Jam Sessions: Analysis and Experimental Evaluation of Advanced Jamming Attacks in MIMO Networks

Liyang Zhang
Northeastern University
Boston, MA, USA
liyangzh@ece.neu.edu

Francesco Restuccia
Northeastern University
Boston, MA, USA
frestuc@northeastern.edu

Tommaso Melodia
Northeastern University
Boston, MA, USA
melodia@northeastern.edu

Scott M. Pudlewski
Air Force Research Laboratory
Rome, NY, USA
scott.pudlewski.1@us.af.mil

ABSTRACT

Recent research advances in wireless security have shown that advanced jamming can significantly decrease the performance of wireless communications. In advanced jamming, the adversary intentionally concentrates the available energy budget on specific critical components (*e.g.*, pilot symbols, acknowledgement packets, etc.) to (i) increase the jamming effectiveness, as more targets can be jammed with the same energy budget; and (ii) decrease the likelihood of being detected, as the channel is jammed for a shorter period of time. These key aspects make advanced jamming very stealthy yet exceptionally effective in practical scenarios.

One of the fundamental challenges in designing defense mechanisms against an advanced jammer is understanding which jamming strategies yields the lowest throughput, for a given channel condition and a given amount of energy. To the best of our knowledge, this problem still remains unsolved, as an analytic model to quantitatively compare advanced jamming schemes is still missing in existing literature. To fill this gap, in this paper we conduct a comparative analysis of several most viable advanced jamming schemes in the widely-used MIMO networks. We first mathematically model a number of advanced jamming schemes at the signal processing level, so that a quantitative relationship between the jamming energy and the jamming effect is established. Based on the model, theorems are derived on the optimal advanced jamming scheme for an arbitrary channel condition. The theoretical findings are validated through extensive simulations and experiments on a 5-radio 2x2 MIMO testbed. Our results show that the theorems are able to predict jamming efficiency with high accuracy. Moreover, to further demonstrate that the theoretical findings are applicable to address crucial real-world jamming problems, we show that the theorems can be incorporated to state-of-art reinforcement-learning based jamming algorithms and boost the action exploration phase so that a faster convergence is achieved.

MobiHoc'19, July 2019, Catania, Italy

ACM ISBN 123-4567-24-567/08/06…$15.00
https://doi.org/10.475/123_4

CCS CONCEPTS

• **Computer systems organization** → **Embedded and cyber-physical systems**; • **Security and privacy** → **Mobile and wireless security**; • **Networks** → *Network experimentation.*

KEYWORDS

Wireless, Security, Jamming, Theory, Optimization, Model, Reinforcement Learning, Testbed, Experiments

ACM Reference Format:
Liyang Zhang, Francesco Restuccia, Tommaso Melodia, and Scott M. Pudlewski. 2019. *Jam Sessions:* Analysis and Experimental Evaluation of Advanced Jamming Attacks in MIMO Networks. In *Proceedings of ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc'19).* ACM, New York, NY, USA, 10 pages. https://doi.org/10.475/123_4

1 INTRODUCTION

Wireless jamming is widely recognized as one of the most crucial topics in wireless security [35]. To understand how harmful a jammer could be, researchers have studied the worst-case jamming attack with fine-tuned temporal pattern (*e.g.*, pulse jamming [25]), frequency-pattern (*e.g.*, frequency-hopping jamming [26]), and so on. During the last few years, a new family of *advanced jamming* has gained momentum [2, 7, 8, 15, 27, 28], where the *target component* of the jammed transmission is the main objective to optimize.

The key intuition behind advanced jamming is that, although some components of the wireless transmission do not carry payload information, they nevertheless constitute the "Achilles' heel" of the entire communication process. For example, at the physical (PHY) layer, *pilot jamming* has been proposed to disrupt orthogonal frequency-division multiplexing (OFDM) and multiple-input and multiple-output (MIMO), since they heavily rely on accurate channel estimation through pilot symbols [5, 10, 15, 23]. Furthermore, at the link (MAC) layer, acknowledgement (ACK) jamming has been proposed to disrupt medium access control operations [3, 21, 24]. Since advanced jamming activities are restricted to a specific period of time, an advanced jammer can degrade the network throughput with a comparatively lower energy budget yet with lower probability of being detected [7, 8].

The key limitation of existing work (discussed in details in Section 2) is that *it does not provide the analytic tools to thoroughly investigate the quantitative relationship between the jamming energy*

and the jamming outcome. As a consequence, literature still lacks a mathematical model to *compare advanced jamming schemes, each with a different target component, to understand which one yields the highest jamming efficiency.* We point out that a mathematical model of advanced jamming attacks is of fundamental importance not only from a theoretical perspective, but for a number of practical reasons as well. First, it is straightforward to notice that jammers usually do not want their attacks to be discovered. Therefore, to increase stealthiness, advanced jammers need to keep the jamming signal energy below a certain threshold and corrupt the channel for as little time as possible. Furthermore, it is well known that jammed nodes usually react with strategies such as rate adaptation and rerouting [14, 29–31]. Therefore, real-world advanced jammers necessarily face dynamic, time-varying scenarios, where adaptive jamming strategies are almost mandatory. In these circumstances, analytic tools on most efficient advanced jamming schemes are extremely valuable for an advanced jammer.

As a first attempt to address the existing research gap, this article focuses on widely-used MIMO networks, and makes the following contributions:

(1) We select three most viable schemes in MIMO networks, each with a different target component, and rigorously model them at the signal-processing level, so that a mathematical relationship between the *amount of jamming energy* used and the *throughput degradation* is established for an arbitrary channel condition. Based on the model, we compare the jamming schemes and derive theorems on optimal jamming strategies in various scenarios. Our main theoretical results conclude that (i) for a given data packet, the relative efficiency for barrage jamming and pilot jamming is decided by the pilot sequence length and the number of transmitting antennae (Theorem 1); (ii) under some conditions, pilot jamming will lead to a lower average signal-to-inteference-and-noise ratio (SINR) of the ACK packet than directly jamming the ACK packet itself (Theorem 2); and (iii) ACK jamming can be compared with pilot and barrage jamming, on the basis of packet error rate (PER) lower bound, and the result is decided by a number of factors including pilot length, modulation and coding scheme (MCS) used for both data and ACK packets, and the distances from the jammer to the transmitter and receiver (Theorem 3). Our theoretical foundations are validated through extensive simulations and experiments on MIMO testbed made up by 5 USRP software-defined radios. Results indicate that our model is able to capture the behavior of advanced jamming strategies in complex MIMO scenarios accurately.

(2) To demonstrate that applicability of our theoretical results in real-world scenarios, where the theorems' assumptions cannot be validated either due to missing information or dynamic environment, we propose a way to incorporate the theorems to state-of-art reinforcement-learning [9, 18, 20] based jamming algorithms, such as [1]. Specifically, *we show that the theorems on jamming efficiency can be used by the jammer to improve the efficiency of the action space exploration up to a significant extent.* Indeed, extensive simulations show the effectiveness of our approach and show significant improvement in both convergence speed and the total reward.

Scope and Limitations

We point out that the objective of our study is not to investigate every possible component in a wireless network that an advanced jammer can target. Instead, we focus on the modeling and analysis on a typical wireless scenario (MIMO networks) and several typical advanced jamming schemes, and show that (i) theoretical analysis based on rigorous model can predict the worst-case jamming results for certain scenarios; (ii) the theorems can be incorporated with practical, reinforcement-learning based jamming algorithms, thus extending their applicability to scenarios with limited information or dynamic environment.

Moreover, we do not claim that our results provide jamming strategies that are optimal on *every* possible aspect. As we have pointed out earlier, there are also other aspects of jamming strategy that can be optimized, such as temporal pattern and frequency pattern of the jamming signal. We have focused on the relatively under-explored aspect of jamming target, and provided insights on how to optimally choose the component of the wireless transmission to jam.

2 RELATED WORK

As far as the physical layer is concerned, Clancy [5] proposes to disrupt OFDM links using pilot jamming and pilot nulling attacks. La Pan *et al.* [10] consider false preamble timing and preamble nulling attacks in OFDM. Moreover, Rahbari *et al.* [17] study the impact of jamming on OFDM frequency offset (FO). Specific to MIMO systems, Sodagari *et al.* [19] propose an attack where the jamming signal transforms the estimated MIMO channel matrix into a singular matrix. Pirzadeh *et al.* show in [15] that the MIMO spectral efficiency can be significantly degraded by jamming both the training phase and the data transmission phase. Wang *et al.* [23] propose a random channel training (RCT)-based secure transmission framework to address MIMO jamming.

