE)" of the Markov game and show that the equilibrium defense strategy of the transmitter is deterministic.

**Paper Organization**–The rest of the paper is organized as follows. In Section II, we present the transmission and jamming models. In Section III, the interactions between the transmitter and jammer are modeled as a zero-sum Markov game. We study the zero-sum Markov game and derive the optimal defense strategies in Section IV. Our conducted numerical experiments are discussed in Section V. Finally, we conclude the paper in Section VI. Due to space limitation, the proofs of some of the results have been omitted. They can be found online at [7].

## II. Transmission and Jamming Models

Consider a legitimate transmitter that communicates with its receiver in the presence of a jammer. The jammer is within the transmission range of the receiver, but outside the transmission range of the transmitter (i.e., the jammer is a hidden node to the transmitter). The jammer can overhear the receiver's messages, but not the transmitter's messages. This scenario arises in various wireless communication systems, such as satellite communications [16]. In a satellite communication system, a ground station on the earth communicates with a satellite. The footprint of the satellite beam on the earth is typically large. Therefore, a jammer on the earth close to the ground station can overhear the messages from the satellite to the ground station. However, because of the directionality of the uplink transmission, the jammer cannot overhear the messages from the ground station to the satellite.

### A. Transmission Model

We consider a time-slotted system. Transmissions are assumed to be packet-based, i.e., transmissions happen on disjoint intervals. During each interval, the states of the transmitter and jammer are assumed to remain unchanged. The transmitter can communicate on any one of $K$ available channels in each time slot. Let $\mathcal{F} = \{f_1, \ldots, f_K\}$ denote the set of non-overlapping channels. We assume that the transmitter supports $M+1$ different rates on any given channel. The set of rates is denoted by $\mathcal{R} = \{R_0, \ldots, R_M\}$. Without loss of generality, we assume that $R_0 < \ldots < R_M$. The transmitter and receiver are equipped with a single radio. The transmitter can communicate on one channel in a given slot, and can either switch to another channel or stay on the same channel in the next slot. Each channel experiences additive white Gaussian noise with a fixed noise variance. Without loss of generality, we assume that the noise variance is the same across all channels, and denote it by $\sigma^2$. The jammer injects additive interference into the channels to degrade the SINR. On each channel, the rate achieved by the transmitter depends on the received SINR. Consider a particular channel. Let the received power from the transmitter over this channel be $P_R$, and let $P_J$ be the injected power by the jammer. The received jamming power at the receiver will be attenuated by factor $\alpha, 0 \leq \alpha \leq 1$. Then, the SINR at the receiver, $\eta$, is given by:

$$\eta = \frac{P_R}{\alpha P_J + \sigma^2}. \tag{1}$$

For a given SINR, only certain rates are supported at the receiver. The relationship between the achievable rates and the SINR is shown in Figure 1. When the SINR is between $\gamma_{i-1}$ and $\gamma_i$, only rates $R_0, R_1, \ldots, R_i$ are achievable. If the transmitter transmits at a rate higher than $R_i$ while the SINR at the receiver is less than $\gamma_i$, the transmitted packet is completely lost. We assume that there is a feedback mechanism from the receiver to the transmitter. When the transmission is successful, the receiver sends an ACK message to the transmitter. On the other hand, when jamming is successful, the receiver sends a negative ACK (NACK) to the transmitter. The ACK/NACK messages can be overheard by the jammer.

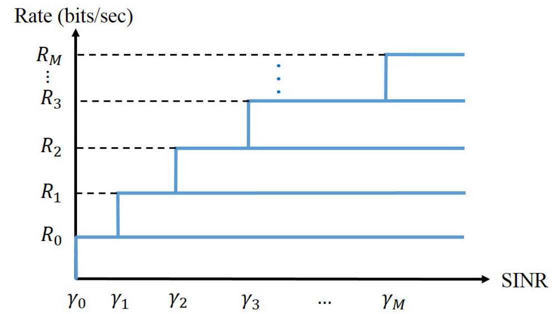


Fig. 1: Rate vs. SINR relationship.

### B. Jamming Model

We consider a time-slotted multi-channel reactive-sweep jammer. The jammer sweeps through channels, jamming $m$ channels at a time according to a random pattern. After injecting its interference power into the channel, the jammer listens to overhear any messages over the jammed channel. The jammer can overhear (i) an ACK, (ii) a NACK, or (iii) nothing. Detecting an ACK means that the jammer is on the same channel as the transmitter-receiver but its jamming power was not sufficient to destroy the legitimate communication. If the jammer detects a NACK, then the legitimate transmission was successfully jammed. If no messages were detected by the jammer, then the jammer is tuned to a different channel than the transmitter-receiver channel. Based on its listening outcome, the jammer decides about its actions.

The jammer is power constrained, as in [5] (recall that only RA was considered in [5]). It can emit on each of the $m$ channels a maximum power of $P_{\max}$ in each time slot. The jammer also has a constraint $P_{\text{avg}}$ on its average power, where $P_{\text{avg}} < P_{\max}$. In each time slot, the jammer can choose from $M+1$ discrete power levels in $\mathbf{P}_J = \{P_{J_0}, P_{J_1}, \ldots, P_{J_M}\}$. Without loss of generality, we assume that the jammer emits the same power level on all the $m$ channels. $P_{J_i}, i \in \{0, 1, \ldots, M\}$ is calculated by setting $\eta$ in (1) to $\gamma_{M-i}$ in Figure 1 and finding the corresponding $P_J$ in (1). Let $\mathcal{M} \stackrel{\text{def}}{=} \{0, 1, \ldots, M\}$. $P_{J_i}$ is given by:

$$P_{J_i} = \frac{\frac{P_R}{\gamma_{M-i}} - \sigma^2}{\alpha}, i \in \mathcal{M}. \tag{2}$$

