Cognitive Networks with In-band Full-duplex Radios: Jamming Attacks and Countermeasures

Manjesh K. Hanawal[1], Diep N. Nguyen[2], and Marwan Krunz[2,3]
[1] IEOR, IIT Bombay, Mumbai, MH-400076, India
[2]Faculty of Engineering and Information Technology, University of Technology Sydney, Australia
[3] Department of Electrical and Computer Engineering, University of Arizona, USA
Email: mhanawal@iitb.ac.in, diep.nguyen@uts.edu.au, krunz@email.arizona.edu

Abstract—Although in-band full-duplex (IBFD) radios promise to double the throughput of a wireless link, they are more vulnerable to jamming attacks than their out-of-band full-duplex (OBFD) counterparts. For two communicating OBFD nodes, a jammer needs to attack both the uplink and the downlink channels to completely break the communication link. In contrast, only one common channel needs to be jammed in the case of two IBFD nodes. Even worse, a jammer with self-interference suppression (SIS) capabilities (the underlying technique of IBFD radios) can learn the transmitters' activity while injecting interference, allowing it to react instantly to the transmitter's strategies. In this work, we consider a power-constrained IBFD "reactive-sweep" jammer that sweeps through the set of channels by jamming a subset of them simultaneously. We model the interactions between the IBFD radios and the jammer as a stochastic constrained zero-sum Markov game in which nodes adopt the frequency hopping (FH) technique as their strategies to counter jamming attacks. Beside the IBFD transmission-reception (TR) mode, we introduce an additional operation mode, called transmission-detection (TD), in which an IBFD radio transmits and leverages its SIS capability to detect jammers. The aim of the TD mode is to make IBFD radios more cognitive to jamming. The nodes' optimal defense strategy that guides them when to hop and which operational mode (TD or TR) to use is then established from the equilibrium of the stochastic Markov game. We prove that this optimal policy has a threshold structure, in which IBFD nodes stay on the same channel up to a certain number of time slots before hopping. Simulation results show that our policy significantly improves the throughput of IBFD nodes under jamming attacks.

Index Terms—Jamming attack, dynamic frequency hopping, in-band full-duplex radio, Markov games.

1 Introduction

Self-interference suppression (SIS) techniques (e.g., [2], [3]) allow a transmitting device to suppress its self-interference up to the noise floor, thus enabling wireless radios to simultaneously transmit and receive on the same frequency channel. This in-band full-duplex (IBFD) capability not only boosts the link throughput but it also helps solve various channel-access issues (e.g., Tx deafness, hidden/exposed nodes) [4]. More applications of IFDB to cognitive radio communications can be found in [5] and references therein. A network of IBFD radios has the potential to double the network throughput, compared with half-duplex (HD) radios that have to alternate in time/frequency/code between transmit and receive modes. However, such a network is also more vulnerable to jamming attacks. In this work, we identify some of these jamming threats and investigate anti-jamming techniques that make IBFD radios more cognitive to these threats. Our techniques optimally leverage the simultaneous transmit-and-receive capability of IBFD devices.

A jammer can hinder legitimate transmissions in one of two ways: (i) the jammer can inject interfering power into the wireless medium, thus degrading the signal-to-interference-plus-noise ratio (SINR) at a legitimate receiver, and (ii) in carrier-sensing systems, a persistent jammer can prevent a legitimate transmitter from accessing the medium, effectively creating a denial-of-service (DoS) attack. These stealthy attacks can be easily launched by an adversary using commercial off-the-shelf (CoTS) products [6]–[8]. In this work, we focus on the former type of jamming.

The implications of jamming on IBFD radios can be particulary acute. First, compared with out-of-band full-duplex (OBFD) systems, which include HD devices as a special case, a jammer can *simultaneously* interfere with both the uplink and downlink, as both communicating IBFD share the same frequency channel in the same vicinity, and hence are likely to suffer the same jamming effect. Second, unlike OBFD radios, operating as IBFD devices hinders nodes from detecting a jamming attack, especially under fading. Specifically, under fading, a transmission failure is not always caused by jammers when the jamming interference (if any) is distorted by self-interference due to imperfect SIS. Third, a jammer with IBFD capability can discern the outcome of its jamming instantaneously, while continuously attacking legitimate transmissions.

Several physical-layer techniques have been developed to mitigate the jamming of OBFD devices. These include

Manjesh K. Hanawal would like to thank funding support from IIT-Bombay and the Inspire faculty fellowship (DST, Govt. of India). Diep N. Nguyen was supported by Australian Research Council (Discovery Early Career Researcher Award DE150101092). Marwan Krunz was supported in part by NSF (grants CNS-1409172, CNS-1513649, CNS-1563655, CNS-1731164, and IIP-1822071) and by the Broadband Wireless Access & Applications Center (BWAC). 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. An abridged version of this paper was presented at the IEEE INFOCOM Conference, May, 2016 [1].

spread spectrum, including frequency hopping (FH), directional antennas, and adaptive
power/coding/modulation. Jammer-specific techniques have also been developed [6]. Common jamming models in the literature include random, persistent, proactive, and reactive [7], [9]. This classification is based on the jammer's capabilities. Persistent jammers always emit power into the medium. Proactive jammer can vary the power to meet various constraints. A reactive jammer exhibits a more sophisticated behavior, and emits power only when it detects a legitimate transmission [8]. In this paper, we consider a power-constrained jammer with IBFD capability, referred to as "IBFD reactive-sweep" jammer. Specifically, this jammer sweeps through blocks of m channels in each slot. While jamming, the jammer can simultaneously learn the outcome of the attack and accordingly adapt its strategy (thanks to its SIS capability). Our jamming model explained in detail in Section 3.

We consider a FH-based transmission link for which the hopping sequence can be adapted on the fly. If a transmission fails, the legitimate nodes should not always hop to a new channel. This is because radio channels are inherently subject to fading that may also be the cause of the failure. There is a chance nodes may hop to a channel that is also experiencing fading. Moreover, too frequent hopping reduces the effective throughput due to the time overhead required for oscillators to settle down after changing the frequency [10] (e.g., Anthros chipset has settling time of about 7.6 ms). Of course, if a node can reliably detect the presence of a jammer on a given channel, it should hop to evade this jammer.

Accordingly, to effectively counter a jammer, nodes must detect its presence. As noted in [6], it is not possible to reliably identify the presence of a jammer through measurements only. Hence it is necessary to execute consistency checks to ascertain such presence. To that end, beside the classical IBFD *Transmission Reception* (TR) mode, we introduce the *Transmission Detection* (TD) in which an IBFD node transmits data and simultaneously leverages its SIS capability to detect jamming (instead of receiving data as in the TR mode). The aim of the TD mode is to make IBFD radios more cognitive to jamming. The jamming mitigation technique that we develop uses the packet delivery ratio[1](PDR) as an indicator of a potential jamming activity, while leveraging the TD mode to reliably confirm that.

In the TD mode, one node acts as a receiver while the other node transmits and receives (listens) simultaneously. If the PDR is low, one of the nodes can switch to TD mode. This node can then assess the link quality by measuring the Received Signal Strength (RSS). If the RSS is high and the PDR is low, it would be an indication of the presence of a jammer. On the other hand, if both the PDR and the RSS are low, then the deterioration link quality is related to fading. Note that the above cognition to jamming of IBFD radios comes at the cost of lower throughput in the TD mode, compared to the TR mode. PDR measurements are conducted over a frame duration. RSS measurements need to be done only when the PDR goes below a certain threshold for a short time period. The duration of the sampling window for RSS measurements should be carefully tuned based on the traffic rate, the measuring accuracy, and the desired detection confidence level. For details about how to choose these values, we refer to [6].

1. PDR is the ratio of the number of successfully decoded packets to the number of received packets.

The interactions between IBFD nodes and the jammer are modeled as a a constrained zero-sum Markov game [11]. The nodes' optimal defense strategy that guides them when to hop (equivalently, how long to remain on the same frequency channel) and which operational mode (TD or TR) to use is established from the equilibrium of the Markov game such that the aggregate throughput is maximized. The major contributions of the paper are as follows:

- We identify the severe susceptibility of IBFD radios to attacks by IBFD-capable jammers.
- We formulate the interactions between IBFD radios and power-constrained "IBFD reactive-sweep" jammers as a stochastic zero-sum Markov game.
- We make IBFD radios more cognitive to jamming by defining two operational modes for IBFD nodes that help them improve jamming detection. We derive the optimal counter-jamming strategy for IBFD nodes using either the total discounted reward or the average reward criteria.
- We compare the performance of the jointly optimal FH and cognitive mode-switching strategy with the optimal strategy based on FH only (i.e., without switching between the two operating modes). Through numerical simulations we show that by jointly optimizing FH and TR/TD mode switching, IBFD nodes are much more resistant to jamming than using adaptive FH only.

Related work: As mentioned before, the SIS capability of IBFD radios has been leveraged to solve various problems in wireless networking and cognitive communications. In [12], the authors developed a "Listen and Talk" protocol that allows radios to simultaneously access and sense the spectrum/medium. The adaptation of IBFD radios to cognitive communications was investigated in [13] [14]. In [15], the authors provided solutions to secure full-duplex spectrum-sharing wiretap networks using different antenna reception schemes. The authors of [16] used game theory to formulate the interactions between IBFD devices and eavesdroppers. In [17], channel training was leveraged to enhance IBFD physical-layer security under a wiretap channel.

Jamming and anti-jamming techniques have been well-studied in wireless networks for HD and OBFD devices (see [18] [19] [20] [21] [22] and therein references). Below, we only discuss the papers that are most related to our work in terms of attack model and defense strategies. In [23] the authors developed an FH strategy against a "sweep jammer" in 802.11 networks. Their hopping strategy optimizes the channel residence time. A similar hopping strategy was developed in [24] using Markov Decision Process for a cognitive radio network. The authors in [25] developed a defense strategy that combines FH and rate adaption techniques. Another direction is to leverage the latest advances in deep learning and artificial intelligence to combat jammers, e.g., [26]

Recently, several authors proposed protocols that leverage IBFD capabilities to improve the performance of ad hoc and cellular networks [27], [28], [29]. However, the vulnerability of IBFD nodes to jamming attacks was not taken into account. In [30], [31], IBFD nodes were treated as jammer-cum-receiver devices, whereby eavesdroppers are prevented from listening to the communication through *friendly jamming*. The issue of "non-friendly" jamming attacks on the IBFD nodes was not considered.

To the best of our knowledge, this paper is the first to study jamming attacks on IBFD devices and to develop jamming mitigation techniques that exploit the SIS capability of these devices to counter jammers.

