An Active Identification Overriding Attack against RFID: Attack Strategy and Defense Design

Jiahao Xue*, Xiao Han*, Shangqing Zhao†, Yao Liu* and Zhuo Lu*
*University of South Florida, Tampa, FL, USA.
†University of Oklahoma, Norman, OK, USA.

***Abstract*—The security of radio frequency identification (RFID) has been extensively studied in terms of eavesdropping, jamming, relay and tag cloning attacks in the literature. In this paper, we aim to explore a new type of attack, called identification overriding (IDO), in which an attacker tries into inject malicious signals to a tag's reflected signal to override the unique information transmitted by the tag. The IDO attack is designed without the knowledge of the tag's transmitted data and has low computational complexity to inject the malicious data. In addition, the attacker also minimizes the overall energy of the injection signal to make the received signal look normal to the reader. Extensive experiments show that the IDO attack maintains a high success probability generally ranging from 60% to 99% in various evaluation scenarios. Finally, we provide a defense method for the reader to detect the presence of the IDO attack from the received signal.**

***Index Terms*—Radio frequency identification (RFID), Internet of things (IoT), Security, Active attacks**

I. Introduction

Radio frequency identification (RFID) is a wireless communication technique for digit identification that has remarkably low power consumption and cost. A basic RFID system consists of two main components, a reader and a tag [1]. The reader is able to send an RF signal to a reflective wireless device called a tag. The tag modifies and reflects the RF signal to the reader for delivering its identification (and data). RFID is an important part of the Internet of Things (IoT) [2], such as product labels in supermarkets and inventories, wearable devices for healthcare, and highway toll systems.

RFID may be vulnerable to security threats because many reader devices are of small size and limited capabilities, and RFID tags are usually passive (powered by a reader's transmit signal) and cannot execute complex communication and computational tasks, such as cryptography-based encryption or authentication [1]. Passive eavesdropping attacks can simply intercept communication between an RFID reader and a tag to capture confidential data or analyze traffic to infer such data [3], [4]. There are also several types of active attacks that can transmit signals to compromise RFID at the physical layer, including relay attacks, tag cloning, and jamming [5]–[9].

By closely examining existing attack methods, we find that they pay less attention to altering the identification or data of tags on the fly [3]–[9]. For example, the work [7] studied how jamming attacks use high power interference to overwhelm the tag's signal; and cloning attacks only replicate and use genuine tags [9]. We aim to create a new form of active attack that can arbitrarily override a tag's original signal and make the reader accept a falsified identification or other data from the tag in a stealthy way. The attack can disrupt the inventory and logistics (e.g., changing the item label to another label during tag interrogation) while reducing the chance of being detected.

In particular, we create a new attack named identification overriding (IDO) attack against RFID. The IDO attack is designed based on the Electronic Product Code (EPC) RFID Protocols Generation-2 (EPC Gen 2) standard [10], a widely-used protocol today. The IDO attack aims to exploit the RFID decoding process [10], [11] that is based on the average power of a received signal. When a reader first receives the reflected signal by a tag, it calculates the average power as the decoding threshold, then decides a signal symbol as bit 1 if its power is higher than the threshold and 0 otherwise. The basic attack intuition is that the decoding process is vulnerable to the impact of an intentional energy offset, because the power level of 0 may be raised by this offset and decoded as 1. By carefully injecting a signal to the tag's original signal reflected back to the reader, the attacker can tamper with the average power threshold and the power of every bit to affect the decisions of bits 1 and 0. At the same time, the overall energy of the injected signal is minimized by the attacker to make the injection signal look normal (instead of being a high-power signal). We also design the tampering process without the knowledge of the original identification data of the tag, making the attacker more flexible to deploy in practice. Real-world experimental evaluation results show that the IDO attack can alter identification bits on different tags with success probabilities of 60% – 99% in various scenarios.