Similar to [5], under an average-power constraint, the attack strategy is to choose a distribution on the set of available powers that satisfies the average-power constraint. Let $\mathbf{P}_J$ represents the jammer's set of pure strategies. Let $J_s$ denote the strategy space of the jammer and $\mathbb{Y}$ be an $(M+1)$-probability simplex. Then, $J_s \subset \mathbb{Y}$ and is given by:

$$J_s = \left\{ \mathbf{y} = (y_0, \ldots, y_M), \sum_{i=0}^{M} y_i = 1, \mathbf{y}\mathbf{P}_J^T \leq P_{\text{avg}} \right\}. \tag{3}$$

**Switching and Jamming Costs**–The transmitter hopping from one channel to another results in an *outage* period, the duration of which depends on the device. For example, the average latency of a channel hop for the Anthros chipset card is measured to be 7.6 ms [10]. The outage results due to the time required to reconfigure the devices on the new channel, and also due to any lack of synchrony between the transmitter and receiver's hopping instances. The throughput during the outage period is zero. We denote the average loss in throughput due to hopping by $C$, and refer to it as *hopping cost*. Outage periods also occur when the transmitter is jammed. Jamming disrupts the link between the transmitter and the receiver, which needs to be re-established through exchanging several packets, that do not contribute toward the throughput. We denote the average loss in throughput due to jamming by $L$, and refer to it as *jamming cost*. We account for $C$ and $L$ in deriving the optimal defense policy of the transmitter.

## III. Dynamic FH Game with Rate Adaptation

In this section, we develop a repeated game model between the reactive-sweep jammer and the transmitter, and derive the optimal attack (defense) strategy of the jammer (transmitter). We first discuss the attack (defense) strategies that can be adopted by the jammer (transmitter) and setup a game between them. As noted in [17], the attack and defense strategies adopted by the jammer and transmitter are like an arms race. If the jammer improves its attack strategy, the transmitter counters it with an improved defense strategy, and vice versa. The strategies adopted by each player also depend on its hardware and computational capability. Below we discuss the attack and defense strategies assuming the jamming and transmission models explained in Section II.

### A. Attack and Defense Strategies

For each time slot, the jammer emits power at the beginning of the slot and listens for an ACK/NACK message at the end of the slot. If the jammer receives an ACK/NACK on one of the $m$ jammed channels, it learns that the transmitter is on that channel[1]. If there exists only one channel, i.e., $K = 1$, the only way for the transmitter to escape from the jammer is to adapt its rate. In this case, it is shown in [5] that by randomizing its power levels, the jammer can enforce the transmitter to use the lowest rate. Therefore, when multiple channels are available, it is better for the transmitter to hop from one channel to another and evade the jammer. Accordingly, the jammer will also hop between channels in searching for the transmitter.

In [17], few rounds of arms race have been discussed. It is argued that the best strategy for the jammer is to sweep through all the $K$ channels sequentially, jamming $m$ channels in each slot, and restart the sweep cycle with a randomly reordered sweep pattern after completing each cycle. The jammer can further aggravate its attack strategy making use of its listening capability. When it overhears either an ACK or a NACK on a channel, it learns that the transmitter is operating on that channel. Accordingly, the jammer attacks the detected channel, allocating all of its power into it, until the transmitter switches to another channel. Unlike the jammer, the transmitter does not always learn the presence of the jammer based on the ACK/NACK messages. If a NACK is received, the transmitter learns the presence of the jammer on its channel. However, if an ACK is received, the transmitter does not have this information[2]. Therefore, when a NACK is received, it is better for the transmitter to hop to a new channel; because otherwise it will be jammed again in the subsequent slot. Being aware of this, the jammer also leaves the channel after receiving a NACK and starts a new sweep cycle randomly reordering its sweep pattern. When the jammer receives an ACK, it continues to stay on the same channel until it either receives a NACK or overhears nothing. If a NACK is received, the jammer begins a new random sweep cycle as earlier, otherwise (i.e., nothing is received) it continues with the current sweep cycle. We refer to the jammer that adopts the channel hopping strategy derived in the last round of the arms race discussed above as a *reactive-sweep jammer*.

We consider a *jamming game* between the transmitter and the reactive-sweep jammer. Although the hopping decisions of the jammer are fixed, the jammer still needs to decide about the amount of power to emit in each slot while satisfying its average and maximum power constraints. The transmitter's decision consists of what transmission rate to use and also whether to *stay* on the same channel or to *hop* to a new channel. We model the interactions between the transmitter and the reactive-sweep jammer as a zero-sum game.

### B. Frequency Hopping Strategies

As explained above, the hopping pattern of the reactive-sweep jammer is as follows. The jammer sweeps through the $K$ channels sequentially, jamming $m$ non-overlapping channels in each slot. At the end of each slot, the jammer will take one of the following actions based on what it is overheard: (i) If nothing is overheard, it continues to jam the next $m$ channels in the sweep cycle. (ii) If an ACK is overheard on a particular channel, it continues to jam only that channel in the next slot, changing its attack from a multi-channel attack to a single-channel attack. (iii) If a NACK is received or the sweep cycle ends, a new random cycle is restarted immediately.

For the transmitter, we assume that it does not have any mean to know the quality of various channels and it does not assign any priority to any channel. The transmitter-receiver pair follows a common FH pattern, generated by a pseudorandom noise (PN) sequence. We note that our optimization of the transmitter's channel hopping policy is in terms of how long it stays on a channel (in number of slots) before it hops to a new channel[3].

### C. Reward

Recall that the transmission at rate $R_i$ is successful only if the SINR at the receiver is at least $\gamma_i$. If an ACK is received after transmitting at rate $R_i$, the transmitter obtains a reward of $R_i$ units. In line with [17], we define the transmitter payoff in a given slot as the difference between the reward and the costs it incurs in that slot. Let $U_n$ denote the transmitter's payoff in

[1] A NACK is generated when the transmission fails either due to a jamming attack or due to the channel being bad (high fading, attenuation, etc.). Since our focus is on jamming attacks, we restrict our attention to transmission failures due to jamming attacks.

[2] The transmission is successful in the presence of the jammer if the transmitter uses a rate that is de-codable at the interference power chosen by the jammer.