Paper organization: In Section 2 we describe the problem setup. In Section 3 we study the attack and defense strategies of the jammer and the IBFD nodes, respectively. The optimal defense strategies of the transmitter are derived in Section 4 using MDPs. Performance evaluation through simulations is given in Section 5. Finally, in Section 6 we discuss future work and give concluding remarks.

2 Model and Setup

We consider two IBFD nodes A and B that communicate in the presence of jammers. The two nodes have IBFD radios that can operate on any one of K non-overlapping channels $\mathcal{F} = \{f_1, \ldots, f_K\}$. Each channel experiences additive white Gaussian noise (AWGN), which is independently and identically distributed (i.i.d.) across all channels.

2.1 Link and Channel Models

We assume time-slotted transmissions. In each time slot, several packets can be transmitted. During a time slot, the states of the transmitter and the jammer remain unchanged. Nodes transmit at a fixed power in each time slot, and the jammer injects additive interference into the channels to degrade the received signal.

The two-state Gilbert-Elliot channel model is used to characterize the fading process. At a given time, each channel can either be in a fading state with probability 1-p. We assume that when the channel is not fading, a transmission always succeeds in the absence of a jamming attack. In contrast, if the channel is in a fading state, the transmission always fails irrespective of jamming. We further assume that the fading process is independently and identically distributed (i.i.d) across time and all channels.

Under the IBFD capability, when both uplink and downlink are active, the net throughput achieved in the absence of jamming is $(\xi_1 + \xi_2)R$, where R denotes the throughput obtained when only the uplink or downlink is active and $\xi_i \in [0, 1]$ denotes the fraction of throughput lost due to imperfect SIS at node i. ξ_i is referred to as the *SIS factor*. The value of ξ_i depends on the hardware capabilities of the nodes to suppress its own self interference. Recent developments indicate that this value is $\xi_i > 0.5$, ensuring higher rates than that is achievable for a HD link [2].

Modes of Operation: In the TR mode, a node transmits and receives simultaneously on a link. In the TD mode, one of the nodes only receives data packets from the other. With some abuse of terminology, we say that a pair of nodes operate in this TD mode when one of them operates in the TD mode while the other operates in the TR mode. When one node operates in the TD mode, the other node only receives the ambient noise and can measure its strength. If both nodes are under a jamming attack, then any of them can operate in the TD mode and the other node can measure the strength of the ambient noise over the same channel. If the jammer can attack only one of the nodes and not the other[2], then the node that is in close proximity to the jammer suffers low PDR. Hence this node can measure the ambient noise power if the other node operates in the TD mode. In the rest of the paper, we assume that the node that enters into the TD is agreed upon a priori, and we focus on the scenario where the jammer can attack both nodes simultaneously. Our jamming detection strategy can be easily tailored to other scenarios. For example, if the jammer attacks only one node and the not the other, then only one node will experience high PDR . In this case the node which doesn't experience high PDR can switch to the PD node to facilitate attack detection by the other node. Since the defence strategy we discuss later depends only on whether or not a jammer is detected and not on how it is detected, this defence strategy remains optimal for this scenario.

Switching and Transmission Cost: When a node wishes to hop between channel, it must first reconfigure its RF components before it can start the transmissions. The duration of this *settling time* depends on the device (e.g., for the Atheros chipset card, this time is about 7.6 ms [23]), and presents throughput loss. Additional loss occurs due to the lack of synchronization between the Tx and Rx's hopping time. Collectively, we denote the average loss in throughput due to hopping by C, and refer to it as *switching cost*. Outage periods can also occur when the nodes are jammed. To re-establish communications following a jamming attack, several control packets must be exchanged, which do not contribute to throughput. We denote the average loss in throughput due to jamming by L, and refer to it as *transmission cost*. We account for C and L in deriving the optimal defense policy of the Tx.

2.2 Jamming Model

We consider multiple IBFD-capable jammers. Each jammer attacks one of the channel in the set of available channels $\mathcal{F}$ in each time slot and simultaneously observes the activity of legitimate nodes (if any) on that channel, allowing it to learn the outcome of its attack. Jammers can co-ordinate among themselves by attacking nonoverlapping channels to increase their chances of success. Further, when a jammer detects activity on a channel, it and other jammers can all simultaneously attack the same channel causing the maximum possible degradation of the link quality. This multi-jammer attack is equivalent to a single jammer that attacks m channels sequentially in a time slot. Henceforth, we can consider a single jammer that sequentially attacks m channels, $m < K$, in each slot. Furthermore, the jammer should attack each channel for a sufficiently long time to be

2. This scenario arises when a jammer is in the proximity of one node but is 'hidden' from the other.

effective; otherwise attacked nodes can easily recover lost packets from brief outages. We assume that the jamming power is sufficiently high to ensure PDR is low and zero throughput for the attacked link (on the same channel).

In this paper, we consider an IBFD reactive jammer. At the beginning of a time slot, the jammer continuously emits white Gaussian noise into the channel. At the same time, it uses its IBFD capability to listen for nodes' activity on the channel[3]. If the jammer detects transmission activity, it continuously attacks that channel until no such activities is detected. If the jammer does not detect node activity for a while, it moves to attack other channels.

We assume a power-limited jamming model (as in [32]) with a finite number of radios. The have capacity is similar to that of other nodes in the network. Specifically, in each time slot, the jammer attacks m channels at a maximum power of P. The jammer has a constraint P_{avg} on its time-averaged power, where $P_{\text{avg}} \leq P$. Due to this power constraint, the jammer may not be able to emit power in all time slots. Instead, for each time slot, the jammer can choose to jam or not to jam power, i.e., $\mathbf{P}_J = \{0, P\}$ [4]. Let ω_P and ω_0 denote the probability of jamming with power P and 0 (i.e., not jamming) in a time slot, respectively[5]. Similar to [32], under an average-power constraint, the jammer's strategy space is the distribution (ω_0, ω_P) on the set of available powers that satisfies the average-power constraint. Formally, let $\mathbf{\Omega}$ be the set of feasible jamming strategies ω, defined as follows

$$\mathbf{\Omega} \stackrel{\text{def}}{=} \left\{ \omega \stackrel{\text{def}}{=} (\omega_0, \omega_P) | \omega_0 + \omega_P = 1 \text{ and } \omega \mathbf{P}_J^T \leq P_{\text{avg}} \right\}. \quad (1)$$

We enforce the average-power constraint only on the jammer and not on the nodes, as the jammer aims to minimize the risk for being detected, which is not the case for the legitimate nodes [32].

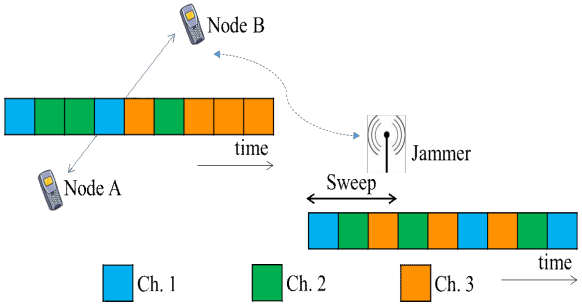

Fig. 1: System model.

3. The jammer could first listen for channel activity and jam it only if some activity is detected. This conserves the jamming power but reduces its effectiveness.

4. To be power efficient, P should be the minimum power/interference level that the jammer believes can disrupt the link transmission. If $P = P_{avg}$, jammer can attack in every round.

5. The jammer can use more than two power levels. In this case, the jammer can better control his power. Our analysis can be extended for this case provided nodes are aware of which power levels the jammer can use.

3 Jamming Game: Attack and Defense Strategies

In this section, we discuss the attack and defense techniques for the jammer and legitimate nodes respectively. As discussed in [24], the attack and defense strategies follow an *arms race* between the jammer and nodes. The best attack (defense) strategy of the jammer (nodes) depends on the strategy adopted by its opponent. Below we discuss a few rounds of this arms race.

3.1 Attack Strategy

Assuming that the jammer is aware that nodes can hop between channels to evade jamming, one naive attack strategy is to randomly jam m out of the K channels, chosen with equal probabilities in each time slot. In this case, as argued in [24], the nodes should stay on the same channel as they are equally vulnerable on all channels. In this case, the probability of getting jammed is m/K in each time slot, whether nodes hop or not. Anticipating nodes' response, the jammer may now go through all K channels sequentially, jamming m of them in each slot and sweeping through the next m channels, and so on. If the jammer follows a deterministic sweep pattern, nodes can learn and effectively counter the jamming attack by avoiding attacked channels in a given slot. Aware of this response from nodes, the jammer could further randomize its sweep pattern after completing a sweep cycle.

Sweep Jammer: In the next round of arms race, legitimate nodes update their strategy as follows. Once they are jammed in a given sweep cycle, they can simply turn off their transmitters. Not finding any activity on the channel, the jammer would leave the channel and continue the current sweep cycle. Nodes can then resume transmission on the same channel because they will not be jammed again till the end of current sweep cycle. Since a node can be jammed at most once in each sweep cycle, the average throughput achieved by nodes operating in the TR mode is $2(K-m)\xi R/K$. If the transmitters were to use FH, they can improve the throughput utmost by $2\xi m R/K$ but may also suffer throughput loss due to channel switching. When the gain is small compared to the switching loss, nodes may prefer to stay on the same channel and tolerate the small loss in throughput due to a jamming attack.

IBFD Reactive Sweep Jammer: Aware of nodes' response to the fixed sweep pattern in each round, the jammer then can revise its strategy so as to use a randomly ordered sweep pattern at the state of a new sweep cycle. The node's strategy would be not to use the same channel but to switch channels in anticipating the jamming attack on the current channel (nodes may evade jamming if they leave their channel when the jammer is about to attack and hop on to a channel that has already been swept by the jammer in that cycle). With the IBFD capability, the jammer can discern nodes' activity, if any, on channels being jammed. Hence, if the jammer and the nodes are on the same channel, the jammer can continuously jam it until the nodes hop to a different channel. We refer to a jammer that applies this strategy, derived in last round of arms race, as IBFD reactive sweep jammer. This is the jammer's strategy that

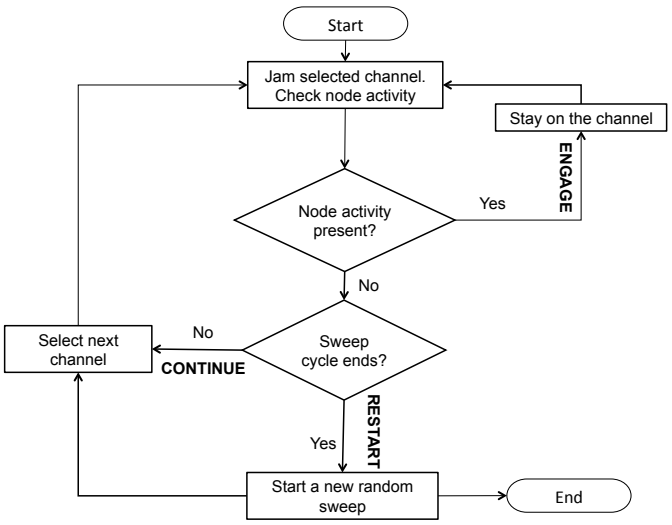

Fig. 2: **IBFD Reactive Sweep Jammer:** In each time slot, the jammer attacks one selected channel and observes channel activity. If some activity is observed, it stays on the channel and continuously attacks it (ENGAGE state). Otherwise, it RESTARTs a new sweep pattern if this is the end of the current cycle, else it attacks the next channel in the sweep cycle (CONTINUE state).

is considered in the rest of the paper. The operation of the jammer in each time slot is depicted in Figure 2.