Regarding the link layer, an advanced jammer can exploit the temporal pattern between transmissions as well as critical control packets. Bayrak *et al.* [3] consider a scenario where the jammer exploits the exponential backoff mechanism of IEEE 802.11, and show that the jammer can achieve higher efficiency if aware of the current backoff states of the users. Jamming against IEEE 802.11b is discussed by Thuente *et al.* in [21], based on launching CTS jamming, ACK jamming, or DIFS wait jamming attacks. ACK jamming in IEEE 802.15.4 is also considered by Wilhelm *et al.* in [24].

The core limitation of the above mentioned pioneering works is the approach of investigating an individual jamming scheme for a fixed environment. While the potential threat of various advanced jamming schemes have been revealed, the important question of optimal advanced jamming strategy, when multiple jamming schemes are available and the environment is time-varying, is still open.

The closest work to ours is [7, 8], where the authors investigate energy-optimal jamming strategies to achieve stealthiness and increase effectiveness. Specifically, DeBruhl and Tague [8] show that energy-efficient short-form periodic jamming can seriously degrade communication capabilities without compromising the jammer's activity. DeBruhl *et al.* [7] investigate finite-energy jamming games, where jammers choose among different actions (i.e., sleep, power and channel). The authors compute the games' Nash equilibria, and test the performance of the optimal strategies against random and adaptive strategies. These works, however, have not focused on jamming schemes directed at critical components of a transmission, and therefore differ significantly from ours.

3 SYSTEM MODEL

In this section, we illustrate a typical model for MIMO wireless communication, upon which three jamming schemes are modeled. Some commonly used notations: we use $\{*\}^T$, $\{*\}^H$ and $\{*\}^\dagger$ to denote the transpose, Hermitian, and pseudo inverse of a matrix $\{*\}$; $\{\hat{*}\}$ and $\{\tilde{*}\}$ represent the estimate and estimation error of $\{*\}$; $I_{\{*\}}$ is used to denote a $\{*\} \times \{*\}$ unity matrix; and $\mathbb{E}\{*\}$ denotes the expectation of $\{*\}$.

Physical Layer. We consider a MIMO link with forward (*i.e.*, data) and backward (*i.e.*, ACK) transmissions as illustrated in Fig. 1.

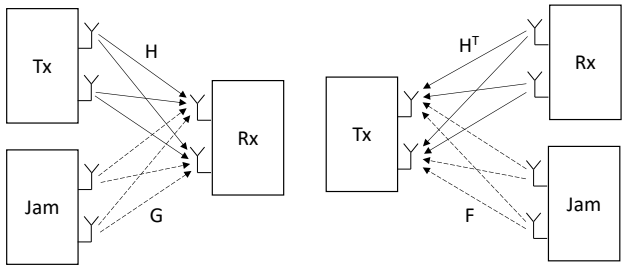

Figure 1: MIMO link for (a) forward and (b) backward transmissions under jamming. Tx: transmitter; Rx: receiver; Jam: jammer.

Without loss of generality, we assume the transmitter, receiver, and the jammer are equipped with M, N, and L antennae, respectively.[1] Therefore, the channel from the transmitter to the receiver, from the jammer to the receiver, and from the jammer to the transmitter can be denoted with matrices $H = \{h_{nm}\}_{1\le n\le N, 1\le m\le M}$, $G = \{g_{nl}\}_{1\le n\le N, 1\le l\le L}$, and $F = \{f_{ml}\}_{1\le m\le M, 1\le l\le L}$. Vectors $\mathbf{x} = \{x_1, \ldots, x_M\}^T$, $\mathbf{y} = \{y_1, \ldots, y_N\}^T$ and $\mathbf{z} = \{z_1, \ldots, z_L\}^T$ are used to represent the transmitted, received and jamming signals, respectively. while vector $\mathbf{w} = \{w_1, \ldots, w_N\}^T$ represents Gaussian noise at the receiver, with spectral density N_0.

We assume that the channels are subject to path loss and Rayleigh fading. The path loss is determined by the physical distance of the communicating parties, and therefore remains the same for every entry in the same channel matrix. We will use $\theta_{\{*\}}$ to denote the path loss of the channel $\{*\}$. Combined with the Rayleigh fading component, it follows that

$$h_{nm} \sim \mathcal{CN}(0, \theta_H), \quad \forall 1 \le n \le N, \forall 1 \le m \le M, \tag{1}$$

$$g_{nl} \sim \mathcal{CN}(0, \theta_G), \quad \forall 1 \le n \le N, \forall 1 \le l \le L, \tag{2}$$

$$f_{ml} \sim \mathcal{CN}(0, \theta_F), \quad \forall 1 \le m \le M, \forall 1 \le l \le L, \tag{3}$$

i.e., the entries of the channel matrix are independent and identically distributed (i.i.d.) complex Gaussian random variables, with zero mean and variance equal to the path loss.

MIMO can be used to achieve multiplexing or diversity gain [33]. For the data transmission on the forward link, since high throughput is usually required, we focus on spatial multiplexing, where different bits are transmitted on the M antennae simultaneously (we will refer to them as M spatial "channels"). Following [16], the baseband model for this scheme is

[1]For simplicity, we will refer to the transmitter of the forward link as *transmitter*, even when it is the receiver of the backward link. Similarly, the term *receiver* will be used solely to address the receiver of the forward link.

$$\mathbf{y} = \begin{cases} H\mathbf{x} + \mathbf{w}, & \text{not jammed,} \\ H\mathbf{x} + G\mathbf{z} + \mathbf{w}, & \text{jammed.} \end{cases} \tag{4}$$

Let E_s and E_j denote the average symbol energy for the transmitted and jamming signals. Then, it follows that

$$\mathbb{E}\{\mathbf{x}\mathbf{x}^H\} = E_s I_M, \tag{5}$$

$$\mathbb{E}\{\mathbf{z}\mathbf{z}^H\} = E_j I_L. \tag{6}$$

For the ACK transmission on the backward link, since high reliability is often required, we focus on MIMO schemes that achieve diversity instead of spatial multiplexing. To this end, we consider a scheme where diversity is achieved via beamforming and Maximal Ratio Combining (also called as MIMO-MRC). Specifically, the receiver leverages the channel information acquired in the previous forward transmission, and beamforms using a vector $\mathbf{u}$ satisfying

$$(\hat{H}^T)^H \hat{H}^T \mathbf{u} = \lambda_{\max} \mathbf{u}, \tag{7}$$

i.e., the eigenvector corresponding to the maximal eigenvalue ($\lambda_{\max}$) of the matrix $(\hat{H}^T)^H \hat{H}^T$. A symbol x is precoded by $\mathbf{u}$, and the resultant $\mathbf{u}x$ is transmitted over the antennae. Therefore, we have the baseband model

$$\mathbf{y} = H^T \mathbf{u} x + F\mathbf{z} + \mathbf{w}. \tag{8}$$

Without loss of generality, we assume that $\|\mathbf{u}\|^2 = N$, so that the total transmitting energy across the N antennae is NE_s, for $\mathbb{E}\{x^2\} = E_s$.

Link Layer. Received data packets are acknowledged by the receiver on link layer. To establish a relationship between physical layer metrics and the PER, we borrow the model in [12]. Specifically, suppose a MCS z is chosen from a set $\mathcal{Z}$, the PER e, as a function of SINR γ, can be approximated as

$$e(\gamma|z) = \begin{cases} 1, & \gamma \le \gamma_{\text{th}}, \\ a_z e^{-b_z \gamma}, & \gamma > \gamma_{\text{th}}, \end{cases} \tag{9}$$

with MCS-dependent parameters a_z and b_z. The threshold γ_{th} also varies with different MCSs.

3.1 Advanced Jamming Schemes

The objective of the jammer is to efficiently degrade the throughput of the jammed link. To this end, the jammer may (i) directly inject interference to the entire data packet and lower the achievable SINR (we will refer to this scheme as *barrage jamming*, following [5]); (ii) jam the pilot symbols to invalidate the estimated channel, also called *pilot jamming* [32]; or (iii) prevent ACKs from being delivered, also called *ACK jamming* [11]. We will formally model each of these cases, and analyze the resulting effects.