[3] This duration is referred to as the channel residency time in [10].

slot $n$[4]. Then,

$$\begin{aligned} U_n &= \sum_{i=0}^{M} R_i \cdot \mathbf{1}[\text{successful transmission at rate } R_i \text{ in slot } n] \\ &- L \cdot \mathbf{1}[\text{successful jamming}] - \mathbf{C} \cdot \mathbf{1}[\text{transmitter hops}] \quad (4) \end{aligned}$$

where $\mathbf{1}[\cdot]$ is the indicator function. We note that only one term can be positive in the summation above, as the transmitter can use only one rate in each slot. An action taken by the transmitter in a given time slot affects its payoff in the future slots. Thus, we will consider a total discounted payoff with a discount factor $\delta \in (0,1)$, which indicates how much the transmitter values its future payoff over its current payoff. Let $\bar{U}$ denote the total discounted payoff of the transmitter. Then,

$$\bar{U} = \sum_n \delta^{n-1} U_n. \quad (5)$$

In the next section, we model the interactions between the transmitter and jammer as a zero-sum Markov game and derive the constrained NE (recall that the jammer is average-power constrained). We also characterize the properties of the optimal policies using Markov decision processes (MDPs).

## IV. Zero-sum Markov Game

A Markov game is characterized by a state space, an action space, an immediate reward for each player, and transition probabilities. The decision epochs are taken at the end of each time slot, and the effect of the decision takes place at the beginning of the next slot. The state of the system identifies the status of the transmitter. First, note that while the transmitter is operating on a channel, say $f$, it does not know which channels the jammer is currently sweeping unless it receives a NACK. If the transmitter is successful on channel $f$ for $k$ successive slots, it can only infer that the jammer did not sweep channel $f$ in the last $k$ slots. Therefore, keeping track of the channels that the transmitter used in the past, and how many slots it stayed on each of them is not helpful to the transmitter. We use these observations to define the state space and derive the transition probabilities below.

**State Space**–The state is defined as the number of consecutive slots that the transmitter has been successful on a channel since it last hopped into it. Let $X$ denote the state space. Then,

$$X = \left\{ J, 1, 2, \ldots, \left\lfloor \frac{K}{m} \right\rfloor \right\} \quad (6)$$

where $J$ denotes that the transmitter is jammed and $i = 1, 2, \ldots, \lfloor \frac{K}{m} \rfloor$ denotes that the transmitter is successful on the current channel for the last $i$ slots.

**Action Space**–At the end of each slot, the transmitter decides whether to stay on the current channel or hop to a new channel. It also decides which rate to use from $\mathcal{R}$. Therefore, the set of actions available to the transmitter for any state in $X$ is as follows:

$$A = \{(s, R_1), \ldots, (s, R_M), (h, R_1), \ldots, (h, R_M)\} \quad (7)$$

where $(s, R_i)$ represents the decision to stay on the current channel and use rate $R_i$, and $(h, R_i)$ represents the decision to hop to a new channel and use rate $R_i$. For notational convenience, let $s_i \stackrel{\text{def}}{=} (s, R_i)$ and $h_i \stackrel{\text{def}}{=} (h, R_i), \forall i \in \mathcal{M}$.

**Immediate Reward**–$U_n = U_n(x, a_1, a_2, x')$ represents the immediate reward the transmitter receives after going from state $x$ to state $x'$ when the actions taken by the transmitter and the jammer are $a_1 \in A$ and $a_2 \in \mathbf{P}_J$, respectively. This reward does not depend on the slot index, hence we drop the subscript $n$. For any $(a_1, a_2, x) \in A \times \mathbf{P}_J \times X$, the immediate payoff of the transmitter is given by:

$$U(\cdot, a_1, a_2, x') = \begin{cases} -L - C, & \text{if } x' = J,\ a_1 = h_i,\ a_2 = P_{J_j},\ j > M - i \\ R_i - C, & \text{if } x' = 1,\ a_1 = h_i,\ a_2 = P_{J_j},\ j \le M - i \\ -L, & \text{if } x' = J,\ a_1 = s_i,\ a_2 = P_{J_j},\ j > M - i \\ R_i, & \text{if } x' \neq J,\ a_1 = s_i,\ a_2 = P_{J_j},\ j \le M - i \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$

Note that the reward of the transmitter depends only on the action it takes and the new state it enters, and not on its current state. Since the jammer cannot observe the state of the transmitter, it just chooses its power level such that its average-power constraint is satisfied.

**Transition Probabilities**–Let $P(x'|x, a_1, a_2)$ denote the transition probability to state $x'$ when the current state is $x$ and the transmitter chooses action $a_1 \in A$ while the jammer chooses action $a_2 \in \mathbf{P}_J$. First, let us consider the case where the transmitter's action involves hopping to a new channel. Let the transmitter be on channel $f$ after hopping. When action $h_i$ is taken in any state, the state on the new channel can be either $J$ or 1, $\forall i \in \mathcal{M}$. Let $x = J$. Then, on taking action $h_i$ the system enters state $J$ again only if the jammer also hops into channel $f$ and uses a power level that does not allow the transmitter to succeed at rate $R_i$. Recall that on each successful jamming the transmitter hops to a new channel, and the jammer repeats the jamming process with a new sweep pattern that is independent of its past sweep pattern. Then, the transmitter and jammer hop to the same channel with probability $m/(K-1)$. Hence, $P(J|J, h_i, P_{J_j}) = 1 - P(1|J, h_i, P_{J_j}), i \in \mathcal{M}$, is given by:

$$P(J|J, h_i, P_{J_j}) = \begin{cases} m/(K-1), & \text{if } j > M - i \\ 0, & \text{otherwise.} \end{cases} \quad (9)$$

Taking action $h_i, i \in \mathcal{M}$, in state $x \neq J$, say $x = \tilde{x}$, the transmitter's next state can be either $x' = J$ or 1. $x' = 1$ if any of the following happens: (i) channel $f$ is already swept by the jammer, (ii) channel $f$ is not swept by the jammer and the jammer does not hop to it in the next time slot, or (iii) the jammer hops to $f$ and uses a power level that does not disrupt the transmission at rate $R_i$. Let $a_1 = h_i, a_2 = P_{J_j}$ and $j > M - i$, then the transmitter is successful on $f$ if the jammer does not hop into $f$.

$$\begin{aligned} P(1|x, h_i, P_{J_j}) &= 1 - P(J|x, h_i, P_{J_j}) \\ &= \frac{mx}{K-1} + \frac{K-1-mx}{K-1}\left\{1 - \frac{m}{K-1-mx}\right\} \\ &= 1 - m/(K-1). \end{aligned} \quad (10)$$

[4]We specify how the payoff depends on the actions in the next section after defining the state space.