3.2 Defense Strategy

If nodes observe a low PDR, hopping to a different channel would not a good strategy, because the new channel could be in fading, and also switching channels would result in throughput loss. However, if nodes can verify the presence of a jammer, they must hop to a different channel, otherwise, they will be continuously jammed (due to the IBFD capability of the jammer). Thus, when the PDR is low, nodes can switch to the TD mode (if not already in this mode) to ascertain the cause of failure before taking action.

Note that while transmission fails in TR mode due to jamming attack, the nodes will be jammed again if they switch to TD mode in the next time slot. However, if the nodes are already in the TD mode when they are jammed, then they know the cause of the transmission failure, and hence leave the channel. Also when the nodes are jammed while operating in the TR mode, they have to re-establish both the links and suffer throughput loss of $2L$. This is contrast to operating in the TD mode, where the throughput loss due to jamming is L as the nodes need to re-establish only one link. Thus, operating in TR mode gives higher throughput of $2\xi R$, but jamming cannot be detected in this mode and transmission loss is high. We define the reward for each node and in a given round n as

$$R(n) = \begin{cases} 2\xi R \cdot 1[\text{Success}] - 2L \cdot 1[\text{Jammed}] - 2C \cdot 1[\text{Hop}] & \text{in TR mode} \\ R \cdot 1[\text{Success}] - L \cdot 1[\text{Jammed}] - C \cdot 1[\text{Hop}] & \text{in TD mode} \end{cases} \quad (2)$$

The indicator $1[\text{Event}]$ is 1 when 'Event' occurs otherwise, it is zero. Note that the reward is negative when nodes are being jammed or when they hop. This accounts for throughput loss in the next slot where nodes have to resynchronize and cannot transmit immediately after hopping or incurring jamming. We consider the worst case scenario, where any successful jamming attack is assumed to cause throughput loss to both the nodes.

Performance metric: If nodes have full knowledge of the jammer's sweep pattern, they can evade the jammer at the minimal switching cost: In each sweep cycle, nodes hop when the jammer is about to sweep their channel and operate on a new channel for the rest of the time in the TR mode. If the nodes hop to a channel that was previously swept by the jammer in each sweep cycle, then the nodes were never jammed. By repeating this process, nodes can achieve the highest average throughput per round, which can be computed as

$$R_m = p(2\xi R) - (1-p)(2L) - 2mC/K, \quad (3)$$

where p is the probability that a channel is not in fading, as defined in Section 2.

In the absence of such knowledge, nodes should use a policy that achieves average throughput (per round) that is as close as possible to (3). Let $R_\pi(n)$ denote the throughput in round n based on policy π. We define the regret of policy π over period T as follows

$$Reg(T,\pi) = R_m - \frac{1}{T}\sum_{n=1}^{T} R_\pi(n). \quad (4)$$

The goal of the nodes is to use a policy that minimizes the regret. Given that the nodes can learn the strategy of the smart sweep jammer (but not the sweep pattern itself), the nodes can estimate the likelihood of jamming attack in the current slot and pro-actively decide to leave the channel. Then, in each time slot, the nodes have to decide whether to stay on or leave the current channel and also the mode of operation. We refer to the number of time slots the nodes operate on a channel before they leave it without being jammed as *channel residence time*. The channel residence time indicates frequency of node hops, and thus influences the throughout loss due to channel switching.

Intuitively, the best policy for the nodes is to operate in the TR mode on a channel and switch to the TD mode after certain time slots to detect presence of any jamming activity. The nodes leave the current channel, either after detecting the jammer, or when the jammer is likely to arrive on the current channel. Note that current decision of the nodes influences their throughput in the subsequent slots. In the next section, we formalize this intuition by defining appropriate state and action space and derive the optimal strategy using Markov decision process under both total discounted and average reward criteria.

4 Zero-sum Markov game

We begin by defining the state space, action space and derive the transition probabilities of the Markov game that models the interaction between the jammer and the nodes. First, note that while nodes operate on a channel, say f, they do not know which channels the jammer is currently sweeping. If the nodes succeed on $f \in \mathcal{F}$ for k slots, they can only infer that the jammer hasn't swept through f over the last k slots. We use these observations in defining the state

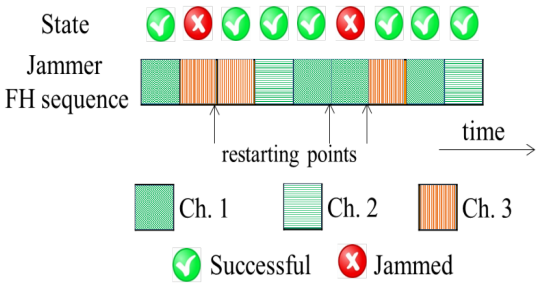

Fig. 3: IBFD reactive sweep jammer.

space and derivation of the transition probabilities below. For ease of notation, write $\tilde{K} = \lceil \frac{K}{m} \rceil$.

States of the Markov chain: The state denotes outcome of the nodes' transmission at the end of a time slot. Let $\mathcal{X}$ denote set of states given by

$$\mathcal{X} = \{J, y_1, y_2, \cdots, y_{\tilde{K}-1}, u_1, u_2, \cdots u_{\tilde{K}-1}\}.$$

The state space contains two classes of states: detected jamming and undetected jamming. The former contains only state J, denoting the transmission failure due to the jamming attack (i.e., without ambiguity, jamming is detected)[6]. Since the nodes can resolve the cause of transmission failure in the TD mode, the transmitter takes state J while operating in the TD mode only. The second class of states (jamming not detected) has two subclasses, namely $\mathcal{Y} := \{y_1, y_2, \cdots, y_{\tilde{K}-1}\}$ and $\mathcal{U} := \{u_1, u_2, \cdots, u_{\tilde{K}-1}\}$. State $y_k \in \mathcal{Y}$ denotes that the nodes has been staying on a channel continuously for k time slots (since they last hopped onto that channel including failed transmissions and the current slot) and has not detected presence of jammer unambiguously and the current transmission succeeds. State $u_k \in \mathcal{U}$ denotes that the nodes has been staying on a channel continuously for k time slots (since they last hopped onto that channel including failed transmissions and the current slot) and has not detected presence of jammer unambiguously and the current transmission fails. The subclasses $\mathcal{Y}$ and $\mathcal{U}$ both have $\tilde{K}-1$ states as m channels are jammed in each time slot and the nodes can then stay on the same channel without being unjammed for at most $\tilde{K}-1$ slots. A state u_k distinguishes from a state y_k by checking the transmission's failure (u_k) or success (y_k) in the current slot. Note that the current state of the Markov chain is observable only to the nodes. We use $x \in \mathcal{X}$ to denote a generic state.

Nodes' Actions: The set of actions available to the nodes is denoted as $\mathcal{A}$ and is given by

$$\mathcal{A} = \{(s, \text{TD}), (h, \text{TD}), (s, \text{TR}), (h, \text{TR})\}.$$

We assume that the nodes take action at the end of each time slot after observing its current state, resulted from the effect of its previous action. Action $s_1 := (s, \text{TD})$ denotes that the nodes stay on the same channel it used in the previous slot and operate in the TD mode. Action $h_1 := (h, \text{TD})$ denotes that the nodes hop to a new randomly selected channel and operate in the TD mode. Similarly, $s_2 := (s, \text{TR})$ denotes that the nodes stay on the same channel and operate in the TR mode and $h_2 := (h, \text{TR})$ denotes that they hop to a randomly selected channel and operate in the TR mode. Note that after observing a failure in the TR mode (i.e., u_k states), the nodes should either hop (using h_1 or h_2) or stay in the TD mode to detect the nature of failure but not stay in the TR mode (i.e., action s_2 should not be used in u_k). This allows the nodes to come out of u_k states sooner in case the channel is under jamming. We use $a \in \mathcal{A}$ to denote a generic action.

Nodes' Reward: Let $U(x, a_1, a_2, x')$ denote the reward to the transmitter while moving from state $x \in \mathcal{X}$ to state $x' \in \mathcal{X}$ after it takes action $a_1 \in \mathcal{A}$ and the jammer takes action $a_2 \in \mathbf{P}_J$. Using (2) we define rewards of the nodes in different states as follows: for all $a_2 \in \mathbf{P}_J$

$$U(\cdot, a_1, a_2, x') = \begin{cases} R, & \text{if } x' = y_k, a_1 = s_1, k = 1, 2, \cdots, \tilde{K}-1 \\ -L, & \text{if } x' \in \{J, u_k\}, a_1 = s_1, k = 1, 2, \cdots, \tilde{K}-1 \\ 2\xi R, & \text{if } x' = y_k; a_1 = s_2; k = 1, 2, \cdots, \tilde{K}-1 \\ -2L, & \text{if } x' = u_k; a_1 = s_2; k = 1, 2, \cdots, \tilde{K}-1 \\ R - C, & \text{if } x' = y_1, a_1 = h_1 \\ -L - C, & \text{if } x' = \{J, u_1\}, a_1 = h_1 \\ 2\xi R - 2C, & \text{if } x' = y_1, a_1 = h_2 \\ -2L - 2C, & \text{if } x' = u_1, a_1 = h_2. \end{cases}$$

Transition probabilities: As the nodes take action based only on its current state, the state evolves according to a Markov chain on $\mathcal{X}$. Let $P(x'|x, a_1, a_2)$ denote the probability that the nodes enter state $x' \in \mathcal{X}$ from state $x \in \mathcal{X}$ after nodes take action $a_1 \in \mathcal{A}$ and the jammer takes action a_2. We next calculate the transition probabilities for all possible state-action pairs.