Barrage Jamming. In this case, the jammer emits Gaussian noise uniformly on the entire data packet, lowering the resulting receiver SINR and consequently degrading the throughput of the jammed link. The baseband model in presence of jamming is shown in (4). We assume that when the jamming energy is uniformly allocated upon the entire data packet, the channel estimation error is negligible, *i.e.*, $\hat{H} \approx H$. Therefore, the transmitted signal is recovered as follows:

$$\hat{\mathbf{x}} = \hat{H}^\dagger \mathbf{y} \approx \mathbf{x} + H^\dagger G\mathbf{z} + H^\dagger \mathbf{w}, \tag{10}$$

where $H^\dagger = (H^H H)^{-1} H^H$ is the left pseudo inverse of H.

The SINR per symbol on the m-th spatial channel is then

$$\gamma_m = \frac{E_s}{E_j[H^\dagger GG^H(H^\dagger)^H]_{mm} + N_0[(H^HH)^{-1}]_{mm}}. \quad (11)$$

Pilot Jamming. The advanced jammer may also aim at jamming the pilot symbols, resulting in channel estimation errors and further impairing data decoding. The pilot signal comprises of a sequence of symbols agreed on by the transmitter and receiver, denoted as $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_K\}$, assuming a sequence of length K is used. Similarly, the received sequence, jamming sequence, and the noise sequence are $Y = \{\mathbf{y}_1, \ldots, \mathbf{y}_K\}$, $Z = \{\mathbf{z}_1, \ldots, \mathbf{z}_K\}$, and $W = \{\mathbf{w}_1, \ldots, \mathbf{w}_K\}$, respectively. Thus, for the pilot, we have

$$Y = HX + GZ + W. \quad (12)$$

Least square estimation gives the estimated channel matrix

$$\hat{H} = H + GZX^\dagger + WX^\dagger, \quad (13)$$

with $X^\dagger = X^H(XX^H)^{-1}$ as the right pseudo inverse of X. The estimation in (13) introduces an error term

$$\tilde{H} = (GZ + W)X^\dagger, \quad (14)$$

which will affect the recovery of the received signal.

According to [4] and [13], to achieve optimal estimation, the pilot must satisfy

$$XX^H = KE_sI_M. \quad (15)$$

With this condition, the statistical characteristic of the channel estimation error is given by Lemma 1.

Lemma 1. *The channel estimation error $\tilde{H}$ satisfies $\mathbb{E}\{\tilde{H}\} = \mathbf{0}_{N\times M}$ and*

$$\mathbb{E}\{\tilde{H}A\tilde{H}^H\} = \frac{1}{KE_s}\mathrm{tr}(A)(E_jGG^H + N_0I_N), \quad (16)$$

for an arbitrary $M \times M$ matrix A.

Proof. Since $\mathbb{E}\{\tilde{H}\} = \mathbf{0}_{N\times M}$ is obvious, we will focus on $\mathbb{E}\{\tilde{H}A\tilde{H}\}$. With the optimal training sequence $XX^H = KE_sI_M$, we have

$$\begin{aligned}\mathbb{E}\{\tilde{H}A\tilde{H}^H\} =& \mathbb{E}\{(GZ+W)X^\dagger A(X^\dagger)^H(GZ+W)^H\}\\ =& \frac{1}{K^2E_s^2}\mathbb{E}\{(GZ+W)X^HAX(GZ+W)^H\}\\ =& \frac{1}{K^2E_s^2}\Big(\mathbb{E}\{GZX^HAXZ^HG^H\}\\ &+\mathbb{E}\{WX^HAXW^H\}\Big)\\ =& \frac{1}{K^2E_s^2}\left(DBD^H + WBW^H\right),\end{aligned} \quad (17)$$

where we denote $D = GZ$, and $B = X^HAX$.

For the interference D, we have

$$\begin{aligned}D =& GZ\\ =& \begin{bmatrix} g_{11} & \cdots & g_{1L}\\ \vdots & \ddots & \vdots\\ g_{N1} & \cdots & g_{NL}\end{bmatrix}\cdot\begin{bmatrix} z_{11} & \ldots & z_{1K}\\ \vdots & \ddots & \vdots\\ z_{L1} & \ldots & z_{LK}\end{bmatrix}\\ =& \begin{bmatrix} \sum_l g_{1l}z_{l1} & \cdots & \sum_l g_{1l}z_{lK}\\ \vdots & \ddots & \vdots\\ \sum_l g_{Nl}z_{l1} & \ldots & \sum_l g_{Nl}z_{lK}\end{bmatrix}\\ \triangleq& (\mathbf{d}_1, \ldots, \mathbf{d}_K).\end{aligned} \quad (18)$$

Column k of Z represents the jamming signal at time instant k. Therefore, they are mutually independent. Consequently, the columns of D are also uncorrelated, and

$$\mathbb{E}\{\mathbf{d}_k\mathbf{d}_{k'}^H\} = \begin{cases} E_jGG^H, & k = k',\\ \mathbf{0}_{N\times N}, & k \neq k'.\end{cases} \quad (19)$$

Therefore,

$$\begin{aligned}\mathbb{E}\{DBD^H\} =& \mathbb{E}\left\{(\mathbf{d}_1, \ldots, \mathbf{d}_K)\cdot\begin{bmatrix} b_{11} & \ldots & b_{1K}\\ \vdots & \ddots & \vdots\\ b_{K1} & \ldots & b_{KK}\end{bmatrix}\cdot\begin{pmatrix}\mathbf{d}_1^H\\ \vdots\\ \mathbf{d}_K^H\end{pmatrix}\right\}\\ =& \mathbb{E}\left\{\sum_k b_{kk}\mathbf{d}_k\mathbf{d}_k^H\right\} = E_j\mathrm{tr}(B)GG^H.\end{aligned} \quad (20)$$

Since $B = X^HAX$, we have $\mathrm{tr}(B) = \mathrm{tr}(X^HAX) = \mathrm{tr}(XX^HA) = KE_s\mathrm{tr}(A)$ and $\mathbb{E}\{DBD^H\} = E_jKE_s\mathrm{tr}(A)GG^H$.

For the term WBW^H, since all the entries of W are i.i.d, we have

$$\mathbb{E}\{WBW^H\} = N_0\mathrm{tr}(B) = N_0KE_s\mathrm{tr}(A)I_N. \quad (21)$$

Therefore,

$$\mathbb{E}\{\tilde{H}A\tilde{H}^H\} = \frac{1}{KE_s}\mathrm{tr}(A)\left(E_jGG^H + N_0I_N\right). \quad (22)$$

□

The subsequent data transmission is not affected by jamming, and complies with (4). Therefore, the decoded signal is

$$\hat{\mathbf{x}} = \hat{H}^\dagger\mathbf{y} \approx \mathbf{x} + H^\dagger\mathbf{w} - H^\dagger\tilde{H}\mathbf{x} - H^\dagger\tilde{H}H^\dagger\mathbf{w}. \quad (23)$$

The post-processing noise is then

$$\hat{\mathbf{w}} = H^\dagger\mathbf{w} - H^\dagger\tilde{H}\mathbf{x} - H^\dagger\tilde{H}H^\dagger\mathbf{w}, \quad (24)$$

with an autocorrelation given by Lemma 2.

Lemma 2. *The Autocorrelation of the post-processing noise (24) is*

$$\mathbb{E}\{\hat{\mathbf{w}}\hat{\mathbf{w}}^H\} \approx \left(1 + \frac{1}{K}\right)N_0(H^HH)^{-1} + E_j\frac{M}{K}H^\dagger GG^H(H^\dagger)^H. \quad (25)$$

Proof. The autocorrelation can be derived as follows:

$$\begin{aligned}\mathbb{E}\{\hat{\mathbf{w}}\hat{\mathbf{w}}^H\} =& \mathbb{E}\{(H^\dagger\mathbf{w} - H^\dagger\tilde{H}\mathbf{x} - H^\dagger\tilde{H}H^\dagger\mathbf{w})\\ &\cdot(H^\dagger\mathbf{w} - H^\dagger\tilde{H}\mathbf{x} - H^\dagger\tilde{H}H^\dagger\mathbf{w})^H\}\\ =& \mathbb{E}\{H^\dagger\mathbf{w}\mathbf{w}^H(H^\dagger)^H + H^\dagger\tilde{H}\mathbf{x}\mathbf{x}^H\tilde{H}^H(H^\dagger)^H\\ &+ H^\dagger\tilde{H}H^\dagger\mathbf{w}\mathbf{w}^H(H^\dagger)^H\tilde{H}^H(H^\dagger)^H\}\\ =& N_0(H^HH)^{-1} + \frac{M}{K}H^\dagger\left(E_jGG^H + N_0I_N\right)(H^\dagger)^H\\ &+ \frac{N_0}{KE_s}\mathrm{tr}((H^HH)^{-1})H^\dagger(E_jGG^H + N_0I_N)(H^\dagger)^H\end{aligned} \quad (26)$$