Therefore,

$$P(1|x,h_i,P_{J_j}) = \begin{cases} 1-m/(K-1), & \text{if } j > M-i \\ 1, & \text{otherwise.} \end{cases} \quad (11)$$

For the case where the transmitter decides to stay on its current channel. Suppose that the transmitter is on channel $f$ and state $x \neq J$, say $x = \tilde{x}$, and takes action $s_i, i \in \mathcal{M}$. Then, the transmitter enters into state $x' = J$ or $x' = \tilde{x}+1$. $x' = J$ (i) if the jammer did not sweep $f$ in the last $\tilde{x}$ slots, hops into $f$ in the next slot, and jams at a power that does not allow decoding at rate $R_i$, or (ii) if the jammer is already on $f$ and jams at a power that does not allow decoding at rate $R_i$. The probability of the first event can be computed as $m/(K-m\tilde{x})$ and the probability of the second event is $m\tilde{x}/K$. Note that if the jamming interference power is $P_{J_j}$ in the first case, then it is $mP_{J_j}$ in the second case (i.e., single-channel attack). Let $\gamma(j,m) \overset{\text{def}}{=} \frac{P_R}{\alpha m P_{J_j}+\sigma^2}$. Then,

$$\begin{aligned} &P(J|x,s_i,P_{J_j}) = 1 - P(x|x,s_i,P_{J_j}) \\ &= \begin{cases} \frac{m(x+1)}{K}, & \text{if } x < K/m \text{ and } j > M-i \\ \frac{mx}{K}, & \text{if } x < K/m \text{ and } \gamma(j,m) < \gamma_i \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \quad (12)$$

Next, we introduce the required notations[5] to define the transmitter and jammer strategies and their objective functions. In each time slot, the transmitter takes an action that depends on its past observation. We will restrict to the Markov stationary policies[6][7], where the transmitter takes an action based on its current state only. The set of Markov stationary policies of the transmitter is denoted by $F_s$. Let $\mathcal{M}(A)$ denote the distribution on set $A$ and $\mathbf{f}: X \to \mathcal{M}(A)$ denote the strategy of the transmitter. Let $f(x) \overset{\text{def}}{=} \{f(x,a_1), a_1 \in A\}$, where $f(x,a_1)$ is the probability of choosing action $a_1 \in A$ in state $x \in X$. Similarly, let the jammer's strategy be $\mathbf{g}: X \to \mathbf{P}_J$, and let $g(x) \overset{\text{def}}{=} \{g(x,a_2), a_2 \in \mathbf{P}_J\}$, where $g(x,a_2)$ is the probability of choosing action $a_2 \in \mathbf{P}_J$ in state $x \in X$. Since the jammer does not know the state, for any jammer's strategy $\mathbf{y} = (y_0,\ldots,y_M) \in J_s$, $g(x) = \mathbf{y}, \forall x \in X$.

Let $r : X \times A \times \mathbf{P}_J \to \mathbb{R}$ denote the immediate reward for the transmitter. For any actions $a_1$ and $a_2$ taken by the transmitter and jammer, respectively, and any state $x$, $r(x,a_1,a_2)$ is given by:

$$r(x,a_1,a_2) = \sum_{x'} U(x,a_1,a_2,x')P(x'|x,a_1,a_2). \quad (13)$$

For given $\mathbf{f} \in F_s$ and $\mathbf{y} \in J_s$, the expected discounted payoff of the transmitter when the initial state is $x$ is:

$$\tilde{V}(x,\mathbf{f},\mathbf{y}) = \mathbb{E}^{\mathbf{f},\mathbf{y}} \left\{ \sum_n \delta^n r(X_n,A_{1n},A_{2n}) | \, X_0 = x \right\} \quad (14)$$

where $\{(X_n,A_{1n},A_{2n}) : n = 1,2,\ldots\}$ is a sequence of random variables, denoting the state and the actions of the transmitter and jammer in each slot, respectively. This sequence evolves according to the policy $(\mathbf{f},\mathbf{y})$. The operator $\mathbb{E}^{\mathbf{f},\mathbf{y}}$ denotes the expectation over the process induced by the policies $\mathbf{f}$ and $\mathbf{y}$.

The transmitter's objective is to choose a policy $\mathbf{f}$ that results in the highest expected reward starting from any state $x \in X$, and is defined as:

$$V_T(x,\mathbf{y}) = \max_{\mathbf{f} \in F_s} \tilde{V}(x,\mathbf{f},\mathbf{y}). \quad (15)$$

In contrast, the jammer's objective is to choose a strategy $\mathbf{y}$ that minimizes the transmitter's expected discounted payoff.

$$V_J(x,\mathbf{f}) = \min_{\mathbf{y} \in J_s} \tilde{V}(x,\mathbf{f},\mathbf{y}). \quad (16)$$

Note that the strategy space of the jammer is constrained, whereas the transmitter can choose any stationary policy. A strategy pair $(\mathbf{f}^*,\mathbf{y}^*)$ is constrained NE if $\mathbf{y}^* \in J_s$ and $\forall x \in X, \mathbf{f} \in F_s$, and $\mathbf{y} \in J_s$,

$$\tilde{V}(x,\mathbf{f},\mathbf{y}^*) \leq \tilde{V}(x,\mathbf{f}^*,\mathbf{y}^*) \leq \tilde{V}(x,\mathbf{f}^*,\mathbf{y}). \quad (17)$$

Let $V^*(x) \overset{\text{def}}{=} \tilde{V}(x,\mathbf{f}^*,\mathbf{y}^*)$. Then, $\{V^*(x), x \in X\}$ is referred to as the value of the zero-sum game[8].