Given $(x, a_1, a_2) = (J, s_1, a_2)$ or (J, s_2, a_2): Recall that the jammer can detect activity on the channels while jamming, and hence continues[7] to jam the channel till the nodes leave that channel, i.e., $P(J|J, a_1, P) = 1$ for all $a_1 = s_1, s_2$. Thus, in state J, the nodes should only take action h_1 or h_2, otherwise they will get jammed again in the next slot. Also recall jammer's strategy that once the nodes are jammed on a channel, the jammer continues to emit power into the channel if the nodes continue to stay on the channel. Hence action s_2 is not taken when $x = J$.

Given $(x, a_1, a_2) = (J, h_1, a_2)$: When the nodes take action h_1 in state J they can enter state J or y_1 or u_1. In state J, the nodes leave the channel and the jammer restarts the sweep cycle. The probability they hop onto the same channel in the next slot is $1/\tilde{K}$. We get

$$\begin{aligned} P(J|J, h_1, P) &= 1/\tilde{K}, \quad P(y_1|J, h_1, P) = (1 - 1/\tilde{K})p \\ P(u_1|J, h_1, P) &= (1 - 1/\tilde{K})(1 - p) \\ P(J|J, h_1, 0) &= 0 \\ P(u_1|J, h_1, 0) &= \Pr(\text{fading}) = 1 - p \\ P(y_1|J, h_1, 0) &= 1 - P(u_1|J, h_1, 0). \end{aligned} \tag{5}$$

6. Note that fading and jamming can simultaneously disrupt a transmission, in such a case, we still call it the state J.

7. That also means that if nodes are in state J, the jammer will never play the action $a_2 = 0$.

Given $(x, a_1, a_2) = (J, h_2, a_2)$: The new possible states are y_1 or u_1. The following are straightforward

$$\begin{aligned} P(y_1|J, h_2, P) &= P(y_1|J, h_1, P) \\ P(u_1|J, h_2, P) &= 1 - P(y_1|J, h_2, P) \\ P(y_1|J, h_2, 0) &= P(y_1|J, h_1, 0) \\ P(u_1|J, h_2, 0) &= 1 - P(y_1|J, h_2, 0). \end{aligned} \tag{6}$$

Given $(x, a_1, a_2) = (y_k, s_1, a_2), k = 1, 2, \cdots \tilde{K} - 2$: As the nodes can verify the cause of transmission failure in the TD mode, the nodes enter state J only if jamming attack is successful, otherwise the nodes enter the state y_{k+1} or u_{k+1} depending on state of the channel. The nodes are jammed if the jammer is on the same channel that the nodes are currently using. The following cases are possible: 1) the jammer entered the channel in the previous slot, but did not emit power[8]. 2) the jammer entered the channel only in the current slot. Then, $\forall k = 1, 2, \cdots \tilde{K} - 2$ we have

$$\begin{aligned} P(J|y_k, s_1, P) &= \frac{1-q}{\tilde{K} - (k-1)} + \frac{1}{\tilde{K} - k} \\ P(u_{k+1}|y_k, s_1, P) &= (1 - P(J|y_k, s_1, P))(1-p) \\ P(y_{k+1}|y_k, s_1, P) &= \\ & \qquad 1 - P(J|y_k, s_1, P) - P(u_{k+1}|y_k, s_1, P) \\ P(J|y_k, s_1, 0) &= 0 \\ P(u_{k+1}|y_k, s_1, 0) &= \Pr(\text{fading}) = 1 - p \\ P(y_{k+1}|y_k, s_1, 0) &= 1 - P(u_{k+1}|y_k, s_1, 0). \end{aligned} \tag{7}$$

Given $(x, a_1, a_2) = (y_k, s_2, a_2), k = 1, 2, \cdots \tilde{K} - 2$: The nodes can transit to u_{k+1} or y_{k+1}. New state is u_{k+1} if the transmission fails due to fading or jamming or both. The nodes enter into state y_{k+1} only if channel remains good, and either 1) jammer did not enter the channel the nodes are currently using 2) the jammer entered the channel but did not emit power. We get:

$$\begin{aligned} P(y_{k+1}|y_k, s_2, P) &= \left(1 - \frac{1}{\tilde{K} - k}\right) p \\ P(u_{k+1}|y_k, s_2, P) &= 1 - P(y_{k+1}|y_k, s_2, P) \\ P(u_{k+1}|y_k, s_2, 0) &= P(u_{k+1}|y_k, s_1, 0) \\ P(y_{k+1}|y_k, s_2, 0) &= 1 - P(u_{k+1}|y_k, s_2, 0). \end{aligned} \tag{8}$$

Given $(x, a_1, a_2) = (y_k, h_1, a_2), k = 1, 2, \cdots, \tilde{K} - 2$: When the nodes take action h_1, they can unambiguously determine the cause of transmission failure. Also, when it hops, counting of number of slots spent on the new channel restarts. Thus, if the nodes take action h_1, it enters state J or y_1 or u_1. If $a_2 = P$ and the nodes hop from a channel, say f, to one of the $K-1$ channels chosen uniformly at random, it will get jammed if the nodes hop to one of the $K - 1 - mk$ channels not yet swept by the jammer and the jammer also hops onto that channel in the current time slot[9]. Then, we get:

$$\begin{aligned} P(J|y_k, h_1, P) &= \frac{K - 1 - mk}{K - 1} \frac{1}{\tilde{K} - k}, \\ p(y_1|y_k, h_1, P) &= p\left(1 - P(J|y_k, h_1, P)\right), \\ P(u_1|y_k, h_1, P) &= 1 - P(J|y_k, h_1, P) - P(y_1|y_k, h_1, P). \end{aligned} \tag{9}$$

Similarly, if $a_2 = 0$, we have

$$p(y_1|y_k, h_1, 0) = p = 1 - P(u_1|y_k, h_1, 0). \tag{10}$$

Given $(x, a_1, a_2) = (y_k, h_2, a_2), k = 1, 2, \cdots, \tilde{K} - 1$: The new states can be u_1 or y_1. When nodes hop, the probability of entering state y_1 is the same regardless of being in TD or TR mode (i.e., regardless of the nodes' action h_1 or h_2). We have

$$\begin{aligned} P(y_1|y_k, h_2, P) &= P(y_1|y_k, h_1, P), \\ P(u_1|y_k, h_2, P) &= 1 - P(y_1|y_k, h_2, P), \\ P(y_1|y_k, h_2, 0) &= p = 1 - P(u_1|y_k, h_2, 0). \end{aligned} \tag{11}$$

Given $(x, a_1, a_2) = (u_k, s_1, a_2), k = 1, 2, \cdots \tilde{K} - 2$: The new state can be J or u_{k+1} or y_{k+1}. Given that the nodes are in state u_k implies that the previous transmission failed which could be either due to jamming or fading. In the former case, the new state will be certainly J. In the latter case, the new state can be J only due to jamming in the current slot (if $a_2 = P$). Thus the new state is J either because the Jammer is on the channel in the previous slot and jammed previous and the current slot, or because in the previous slot the channel was in fading and jammer came on the channel only in the current slot and emitted power. The new state is y_{k+1} if the transmission is successful in the current slot. We have[10]:

$$\begin{aligned} P(J|u_k, s_1, P) &= \frac{q}{\tilde{K} - k + 1} + \frac{1-p}{\tilde{K} - k} \\ P(u_{k+1}|u_k, s_1, P) &= (1-p)\left(1 - P(J|u_k, s_1, P)\right) \\ P(y_{k+1}|u_k, s_1, P) &= \\ & \qquad 1 - P(J|u_k, s_1, P) - P(u_{k+1}|u_k, s_1, P) \\ P(y_{k+1}|u_k, s_1, 0) &= p = 1 - P(u_{k+1}|u_k, s_1, 0). \end{aligned} \tag{12}$$

Recall that when the nodes enters state u_k, action s_2 is not applied. Hence we skip computation of $P(\cdot|u_k, s_2)$.

Given $(x, a_1, a_2) = (u_k, h_1, a_2), k = 1, 2, \cdots \tilde{K} - 2$: The nodes can transit to state J or u_1 or y_1. If the failure happens in the previous slot due to jamming, the nodes enter state J after h_1 when the new channel is jammed. Since jammer restarts its sweeping cycle, this probability is $1/\tilde{K}$. If the failure in the previous slot is due to fading, the probability that the nodes get jammed in the next slot is the same as

8. In this case the jammer will stay on the channel and attack in the current slot.

9. If the jammer and nodes are on the same channel in the previous slot but the jammer did not emit power, then in the current slot the jammer simply continues the sweep cycle if it finds that nodes have left the channel.

10. The transition to state u_k depends on action taken previously. For example, if last action was s_1 or h_1, the current state u_k must be due to fading in the previous slot. And it is due to either fading or jamming if the last action was s_2 or h_2. For simplicity, we ignore this dependency on the previous action and compute the probabilities assuming that state u_k is entered either due to jamming or fading only.

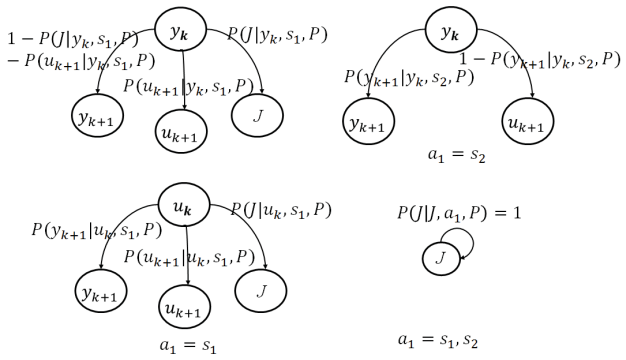

Fig. 4: State transition diagram when the nodes "stay" and the jammer jams

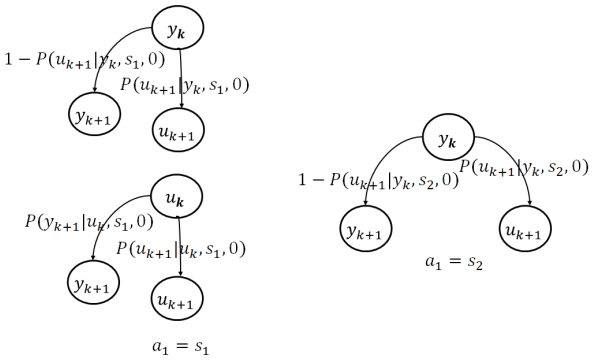

Fig. 5: State transition diagram when the nodes "stay" and the jammer does not jam

