Jamming Attack on In-band Full-duplex Communications: Detection and Countermeasures

Manjesh K. Hanawal
IEOR Group, IIT-Bombay
Mumbai-460076, India
Email: mhanawal@iitb.ac.in

Diep N. Nguyen
FEIT, UTS
Sydney, Australia
Email: diep.nguyen@uts.edu.au

Marwan Krunz, *Fellow IEEE*
Dept. of ECE, University of Arizona
Tucson, AZ 85721, USA
Email: krunz@email.arizona.edu

***Abstract*—Recent advances in the design of in-band full-duplex (IBFD) radios promise to double the throughput of a wireless link. However, IBFD-capable nodes are more vulnerable to jamming attacks than their out-of-band full-duplex (OBFD) counterparts, and any advantages offered by them over the OBFD nodes can be jeopardized by such attacks. A jammer needs to attack both the uplink and the downlink channels to completely break the communication link between two OBFD nodes. In contrast, he only needs to jam one channel (used for both uplink and downlink) in the case of two IBFD nodes. Even worse, a jammer with the IBFD capability can learn the transmitters' activity while injecting interference, allowing it to react instantly with the transmitter's strategies. In this paper, we investigate frequency hopping (FH) technique for countering jamming attacks in the context of IBFD wireless radios. Specifically, we develop an optimal strategy for IBFD radios to combat an "IBFD reactive sweep jammer". First, we introduce two operational modes for IBFD radios: transmission-reception and transmission-detection. These modes are intended to boost the anti-jamming capability of IBFD radios. We then jointly optimize the decision of when to switch between the modes and when to hop to a new channel using Markov decision processes. Numerical investigations show that our policy significantly improves the throughput of IBFD nodes under jamming attacks.**

I. Introduction

Recent advances in self interference suppression (SIS) (e.g., [1], [2]) allows a transmitting device to suppress its self-interference up to the noise floor, enabling wireless radios to simultaneously transmit and receive on the same channel. This in-band full-duplex (IBFD) capability not only doubles link throughput but also helps solve various issues (e.g., Tx deafness, hidden/exposed nodes) at the MAC and network layers [3]. 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 such jamming threats and investigate mitigation techniques that optimally leverage the simultaneous transmit-and-receive capability of IBFD devices.

In a jamming attack, an adversary (jammer) can hinder legitimate transmissions in one of two ways: (i) he 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 attack (DoS). Such stealth jamming attacks can be easily launched by an adversary using commercial off-the-shelf (CoTS) products [4]–[6]. In this work, we focus on the former type of jamming attack.

The jamming consequences on IBFD radios is particulary acute. First, compared with out-of-band full-duplex (OBFD) systems, which include HD devices as a special case, a jammer can interfere with both the uplink and downlink *simultaneously* (as both IBFD radios that are within the same vicinity are likely suffer the same jamming effect). Second, unlike OBFD radios,

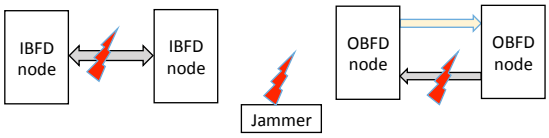

Fig. 1: Effect of jamming on IBFD and OBFD nodes.

operating in IBFD mode hinders the nodes' jamming detection capability, especially under the fading effect. Specifically, under fading, a transmission failure is not always caused by jammer while the jamming interference (if any) perceived by RF receiving chains 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 jamming. These include spread spectrum (particularly, frequency hopping (FH)), directional antennas, and adaptive power/coding/modulation. Jammer-specific techniques have also been developed [4]. Common jamming models in the literature include random, persistent, proactive, and reactive jammers [5], [7]. This classification is based on the channel behavior and the jammer's transmission capabilities. Persistent jammers always emit power into the medium. Proactive jammer can vary the power to meet various constraints. A reactive jammer exhibits more sophisticated capabilities, and emits power only when it detects a legitimate transmission [6]. In this paper, we consider a jammer with an IBFD capability, referred to as "IBFD reactive sweep jammer". This jammer sweeps through blocks of m channels in each slot. Further, while jamming, the jammer can simultaneously learn if the attack is successful and accordingly adapts its strategy (thanks to its SIS capability). Our jamming model explained in detail in Section III.

If a transmission fails, the sending nodes should not always hop to a new channel as this channel can be experiencing fading and hopping involves reduction in throughput due to the need to allow the oscillator to settle down after changing its operating frequency. However, if a node can reliably detect the presence of a jammer, it should hop to evade this jammer. As noted in [4] it is not possible to reliably identify the presence of a jammer through measurements only. Hence it is necessary to use some consistency checks to ascertain such presence. The jamming mitigation technique we develop uses the packet delivery ratio (PDR[1]) as an indicator of potential jamming activity, while leveraging the simultaneous transmit and receive capability of IBFD nodes to reliably confirm that.