According to [22], $\mathrm{tr}((H^HH)^{-1})$ is a small value typically no larger than M. Since we are focusing on high SNR (by SNR we mean signal to noise ratio, the SIR or SJR is not necessarily high) scenario, this implies that ${}^{N_0}/_{E_s}\mathrm{tr}((H^HH)^{-1}) << M$. In other words, the last term is much smaller than the second term and can be neglected. Therefore, we have

$$\mathbb{E}\{\hat{\mathbf{w}}\hat{\mathbf{w}}^H\} \approx \left(1 + \frac{1}{K}\right)N_0(H^HH)^{-1} + \frac{M}{K}E_jH^\dagger GG^H(H^\dagger)^H. \quad (27)$$

□

Therefore, the "effective" SINR on the m-th spatial channel is

$$\gamma_m = \frac{E_s}{E_j \frac{M}{K}[H^\dagger G G^H (H^\dagger)^H]_{mm} + \left(1 + \frac{1}{K}\right) N_0 [(H^H H)^{-1}]_{mm}}, \quad (28)$$

ACK Jamming. With the baseband model for ACK transmission in (8), the signal can be recovered by multiplying $\mathbf{u}^H(\hat{H}^T)^H$ with the received signal and normalizing. Therefore, at the transmitter side, we have

$$\begin{aligned} \hat{x} &= \frac{\mathbf{u}^H(\hat{H}^T)^H(H^T\mathbf{u}x + F\mathbf{z} + \mathbf{w})}{\mathbf{u}^H(\hat{H}^T)^H\hat{H}^T\mathbf{u}} \\ &= x - \frac{\mathbf{u}(\hat{H}^T)^H\tilde{H}^T\mathbf{u}x - \mathbf{u}(\hat{H}^T)^H(F\mathbf{z} + \mathbf{w})}{\mathbf{u}^H(\hat{H}^T)^H\hat{H}^T\mathbf{u}}, \end{aligned} \quad (29)$$

and the post-processing SINR is as shown in (30), with $E_{j,p}$ and $E_{j,a}$ denoting the jamming energy per symbol in pilot jamming and ACK jamming.

$$\begin{aligned} \gamma_{\text{ACK}} = \|\mathbf{u}^H(\hat{H}^T)^H\hat{H}^T\mathbf{u}\|^2 E_s / \{ &\left[\frac{E_{j,p}}{L}\text{tr}(G^H G) + NN_0\right] \\ \cdot &\mathbf{u}^H(\hat{H}^T)^H\hat{H}^T\mathbf{u} + \mathbf{u}^H(\hat{H}^T)^H(E_{j,a}FF^H + N_0 I_M)\hat{H}^T\mathbf{u}\} \end{aligned} \quad (30)$$

4 THEORETICAL ANALYSIS OF OPTIMAL JAMMING TARGET

In this section, we will quantitatively compare the jamming schemes modeled in Section 3 for the optimal target in the sense of jamming efficiency. Since different targets (*i.e.*, data packet, pilot, and ACK packet) differ in terms of their lengths, it is unfair to directly compare the jamming power. Instead, we compare the jamming effects caused by a *unit* of jamming energy spent on each specific target, regardless of their lengths (*e.g.*, we compare the effect of spending one unit of energy on pilot jamming vs ACK jamming). Notice that since the analysis is based on a unit of jamming energy, the results are extendable to arbitrary amount of jamming energy.

4.1 Physical Layer Jamming

At the physical layer, the objective of degrading throughput is equivalent to reducing the achievable SINR, for which the following theorem holds.

Theorem 1. *For a unit of jamming energy, pilot jamming yields a lower SINR than barrage jamming if*

$$K < \sqrt{D \cdot M}, \quad (31)$$

where D and K are the lengths of the data packet and pilot in symbols, and M is the number of antennae at the transmitter.

Proof. For a unit of jamming energy, the values of jamming energy per symbol can be expressed as $E_j = 1/(L \cdot K)$ and $E_j = 1/(L \cdot D)$, with L as the number of antennae at the jammer, for pilot jamming and barrage jamming, respectively. Plugging these into the SINR expressions in (11) and (28), the theorem follows immediately. □

Remark 1. *According to Theorem 1, for a given data packet and the number of transmitting antennae, the optimal choice between barrage jamming and pilot jamming is decided solely by the pilot sequence length. This matches intuition, since longer pilot sequences are more robust to jamming. As long as the pilot length $K \geq \sqrt{DM}$, there is no incentive for the jammer to launch a pilot jamming attack, and the optimal jamming strategy is reduced to barrage jamming.*

Note, here we are comparing on the basis that every data packet has a pilot. In reality, a pilot is supposed to cover a channel coherence period, during which there might be multiple data packets transmitted. In this case, the data packets can be treated as one single payload, in the sense of pilot jamming, since they share the same pilot.

Remark 2. *The square-root form of data packet length D in (31) may look anti-intuitive. A better interpretation is available if we rewrite (31) to the equivalent form of*

$$\frac{M}{K} \cdot \frac{1}{K} > \frac{1}{D}.$$

Note, with a unit jamming energy, $\frac{1}{K}$ and $\frac{1}{D}$ are the per-symbol jamming energy for pilot and barrage jamming, respectively. So Theorem 1 essentially means that, the interference caused on the pilot signal is equivalently scaled by $\frac{M}{K}$ during the signal recovering phase for the following data packet. (For a in-depth understanding of this scale, please refer to the proofs of Lemma 1 and 2.)

This scaling reflects the way how pilot jamming works. The jamming signal added to the pilot introduces errors to the estimated channel matrix, which are "transformed" during the signal processing in signal recovering phase. As a result, for the recovered signal, the interference is equivalently scaled by $\frac{M}{K}$. Interestingly, the scaling effect is proportional to the number of transmitting antennae, and inversely proportional to the pilot length. Therefore, while increasing number of transmitting antennae increases the total throughput, it does not alleviate pilot jamming. On the contrary, it makes the situation worse.

4.2 Link Layer Jamming

At the link layer, the jammer aims at disrupting the transmission of ACK packets. Since a corrupted ACK leads to retransmission of the corresponding data packet, it is equivalent to a corrupted data packet in terms of the effective link throughput. To evaluate the jamming effect, we need to compare the PER for the data and ACK packets. However, substituting the complex forms of *instantaneous* SINR of (11), (28) and (30) into the PER-SINR function (9) produces intractable equations. To ease the analysis, we apply the *expected* SINR in (9) instead, which produces a lower bound for the PER, as shown below.

With Rayleigh fading model, the entries of channel matrices H, G, and F are i.i.d. complex Gaussian random variables, as shown in Eq. (1), (2), and (3). The expected SINRs for the three jamming schemes in Eq. (11), (28), and (30) are then derived as

$$\bar{\gamma}_{\text{b}} = \frac{E_s \cdot \theta_H}{E_{j,a} \cdot \theta_G + N_0}, \quad (32)$$

$$\bar{\gamma}_{\text{p}} = \frac{E_s \cdot \theta_H}{\frac{M}{K} \cdot E_{j,p} \cdot \theta_G + (1 + \frac{1}{K}) \cdot N_0}, \quad (33)$$

$$\bar{\gamma}_{\text{a}} = \frac{\mathbb{E}\{\lambda_{\max}\} \cdot N \cdot E_s \cdot \theta_H}{\frac{N^2}{K}(E_{j,p} \cdot \theta_G + N_0) + (L \cdot E_{j,a} \cdot \theta_F + N_0)}, \quad (34)$$

where we use subscripts b, p, and a in $E_{j,\{*\}}$ to distinguish the per-symbol jamming energy for barrage jamming, pilot jamming, and ACK jamming, respectively. With the PER model in (9), we

approximate the PER as

$$e \approx 1 - \mathbb{P}\{\gamma \geq \gamma_{\text{th}}\} \geq 1 - \frac{\bar{\gamma}}{\gamma_{\text{th}}}, \tag{35}$$

where the inequality comes from Markov's inequality. Apparently, (35) establishes a PER lower bound

$$e_{\text{lwr}} = 1 - \frac{\bar{\gamma}}{\gamma_{\text{th}}}. \tag{36}$$

Theorem 2. *When used exclusively, pilot jamming yields a higher PER lower bound for the following ACK packet than ACK jamming, with a unit jamming energy, if the following holds:*

$$\frac{A \cdot \theta_G}{K \cdot \theta_F} \geq \frac{K \cdot L}{N^2} \tag{37}$$

where A and K are lengths of the ACK packet and the pilot in symbols; θ_G and θ_F denote the path loss from the jammer to the receiver and transmitter; L and N are the number of antennae at the jammer and the receiver, respectively.