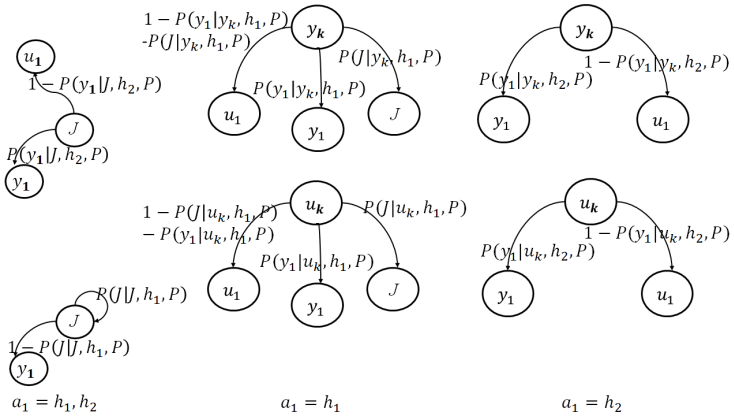

Fig. 6: State transition diagram when the nodes "hop" and the jammer jams

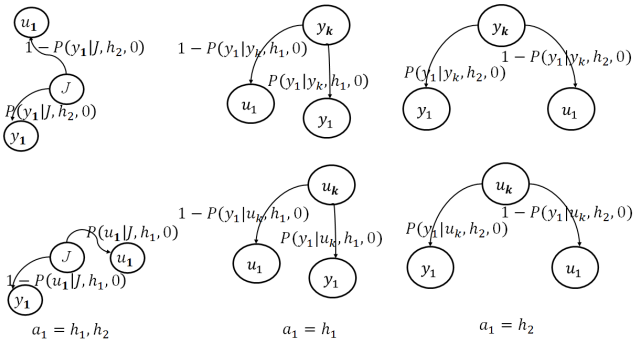

Fig. 7: State transition diagram when the nodes "hop" and the jammer does not jam

$P(J|y_k,h_1,P)$. Additionally, the nodes move to state y_1 if the new channel is not jammed and not in fading. We have

$$\begin{aligned}
P(J|u_k,h_1,P) &= \frac{q}{(\tilde{K}-k+1)\tilde{K}} + (1-p)P(J|y_k,h_1,P)\\
P(u_1|u_k,h_1,P) &= (1-p)(1-P(J|u_k,h_1,P))\\
P(y_1|u_k,h_1,P) &= 1-P(J|u_k,h_1,P)-P(u_1|u_k,h_1,P)\\
P(y_1|u_k,h_1,0) &= p = 1-P(u_1|u_k,h_1,0).
\end{aligned} \tag{13}$$

Given $(x,a_1,a_2)=(u_k,h_2,a_2), k=1,2,\cdots\tilde{K}-2$: The node can move to state u_1 or state y_1. It enters state y_1 if the channel is not jammed and not under fading. We have

$$\begin{aligned}
P(y_1|u_k,h_2,P) &= P(y_1|u_k,h_1,P)\\
P(u_1|u_k,h_2,P) &= 1-P(y_1|u_k,h_2,P)\\
P(y_1|u_k,h_2,0) &= P(y_1|u_k,h_1,0)\\
P(u_1|u_k,h_2,0) &= 1-P(y_1|u_k,h_2,0).
\end{aligned} \tag{14}$$

Possible transitions are shown in the Figure 4 and 5.

Lemma 1. The longer the nodes succeed on a channel, the higher the chance of success on the new channel when it hops.

Proof: The proof follows by verifying that regardless of a_2, $P(y_1|y_k,h_1,a_2)$ is increasing w.r.t. k, i.e.,

$$\begin{aligned}
P(y_1|y_{k+1},h_2,a_2) &\geq P(y_1|y_k,h_2,a_2)\\
P(y_1|y_{k+1},h_1,a_2) &\geq P(y_1|y_k,h_1,a_2).
\end{aligned} \tag{15}$$

□

Intuitively, the longer the nodes stay on a channel and their current transmission succeeds, the higher the number of channels that the jammer swept on which the nodes did not operate (in the current sweeping cycle). Hence, when it hops, it is likely that the new channel was already swept by the jammer and will be not attacked in the current sweep cycle. However, longer the nodes stay on a channel, the probability they get jammed on the channel increases. This implies that the nodes should balance the probability of getting jammed on the current channel and the probability of not getting jammed when they hop by a proper choice of channel residence time.

Policy: Policy of each player determines the action they take in each state. We shall be interested in Markov stationary policies where the nodes take an action based on current state and follows the same policy in each time slot. Following the notational convention of [11], let $\mathcal{M}(\mathcal{A})$ denote the distribution on set $\mathcal{A}$ and $\pi:\mathcal{X}\to\mathcal{M}(\mathcal{A})$ denote such a policy. $\pi(x)\stackrel{\text{def}}{=}\{\pi(x,a_1), a_1\in\mathcal{A}\}$, where $\pi(x,a_1)$ is the probability of choosing action $a_1\in\mathcal{A}$ when the nodes are in state $x\in\mathcal{X}$. We denote the collection of such policies as $\mathbf{\Pi}$. Similarly, let the jammer's strategy be $\mathbf{g}:\mathcal{X}\to\mathcal{M}(\mathbf{P}_J)$, and let $\mathbf{g}(x)\stackrel{\text{def}}{=}\{g(x,a_2), a_2\in\mathbf{P}_J\}$, where $g(x,a_2)$ is the probability of choosing action $a_2\in\mathbf{P}_J$ in state $x\in\mathcal{X}$. Since the jammer does not know the state, for any jammer's

strategy $\omega = (\omega_0, \omega_P) \in \mathbf{\Omega}$, $\mathbf{g}(x) = \omega, \forall x \in \mathcal{X}$.

4.1 Game with Discounted Reward Criterion

From state x, the immediate reward of the transmitter $r(x, a_1, a_2)$ where $r : \mathcal{X} \times \mathcal{A} \times \mathbf{P}_J \to \mathbb{R}$ and (pure) actions a_1 and a_2 are taken by the nodes and the jammer, respectively, is

$$r(x, a_1, a_2) = \sum_{x' \in \mathcal{X}} U(x, a_1, a_2, x') P(x'|x, a_1, a_2). \quad (16)$$

For a given $\pi \in \Pi$, and $\omega \in \mathbf{\Omega}$ the expected discounted reward/payoff of the nodes with an initial state $x \in \mathcal{X}$ is

$$\tilde{V}(x, \pi, \omega) = \mathbb{E}^{\pi,\omega} \left[\sum_{n=1}^{\infty} \delta^{n-1} r(X_n, A_{1n}, A_{2n}) | X_0 = x \right],$$

where $\{(X_n, A_{1n}, A_{2n}), n = 1, 2, \cdots, \infty\}$ is a sequence of random variables, denoting the state and the actions of the nodes and the jammer in each time slot. This sequence evolves according to the initial state and the policy (π, ω). The operator $\mathbb{E}^{\pi,\omega}$ denotes the expectation over the policies (π, ω).

The objective of the FD link is to choose a policy π that maximizes its expected reward $V(x, \pi, \omega)$ starting from any initial state $x \in \mathcal{X}$, and is defined as

$$\tilde{V}_L(x, \omega) = \max_{\pi \in \Pi} V(x, \pi, \omega). \quad (17)$$

On the other hand, the jammer's objective is to choose a policy ω that minimizes the nodes' expected discounted reward

$$\tilde{V}_J(x, \pi) = \min_{\omega \in \mathbf{\Omega}} V(x, \pi, \omega). \quad (18)$$

Note that the strategy space of jammer is constrained, whereas it is unconstrained for the nodes. A strategy pair (π^*, ω^*) is a constrained-Nash equilibrium (NE) if the following holds

- $\omega^* \in \mathbf{\Omega}$
- for all $x \in \mathcal{X}$, $\pi \in \Pi$ and $\omega \in \mathbf{\Omega}$

$$\tilde{V}(x, \pi, \omega^*) \leq \tilde{V}(x, \pi^*, \omega^*) \leq \tilde{V}(x, \pi^*, \omega).$$

Note that the Jammer cannot observe the state of the nodes. Hence its policy is independent of the states[11].

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

Proof: The proof follows by verifying strong Slater condition in [33][Thm 2.1]. Indeed, for all $\omega \in \mathbf{\Omega}$ and $\pi \in \mathbf{\Pi}$ jammer's average power constraint can be expressed the following discounted cost criteria

$$C(\omega, \pi) = (1 - \delta) \mathbb{E}_{\beta}^{\pi,\omega} \left[\sum_{n=1}^{\infty} \delta^{n-1} C_J(X_n, A_{1n}, A_{2n}) \right],$$

for any $\delta = (0, 1)$, and $C_J(X_n, A_{1n}, A_{2n}) = A_{2n}$. Where β denotes the initial distribution of states. Then $C(\omega, \pi) \leq P_{\text{avg}}$. Further, for $\tilde{\omega}$ such that $\tilde{\omega}_0 = 1$, $C(\tilde{\omega}, \pi) < P_{\text{avg}}$ for all π. Hence the strong slater condition holds and the constrained NE exists. □

The optimal strategy for the jammer is to attack the nodes as frequently as possible utilizing all of its power. Then, optimal strategy for the jammer is to emit power into the selected channel with probability P_{avg}/P in each time slot. In the following we compute the optimal strategy of nodes against this fixed optimal strategy of the jammer.

Optimal defense strategy of nodes: For a fixed jammer's strategy $\omega \in \mathbf{\Omega}$, the average reward of the nodes in state x for playing action a_1 is given as

$$r_\omega(x, a) = \omega_0 r(x, a_1, 0) + \omega_1 r(x, a_1, P) \quad (19)$$

and the probability of state transitioning to state x' is give as

$$P_\omega(x'|x, a) = \omega_0 P(x'|x, a_1, 0) + \omega_1 P(x'|x, a_1, P). \quad (20)$$

Since the jammer's strategy is fixed and does not depend on the state, the optimal policy for the nodes can be obtained by solving the resulting MDP with reward and transition probability computed in (19) and (20) against jammer's policy, and in particular against jammer's optimal policy. Let f_ω^* denote the optimal policy of the jammer against ω. Since f_ω^* can be obtained as optimal policy of an MDP, it is a deterministic strategy, i.e., $f_\omega^* : \mathcal{X} \to \mathcal{A}$. For notational convenience we drop the dependence of r_ω, P_ω and f_ω^* on ω. We also write $V(x) := \tilde{V}_L(x, \omega)$.

The well-known Bellman equations for the expected discounted utility maximization problem in (17) are as follows

$$Q(x, a) = r(x, a) + \delta \sum_{x' \in X} p(x'|x, a) \{V(x')\} \quad (21)$$

$$V(x) = \max_{a \in A} Q(x, a). \quad (22)$$

We can then use the value iteration [34][Ch. 6] method to derive the optimal defence strategy and its properties.