Our proposed method defines two operational modes for each IBDF node: *Transmission Reception* (TR) and *Transmission Detection* (TD). In the TR mode, a node can transmit and receive

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

978-1-4673-9953-1/16/$31.00 ©2016 IEEE

data simultaneously. In the TD mode, a node does not transmit any data packets but only receives, i.e., acts as a receiver. When the PDR between a pair of nodes is low, one of the nodes can switch to TD mode. This node then receives only ambient noise and can thus identify the link quality by measuring signal strength (RSS). This information in turn can be used to reliably confirm the cause of low PDR. Specifically, if the RSS is high but the PDR is low, then the high RSS is likely from the jammer's interference power. On the other hands, if the low PDR is caused by fading, then the RSS will be low. Note that in the TD mode, the higher throughput of IBFD radios is not possible, but any jamming activity can be reliably detected. On the other hand, if the nodes operate in the TR mode, they get higher spectral efficiency, but jamming activity cannot be detected.

In this work, the behavior of IBFD radios to combat a IBFD reactive sweep jammer is captured by a Markov decision process [8]. The strategic decisions of IBFD radios regarding which mode to operate and when to hop to a different channel (equivalently, the duration of the channel residency time) are jointly optimized so that the aggregate throughput is maximized (under both discounted and average reward criteria). The main contributions of the paper include:

- We identify the severe susceptibility of IBFD radios to jamming attacks from IBFD-capable jammers.
- We analyze a jamming scenario in a network with IBFD radios and discuss the attack and defence strategies of the jammer and the nodes, respectively.
- We define two operational modes for the nodes that help them identify the cause of poor link quality. We then derive the optimal strategy based on the total discounted reward and the average reward criteria. Such a strategy informs an IBFD node when to hop to a new channel and which operational mode to use.
- We compare the performance of the jointly optimal FH and mode-switch strategy with the optimal strategy that is based on FH only (i.e., without switching between the two proposed operating modes). Through numerical simulations we show that the defense strategy obtained by jointly optimizing FH and mode switching results in better performance than that based on FH only.

Related work: Jamming and anti-jamming techniques are well studied in wireless networks with HD and OBFD devices (see [9], [10] 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 [11] 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 [12] using MDP for a cognitive radio network. The authors in [13] developed a defense strategy, combining FH and rate adaption techniques.

Recently, several authors proposed protocols that leverage IBFD capabilities to improve the performance of ad hoc and cellular networks [14], [15], [16]. However, they did not take into account the vulnerability of IBFD nodes to jamming attacks. In [17], [18], IBFD nodes are 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 simultaneous transmit and receive capabilities of these devices.

Paper organization: In Section II we describe the problem setup. In Section III 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 IV using the MDP. Its performance evaluation through numerical simulation is given in Section V. Finally, in Section VI we discuss future work and give concluding remarks.

II. Model and Setup

Consider two IBFD nodes A and B that communicate in the presence of a jammer, as shown in Figure 1. The two nodes have IBFD radios that can operate on any of K available channels. Let $\mathcal{F} = \{f_1, \ldots, f_K\}$ denote the set of non-overlapping channels. Each channel experiences additive white Gaussian noise (AWGN), which is independently and identically distributed (i.i.d.) across all channels.

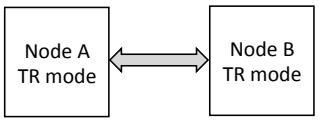

Fig. 2: TR mode

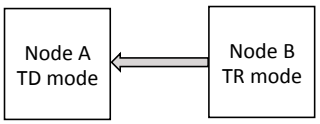

Fig. 3: TD mode

A. Transmitter and Channel Models

We assume that time is slotted, and transmissions are packet based. In each time slot, several packets can be transmitted. During this time, 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 SINR.

The two-state Gilber-Elliot channel model [19] is used to characterize the fading process. At a given point in time, each channel can either be in a fading state with probability (1-p) or not with probability p. We assume that when the channel is not under fading, a transmission always succeeds in the absence of jamming attack. In contrast, if the channel is in the fading state, the transmission always fails irrespective of jamming. We further assume that the fading status is independently and identically distributed (i.i.d) across time and all channels. Note that our analysis here can also be manipulated to accommodate (a more realistic) finite-state Markov channel model.

With IBFD capability, when both uplink and downlink are active, the net throughput achieved in the absence of jamming is $\xi_1 R + \xi_2 R$, where R denotes the throughput obtained when only one of the link (uplink or downlink) is used and $0.5 < \xi_i \leq 1$[2] for $i = 1, 2$ denotes the fraction of throughput loss due to imperfect SIS at node i. ξ_i is referred to as the *SIS factor*. We assume that both nodes are identical in their SIS capabilities, hence set $\xi \stackrel{\text{def}}{:=} \xi_1 = \xi_2$.