Proof. Since we are comparing the PER lower bound of the same packet (ACK), the term γ_{th} in (36) holds the same for both jamming schemes. Therefore, we only need to compare the expected SINR in (34).

For a unit of jamming energy, the jamming energy per symbol is $1/(L \cdot K)$ and $1/(L \cdot A)$, for pilot jamming and ACK jamming, respectively. Plugging these into the first and second terms in the denominator of Eq. (34) and ignore the noise, it follows that, when (37) holds:

$$\frac{N^2}{K} \cdot E_{j,p} \cdot \theta_G > L \cdot E_{j,a} \cdot \theta_F,$$

i.e., when used exclusively, pilot jamming yields a higher denominator than ACK jamming. A lower (34) follows. Hence a higher (36). □

Remark 3. *Theorem 2 states that when the condition holds, pilot jamming alone will lead to a higher PER lower bound of the ACK packet than directly jamming the ACK packet itself. This is caused by incorrect CSI being propagated directly to the ACK transmission.*

Note the primary target of pilot jamming is the data packet. However, as long as the conditions in Theorem 2 holds, the optimal jamming scheme to disrupt the following ACK packet is also pilot jamming (on the prior data packet). Therefore, under this condition, ACK jamming cannot be the optimal jamming scheme, and no jamming energy should be spent on it. Partly based on this theorem, the following theorems will give the optimal jamming schemes.

Theorem 3. *Let us assume that noise is negligible. With a unit of jamming energy, ACK jamming yields a higher PER lower bound than barrage jamming if*

$$\frac{A \cdot \theta_G}{D \cdot \theta_F} < \frac{L \cdot \gamma_{\text{th},a}}{\mathbb{E}\{\lambda_{\max}\} \cdot N \cdot \gamma_{\text{th},d}}; \tag{38}$$

and higher lower bound than pilot jamming if

$$\frac{A \cdot \theta_G}{K \cdot \theta_F} < \frac{K \cdot L \cdot \gamma_{\text{th},a}}{\mathbb{E}\{\lambda_{\max}\} \cdot M \cdot N \cdot \gamma_{\text{th},d}}, \tag{39}$$

where A, D, and K are the lengths of the ACK packet, data packet, and pilot, respectively; M, N, and L are the numbers of antennae at the transmitter, receiver, and jammer, respectively; θ_F and θ_G are the path loss from the jammer to the transmitter and receiver; $\gamma_{\text{th},d}$ and $\gamma_{\text{th},a}$ are the SINR thresholds for data and ACK packets. $\lambda_{\max}$ is the maximum eigenvalue of $(\hat{H}^T)^H \cdot \hat{H}^T$.

Proof. With a unit of jamming energy, the jamming energy per symbol is $E_{j,b} = 1/(LD)$, $E_{j,p} = 1/(LK)$, and $E_{j,a} = 1/(LA)$, with L as the number of antennae at the jammer, for barrage jamming, pilot jamming, and ACK jamming, respectively. Omitting the noise components and plugging them in (32), (33) and (34), and applying them to (36), the theorem follows. □

Remark 4. *Although (39) is sufficient to guarantee a higher PER lower bound of the ACK packet (as a result of ACK jamming) than that of the data packet (as a result of pilot jamming), it does not guarantee that ACK jamming is the best choice. Recall Theorem 2 and Remark 3, pilot jamming also affects the PER of the following ACK packets, and sometimes yields a higher PER lower bound (on the ACK packet) than ACK jamming per unit jamming energy. Therefore, only when conditions in (37) and (39) both hold, is ACK jamming a better choice than pilot jamming.*

Remark 5. *The major factors affecting (38) and (39) include the SINR thresholds θ_{th} for the MCSs, the component lengths K, D and A, and the path loss $\theta_{\{*\}}$ for G and F. This means the optimal jamming target varies with the MCSs used by the data and ACK packets, the pilot, payload, and ACK lengths, as well as the distance from the jammer to the transmitter/receiver. In a scenario where any of these factors are dynamic (e.g., in IEEE 802.11 with rate adaptation), optimal jamming strategy needs to be dynamic, too.*

5 APPLICATION TO PRACTICAL SCENARIOS

The theorems presented in the previous section reveals that the optimal jamming target is decided by a set of variables. Obviously, if perfect information about these variables is available, the jammer can at all times optimize its strategy to maximize the impairment generated to the legitimate communication process. However, in practical scenarios, information describing these variables (*e.g.*, pilot length, path loss, channel information, etc.) may be unknown or dynamic. Therefore, it remains unclear whether the theoretical results are applicable to practical scenarios. To answer this question, we resort to machine learning and show that the theorems can be incorporated to and enhance learning-based algorithms.

In the following, we will first introduce a variant of state-of-art reinforcement-learning based jamming algorithm, which does not rely on perfect knowledge on pilot length, channel information and such; then, we design a novel method to improve the algorithm by using the theorems to boost the action exploration. This way, we show how to apply the theoretical results to practical scenarios with unknown and dynamic environment.

5.1 Reinforcement-Learning Based Jamming

Reinforcement learning [20] allows the learning agent to adapt to the optimal action to maximize the reward in a certain environment through trial-and-error. It has drawn attentions from researchers and jamming algorithms based on it have been proposed in [1, 34]. We will derive a similar reinforcement learning algorithm for a dynamic and interactive scenario described below.

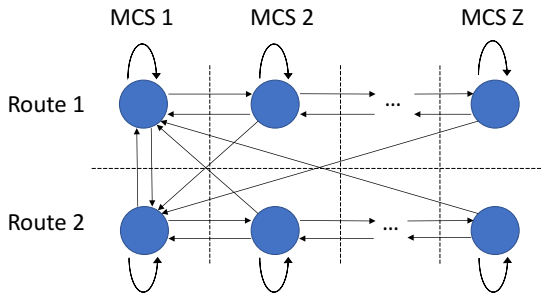

Figure 2: The advanced jamming problem as an MDP.

Without loss of generality, we consider that before each transmission, the transmitter may choose (i) a MCS; and (ii) a route. The choice of MCS is guided by a rate adaption scheme that adjusts the MCS according to the achieved PER. The transmitter may also choose among a set of possible receivers for the next hop, depending on the link quality to each of them. As a result, the jammer faces a dynamic unknown environment where information such as MCS and path loss changes interactively with its own actions.

The scenario can be modeled as a Markov Decision Process (MDP), as shown in Fig. 2 . Formally, it can be defined as a 5-tuple $\{\mathcal{S}, \mathcal{A}, \mathbb{P}(s,s'), R(s,a), \beta\}$, where $\mathcal{S}$ and $\mathcal{A}$ represent the *state* set and *action* set, $\mathbb{P}(s,s')$ is the transition probability from state s to s', $R(s,a)$ is the reward obtained by the decision agent for taking action a at state s, and β is a discount factor.

In our model, the user's degree of freedom is composed of MCS and route, therefore the state set of the MDP is the Cartesian product of the MCS set and route set. To capture the different jamming schemes discussed in Sections 3, we set the action space to a tuple $a = \{J_b, J_p, J_a, T_p\}$, where J_b, J_p, and J_a denote the energy spent on barrage jamming, pilot jamming, and ACK jamming, respectively. We denote by T_p the length of the pilot jamming signal, since the real length of the pilot K is not directly observable.

The objective of the jammer is to reduce the wireless node's throughput with maximum energy efficiency. Therefore, the reward should increase with the throughput degradation and decrease with the energy consumption. For simplicity, we use the degraded throughput subtracted by a "price" paid for the energy spent for the degradation as the reward. Specifically, $R(s,a) = n_e(a) - p(J_b + J_p + J_a)$, where $n_e(a)$ is the number of data packets lost as a result of the jamming action a and p is the price for a unit of energy[2]. The objective of the jammer is to identify a policy $\pi : s \rightarrow a, \forall s \in \mathcal{S}, \forall a \in \mathcal{A}$, that maximizes the discounted sum of the instant and future rewards. We solve this Markov decision process (MDP) problem using the well-known Q-learning algorithm [20], in which the optimum policy is found by iteratively updating the value function for each combination of state and action. We will omit the detailed learning algorithm here for space limit, and focus on the novel action exploration scheme.

[2]Notice that the reward can be inferred by the jammer by (i) computing the amount of jamming energy spent; and (ii) by estimating the throughput degradation through eavesdropping on the number of retransmissions.