Theorem 2. The optimal policy π^* satisfies

- There exists a constant $K^* \in \{1, \ldots, \tilde{K} - 1\}$ and $i^* \in \{1, 2\}$ such that:

$$\pi^*(y_k) = h_{i^*} \text{ for } K^* \leq k \leq \tilde{K} - 1 \quad \text{and} \quad \pi^*(J) = h_{i^*}.$$

- There exits a constant $K_1^* \leq K^*$ such that $\pi(y_k) = s_2$ for all $1 \leq k \leq K_1*$ and $\pi(y_k) = s_1$ for all $K_1* < k < K^*$.

Lemma 2. $V(y_k) \geq V(y_{k+1})$ for all $k \leq K^*$ and $V(y_k) < V(y_{k+1})$ for $k \geq K^*$.

Proof: Using the Bellman relation in (22), we have

$$V(y_k) \geq Q(y_k, a) \text{ for all } a \in \mathcal{A}.$$

From the value iteration algorithm we know that if we start with any initialization of $V_0(x), x \in \mathcal{X}$ converges to the values $V(x)$. Without loss of generality assume that at some iteration i, $v_i(y_k) \geq V(y_{k+1})$ for $k \leq K^*$ and $V_i(y_k) < V_i(y_{k+1})$ for all $k \geq K^*$. Then using the last inequality we can show that the same ordering holds for $(i + 1)$th iteration as well. □

Outline of the proof of Theorem 2 is as follows: using Lemma 2, we show that $Q(y_k, s_1)$ and $Q(y_k, s_2)$ are decreasing in k, while $Q(y_k, h_1)$ and $Q(y_k, h_2)$ are increasing

11. The jammer knows the state of the nodes only when they are jammed. But this is of no use as the nodes will leave the channel in the next slot. The jammer can also infer the state of the nodes in the round subsequent to jamming (as it can be either y_1 or u_1). However, we ignore this information as it does not changes the final analysis.

in k. The structure of the policy then follows by noting that optimal action is each state is selected greedily according to (22). Detailed proof is given in the appendix A.

The above result suggests the following optimal strategy: If the nodes are successful on a channel for K^* number of slots, they should leave the channel. On the new channel, nodes should operate for the next K^* slots unless they are jammed. While they stay on the new channel, the nodes should operate in the TR for the first K_1^* slots, and then switch to the TD mode for the next $K^* - K_1^*$ slots. If they are jammed while operating in the TD mode they should hop immediately. We note that, for some set of parameters, the optimal policy could be such that $K_1^* = 1$, in which case the nodes never use the TR mode, and in some cases $K_1^* = K^*$, in which case the nodes never use the TD mode. In state u_k, the nodes use either s_1 or hop depending on C and L.

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

Proof: The proof follows by noting that for any $k' > k$, $Q(y_{k'}, s_i) - Q(y_k, s_i)$ is increasing in L and decreasing in $K, \forall i \in \{1, 2\}$. Moreover, $Q(x, h_i)$ is decreasing in $C, \forall i \in \{1, 2\}, x \in X$. This verifies that K^* is decreasing in C. □

4.2 Optimal Defense Strategy in Average Reward Criterion

In this section we compute the optimal defense strategy when the utility of nodes is the average reward. In the sequel, we follow conventional notations in [11]. Lets define a $|X| \times |X|$ stochastic transition probability matrix $\boldsymbol{P}$ where its element $P(x, x')$ is the transition probability from state x to state x' when the stationary policy Π is employed. Let $r^{(t)}(x, \Pi)$ be the expected reward at time t when the transmitter starts with an initial state x, and $\boldsymbol{r}^{(t)}(\Pi) \stackrel{\text{def}}{=} (r^{(t)}(1, \Pi), \cdots, r^{(t)}(|X|, \Pi))$ be the expected reward vector for all initial states $x \in A$. We have

$$\boldsymbol{r}^{(t)}(\Pi) = \boldsymbol{P}^t \boldsymbol{r}(\Pi), \tag{23}$$

where $\boldsymbol{r}(\Pi) \stackrel{\text{def}}{=} (r(1, \Pi), \cdots, r(|X|, \Pi))$ is the vector of immediate expected reward of the transmitter for all $|A|$ initial states.

If the underlying Markov chain for a given stationary policy $\pi \in \Pi$ is irreducible, the following average reward (or reward rate) of the transmitter (while starting from state x) exists

$$V^{av}(x, \Pi) \stackrel{\text{def}}{=} \lim_{T \to \infty} \frac{1}{1+T} \sum_{t=0}^{T} r^{(t)}(s, \Pi). \tag{24}$$

From the above transition probabilities, for any stationary policy $\pi(x, :)$ that implements either action h_1 or s_1 with non-zero probability, the transmitter can visit state J and from state J, it can recover to move to any state y_k and u_k with non-zero probability. In such cases, the underlying Markov chain is irreducible and the above average reward is well-defined. However, the irreducibility of the Markov chain is not always guaranteed. For example, when chance of being under fading $1 - p$ is so small that a transmission failure likely suggests it is under jamming, it is not necessary for the transmitter to determine the cause of failure but to hop onto another channel. In such cases, state J is not visited. Hence, the selection of using the average reward criterion should depend on the channel quality.

According to Theorem 2.4.4 and Corollary 2.4.5 in [11], there exists an optimal NE pure strategy π^* to maximize the transmitter's average reward. The pure strategy is formally stated as follows.

Let a $4|X| \times 1$ vector $\boldsymbol{f} \stackrel{\text{def}}{=} [f_{x,a}] = ([f(1, s_1), \cdots, f(1, h_2)], \cdots, [f(|X|, s_1), \cdots, f(|X|, h_2)])$ be the solution of the following programming

$$\begin{aligned} \underset{\{f_{x,a}\}}{\text{maximize}} \quad & \sum_{x \in X} \sum_{a \in A} r(x, a) f_{x,a} \\ \text{s.t.} \quad & \text{C1: } \boldsymbol{W}\boldsymbol{x} = 0 \\ & \text{C2: } \boldsymbol{1}^T \boldsymbol{f} = 1 \\ & \text{C3: } f_{s,a} \geq 0, \forall x, a \end{aligned} \tag{25}$$

where $\boldsymbol{1}$ is a all-one $1 \times 4|X|$ vector; $\boldsymbol{W}$ is an $4|X| \times |X|$ matrix whose element $W(x', (x, a)) = -P(x'|x, a)$ if $x' \neq x$ or $W(x', (x, a)) = 1 - P(x'|x, a)$ if $x' = x$.

Then, $\pi^*(x, a) = 1$ if $f(x, a) > 0$, otherwise $\pi^*(x, a) = 0$.

5 Performance Evaluation

To demonstrate the benefits of using the TR and TD operational modes we compare the performance of our policy, referred to as "Jointly Optimal" against two strategies which we refer as "Optimal FH" and "Random FH". Optimal FH policy is obtained by optimizing the channel residence time restricting the mode of operation to TR only. In the Random FH policy, nodes always hop and use only the TR mode. The jointly optimal and the optimal FH policy are computed solving Bellman equations in (21) using value iteration algorithm, allowing both TR and TD modes for the former and allowing only TR mode for the later. For all the policies we compute the average goodput (in Mbps), the success rate (percentage of un-jammed transmissions), and the average regret (in Mbps) as in equation (4). The parameters of study are K, C, L, q and ξ. We set $R = 25$ Mbps and $p = .8$, and $m = 2$. Unless stated otherwise, we use the following parameters: $K = 8$, $L = 6$ Mbps, $C = 8$ Mbps, $\xi = .7$ and $q = 1$, in the plots. We show the performance for the policy for the case of discounted rewards. Similar behavior is expected for the case of average reward.

It is easy to see that if the nodes hop in every slot, then the probability of them getting jammed in a slot is the lowest. Hence, if $C = 0$ and $L = 0$, the optimal policy for the nodes is to hop in each slot, i.e., the Random FH policy, and it results in highest success rate. We thus selected Random FH for comparison with the success rate of the optimal policies.

Effect of number of channels (K): Figures (8), (9), and (10) plot the regret, success rate and throughput of the three algorithms vs. the number of channel K. As seen, the Jointly optimal policy has much lower regret and higher throughput than the Optimal FH (TD mode is not used) or Random FH. This is due to the effective utilization of TD to increase the channel residence time (which in turn increases probability of success on hopping). Since the nodes hop in every time slot in the Random FH policy, its success rate is higher than that of the optimal hopping policy (Fig.(9)).