Modes of Operation: In the TR mode, a node transmits and receives simultaneously on a link. In the TD mode, a node does not transmit, but only receives data packets from the other. With some abuse of terminology, we say that a pair of nodes operate in TR mode when both of the IBFD nodes operate in the TR mode (Fig. 2), and we say that they operate in TD mode when one of them operates in the TD mode while the other operates in the TR mode (Fig. 3). When one of the nodes, say node A in Figure 3, operates in the TD mode, node B only receives the ambient noise and can measure its strength. If both the nodes are under jamming attack, then any of the nodes 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

[2] $\xi_i > 0.5$ to make sure that operating in the full-duplex TR mode yields higher throughput than the half-duplex mode.

only one of the nodes and not the other[3], then the node that is in close proximity to the jammer suffers low PDR, hence this node can measure the power of ambient noise if the other node operates in the TD mode. In the rest of the paper, we assume that which node enters into the TD is agreed upon a priori and focus on the scenario where the jammer can attack both the nodes simultaneously, as depicted in Figure 1. Our strategy can be easily adopted to other cases.

Switching and Transmission Cost: When the nodes hop, they are first required to reconfigure the device on the new channel and cannot immediately start the transmissions. The duration this *settling time* depends on the device (e.g., for the Anthros chipset card, this time is about 7.6 ms [11]). Additional loss in throughput occurs due to the lack of synchronization between the Tx and Rx's hopping instances. Collectively, we denote the average loss in throughput due to hopping by C, and refer to it as *switching cost*. Outage periods also occur when the nodes are jammed. Jamming disrupts the link between the nodes, which needs to be re-established through exchanging several control packets that do not contribute to data 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.

B. Jamming Model

A jammer with unrestricted resources can attack all the channels simultaneously with sufficient power to make it infeasible for the nodes to operate. However, often the jammers are subject to power limitation and are of the similar configurations as the legitimate nodes, making it feasible to launch a defense. We consider attack from multiple jammers that are IBFD-capable. Specifically, each jammer attacks one of the channel in $\mathcal{F}$ in each time slot and simultaneously observes the activity of legitimate nodes (if any) and learns jamming outcome using its RF receiving chains.

The jammers can co-ordinate among themselves by attacking non-overlapping channel to increase their chance of success. Further, when a jammer detects activity on a channel, they can all simultaneously attack the same channel causing maximum degradation in the link quality. Then, multi-jammer attack is equivalent to a single jammer that attacks m channels sequentially in a time slot. Also, the jammer should attack each channel sufficiently long to be more effective, otherwise nodes can easily recover lost packets from brief outages. We thus consider a single jammer that sequentially attacks $m < K$ channels in each slot. We assume that the jammer transmits at sufficiently high power such that whenever both the jammer and the nodes are on the same channel it drives PDR to a low value resulting in zero throughput for the nodes.

In this paper we consider a IBFD reactive jammer. At the beginning of a time slot, the jammer continuously emits white Gaussian noise into the channel and, at the same time, uses its IBFD capability to listen for nodes' activity on the channel. If the jammer detects nodes' activity, it continuously attacks the channel until nodes leave that channel. If the jammer does not detect node activity for a while, it moves to attack other channels.

III. Jamming Game: Attack and Defense Strategies

In this section we discuss attack and defense techniques for the jammer and the nodes respectively. As discussed in [12], attack and defense strategy update is like 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. As a bare minimum, the nodes must hop onto different channel to avoid jamming. Below we discuss few rounds of arms race.

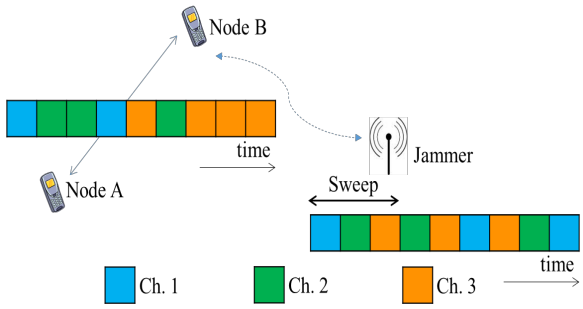

Fig. 4: System model.

A. Attack Strategy

When the jammer is aware that the nodes can hop channel, one naive attack strategy is to randomly choose m out of the K channels with equal probabilities in each time slot. In this case, as argued in [12], the nodes should stay on the same channel as they are equally vulnerable on all channels[4]. Anticipating the nodes' response, the jammer may now go through all the K channels sequentially, jamming m channels in each slot without overlap and repeat the sweeping process continuously. If the jammer follows a deterministic sweep pattern, the nodes can effectively counter the jamming attack by avoiding the channels the jammer attacks in a given slot. Aware of this nodes' response, the jammer could further randomize its sweep pattern after completing a sweep cycle.

Sweep Jammer: In the next round of arms race, the nodes can update their strategy as follows. Once the nodes are jammed in a sweep cycle, they can simply turn off their transmitters. Not finding any activity on the channel, the jammer leave the channel and continue the current sweep cycle. The nodes can then restart operation on the same channel and they will not be jammed again till the end of current sweep cycle . As the nodes are jammed at most once in each cycle, the average throughput achieved by the 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 $2\xi mR/K$ but may also suffer throughout loss due to channel switching. When the gain is small compared to the switching loss, the nodes may prefer to stay on the same channel and tolerate the small loss in throughput due to jamming attack.