To combat the IDO attack, we propose a defense method that neither changes tags, nor requires the reader to collect fingerprint data about tags. This defense method is based on a statistic of the energy level changes in the signal received by the reader. We show via experiments that it has a high efficiency to detect the presence of the IDO attack and minimize the impact of the IDO attack on legitimate RFID communication. Our main contributions in this paper are summarized as follows.

- We provide a new RFID attack, called the IDO attack, to actively and stealthily override the tag's data by an attacker's intended data. We provide a practical design to deploy the attack in real-world scenarios.
- We design a defense method that can be easily implemented at a reader to detect the presence of the IDO

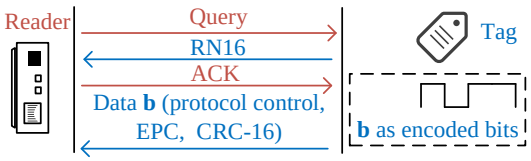

Fig. 1: Communication procedure in EPC Gen-2.

attack. Experimental evaluations show that the method is effective in detecting the attack.

- We conduct experiments to show the impacts of the proposed new attack. We also show that the proposed defense is effective against the attack in experimental evaluations.

The organization of the remainder of this paper is as follows. Section II presents the background of RFID communication and explains our motivation. Section III formulates the attack model and provides practical solutions. Section IV discusses the experiment evaluation results. Section V describes the countermeasure. Sections VI and VII present the related work and the conclusion, respectively.

II. Background and Motivation

In this section, we first introduce RFID communication, then model its signal reception process, and present what motivates us to create the IDO attack.

A. RFID Communication

In this paper, we consider a typical EPC Gen-2 [10] based RFID communication system between a reader and a passive tag, which contains a microchip and an antenna to store and transmit identity information and is powered up by a continuous wave (CW) signal from the reader. According to the standard, as shown in Fig. 1, to identify the identity data stored in the tag, the reader periodically sends a Query command to scan nearby tags. When a tag enters the interrogation zone, it responds by modulating the carrier signal into a random 16-bit data sequence (RN16). The reader then sends an ACK command with the same RN16 to verify communication reliability. Upon successful verification, the tag transmits the signal containing essential information such as protocol control, identity data as EPC, and CRC-16 for error detection, using encoding schemes like FM0 or Miller encoding [10]. If the RN16 does not match, indicating a communication error, the reader reschedules another transmission.

B. RFID Decoding

Let a bit sequence $\mathbf{b} = \{b_1, ..., b_N\}$, where $b_n \in \{0, 1\}$, denote the encoded essential information bits with length N stored on the tag. The tag adopts the Amplitude-Shift Keying (ASK) scheme and modulates the incoming CW signal from the reader by varying its reflection coefficient based on the stored bit sequence. The tag reflection gain $s(b_n)$ for each bit b_n can be defined as a function of the bit value. Typically, $s(b_n) = s_0$ if $b_n = 0$ and s_1 otherwise, where s_0 and s_1 are the tag gains corresponding to the bit values 0 and 1, respectively, satisfying $|s_0| \ll |s_1| \leq 1$ as passive tags never amplify a signal [12].

To interrogate the tag, the reader first transmits a CW signal $x(t) = A\cos(2\pi ft + \phi)$, where A is the amplitude, f and ϕ are the frequency and phase, respectively. Denoted by T_n the set of sampling time slots corresponding to the bit b_n. Then, the tag modulates each bit b_n within T_n based on the reflection gain $s(b_n)$, resulting in the modulated signal $r(t) = s(b_n)h_{RT}A\cos(2\pi ft + \phi)$, for $t \in T_n$, where h_{RT} is the channel coefficient from the reader to the tag. Due to the short time duration of tag reflection [10], we consider the channel remains constant over every T_n, $n \in [1, N]$. Then, the backscattered signal from the tag travels through the channel with coefficient h_{TR} back to the reader. We assume the reader is calibrated to down-convert the signal leakage A from its transmitter to receiver by using RFID carrier leakage cancellation technique [13]. Therefore, the received signal $y(t)$ at the reader can be expressed as:

$$y(t) = h_{RT}h_{TR}s(b_n)A\cos(2\pi ft + \phi) + w(t), \quad (1)$$

where $w(t)$ represents the Additive White Gaussian Noise (AWGN) with zero mean and variance σ^2.