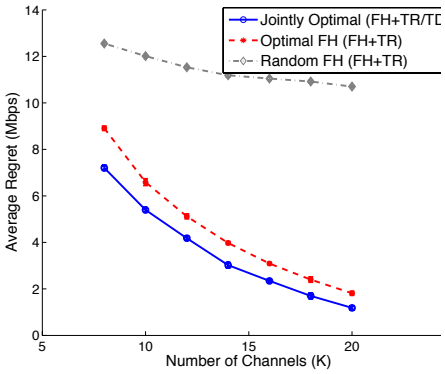

Fig. 8: Avg. Regret vs. K

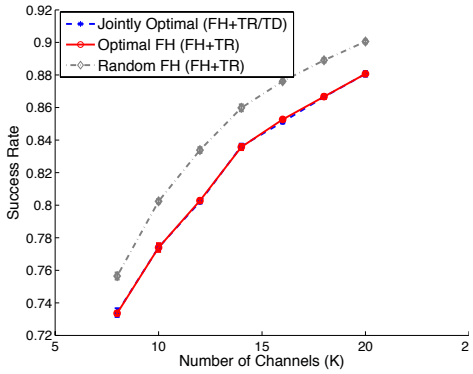

Fig. 9: Success rate vs. K

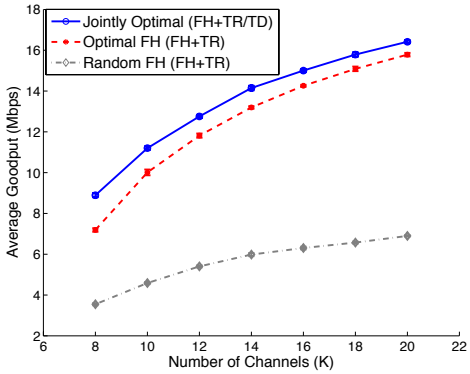

Fig. 10: Avg. goodput vs. K

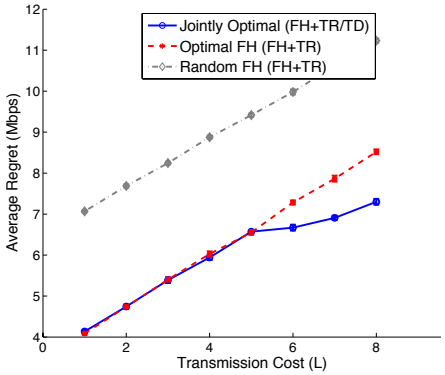

Fig. 11: Avg. Regret vs. L

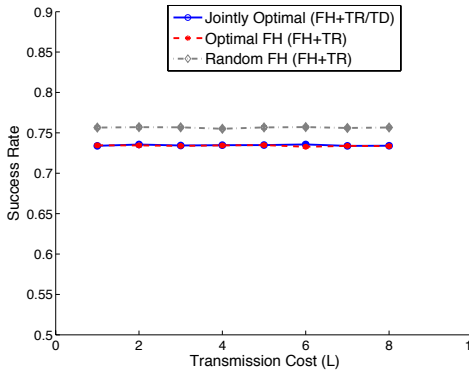

Fig. 12: Success rate vs. L

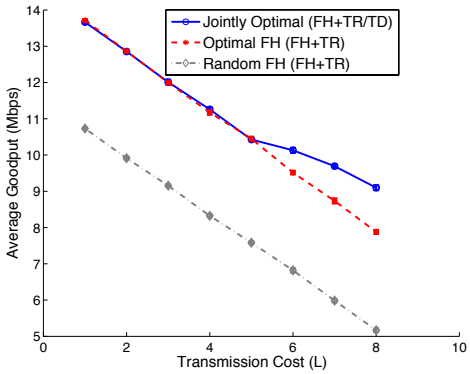

Fig. 13: Avg. goodput vs. L

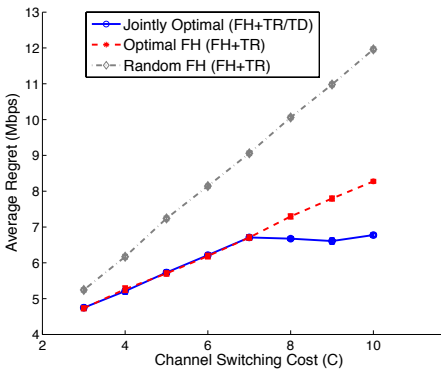

Fig. 14: Avg. Regret vs. C

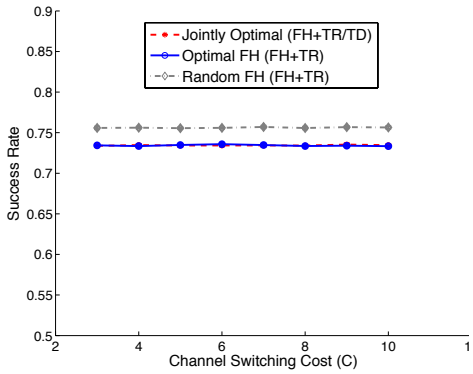

Fig. 15: Success rate vs. C

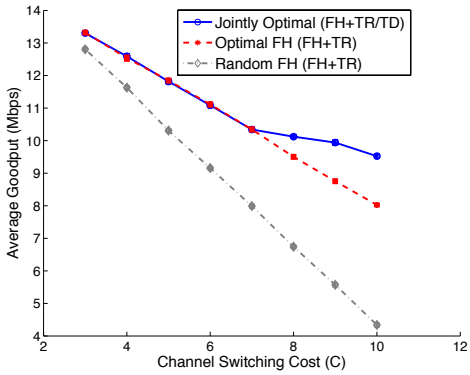

Fig. 16: Avg. goodput vs. C

However, hopping too frequently makes the random hopping policy's regret and throughput worse than that of the Jointly optimal policy (Figs. (10) and (8)). This is because the Jointly optimal policy efficiently avoids unnecessary hops to reduce switching costs. Additionally, as the number of channel increases, the Jointly optimal policy becomes more efficient in combating the jammer in all performance metrics.

Effect of transmission cost (L): Figures (11), (12), and (13) plot the regret, success rate and throughput of the three algorithms as L varies. The regret of Optimal policy is significantly lower (more than 43%) than the other policies, especially when the transmission cost is higher. This is because with higher transmission cost, the loss due to failure/jamming is also higher, that makes the optimal decisions in hopping or switching TD/TR mode of the jointly optimal policy more pronounced. As we can see in Fig. (12), the success rate of the policies remain the same as we vary L. Notice that the plot is for a fixed K and the nodes have the same amount of room (channels) to escape form the jammer for any value of L. Jointly optimal policy is almost the same as that when the nodes hop after every time slot while still attaining higher throughput as seen in (13). As we increase L on the x-axis it contributes linearly to the regret as every

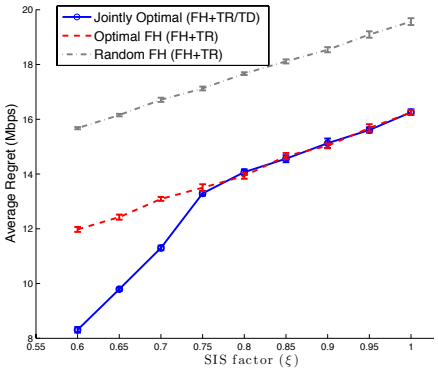

Fig. 17: Avg. regret vs. ξ

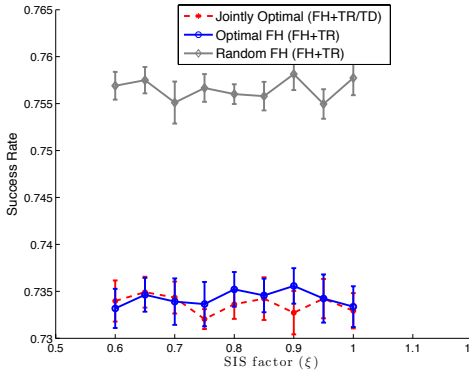

Fig. 18: Success rate vs. ξ

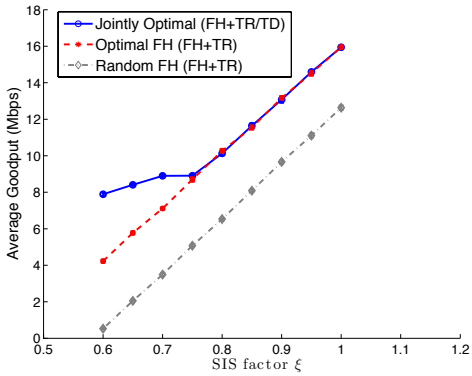

Fig. 19: Avg. goodput vs. ξ

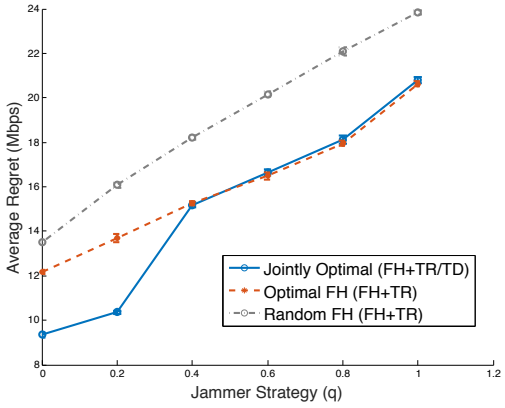

Fig. 20: Avg. Regret vs Jammer strategy (q)

time jamming occurs more amount of loss is incurred with increase in L. Also, the jointly optimal policy has incurs lesser number of jammed slots as Optimal FH snd hence achieves a better regret performance for larger values of L.

Effect of hopping cost (C): In Figures (14), (15), and (16) we compare regret, success rate and throughput of Jointly optimal policy against Random FH and Optimal FH as C varies. The effect of C is similar to that of L in Figures (10)-(11), and the jointly optimal policy is the most robust algorithm to jamming (with more than 20% lower regret, on average, than others').

Effect of imperfect SIS ξ: In Figures (17), (18), and (19) we depict regret, success rate and throughput vs. the effect of imperfect SIS (ξ) of the Jointly optimal policy against Random FH and Optimal FH. As can be seen, the less perfect SIS (i.e., lower ξ), the more effective in combating jamming of the Jointly optimal policy. On average, the Jointly optimal policy yields less than 45% regret than the Random FH policy. Similar to the above, the success rate of the Jointly optimal policy is almost the same as that when nodes hop in every time slot.

Effect of Jammer's strategy q: Figure (20) show the effect of jammer's strategy on the regret. As expected, the regret is monotonically increasing in q. Notice that when the value of q is low, the use of both TD and TR mode helps in improving the throughput. The benefit of using both TR and TD mode diminishes as q is increased comparing to the case when only TR mode is used.

6 Conclusion

In this work, we identified the severe susceptibility to jamming attack of wireless nodes that are equipped with in-band full-duplex radios (IBFD). To combat jammers, we then defined two operational modes for the IBFD radios: *Transmission Reception* (TR) and *Transmission Detection* or half-duplex (TD). Together with the low Packet Detection Rate, jamming can be effectively detected by allowing IBFD radios to switch to the TD mode. We developed an optimal strategy against an "IBFD reactive sweep jammer". The nodes' optimal defense strategy that guides them when to hop and which operational mode to use is established from the equilibrium of the stochastic Markov game. We prove that this optimal policy has a threshold-structure, in which the IBFD nodes stay on the same channel up to a certain number of time slots before hopping.

References

[1] M. K. Hanawal, D. N. Nguyen, and M. Krunz, "Jamming attack on in-band full-duplex communications: Detection and countermeasures," in *The 35th Annual IEEE International Conference on Computer Communications INFOCOM*, April 2016, pp. 1–9.

[2] D. Bharadia, E. McMilan, and S. Katti, "Full duplex radios," in *Proc. of the ACM SIGCOM'13 Conf.*, Hong Kong, China, Aug 2013.

[3] A. Sabharwal, P. Schniter, D. Guo, D. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1637–1652, Sept 2014.

[4] Y. Zhang, L. Lazos, K. Chen, B. Hu, and S. Shivaramaiah, "FD-MMAC: combating multi-channel hidden and exposed terminals using a single transceiver," in *Proceedings of the IEEE INFOCOM Conference*, 2014.

[5] M. Amjad, F. Akhtar, M. H. Rehmani, M. Reisslein, and T. Umer, "Full-duplex communication in cognitive radio networks: A survey," *IEEE Communications Surveys Tutorials*, vol. 19, no. 4, pp. 2158–2191, 2017.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

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

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

[12] Y. Liao, T. Wang, L. Song, and Z. Han, "Listen-and-talk: Protocol design and analysis for full-duplex cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 1, pp. 656–667, 2017.

[13] D. Li, J. Cheng, and V. C. M. Leung, "Adaptive spectrum sharing for half-duplex and full-duplex cognitive radios: From the energy efficiency perspective," *IEEE Transactions on Communications*, pp. 1–1, 2018.

[14] V. Towhidlou and M. Shikh-Bahaei, "Adaptive full-duplex communications in cognitive radio networks," *IEEE Transactions on Vehicular Technology*, 2018.