IBFD Reactive Sweep Jammer: Aware of nodes' response to sweep jamming attack, the jammer may then update its strategy to restart a new sweep cycle with randomly ordered sweep pattern each time the nodes leave the channel after being jammed. This increases the rate of jamming the nodes if they always stay on the same channel. Then, the strategy of the nodes is not to use the same channel but to switch channels. Moreover, with the IBFD capability, the jammer can discern the nodes' activity, if any, on channels it is jamming. Hence, the jammer can continue jamming until the nodes hop. Its operation in each time slot is depicted in Figure 5.

[3]This scenario arises when jammer is in the proximity of one node but 'hidden' from the other.

[4]The probability of getting jammed is m/K in each time slot, whether the nodes hop or not

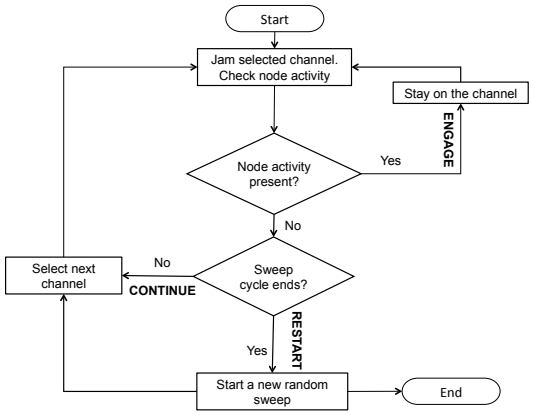

Fig. 5: **IBFD Reactive Sweep Jammer:** In each time slot, the jammer attacks currently selected channel and observes channel activity. If activity is observed, it stays on the channel and continuously attack it (ENGAGE). Otherwise, it RESTARTs a new sweep patten if the current cycle ends, else attacks next channel in the sweep cycle (CONTINUE).

B. Defense Strategy

When the nodes observe low PDR, hopping to a different channel is not a good strategy, as (i) the new channel could be in fading, and (ii) switching channels results in throughput loss. However, if the nodes can verify presence of a jammer, they must hop to a different channel, otherwise they will be continuously jammed. Thus, when the PDR is low, nodes can switch to TD mode (if not already in this mode) to ascertain the cause of failure before taking an action.

Note that if 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 transmission failure and leave the channel. Also note that when the nodes are jammed in the TR mode, they have to re-establish both the links and suffer throughput loss of $2L$. Whereas in TD mode, 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 no jamming can be detected and transmission loss is high. In the TD mode throughput is only R, but any jamming activity can be detected and transmission loss is less. We define the reward for each node and in each round n as

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

where $1[\cdot]$ denotes the indicator function.

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

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

In the absence of such knowledge, nodes like to use a policy that achieves average throughput per round as close as to (2).

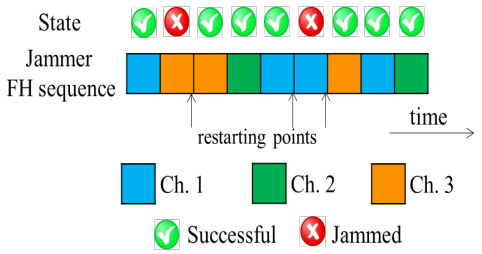

Fig. 6: Smart sweep jammer.

Let $R_\pi(n)$ denote the throughput in round n from policy π. We define regret of a policy π over period T as follows:

$$Reg(T,\pi) = R_m - \frac{1}{T}\sum_{n=1}^{T} R_\pi(n) \quad (3)$$

The goal of the nodes is to use a policy that minimizes the regret. Given that the nodes know 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. We will be interested in both total discounted and average reward criteria.

IV. Markov Model

We begin by defining the state space, action space and derive the transition probabilities of the Markov chain. 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 for k slots, they can only infer that the jammer did not sweep f in the last k slots. Thus, keeping track of which channels the nodes operated in the past and how many slots they stayed on them is not helpful. We use these observations in defining the state space and derivation of the transition probabilities below. For ease of notation, we write $\tilde{K} = \lceil \frac{K}{m} \rceil$.

States: The state denotes the transmission status of the nodes 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: jamming detected and jamming undetected. The former contains only state J, denoting the transmission failure due to the jamming attack (i.e., without ambiguity, jamming is detected)[5]. 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 $\mathcal{Y} := \{y_1, y_2, \cdots, y_{\tilde{K}-1}\}$

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

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, i.e., 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, i.e., 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, because m channels are jammed in each time slot and then the nodes can stay on the same channel 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.

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

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

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 $s_1 := (s, TD)$ denotes that the nodes stay on the same channel it used in the previous slot and operate in the TD mode. Action $h_1 := (h, TD)$ denotes that the nodes hop to a new randomly selected channel and operate in the TD mode. Similarly, $s_2 := (s, TR)$ denotes that the nodes stay on the same channel and operate in the TR mode and $h_2 := (h, TR)$ 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[6]. We use $a \in \mathcal{A}$ to denote a generic action.