**Theorem 1:** The zero-sum game has a stationary constrained NE.

*Proof:* While the jammer aims to minimize the transmitter's payoff, it needs also to meet its average-power constraint. Since the jammer does not know the value of the current state, its average-power constraint for any strategy $\mathbf{y}$ (i.e., $\mathbf{y}\mathbf{P}_J^T \leq P_{\text{avg}}$) can be equivalently written as a constraint on an expected discounted cost as follows:

$$C_\beta(\mathbf{f},\mathbf{y}) = (1-\beta)\mathbb{E}^{\mathbf{f},\mathbf{y}} \left\{ \sum_n \beta^{n-1} C(X_n,A_{1n},A_{2n}) \right\} \leq P_{\text{avg}}$$

for some $\beta \in (0\ 1)$. $C(X_n,A_{1n},A_{2n})$ denotes the cost for the jammer, which is the power it chooses in slot $n$, i.e., $C(\cdot,\cdot,A_{2n}) = A_{2n}$. Further, by choosing a strategy $\mathbf{y}'$ such that $y'_0 = 1$, the constraint on the expected discounted cost is strictly met. Thus, strong Slater condition in [2] is verified and the existence of stationary constrained NE follows from Theorem 2.1 in [2]. ∎

### A. Transmitter Optimal Defense Strategy

In this section, we study the properties of the transmitter's optimal defense strategy against a fixed jammer's strategy. The expected reward of the transmitter when the jammer's strategy is $\mathbf{y}$ is denoted by $r_\mathbf{y} : X \times A \to \mathbb{R}$. For a given state-action pair $(x,a)$, $r_\mathbf{y}(x,a)$ is given by:

$$r_\mathbf{y}(x,a) = \sum_{i=0}^{M} y_i r(x,a,P_{J_i}). \quad (18)$$

Let $P_\mathbf{y}(x'|x,a)$ denote the probability that the transmitter

[5]We follow the notational convention of [4].

[6]Note that the state updates according to the past history.

[7]For any given history-dependent policy, there exists a Markov policy that is equally good [15][Ch. 4].

[8]If $(\tilde{\mathbf{f}},\tilde{\mathbf{y}})$ is another equilibrium, it also results in the same value of the game [4][Sec. 3.1].

enters state $x'$. Then,

$$P_{\mathbf{y}}(x'|x,a) = \sum_{i=0}^{M} y_i P(x'|x,a,P_{J_i}). \quad (19)$$

Let $f^*_{\mathbf{y}}(X)$ denote the policy that maximizes the expected discounted reward function when the jammer uses strategy $\mathbf{y}$. Since $\mathbf{y}$ does not depend on the state, the optimal policy $f^*_{\mathbf{y}}(X)$ can be obtained by solving a single player MDP with the reward and transition probabilities defined in (18) and (19), respectively. Then, $f^*_{\mathbf{y}}(X)$ is a deterministic policy [15], i.e, $f^*_{\mathbf{y}} : X \to A$. For notional convenience, we do not explicitly mention this dependency on $\mathbf{y}$, and write $V(x) \stackrel{\text{def}}{=} V_T(x,\mathbf{y})$.

We use the value iteration [15][Ch. 6] method to derive the optimal defense strategy and its properties. The well-known Bellman equations for the expected discounted utility maximization problem in (15) are written as follows:

$$\begin{aligned} Q(x,a) &= r_{\mathbf{y}}(x,a) + \delta \sum_{x' \in X} P_{\mathbf{y}}(x'|x,a) V(x') \\ &= \sum_{x' \in X} P_{\mathbf{y}}(x'|x,a) \left( r_{\mathbf{y}}(x,a,x') + \delta V(x') \right) \quad (20) \\ V(x) &= \max_{a \in A} Q(x,a). \end{aligned}$$

Note that in our formulation states $J$ and $K$ are equivalent, because the jammer will start the sweep cycle afresh and the transmitter can only take hop decisions in these states. Hence, when the transmitter begins in either state ($J$ or $K$), it will get the same total discounted reward, i.e., $V(J) = V(K)$. From (20), for any $x = J, 1, \ldots, K-1$, $V(x)$ is expressed in terms of $V(J)$ and $V(x+1)$. Below we establish the monotonicity of $V$ on the state space $X \setminus J$ by restricting the transmitter's reward in state $K$, and use this monotonicity property to establish the structure of the optimal policy[9].

**Lemma 1:** $V(\cdot)$ is a decreasing function over $\{1, 2, \ldots, K\}$.

From (9) and (10), we note that when the transmitter takes action $h_i, i \in \mathcal{M}$, the probability of entering into state $J$ or 1 does not depend on the current state. We make use of this observation and the monotonicity of the function $V(\cdot)$ to derive the following structure of the optimal policy.

**Proposition 1:** The optimal policy $f^*$ satisfies:

- $\exists$ constants $K^* \in \{1, \ldots, K-1\}$ and $i^* \le M$ such that:

  $f^*(x) = h_{i^*}$ for $K^* \le x \le K-1$ and $f^*(0) = s_{i^*}$.
- For any integers $x$ and $y$, if $1 \le x < y < K^*$, $f^*(x) = s_j$, and $f^*(y) = s_k$, then $j \ge k$.
- If $r_{\mathbf{y}}(J, s_i)$ is increasing in the index $i$, then $i^* = M$.