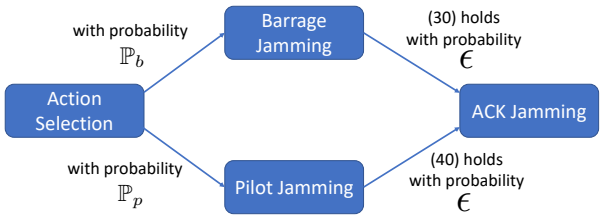

Figure 3: Decision tree for action selection.

5.2 Theorem-Enhanced Action Exploration

Reinforcement learning algorithms are guaranteed to converge, as long as certain conditions are met, **but the convergence speed heavily depends on the action exploration method [20].** Indeed, at each iteration, the learning engine may choose either to (i) *explore* the under-explored actions, or (ii) *exploit* the actions already known to yield a good reward. Thus, an effective exploration method may improve the convergence of the learning process significantly [6]. **Since the theorems derived in Section 4 establish that certain actions are more favorable than others in certain states, we use them to improve the learning exploration phase**.

According to Theorem 1, the jammer should limit the action space to pilot jamming with $T_p \leq \sqrt{D_{\max} \cdot M}$ and barrage jamming. Moreover, Theorem 1 also provides a criterion to choose between pilot jamming and barrage jamming. Suppose the jammer has an estimate on both the data packet length D and pilot length K, in the form of probability distribution $\mathbb{P}\{K = k\}$. Then, it can compute the probability that pilot jamming is more efficient than barrage jamming, as

$$\mathbb{P}_b = \sum_{d \in \mathcal{D}, k \in \mathcal{K}} \mathbb{P}\{D = d\} \cdot \mathbb{P}\{k > \sqrt{dM}\}. \tag{40}$$

Consequently, the jammer should choose barrage jamming with probability $\mathbb{P}_b$, and pilot jamming with $T_p = k$ with probability

$$\mathbb{P}_p = (1 - \mathbb{P}_b) \cdot \mathbb{P}\{K = k\}. \tag{41}$$

At the link layer, Theorem 2 and Theorem 3 specify conditions for the jammer to use ACK jamming, and can be used to prune unfavorable pilot jamming or barrage jamming actions. To be specific, when

$$\frac{A \cdot \theta_G}{K \cdot \theta_F} < \min\left(\frac{K \cdot L}{N^2}, \frac{K \cdot L \gamma_{\text{th},a}}{\mathbb{E}\{\lambda_{\max}\} \cdot M \cdot N \cdot \gamma_{\text{th},d}}\right) \tag{42}$$

holds, ACK jamming is more favorable than pilot jamming; and when Eq. (38) holds, ACK jamming is more favorable than barrage jamming. Since Theorem 2 and 3 are valid for lower bounds on the PER, we let the jammer explore ACK jamming when the conditions hold with probability ϵ. Variables θ_G, θ_F can be estimated by averaging the signal strength received from the transmitter and receiver, given the knowledge of transmitting power. $\mathbb{E}\{\lambda_{\max}\}$, $\gamma_{\text{th},a}$, and $\gamma_{\text{th},d}$ can be computed offline. To summarize, the procedure for enhanced exploration is described as a decision tree in Fig. 3, and reported in detail in Algorithm 1.

Algorithm 1 Enhanced Exploration Algorithm

Initialize $\mathbb{P}\{K = k\}$, $k < \sqrt{D_{\max}M}$;
for each iteration of Q-Learning Algorithm **do**
  Choose barrage jamming with probability (40), and $T_p = k$-length pilot jamming with probability (41);
  if barrage jamming is chosen and (38) holds **then**
    Choose ACK jamming with probability ϵ;
  end if
  if $T_p = k$-length pilot jamming is chosen and (42) holds **then**
    Choose ACK jamming with probability ϵ;
  end if
  Choose jamming energy and set the action;
  Update $\mathbb{P}\{K = k\}$ if pilot jamming is chosen, according to the jamming result;
end for

6 PERFORMANCE EVALUATION

We first validate our theorems in Section 6.1, followed by a testbed evaluation in Section 6.2 and by an evaluation of the learning-based jamming algorithm in Section 6.3.

6.1 Theoretical Validation

To validate the theorems in Section 4, we simulate advanced jamming schemes on a MIMO link. We consider a scenario where the transmitter, receiver, and the jammer all equipped with 2 antennae. The data packet size is set to 1024 bits, and 240 packets are aggregated to one frame, i.e., with one pilot. The pilot length may be 4, 16, 128, or 512 symbols long. ACK packets are assumed to be 512 bits long. The data packets can be modulated with BPSK, QPSK, 16QAM, or 64QAM, while ACK packets are modulated with BPSK.

We evaluate barrage jamming, pilot jamming, and ACK jamming with normalized jamming energy between 0 and 20, where the energy is normalized by the *energy per transmitted symbol*. Note, for barrage jamming on BPSK-modulated signals, a jamming energy of 20 is equivalent to a signal-to-jamming ratio (SJR) of 37.9 dB. In other words, we focus on an energy range that is negligible for traditional barrage jamming.

Due to space limit, we only show the results for BPSK-modulated data packets in Fig. 5. We use bit error rate (BER) as the metric, since it is easy to measure, and is directly decided by SINR, the metric used in both Theorem 1 and 2. Fig. 5 shows that with small pilot length, pilot jamming significantly outperforms barrage jamming. We also notice that the performance gain decreases with increasing pilot length – with pilot length of 512 symbols, pilot jamming becomes similar to barrage jamming on performance. This matches the result predicted by Theorem 1, which states that pilot jamming is more energy-efficient than barrage jamming when $K < \sqrt{DM}$ – in this case, the theoretical crossover point is approximately 495. The effect of pilot jamming on the following ACK packet can also be verified in Fig. 5. For $K = 4$, pilot jamming results in higher BER than ACK jamming with the same energy. However, the advantage becomes negligible for $K = 16$ and ACK jamming becomes better for $K = 128$. This observation validates Theorem 2, which states that pilot jamming is better than ACK jamming if $K < \sqrt{A\theta_G N^2/(L\theta_F)}$ – the point in this case is 32.

Fig. 5 does not precisely match Theorem 3, but we argue that the lower bound of PER is not always reflected well by BER. Actually, with larger jamming energy (not shown due to space limit), the results in BER matches Theorem 3 much better, suggesting that Theorem 3 is more accurate with high jamming energy.

6.2 Experimental Testbed Evaluation

We have implemented a 2×2 MIMO system described in Section 3 using 4 USRP N210s and 1 USRP X310. For the forward (data) transmission, we use the I/Q data from the USRPs to evaluate different jamming schemes. To better control the jamming energy, we manually add barrage and pilot jamming signals to the received samples. The signal processing is run in Matlab, with BER computed as the metric. Due to the limitation on computation speed, we are unable to perform the backward (ACK) transmission in real-time, immediately after a data packet is received. We emulate the ACK transmission and ACK jamming by leveraging channel information estimated from the received samples to perform beamforming, and compute the received ACK signal with jamming signal added. The BER is computed after signal processing of the ACK packet.

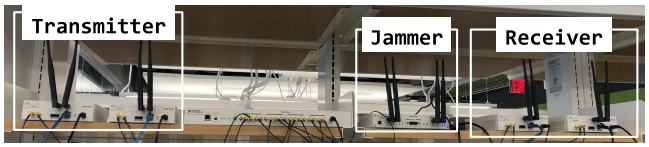

Figure 4: MIMO Experimental Testbed.

For the sake of simplicity, we use QPSK modulation for both data and ACK transmission. Data packets are long 1024 symbols, while the pilot length are long 64, 128, and 512 symbols – we choose long pilots because they are also used for synchronization. The length of ACK is set to 128 symbols. We explore 20 levels of jamming energy, such that the interference level at the receiver side is significant enough to create bit errors. Specifically, the jamming energy is set to $[0.5 : 0.5 : 10] \times 4096$, normalized by the *pre-amplified* symbol energy for the data transmission.

Fig. 6 confirm simulations results. Indeed, for data packets, pilot jamming prevails with short pilots, but becomes less efficient than barrage jamming when the pilot length is increased to a certain level. Although the results do not match precisely Theorem 1 on the crossover point, we point out that Theorem 1 is derived based on the assumption of negligible channel estimation errors, which does not hold in the testbed experiments. Furthermore, the relation between pilot jamming efficiency and pilot length is still valid. For ACK jamming, we observe that ACK jamming effectively destructs ACK transmission, achieving a BER around 0.5.