Rewards: Let $U(x, a, x')$ denote the reward to the transmitter when it takes action $a \in \mathcal{A}$ in state $x \in \mathcal{X}$ and enters into state $x' \in \mathcal{X}$. Using (1) we define rewards of the nodes in different states as follows:

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

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 $P(x'|x, a)$ denote the probability that the nodes enter state $x' \in \mathcal{X}$ when they took action $a \in \mathcal{A}$ in state $x \in \mathcal{X}$. For notational convenience, let A_k denote the event that nodes are not jammed for k time slots since they last hopped. Note that this event precludes that the nodes are not jammed in all the earlier $k-1$ slots. It is straight forward to compute that $\Pr(A_k) = 1 - 1/(\tilde{K}-k)$ and used repeatedly in the following computations.

Given $(x, a) = (J, s_1)$ or (J, s_2): Recall that smart sweep jammer can detect activity on the channels while jamming, and hence continues to jam the channel till the nodes leave that channel, i.e., $P(J|J, a) = 1$ for all $a = s_1, s_2$. 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. Given $(x, a) = (J, h_1)$: 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

$$\begin{aligned} P(J|J, h_1) &= 1/\tilde{K}, \\ P(u_1|J, h_1) &= (1 - P(J|J, h_1)) \times \Pr(\text{fading}), \\ &= \left(1 - \frac{1}{\tilde{K}}\right)(1-p), \\ P(y_1|J, h_1) &= 1 - P(J|J, h_1) - P(u_1|J, h_1). \end{aligned} \tag{4}$$

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

$$\begin{aligned} P(y_1|J, h_2) &= P(y_1|J, h_1), \\ P(u_1|J, h_2) &= 1 - P(y_1|J, h_2). \end{aligned} \tag{5}$$

Given $(x, a) = (y_k, s_1), k = 1, 2, \cdots \tilde{K}-2$: 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 of state of the channel. The nodes are jammed in this case if the jammer enters the channel that the nodes are currently using, and this channel is not swept by the jammer in the last k slots. $\forall k = 1, 2, \cdots \tilde{K}-2$ we have

$$\begin{aligned} P(J|y_k, s_1) &= \frac{1}{\tilde{K}-k}, \\ P(u_{k+1}|y_k, s_1) &= \Pr(A_k) \times \Pr(\text{fading}) \\ &= \left(1 - \frac{1}{\tilde{K}-k}\right)(1-p), \\ P(y_{k+1}|y_k, s_1) &= 1 - P(J|y_k, s_1) - P(u_{k+1}|y_k, s_1). \end{aligned} \tag{6}$$

Given $(x, a) = (y_k, s_2), k = 1, 2, \cdots \tilde{K}-2$: 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 the channel remains good and nodes are not jammed in slot k. We get:

$$\begin{aligned} P(y_{k+1}|y_k, s_2) &= 1 - P(u_{k+1}|y_k, s_2) \\ &= \left(1 - \frac{1}{\tilde{K}-k}\right) p = P(y_{k+1}|y_k, s_1). \end{aligned} \tag{7}$$

Given $(x, a) = (y_k, h_1), k = 1, 2, \cdots, \tilde{K}-2$: When the nodes take action h_1, they are operating in the TD mode and 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. When the nodes hop from a channel, say f, to one of the $K-1$ channels chosen uniformly at random, it will not get jammed if 1) the new channel is already swept by the jammer while it operated on f, i.e., the new channel is one of mk channels swept by the jammer in the last k slots or 2) the nodes hop to one of the other $K-1-mk$ channels not swept

[6]This also helps derive transition probabilities using one step history.

by the jammer yet, but the jammer did not hop to that channel in the current time slot. Then, we get

$$\begin{aligned} P(y_1|y_k,h_1)+P(u_1|y_k,h_1) &= 1-P(J|y_k,h_1) \\ &= \left(\frac{mk}{K-1}+\frac{K-1-mk}{K-1}\left(1-\frac{1}{\tilde{K}-k}\right)\right), \\ p(y_1|y_k,h_1) &= p\left(1-P(J|y_k,h_1)\right). \end{aligned} \tag{8}$$

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

$$\begin{aligned} P(y_1|y_k,h_2) &= P(y_1|y_k,h_1) \\ P(u_1|y_k,h_2) &= 1-P(y_1|y_k,h_2). \end{aligned} \tag{9}$$

Given $(x,a)=(u_k,s_1), k=1,2,\cdots\tilde{K}-2$: The nodes move to state J or u_{k+1} or y_{k+1}. Given that the nodes are in state u_k implies that the previous transmission failed. If the failure happened due to jamming, the new state is J for sure (with probability 1). If the cause of failure was fading, new state can be J only due to jamming in the current slot. The new state is y_{k+1} if the channel is under fading in the previous channel and the transmission is successful in the current slot. We have:

$$\begin{aligned} P(J|u_k,s_1) &= \frac{1}{\tilde{K}-k+1}+(1-p)\frac{1}{\tilde{K}-k}, \\ P(y_{k+1}|u_k,s_1) &= (1-p)p\left(1-\frac{1}{\tilde{K}-k}\right) \\ P(u_{k+1}|u_k,s_1) &= 1-P(J|u_k,s_1)-P(y_{k+1}|u_k,s_1). \end{aligned} \tag{10}$$