*Proof:* The idea of the proof is as follows. Note that $Q(h_i) = Q(x, h_i)$ does not depend on $x, \forall i \in \mathcal{M}$. We show that $Q(x, s_i)$ is decreasing in $x, \forall i \in \mathcal{M}$. Then, for any $i \in \mathcal{M}, \exists x \in X$ such that $Q(x, s_i)$ is smaller than the largest $Q(h_i)$. The detailed proof can be found online at [7]. ∎

[9] This assumption is made only to establish the structure of the policy analytically. Our simulations show that the same property holds in general.

The above proposition says that when the transmitter hops to a new channel it will hop again only after either it is jammed or it spent $K^*$ successive slots on that channel. While it stays on a given channel, the transmitter adapts its transmission rate– the transmitter reduces its rate as the number of successive successful transmissions increases. When the transmitter hops, it always uses a fixed rate. Further, this rate is the maximum rate available ($R_M$) if $r_{\mathbf{y}}(0, s_i)$ is increasing in the index $i$.

Note that since the transmitter hops once it reaches state $K^*$, it never enters into a state larger than $K^*$. Thus, if $K^* < K$, the resulting Markov chain is reducible.

**Corollary 1:** The threshold $K^*$ is decreasing in $L$, and increasing in both $K$ and $C$.

*Proof:* The proof follows by noting that for any $x' > x$, $Q(x', s_i) - Q(x, s_i)$ is increasing in $L$ and decreasing in $K, \forall i \in \mathcal{M}$. Moreover, $Q(x, h_i)$ is decreasing in $C, \forall i \in \mathcal{M}, x \in X$. This verifies that $K^*$ is increasing in $C$. ∎

Next, we return to the study of the Markov game.

### B. Equilibrium of the Markov Game

In this section, we compute the constrained NE of the zero-sum Markov game and study its properties. For a given defense strategy $\mathbf{y} \in J_s$, the following linear program solves the recursive equations in (20) [4][Sec 2.3]:

$$\begin{aligned} \text{minimize} \quad & \sum_x V(x) \\ \text{subject to:} \quad & V(x) \ge r_{\mathbf{y}}(x,a) + \delta \sum_{x' \in X} P_{\mathbf{y}}(x'|x,a) V(x'), \\ & \forall x \in X, a \in A. \end{aligned} \quad (21)$$

From Theorem 1, we know that the zero-sum Markov game has a constrained NE. We use a non-linear version of the above program to compute the equilibria. First, we will develop the necessary notation. Let $R(x) = [r(x,a,p)]_{a \in A, p \in \mathbf{P}_J}$ and $T(x,V) = [\sum_{x'} P(x'|x,a,p) V(x')]_{a \in A, p \in \mathbf{P}_J}$ be the reward and transition probability matrices, respectively. Consider the following non-linear program:

$$\begin{aligned} \text{minimize} \quad & \sum_x \Big\{ V_1(x) + V_2(x) \Big\} \\ \text{subject to:} \quad & V_1(x)\mathbf{1} \ge R(x)\mathbf{y} + \delta T(x, V_1)\mathbf{y}, \forall x \in X \\ & V_2(x)\mathbf{1} \ge -f(x)R(x) + \delta f(x) T(x, V_2), \forall x \in X \\ & \mathbf{P}_J \mathbf{y}^T \le P_{\text{avg}}, \forall x \in X \end{aligned} \quad (22)$$

where $\mathbf{1}$ denotes a vector of all ones of size $M$ when the state is $J$ or $K$, and of size $2M$ for all other states.

**Theorem 2:** Let $(V_1^*(x), V_2^*(x), f^*(x), \mathbf{y}^*)$ denote the minimum of the non-linear program (22). Then, $(f^*(x), \mathbf{y}^*)$ denotes the optimal constrained NE of the game.

*Proof:* The nonlinear program (22) is the same as the one in [4][Sec. 3.7] with the additional average-power constraint on the jammer's strategy. The proof follows from [4][Th. 3.7.2]. ∎

Note that although the optimal strategy of the transmitter for a given jammer's strategy is deterministic, the equilibrium strategy may not be deterministic [4][Ch. 2]. The strategy $\mathbf{y}^*$ is the same for all $x \in X$ as the jammer does not know the state. We know from the previous subsection that the optimal transmitter's strategy against any given $\mathbf{y}$ is deterministic. Therefore, at equilibrium the strategy of the transmitter is deterministic.

**Transmitter-Receiver Rendezvous**–As mentioned earlier, the transmitter and receiver share a common PN sequence. The transmitter follows this PN sequence, however it optimizes how many slots to stay in each channel before switching to the next channel in the PN sequence. As stated in Proposition 1, the optimal policy of the transmitter is to stay on each channel $K^*$ slots as long as it is not jammed. Once it is jammed it will switch to the next channel. The transmitter optimal policy is also shared with the receiver. Therefore, the receiver will follow the common PN sequence staying on each channel $K^*$ slots unless it is jammed, in this case it will switch to the next channel. This way, the rendezvous is ensured.

## V. Performance Evaluation

In this section, we study the performance of the joint FH and RA scheme under different values of the system parameters. Our performance metrics are the average goodput (in Mbps) and the success rate (percentage of un-jammed transmissions). The parameters of study are $K$, $C$, $L$, $P_{\text{avg}}$, and $m$. We use the set of rates adopted by IEEE 802.11a [1], i.e., 6, 9, 12, 18, 24, 36, 48, and 54 Mbps. Unless stated otherwise, we use the following parameters: $K = 4$, $L = 25$ Mbps, $C = 50$ Mbps, $m = 1$, and $P_{\text{avg}} = 0.83P_{\max}$. We implement our game in MATLAB. The 95% confidence intervals are shown. The optimal defense and attack strategies of the transmitter and jammer, respectively, are obtained by solving (22). We compare the joint FH and RA scheme with an FH only scheme, in which the transmitter hops according to the optimal policy without adapting its rate. We implement three variants of the FH only scheme, each with a different fixed rate (6, 24, and 54 Mbps).