6.3 Evaluation of Learning Algorithm

In our experiments, we assume there are two routes available to the transmitter, while the set of MCSs includes BPSK-1/2, QPSK-9/16, 16QAM-3/4, and 64QAM-3/4 – we will refer to them as MCS 1 to 4. All the MCSs comply to the PER-SINR function in (9), with the parameters a_z and b_z borrowed from [12]. The state set is then composed of 8 states, indexed as in Table 1.

The jamming normalized energy is set from 10 to 200 with a step of 10. The jammer needs to decide the pilot jamming length T_p, since the jammer is not aware of the real value, which is fixed to 128 in

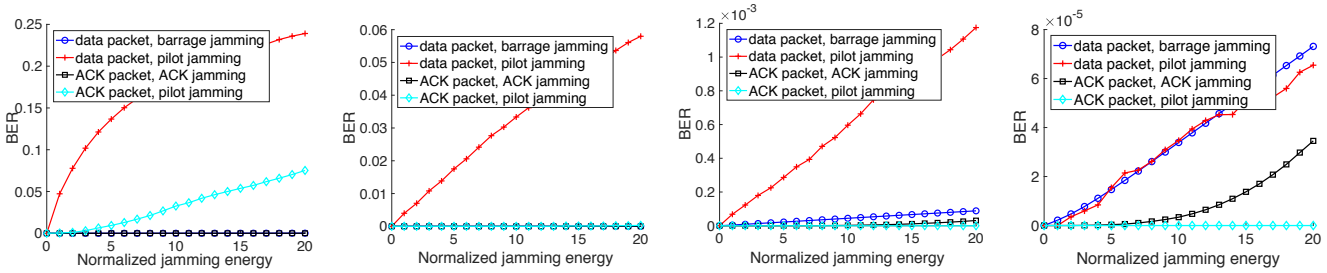

Figure 5: BER for different jamming schemes with pilot length K = (a) 4; (b) 16; (c) 128; (d) 512. Modulation: BPSK

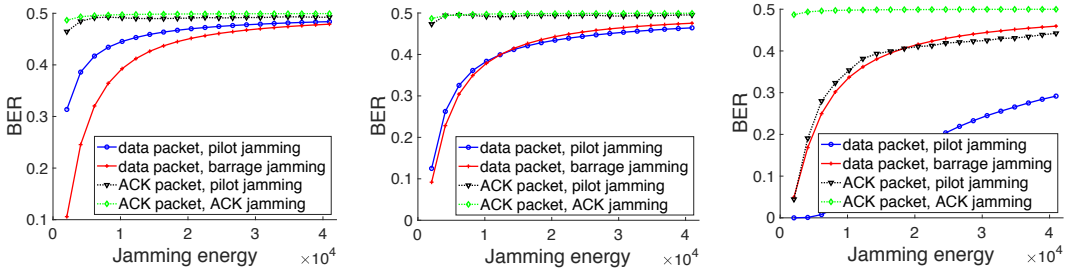

Figure 6: BER as function of jamming energy (normalized to pre-amplified symbol energy) for pilot length (a) 64; (b) 128; and (c) 512.

State Index	Description
1 - 4	Route 1, with MCS 1 - 4
5 - 8	Route 2, with MCS 1 - 4
Action Index	**Description**
1 - 20	Barrage, energy 1 - 20
21 - 40	Pilot, energy 1 - 20, length 4
41 - 60	Pilot, energy 1 - 20, length 16
61 - 80	Pilot, energy 1 - 20, length 128
81 - 100	Pilot, energy 1 - 20, length 512
101 - 120	ACK, energy 1 - 20

Table 1: State and action sets.

the simulation. We discretize the pilot jamming length to values of 4 levels, 4, 16, 128, and 512. We simulate transmission with dynamic MCS and route adaptation (*i.e.*, the transmitter adjusts its MCS and route according to the link throughput). To better illustrate the algorithm convergence in different states, we intentionally let the transmitter perform state transition every 1000 steps.

The resultant state transition is shown in Fig. 7 (a), and the best action at each step is shown in Fig. 7 (b). We can observe that the transmitter starts with route 1 and the highest MCS of 64QAM-3/4. Since receiver 1 is close to the jammer and the pilot length $K = 128$ satisfies (31), the best action is pilot jamming with length 128. This is confirmed by looking at the best action in the initial steps in Fig. 7 (b). After step 1000, the transmitter switches to route 2, and since receiver 2 is far away while the transmitter is close to the jammer, the best action becomes ACK jamming, agreeing with the results shown in Fig. 7 (b). The success of the jammer in degrading throughput is proven by the state adaptation process, which shows that the transmitter adapts a more reliable MCS in each update,

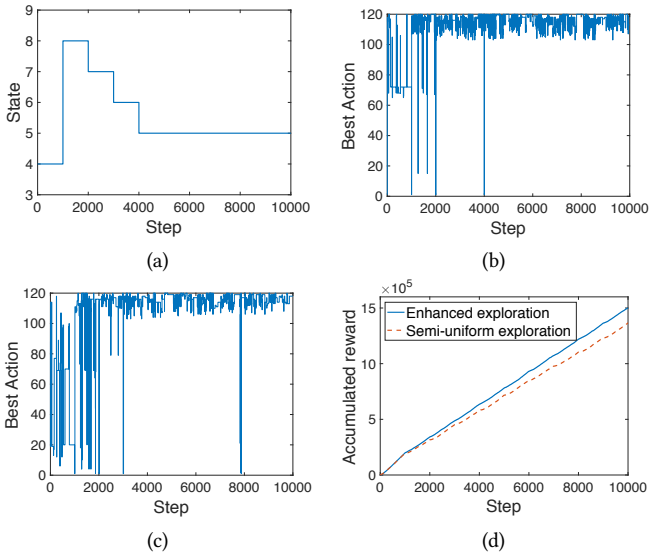

Figure 7: Learning-based jamming: (a) state transition; (b) best action, enhanced exploration; (c) best action, semi-uniform exploration; (d) accumulated reward.

until the most reliable one, *i.e.*, BPSK-1/2 is used. The jammer's best action remains ACK jamming since ACK packets are not affected by the MCS adaption. To show the benefits of the enhanced exploration method discussed in Section 5, we run simulations with the same settings and semi-uniform exploration. State adaptation is the same as in Fig. 7 (a), but the convergence is slower. Fig. 7 (c) concludes that the fluctuation in best action is clearly larger than that for enhanced exploration before step 2000. The benefit of enhanced

exploration is clearer in Fig. 7 (d), where the accumulated reward is shown for both exploration methods.

7 CONCLUSIONS

In this paper, we have identified a set of potential targets for advanced jamming. Then, we have formally modeled jamming schemes aimed at each of these vulnerabilities, and conducted a rigorous analysis resulting in insightful theorems that unveil optimal jamming strategies in scenarios of interest. Them, we have designed a reinforcement learning based algorithm that allows the jammer to adapt its jamming strategy to dynamic environments. The theorems were used to enhance the efficiency of action exploration in the learning process. We have verified the theorems and proved the effectiveness of the proposed algorithm through extensive simulations and experiments.

ACKNOWLEDGEMENTS

We sincerely thank our shepherd Xiaowen Gong and the anonymous reviewers for their constructive feedback, which has helped increase significantly the quality of our manuscript. This material is based upon work funded by the Air Force Research Laboratory (AFRL) under Contract No. FA8750-15-3-6001-NU and by the National Science Foundation (NSF) under Grant CNS-1618727. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the AFRL or the NSF.

REFERENCES

[1] SaiDhiraj Amuru, Cem Tekin, Mihaela van der Schaar, and R. Michael Buehrer. 2016. Jamming Bandits - A Novel Learning Method for Optimal Jamming. *IEEE Transactions on Wireless Communications (TWC)* 15, 4 (April 2016), 2792–2808.

[2] Ahmed H. Anwar, George Atia, and Mina Guirguis. 2017. Dynamic Game-Theoretic Defense Approach Against Stealthy Jamming Attacks in Wireless Networks. In *Proc. of Annual Allerton Conference on Communication, Control, and Computing*. 252–258.

[3] Emrah Bayraktaroglu, Christopher King, Xin Liu, Guevara Noubir, Rajmohan Rajaraman, and Bishal Thapa. 2008. On the Performance of IEEE 802.11 under Jamming. In *Proc. of IEEE Conference on Computer Communications (INFOCOM)*. Phoenix, AZ, USA, 1939–1947.