Given $(x,a)=(u_k,h_1), k=1,2,\cdots\tilde{K}-2$: 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 $P(J|y_k,h_1)$. Additionally, the nodes move to state y_1 if the new channel is not jammed and not in fading. We have:

$$\begin{aligned} P(J|u_k,h_1) &= \frac{1}{(\tilde{K}-k+1)\tilde{K}}+(1-p)P(J|y_k,h_1) \\ P(y_1|u_k,h_1) &= p(1-P(J|u_k,h_1)) \\ P(u_1|u_k,h_1) &= 1-P(J|u_k,h_1)-P(y_1|u_k,h_1). \end{aligned} \tag{11}$$

Given $(x,a)=(u_k,h_2), k=1,2,\cdots\tilde{K}-2$: 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:

$$\begin{aligned} P(y_1|u_k,h_2) &= P(y_1|u_k,h_1), \\ P(u_1|u_k,h_2) &= 1-P(y_1|u_k,h_2). \end{aligned} \tag{12}$$

Possible transitions are shown in the Figure 7 and 8.

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 $P(y_1|y_k,h_1)$ is increasing in w.r.t. k, i.e.,

$$\begin{aligned} P(y_1|y_{k+1},h_2) &\geq P(y_1|y_k,h_2) \\ P(y_1|y_{k+1},h_1) &\geq P(y_1|y_k,h_1). \end{aligned} \tag{13}$$

Intuitively, the longer the nodes stay on a channel and their current transmission succeed, 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 the transmitter is defined as the action it takes 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[7]. Let $\pi:\mathcal{X}\rightarrow\mathcal{A}$ denote such a policy where $\pi(x)$ is the action when the nodes are in state x. We denote the collection of such policies as Π.

A. Optimal Defense Strategy in Discounted Reward Criterion

For a given $\pi\in\Pi$, the expected total discounted payoff of the nodes with an initial state $x\in\mathcal{X}$ is

$$V(x,\pi)=\mathbb{E}^{\pi}\left[\sum_{n=1}^{\infty}\delta^{n-1}r(X_n,A_n)|X_1=x\right],$$

where $\{(X_n,A_n), n=1,2,\cdots,\infty\}$ is a random process of state-action pair, that evolves according to the initial state and policy π. $r(X_n,A_n)$ denotes the immediate payoff in time slot n for taking action A_n in state X_n. The objective of the nodes is to choose a policy that maximizes $V(x,\pi)$ starting from any initial state, i.e., $\forall x\in\mathcal{X}$

$$V(x)=\max_{\pi\in\Pi}V(x,\pi). \tag{14}$$

The well-known Bellman equations for the expected discounted utility maximization problem in (14) are as follows:

$$Q(x,a) = \sum_{x'\in X}p(x'|x,a)\{U(x,a,x')+\delta V(x')\} \tag{15}$$

$$V(x) = \max_{a\in A}Q(x,a) \tag{16}$$

We can then use the value iteration [20][Ch. 6] method to derive the optimal defence strategy and its properties.

Proposition 1: 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 (16), 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. ■

Proof sketch of Proposition 1: (Details in [21]) 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 in k. The structure of the policy then follows by noting that optimal

[7]For the discounted average criteria, any history-dependent policy can be replaced by an Markov stationary policy that is equally good [20][Ch. 4].

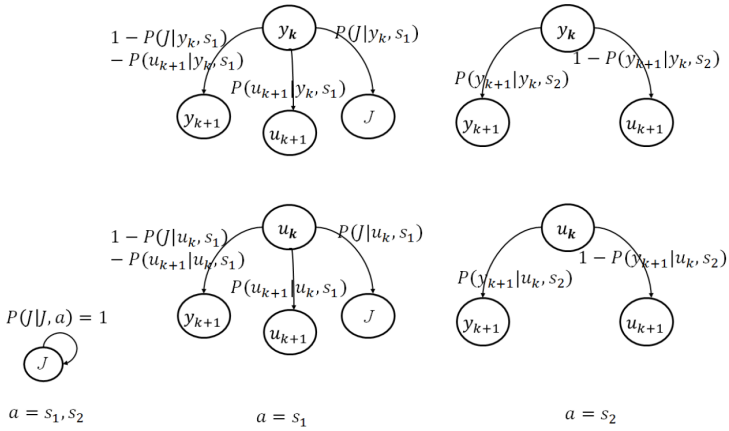

Fig. 7: State transition diagram when action is 'stay'