To decode the received signal $y(t)$ into the original bit sequence $\mathbf{b}$, the reader needs to analyze the amplitude variations over time. In (1), the term $h_{RT}h_{TR}s(b_n)A$ indicates the amplitude variation due to the tag's modulation and the channel effects, and the noise $w(t)$ is independent of the signal [11], thus the power of the received signal $y(t)$ is

$$p_n = \frac{1}{|T_n|}\sum_{t\in T_n}(|h_{RT}h_{TR}s(b_n)A|^2 + |w(t)|^2), \quad (2)$$

where $|T_n|$ is the cardinality of T_n. By measuring p_n, the reader can decode it into the corresponding bit b_n. Based on the optimal decoding rule, the reader first measures the average power p_a of all received signals $p_a = \frac{1}{N}\sum_{n=1}^{N} p_n$. Let D be the decoding function, which can be expressed as [10], [11]:

$$D(p_n) = \begin{cases} 1 & \text{if } p_n > p_a \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

C. Challenges and Motivations

Existing attacks on RFID systems primarily fall into three categories: eavesdropping [3], [4], [14]–[17], relay/cloning attacks [5], [18]–[25], and jamming attacks [6]–[8], [26]. Eavesdropping and relay/cloning attacks can capture confidential information and further copy it to replicate genuine tags and gain unauthorized access. Jamming attacks overwhelm the tag signal by high power interference to prevent user access. These attacks focus on compromising the confidentiality and availability of the system, but there has been less attention paid to active attacks that can jeopardize the integrity of the system by injecting false signals.

The integrity of RFID relies on the unique identity of each tag as the main purpose of RFID is to identify different objects or people by their unique tags. If a tag cannot provide a reliable identity to the reader, the whole RFID system becomes less

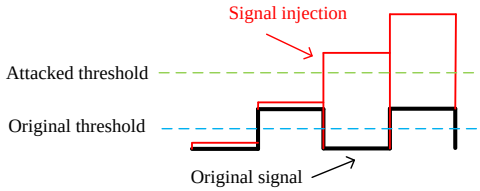

Fig. 2: The impact of injection signal.

trustworthy. Differing from existing attacks, we are interested in a new attack that can actively compromise the integrity of RFID by overriding the identity information. Such an attack should be able to stealthily and arbitrarily override the unique tag information while reducing the chance of being detected.

The intuition of our attack design, as shown in Fig. 2, is that ASK-based RFID adopts a power detection threshold p_a by averaging the signal power over the entire receiving period [10], [11]. By adjusting the received power of malicious injection signals, the attacker can not only manipulate the power of the received signal p_n, but also alter the decoding threshold p_a, which opens a door to influencing the decoding.

Two practical challenges arise in creating such an attack while considering both effectiveness and stealthiness.

- *Input-agnostic design:* The malicious injection signal can not be generated based on a tag's transmit data to override this transmit data. This is because in practice, it can be difficult for the attacker to know the tag's data in advance. In addition, the attack is assumed to be unaware of the reader's and tag's setups, such as CW amplitude A and hardware reflection gains $s(b_n)$.
- *Minimizing the energy of injection signal while maintaining attack effectiveness:* The attack can be quite successful if it significantly overpowers the tag's signal at the reader. However, RFID has a normal range of the received energy and a simple outliner detection can notice the abnormality due to high power injection. The attack needs to make efforts to minimize the injection signal's energy to make the received signal look normal but still result in an effective attack.

In the next section, we will create the IDO attack to address the challenges in practical scenarios.

III. The IDO Attack

In this section, we detail the design of the IDO attack. We first describe the attack model, then formulate the attack, and eventually present the practical attack procedure.