[4] Mehrzad Biguesh and Alex B. Gershman. 2006. Training-Based MIMO Channel Estimation: A Study of Estimator Tradeoffs and Optimal Training Signals. *IEEE Transactions on Signal Processing* 54, 3 (2006), 884–893.

[5] T. Charles Clancy. 2011. Efficient OFDM Denial: Pilot Jamming and Pilot Nulling. In *Proc. of IEEE International Conference on Communications (ICC)*. Kyoto, Japan, 1–5.

[6] Richard Dearden, Nir Friedman, and Stuart Russell. 1998. Bayesian Q-learning. In *AAAI/IAAI*. 761–768.

[7] Bruce DeBruhl, Christian Kroer, Anupam Datta, Tuomas Sandholm, and Patrick Tague. 2014. Power Napping with Loud Neighbors: Optimal Energy-constrained Jamming and Anti-jamming. In *Proc. of ACM Conference on Security and Privacy in Wireless & Mobile Networking*. Oxford, UK, 117–128.

[8] Bruce DeBruhl and Patrick Tague. 2013. How to Jam Without Getting Caught: Analysis and Empirical Study of Stealthy Periodic Jamming. In *Pro. of IEEE International Conference on Sensing, Communication and Networking (SECON)*. New Orleans, LS, USA, 496–504.

[9] Jithin Jagannath, Nicholas Polosky, Anu Jagannath, Francesco Restuccia, and Tommaso Melodia. 2019. Machine Learning for Wireless Communications in the Internet of Things: A Comprehensive Survey. *arXiv preprint arXiv:1901.07947* (2019).

[10] Matthew J. La Pan, T. Charles Clancy, and Robert W. McGwier. 2012. Jamming Attacks against OFDM Timing Synchronization and Signal Acquisition. In *Proc. of IEEE Military Communications Conference (MILCOM)*. Orlando, FL, USA, 1–7.

[11] Marc Lichtman, Jeffrey D Poston, SaiDhiraj Amuru, Chowdhury Shahriar, T Charles Clancy, R Michael Buehrer, and Jeffrey H Reed. 2016. A Communications Jamming Taxonomy. *IEEE Security & Privacy* 14, 1 (2016), 47–54.

[12] Qingwen Liu, Shengli Zhou, and G. B. Giannakis. 2004. Cross-Layer Combining of Adaptive Modulation and Coding with Truncated ARQ over Wireless Links. *IEEE Transactions on Wireless Communications (TWC)* 3, 5 (September 2004), 1746–1755.

[13] Thomas L Marzetta. 1999. BLAST Training: Estimating Channel Characteristics for High Capacity Space-Time Wireless. In *Proc. of Annual Allerton Conference on Communication, Control and Computing*, Vol. 37. 958–966.

[14] Konstantinos Pelechrinis, Ioannis Broustis, Srikanth V. Krishnamurthy, and Christos Gkantsidis. 2009. ARES: An Anti-jamming Reinforcement System for 802.11 Networks. In *Proc. of International Conference on Emerging Networking Experiments and Technologies (CoNEXT)*. Rome, Italy, 181–192.

[15] Hessam Pirzadeh, S Mohammad Razavizadeh, and Emil Björnson. 2016. Subverting Massive MIMO by Smart Jamming. *IEEE Wireless Communications Letters* 5, 1 (2016), 20–23.

[16] John G. Proakis. 2007. *Digital Communications 5th Edition*. McGraw Hill.

[17] Hanif Rahbari, Marwan Krunz, and Loukas Lazos. 2016. Swift Jamming Attack on Frequency Offset Estimation: The Achilles' Heel of OFDM Systems. *IEEE Transactions on Mobile Computing* 15, 5 (May 2016), 1264–1278.

[18] Francesco Restuccia, Salvatore D'Oro, and Tommaso Melodia. 2018. Securing the Internet of Things in the Age of Machine Learning and Software-Defined Networking. *IEEE Internet of Things Journal* 5, 6 (Dec 2018), 4829–4842. https://doi.org/10.1109/JIOT.2018.2846040

[19] Shabnam Sodagari and T. Charles Clancy. 2015. On Singularity Attacks in MIMO Channels. *Transactions on Emerging Telecommunications Technologies* 26, 3 (2015), 482–490.

[20] Richard S Sutton and Andrew G Barto. 1998. *Reinforcement Learning: An Introduction*. Vol. 1. MIT Press Cambridge.

[21] David J. Thuente and Mithun Acharya. 2006. Intelligent Jamming in Wireless Networks with Applications to 802.11b and Other Networks. In *Proc. of IEEE Military Communications Conference (MILCOM)*. Washington, DC, USA, 1075–1081.

[22] Cheng Wang, Edward KS Au, Ross D Murch, Wai Ho Mow, Roger S Cheng, and Vincent Lau. 2007. On the Performance of the MIMO Zero-Forcing Receiver in the Presence of Channel Estimation Error. *IEEE Transactions on Wireless Communications (TWC)* 6, 3 (2007).

[23] Hui-Ming Wang, Ke-Wen Huang, and Theodoros A. Tsiftsis. 2018. Multiple Antennas Secure Transmission under Pilot Spoofing and Jamming Attack. *IEEE Journal on Selected Areas in Communications (JSAC)* (2018).

[24] Matthias Wilhelm, Ivan Martinovic, Jens B. Schmitt, and Vincent Lenders. 2011. Short Paper: Reactive Jamming in Wireless Networks: How Realistic is the Threat?. In *Proc. of ACM Conference on Wireless Network Security (WiSec)*. Hamburg, Germany, 47–52.

[25] Anthoy D. Wood, John A. Stankovic, and Gang Zhou. 2007. DEEJAM: Defeating Energy-Efficient Jamming in IEEE 802.15.4-based Wireless Networks. In *Proc. of IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*. San Diego, CA, USA, 60–69.

[26] Yongle Wu, Beibei Wang, K. J. Ray Liu, and T. Charles Clancy. 2012. Anti-Jamming Games in Multi-Channel Cognitive Radio Networks. *IEEE Journal on Selected Areas in Communications (JSAC)* 30, 1 (Janurary 2012), 4–15.

[27] Liang Xiao, Yanda Li, Canhuang Dai, Huaiyu Dai, and H Vincent Poor. 2018. Reinforcement Learning-Based NOMA Power Allocation in the Presence of Smart Jamming. *IEEE Transactions on Vehicular Technology* 67, 4 (2018), 3377–3389.

[28] Qiben Yan, Huacheng Zeng, Tingting Jiang, Ming Li, Wenjing Lou, and Y Thomas Hou. 2016. Jamming Resilient Communication using MIMO Interference Cancellation. *IEEE Transactions on Information Forensics and Security* 11, 7 (2016), 1486–1499.

[29] Liyang Zhang and Tommaso Melodia. 2015. Hammer and Anvil: The Threat of a Cross-layer Jamming-aided Data Control Attack in Multihop Wireless Networks. In *Proc. of IEEE Conference on Communications and Network Security (CNS)*. Florence, Italy, 361–369.

[30] Liyang Zhang, Francesco Restuccia, Tommaso Melodia, and Scott M Pudlewski. 2017. Learning to detect and mitigate cross-layer attacks in wireless networks: Framework and applications. In *2017 IEEE Conference on Communications and Network Security (CNS)*. IEEE, 1–9.

[31] Liyang Zhang, Francesco Restuccia, Tommaso Melodia, and Scott Puldlewski. 2018. Taming Cross-Layer Attacks in Wireless Networks: A Bayesian Learning Approach. *IEEE Transactions on Mobile Computing* (2018).

[32] Xu Zhang and Edward W Knightly. 2018. Pilot Distortion Attack and Zero-Startup-Cost Detection in Massive MIMO Network: from Analysis to Experiments. *IEEE Transactions on Information Forensics and Security* (2018).

[33] Lizhong Zheng and D. N. C. Tse. 2003. Diversity and Multiplexing: A Fundamental Tradeoff in Multiple-Antenna Channels. *IEEE Transactions on Information Theory* 49, 5 (May 2003), 1073–1096.

[34] Shaoshuai ZhuanSun, Jun-An Yang, Hui Liu, and Keju Huang. 2017. A Novel Jamming Strategy - Greedy Bandit. In *IEEE International Conference on Communication Software and Networks (ICCSN)*. 1142–1146.

[35] Yulong Zou, Jia Zhu, Xianbin Wang, and Lajos Hanzo. 2016. A Survey on Wireless Security: Technical Challenges, Recent Advances, and Future Trends. *Proc. of the IEEE* (2016), 1–39.