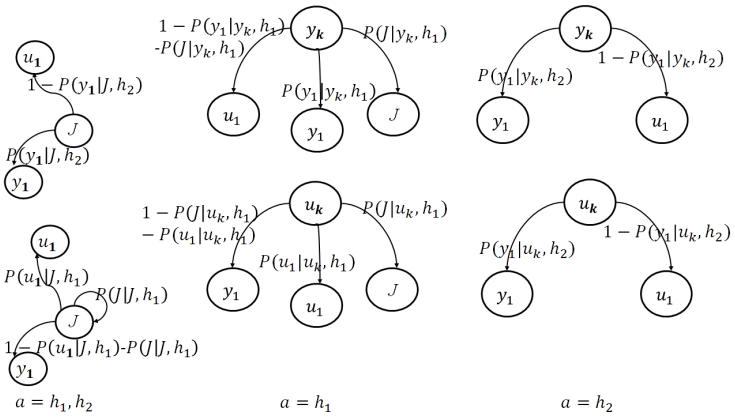

Fig. 8: State transition diagram when action is 'hop'

action is each state is selected greedily according to (16). □

The above result suggest 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. ■

B. Optimal Defense Strategy in Average Reward Criterion

In the sequel, we follow conventional notations in [8]. The immediate reward of the transmitter when taking (pure) action a at state x is:

$$r(x, a) = \sum_{x' \in X} U(x, a, x') P(x'|x, a) \tag{17}$$

The immediate expected reward of the transmitter at state x w.r.t. a stationary strategy Π is:

$$r(x, \Pi) = \sum_{a \in A} r(x, a) \pi(x, a) \tag{18}$$

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:

$$\boldsymbol{r}^{(t)}(\Pi) = \boldsymbol{P}^t \boldsymbol{r}(\Pi) \tag{19}$$

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:

$$V^{av}(x, \Pi) \stackrel{\text{def}}{=} \lim_{T \to \infty} \frac{1}{1+T} \sum_{t=0}^{T} r^{(t)}(s, \Pi) \tag{20}$$

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 [8], 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:

$$\begin{aligned} & \underset{\{f_{x,a}\}}{\text{maximize}} \sum_{x \in X} \sum_{a \in A} r(x, a) f_{x,a} \\ \text{s.t.} \quad & \text{C1: } \boldsymbol{W}\boldsymbol{x} = 0 \\ & \text{C2: } \mathbf{1}^T \boldsymbol{f} = 1 \\ & \text{C3: } f_{s,a} \geq 0, \forall x, a \end{aligned} \tag{21}$$

where $\mathbf{1}$ 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 $\pi^*(x, a) = 0$.

V. Performance Evaluation

To demonstrate the benefits of using the TR and TD operational modes we compare the performance of our policy, referred to as "Joinlty 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