### A. Effect of $K$

We plot in Figures 2 and 3 the average goodput and success rate, respectively, vs. $K$ for various schemes. When $K$ is small ($K < 5$ in Figure 2), the joint FH and RA scheme achieves higher goodput than all other FH only schemes. However, the FH only scheme with rate fixed at 54 Mbps achieves a slightly higher goodput than the joint scheme when $K$ is sufficiently large ($K > 5$ in Figure 2). The reason is that the success rate of the FH only scheme improves when $K$ increases, as shown in Figure 3. Although the FH only scheme with highest rate achieves a slightly higher goodput than the joint scheme when $K$ is sufficiently large, it has a much lower success rate, as shown in Figure 3. Even though the FH only scheme with rate fixed at 6 Mbps achieves a 100 % success rate, its average goodput is zero. This is because $C$ is greater than 6 Mbps. The FH only scheme with rate 24 Mbps provides a balance between the average goodput and success rate, as shown in Figures 2 and 3.

### B. Effect of $C$

The effect of $C$ on the average goodput and success rate is shown in Figures 4 and 5, respectively. As shown in Figure 4,

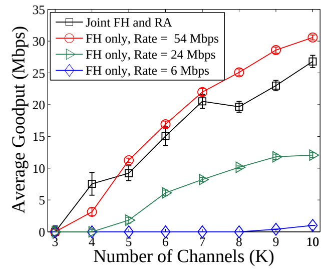


Fig. 2: Average goodput vs. $K$.

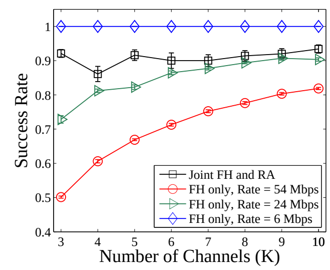


Fig. 3: Success rate vs. $K$.

when $C$ is sufficiently large the joint FH and RA scheme achieves higher goodput than the FH only schemes. This is because the FH only scheme with rate 54 Mbps hops more frequently than the joint scheme. In the joint scheme, the transmitter can avoid the hopping cost evading the jammer by adapting its rate (use a sufficiently small rate). Moreover, when $C$ is sufficiently large, the success rate of the joint scheme is significantly higher than that of the FH only scheme with rate 54 Mbps. The reason is that when $C > L$, the transmitter will be more tempted to stay on the channel than hopping, thus it will be more susceptible to jamming. Again, the FH only scheme with rate 24 Mbps provides a balance between the average goodput and success rate.

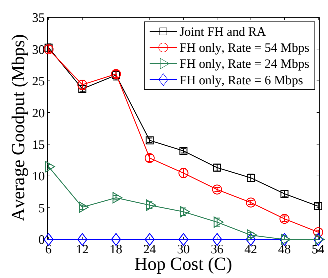


Fig. 4: Average goodput vs. $C$.

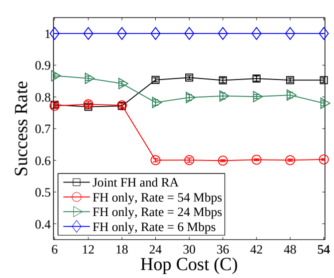


Fig. 5: Success rate vs. $C$.

### C. Effect of $L$

The effect of $L$ is studied in Figures 6 and 7. The average goodput decreases with $L$. It reaches zero when $L$ exceeds a certain value. The average goodput of the FH only scheme with rate fixed at 54 Mbps is smaller than that of the joint scheme, and it reaches zero before the joint scheme.

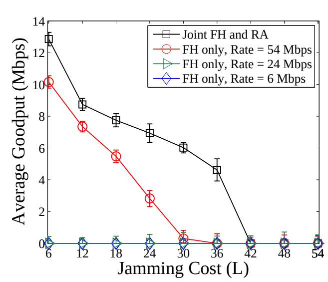


Fig. 6: Average goodput vs. $L$.

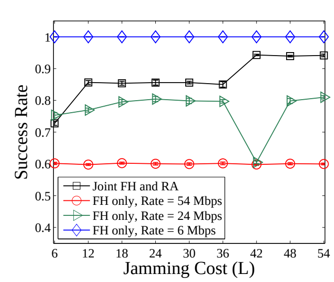


Fig. 7: Success rate vs. $L$.

### D. Effect of $P_{avg}$

The impact of $P_{\text{avg}}$ on the average goodput and success rate is depicted in Figures 8 and 9, respectively. As expected, when $P_{\text{avg}}$ increases the jammer uses high power levels more frequently, which decreases the success rate and the average goodput, especially for the FH only scheme with rate fixed at 54 Mbps.

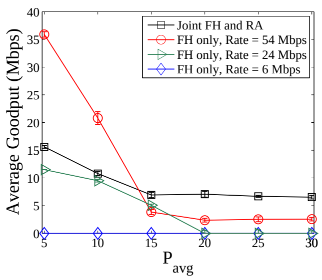


Fig. 8: Average goodput vs. $P_{\text{avg}}$.

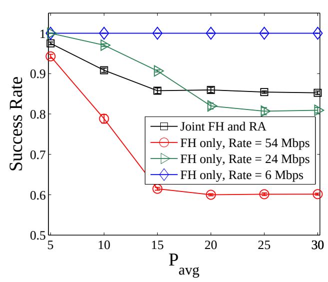


Fig. 9: Success rate vs. $P_{\text{avg}}$.

### E. Effect of $m$

The effect of $m$ on the average goodput and success rate is shown in Figures 10 and 11, respectively, when $K = 16$. For a given $K$, increasing $m$ decreases the length of the sweep cycle, which is $\lceil \frac{K}{m} \rceil$. Increasing $m$ $a$ times is equivalent to decreasing $K$ $a$ times and keeping $m = 1$, e.g., the case when $K = 16$ and $m = 4$ is equivalent to the case when $K = 4$ and $m = 1$ (see Figures 2 and 10). Therefore, as can be seen from Figures 10 and 11, increasing $m$ has the same impact as decreasing $K$ for the case when $m = 1$.