[15] T. Zhang, Y. Cai, Y. Huang, T. Q. Duong, and W. Yang, "Secure full-duplex spectrum-sharing wiretap networks with different antenna reception schemes," *IEEE Transactions on Communications*, vol. 65, no. 1, pp. 335–346, 2017.

[16] X. Tang, P. Ren, and Z. Han, "Combating full-duplex active eavesdropper: A game-theoretic perspective," in *IEEE International Conference on Communications (ICC)*, 2016, pp. 1–6.

[17] S. Yan, X. Zhou, N. Yang, T. D. Abhayapala, and A. L. Swindlehurst, "Channel training design in full-duplex wiretap channels to enhance physical layer security," in *IEEE International Conference on Communications (ICC)*, 2017, pp. 1–6.

[18] Y. E. Sagduyu, R. A. Berry, and A. E. Ephremides, "Jamming games in wireless networks with incomplete information," *IEEE Communications Magazine*, vol. 49, no. 8, pp. 112–118, August 2011.

[19] K. Pelechrinis, I. M, and S. Krishnamurthy, "Denial of service attacks in wireless networks: The case of jammers," *IEEE Journal of Communications Surveys & Tutorials*, vol. 13, pp. 245–257, 2010.

[20] L. Xiao, J. Liu, Q. Li, N. B. Mandayam, and H. V. Poor, "User-centric view of jamming games in cognitive radio networks," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 12, pp. 2578–2590, Dec 2015.

[21] G. Chang, S. Wang, and Y. Liu, "A jamming-resistant channel hopping scheme for cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 10, pp. 6712–6725, Oct 2017.

[22] L. Xiao, T. Chen, J. Liu, and H. Dai, "Anti-jamming transmission stackelberg game with observation errors," *IEEE Communications Letters*, vol. 19, no. 6, pp. 949–952, June 2015.

[23] 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.

[24] 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.

[25] M. K. Hanawal, M. J. Abdel-Rahman, D. Nguyen, 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, Aug. 2013. [Online]. Available: http://www2.engr.arizona.edu/~krunz/

[26] T. Erpek, Y. E. Sagduyu, and Y. Shi, "Deep learning for launching and mitigating wireless jamming attacks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 1, pp. 2–14, March 2019.

[27] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE Journal of Selected Areas in Communication (JSAC)*, vol. 32, pp. 1637–1652, 2014.

[28] Y. Zhang, L. Lazos, K. Chen, B. Hu, and S. Shivaramaiah, "FD-MMAC: Combating multi-channel hidden and exposed terminals using a single transceiver," in *Proceeding of the IEEE INFOCOM'14*, 2014.

[29] J.Kim, M. Alfowzan, and M. Krunz, "Power-controlled channel access protocol for wireless networks with full-duplex and ofdma capabilities," in *Proc. of IEEE SECON'11*, 2015.

[30] S. Gollakota, H. Hassanieh, B. Ransford, D. Katabi, and K. Fu, "They can hear your heartbeats: Non-invasive security for implanted medical devices," in *Proc. of ACM SIGCOMM'11*, Toronto, Canada, Aug 2011.

[31] G. Zheng, I. Krikidis, J. Li, A. P. Petropulu, and B. Ottersten, "Improving physical layer secrecy using full-duplex jamming receivers," in *IEEE Transaction on Signal Processing*, vol. 61, no. 20, 2013.

[32] 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.

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

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

APPENDIX

Proof of Theorem 2: The nodes can take all possible actions in states $y_k \in \mathcal{Y}$. The cost-to-go value of station-action pairs (y_k, a) for all, $k = 1, 2, \cdots, \tilde{K} - 2$ and $a \in \mathcal{A}$ are as follows (after simplification)

$$\begin{aligned}
Q(y_k, s_1) &= -L + \left(p\left(1 - \frac{1}{\tilde{K}-k}\right)\right)(L+R) \\
&+ \delta(V(J) + V(u_{k+1}) + V(y_{k+1})) \\
Q(y_k, s_2) &= -2L + (2L + 2\xi R)\left(1 - \frac{1}{\tilde{K}-k}\right) \\
&+ \delta(V(u_{k+1}) + V(y_{k+1})) \\
Q(y_k, h_1) &= p\left(\frac{K-1-mk}{K-1}\left(1 - \frac{1}{\tilde{K-k}}\right)\right)(R - C + L) \\
&+ \frac{mk}{K-1}(R - C + L) - L \\
&+ \delta(V(J) + V(u_{k+1} + V(y_{k+1}))) \\
Q(y_k, h_2) &= p\left(\frac{K-1-mk}{K-1}\left(1 - \frac{1}{\tilde{K-k}}\right)\right)(2\xi R - 2C + 2L) \\
&+ \frac{mk}{K-1}(2\xi R - 2C + 2L) - 2L \\
&+ \delta(V(u_1) + V(y_1))
\end{aligned}$$

When the nodes are in state J, the Q-values can be calculated as follows

$$\begin{aligned}
Q(J, h_1) &= -L + p\left(1 - \frac{1}{\tilde{K}}\right)(R - C + L) \\
&+ \delta(V(J) + V(u_1) + V(y_1)) \\
Q(J, h_2) &= -2L + p\left(1 - \frac{1}{\tilde{K}}\right)(2R\xi - 2C + 2L) \\
&+ \delta(V(u_1) + V(y_1)).
\end{aligned}$$

First consider the case $Q(J, h_2) > Q(J, h_1)$. Then, the nodes hop in TR mode whenever they are jammed and $V(J) = Q(J, h_2)$. Substituting this value in $Q(y_1, s_1)$ and $Q(y_1, s_2)$, we get $Q(y_1, s_2) \geq Q(y_1, s_1)$. Further, note that $Q(y_k, s_1)$ and $Q(y_k, s_2)$ are decreasing in k. This implies that the longer the nodes stay on the same channel, their vulnerability to jamming increases and the average reward decreases. Next, note that both $Q(y_k, h_1)$ and $Q(y_k, h_2)$ are also decreasing in k. However, the rate at which they decrease is smaller than that of decrease of $Q(y_k, s_1)$ and $Q(y_k, s_2)$. This implies that there exists $1 < K* < \tilde{K} - 1$ such that $Q(y_k, h_2) \leq Q(y_k, s_2)$ and $Q(y_k, h_2) \leq Q(y_k, s_1)$ for $k \leq K^*$ and $Q(y_k, h_2) \geq Q(y_k, s_2)$ and $Q(y_k, h_2) \geq Q(y_k, s_1)$ for $k > K^*$. Hence, the nodes stay on the same channel for K^* slots before they hop.

Consider the difference $Q(y_k, s_2) - Q(y_k, s_1)$. This difference is decreasing in k. This implies that there exists a K_1^* such that for $k \leq \tilde{K}_1^*$, $Q(y_k, s_2) \geq Q(y_k, s_1)$ and for $k > \tilde{K}_1^*$, $Q(y_k, s_2) \leq Q(y_k, s_1)$. Hence, the nodes take action s_2 in state for y_k if $k \leq \tilde{K}_1^*$ otherwise they take action s_1. If $\tilde{K}_1^* < K^*$, we set $K_1^* = \tilde{K}_1^*$, otherwise we set $K_1^* = K^*$.

For the case $Q(J, h_2) \leq Q(J, h_1)$, it is easy to note the nodes never enter the TR mode. Then $\tilde{K}_1^* = K^*$, i.e., nodes take action s_1 in the first K^* if they remain unjammed and then hop.

Manjesh K. Hanawal received the M.S. degree in ECE from the Indian Institute of Science, Bangalore, India, in 2009, and the Ph.D. degree from INRIA, Sophia Antipolis, France, and the University of Avignon, Avignon, France, in 2013. After two years of postdoc at Boston University, he is now an Assistant Professor in Industrial Engineering and Operations Research at the Indian Institute of Technology Bombay, Mumbai, India. His research interests include performance evaluation, machine learning, and network economics. He is a recipient of Inspire Faculty Award from DST and Early Career Research Award from SERB.

Diep N. Nguyen received the M.E. degree in electrical and computer engineering from the University of California at San Diego (UCSD) and the Ph.D. degree in electrical and computer engineering from The University of Arizona (UA). He was a DECRA Research Fellow with Macquarie University and a Member of Technical Staff with Broadcom, CA, USA, ARCON Corporation, Boston, consulting the Federal Administration of Aviation, on turning detection of UAVs and aircraft, and the U.S. Air Force Research Laboratory, on anti-jamming. He is currently a Faculty Member with the Faculty of Engineering and Information Technology, University of Technology Sydney (UTS). His recent research interests include computer networking, wireless communications, and machine learning application, with emphasis on systems performance and security/privacy. He has received several awards from LG Electronics, UCSD, The University of Arizona, the U.S. National Science Foundation, and the Australian Research Council.

Marwan Krunz : is the Kenneth VonBehren Endowed Professor in the ECE Department at the University of Arizona. He is also an affiliated faculty member of the University of Technology Sydney. He directs the Broadband Wireless Access and Applications Center, a multi-university industry-focused NSF center that includes members from industry and government labs. Previously, he served as the UA site director for Connection One, an NSF IUCRC that focuses on wireless communication circuits and systems. In 2010, Dr. Krunz was a Visiting Chair of Excellence at the University of Carlos III de Madrid. He held visiting research positions at UTS, INRIA-Sophia Antipolis, HP Labs, University of Paris VI, University of Paris V, University of Jordan, and US West Advanced Technologies. Dr. Krunzs research interests are in wireless communications and networking, with emphasis on resource management, adaptive protocols, and security issues. He has published more than 280 journal articles and peer-reviewed conference papers, and is a co-inventor on several US patents. He is an IEEE Fellow, an Arizona Engineering Faculty Fellow (2011-2014), and an IEEE Communications Society Distinguished Lecturer (2013 and 2014). He was the recipient of the 2012 IEEE TCCC Outstanding Service Award. He received the NSF CAREER award in 1998. He currently serves as the Editor-in-Chief for the IEEE Transactions on Mobile Computing. He previously served on the editorial boards for the IEEE Transactions on Cognitive Communications and Networks, IEEE/ACM Transactions on Networking, IEEE TMC, IEEE Transactions on Network and Service Management, Computer Communications Journal, and IEEE Communications Interactive Magazine. He was the general vice-chair for WiOpt 2016 and general co-chair for WiSec12. He was the TPC chair for WCNC 2016 (Networking Track), INFOCOM04, SECON05, WoWMoM06, and Hot Interconnects 9. He has served and continues to serve on the steering and advisory committees of numerous conferences and on the panels of several funding agencies. He was a keynote speaker, an invited panelist, and a tutorial presenter at numerous international conferences. See http://www2.engr.arizona.edu/ krunz/ for more details.