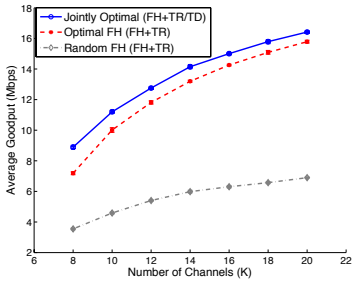

Fig. 9: Average goodput vs. K

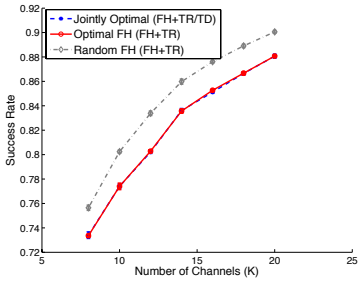

Fig. 10: Success rate vs. K

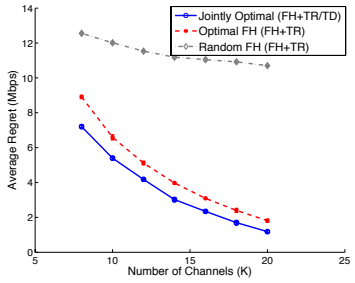

Fig. 11: Average Regret vs. K

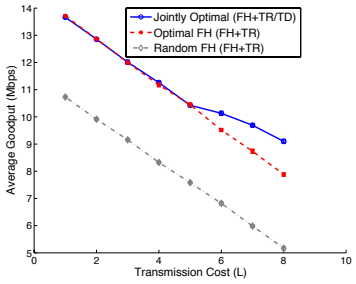

Fig. 12: Average goodput vs. L

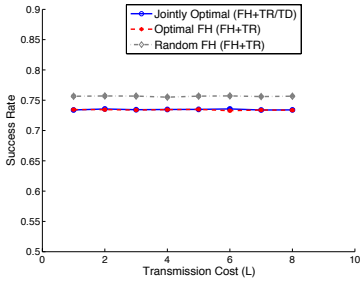

Fig. 13: Success rate vs. L

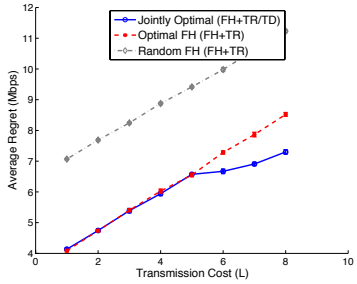

Fig. 14: Average Regret vs. L

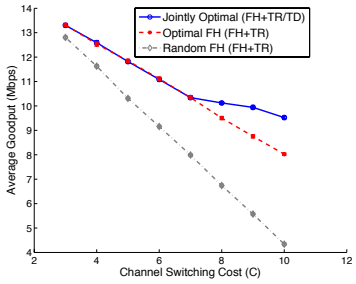

Fig. 15: Average goodput vs. C

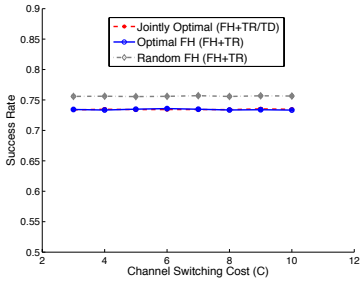

Fig. 16: Success rate vs. C

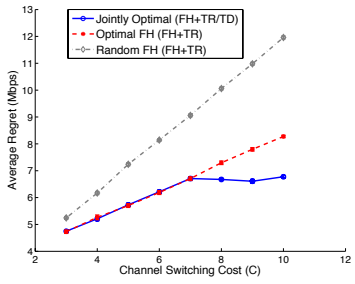

Fig. 17: Average Regret vs. C

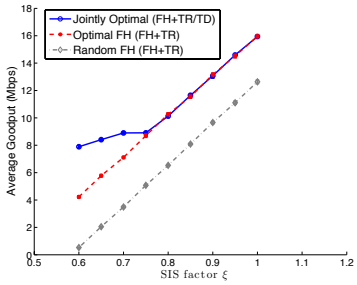

Fig. 18: Average goodput vs. ξ

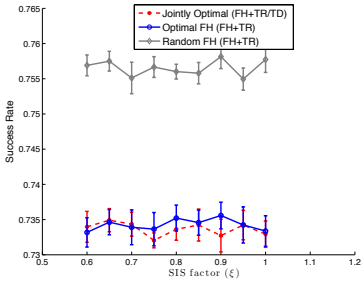

Fig. 19: Success rate vs. ξ

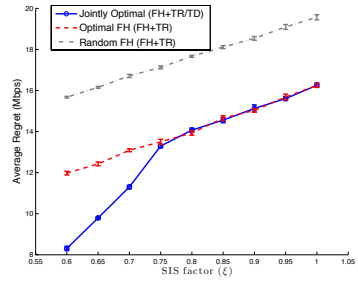

Fig. 20: Average regret vs. ξ

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 (15) 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 (3). The parameters of study are K, C, L, 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, $\xi = .7$, 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. 9, 10, 11 plot the average goodput, success rate, and average regret of the three

algorithms vs. the number of channel K. As seen, the Jointly optimal policy attains much higher goodput (up to 300% when K is large) than the Optimal FH (TD mode is not used) or Random FH. This is due to the effective utilization of TD mode to learn the jammer's behavior 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. 10). However, hopping too frequently makes the random hopping policy's goodput much lower than that under the Jointly optimal policy (Fig. 9). This is because the Jointly optimal policy efficiently avoids unnecessary hops to reduce switching costs. From Fig. 11, Jointly optimal policy has the smallest regret, and as the number of channels increases, its regret approaches zero faster than the others. 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 12, 13, and 14 depict goodput, success rate, and average regret vs. L of the three algorithms. The average goodput of Optimal policy is significantly higher (more than 43%) than the other policies, especially when the transmission cost is higher (Fig. 12). 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. 13, the success rate of the Jointly optimal policy is almost the same as that when the nodes hop after every time slot (this suffers from excessive loss in hopping/failure cost). The average regret of Jointly optimal policy is the smallest and also increases much slower w.r.t. the transmission cost than the random FH and using TR mode only policies (Fig. 14).

Effect of hopping cost (C): In Figures 15, 16, and 17 we compare the performance 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 12, 13, and 14 and the jointly optimal policy is the most robust algorithm to jamming (with more than 20% higher goodput, on average, than others').

Effect of imperfect SIS ξ: Figures 18, 19, and 20 depict the goodput, success rate, and average regret 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 more than 45% goodput 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. In Fig. 20, the Jointly optimal policy's average regret is also lowest among the three policies for various ξ.

VI. Conclusion

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. Using Markov decision processes, we developed an optimal strategy against an "IBFD reactive sweep jammer". The optimal defense strategy informs the nodes when to switch to half-duplex mode and when to hop to a new channel. Numerical investigations showed that the optimal strategy improves the network throughput up to a few times, under jamming attacks.

VII. Acknowledgments

Manjes 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 and IIP-1535573) and the Army Research Office (grant W911NF-13-1-0302). Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of funding agencies/institutes.

References

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

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[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] 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.

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

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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/

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] E. O. Elliot, "Estimates of error rates for codes on burst-noise channels," *Bell System Technical Journal*, vol. 42, 1977.

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

[21] M. K. Hanawal, D. Nguyen, and M. Krunz, "Jamming attack on in-band full-duplex communications: Detection and countermeasures," Tech. Rep., Nov. 2015. [Online]. Available: http://www.u.arizona.edu/~dnnguyen/Papers/TR/TRjammingFD.pdf