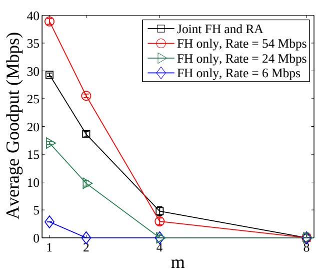


Fig. 10: Average goodput vs. $m$.

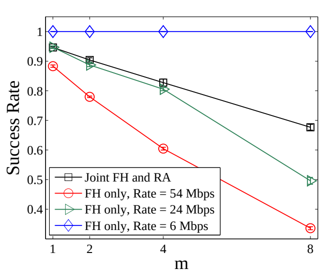


Fig. 11: Success rate vs. $m$.

## VI. Conclusions

In this paper, we analyzed a joint FH and RA defense scheme against a reactive-sweep jammer. We modeled the interaction between the transmitter and jammer as a zero-sum Markov game, and derived the optimal equilibrium defense strategy against the worst attack strategy. Our numerical results showed that joint FH and RA achieves better performance than FH only, especially when (i) the number of channels is small, (ii) the hopping cost is high, or (iii) the average jamming power is high. We studied our joint FH and RA scheme numerically and compared its performance with various FH only schemes. The improvement achieved by the joint FH and RA scheme (in terms of the average goodput and success rate) compared to the FH only schemes depends on the system parameters, and can be very significant.


## Acknowledgment

The authors would like to thank Dr. Diep Nguyen in the department of electronic engineering at the Macquarie University for his helpful discussions. This research was supported in part by the National Science Foundation (grants # IIP-1265960 and CNS-1016943) and in part by the Army Research Office (grant # W911NF-13-1-0302). Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the author(s) and do not necessarily reflect the views of NSF or ARO.


## References


[1] "Wireless LAN medium access control (MAC) and physical layer (PHY) specifications," *IEEE Std. 802.11*, June 2007.

[2] E. Altman and A. Shwartz, "Constrained Markov games: Nash equilibria," *Advances in Dynamic Games and Applications*, vol. 5, pp. 213–221, 2000.

[3] E. Bayraktaroglu, C. King, X. Liu, G. Noubir, R. Rajaraman, and B. Thapa, "On the performance of IEEE 802.11 under jamming," in *Proc. of the IEEE INFOCOM Conf.*, Phoenix, AZ, USA, April 2008, pp. 1939–1947.

[4] J. Filar and K. Vrieze, *Competitive Markov Decision Processes*. New York, USA: Springer-Verlag, 1997.

[5] K. Firouzbakht, G. Noubir, and M. Salehi, "On the capacity of rate-adaptive packetized wireless communication links under jamming," in *Proc. of the ACM WiSec Conf.*, Tucson, AZ, USA, 2012, pp. 3–14.

[6] R. Gummadi, D. Wetherall, B. Greenstein, and S. Seshan, "Understanding and mitigating the impact of RF interference on 802.11 networks," in *Proc. of the ACM SIGCOMM Conf.*, Kyoto, Japan, 2007, pp. 385–396.

[7] M. K. Hanawal, M. J. Abdel-Rahman, and M. Krunz, "Game theoretic anti-jamming dynamic frequency hopping and rate adaptation in wireless systems," University of Arizona, Tech. Rep. TR-UA-ECE-2013-3, Jan. 2014. [Online]. Available: ece.arizona.edu/∼krunz.

[8] S. Khattab, D. Mosse, and R. Melhem, "Jamming mitigation in multi-radio wireless networks: Reactive or proactive?" in *Proc. of the ACM SecureComm Conf.*, Istanbul, Turkey, Sep. 2008.

[9] R. T. Morris, J. C. Bicket, and J. C. Bicket, "Bit-rate selection in wireless networks," *Masters Thesis, MIT*, 2005.

[10] V. Navda, A. Bohra, S. Ganguly, and D. Rubenstein, "Using channel hopping to increase 802.11 resilience to jamming attacks," in *Proc. of the IEEE INFOCOM Conf.*, Anchorage, Alaska, USA, 2007, pp. 2526–2530.

[11] Onoe Rate Control. [Online]. Available: http://madwi.org/browser/trunk/ath rate/onoe.

[12] S. Pal, S. R. Kundu, K. Basu, and S. K. Das, "IEEE 802.11 rate control algorithms: Experimentation and performance evaluation in infrastructure mode," in *Passive and Active Measurement (PAM) Conf.*, Adelaide, Australia, 2006.

[13] K. Pelechrinis, I. Broustis, S. Krishnamurthy, and C. Gkantsidis, "ARES: An anti-jamming reinforcement system for 802.11 networks," in *Proc. of the ACM CoNEXT Conf.*, Rome, Italy, 2009, pp. 181–192.

[14] K. Pelechrinis, C. Koufogiannakis, and S. Krishnamurthy, "Gaming the jammer: Is frequency hopping effective?" in *Proc. of the ACM WiOpt Conf.*, Seoul, Korea, June 2009, pp. 187–196.

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

[16] D. Roddy, *Satellite Communications*. McGraw-Hill, 2006.

[17] Y. Wu, B. Wang, K. J. Liu, and T. C. Clancy, "Anti-jamming games in multi-channel cognitive radio networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 1, pp. 4–15, January 2012.

[18] W. Xu, W. Trappe, Y. Zhang, and T. Wood, "The feasibility of launching and detecting jamming attacks in wireless networks," in *Proc. of the ACM MobiHoc Conf.*, Urbana-Champaign, IL, USA, 2005, pp. 46–57.

[19] W. Xu, T. Wood, W. Trappe, and Y. Zhang, "Channel surfing and spatial retreats: defenses against wireless denial of service," in *Proc. of the ACM WiSe Workshop*, Philadelphia, PA, USA, Oct. 2004, pp. 80–89.

[20] J. Zhang, K. Tan, J. Zhao, H. Wu, and Y. Zhang, "A practical SNR-guided rate adaptation," in *Proc. of the IEEE INFOCOM Conf.*, Phoenix, AZ, USA, April 2008, pp. 146–150.