A. Adversary Model and Assumptions

In this paper, we consider an attacker transmitting a malicious signal $\mathbf{a} = \{a_1, ..., a_N\}$ to the RFID communication. To ensure that the injected signal can be decoded, the frame format and modulation of the malicious signal follow the EPC Gen 2 standard, which includes the CRC field and is modulated by ASK. As previously mentioned, we do not have prior knowledge of $\mathbf{b}$ and only transmit the malicious $\mathbf{a}$ in one shot. We assume that the attacker's injection signal is synchronized and superimposed to the tag's signal transmitted to the reader, altering p_n and p_a during the decoding. In practice, this synchronization is achievable by eavesdropping on the RFID system. When the ACK is detected, the attacker can start transmitting $\mathbf{a}$ after an inter-frame time interval that can be calculated by the symbol length of ACK [10]. For commercial off-the-shelf tags, the bit time length typically exceeds $1\mu s$ [10], [27], making the delay caused by distance negligible and allowing the reader to receive synchronized signals from both the tag and the attacker.

B. IDO Attack Formulation

The objective of the IDO attacker is to transmit the malicious signal $\mathbf{a}$ to the reader to successfully manipulate the received signal without being detected by the CRC check. In ASK, let A_{a_n} be the modulated amplitude of the bit a_n at the attacker, thus $A_{a_n}^2$ is its transmitting power, then the transmitted malicious signal can be written as $x_a(t) = A_{a_n} cos(2\pi ft + \phi)$. There are two propagation paths for the attacker's malicious signal: 1) direct path and 2) reflected path. Let h_{AR} be the channel coefficient from the attacker to the reader, and h_{AT} be the channel coefficient from the attacker to the tag. After adding the malicious signal, the received signal (1) can be rewritten as

$$\begin{aligned} y(t) =& h_{RT}h_{TR}s(b_n)A\cos(2\pi ft+\phi) + w(t) \\ &+ (h_{AR} + h_{AT}h_{TR}s(b_n))A_{a_n}\cos(2\pi ft+\phi), \end{aligned} \tag{4}$$

for $t \in T_n$. Then, the received power (2) can be rewritten as

$$\begin{aligned} p_n = \frac{1}{|T_n|}\sum_{t\in T_n}\big\{&|h_{RT}h_{TR}s(b_n)A \\ &+ (h_{AR} + h_{AT}h_{TR}s(b_n))A_{a_n}|^2 + |w(t)|^2\big\}. \end{aligned} \tag{5}$$

1) Basic Formulation: Based on (5), the attacker can control A_{a_n} such that the decoded signal of each p_n is a_n. To maintain stealthiness and ensure that the received injection signal still resembles a tag-reflected signal, the amplitude A_{a_n} (and the received energy of the malicious signal added to the reflected signal) should be minimized at the reader. Therefore, for $n \in [1, N]$, the attack can be formulated as

$$\text{Objective: } \min_{A_{a_n}} \frac{1}{N}\sum_{n=1}^{N} A_{a_n}^2, \tag{6a}$$

$$\text{Subject to: } D(p_n) = a_n.\ \forall n \in [1, N] \tag{6b}$$

Since we consider an input-agnostic scenario where the attacker has no knowledge of $\mathbf{b}$, obtaining explicit theoretical solutions for directly solving the optimization problem (6) is difficult. Thus, we adopt a statistic model to make the attack formulation input-agnostic (i.e., no need for the exact knowledge of $\mathbf{b}$).

2) Input-Agnostic Attack: The constraint (6b) represents that the attacker aims to successfully inject its intended bits at the reader. To make the solution independent of the knowledge $\mathbf{b}$, we treat b_n in $\mathbf{b}$ as a random variable following the independent Bernoulli distribution with parameter 0.5 (i.e., b_n

is 0 or 1 with equal probability). It enables us to reformulate the constraint (6b) in a probabilistic form as

$$P_{att} = \prod_{n=1}^{N} P_{att}^{n} = \prod_{n=1}^{N} \Pr\left(D(p_n) = a_n\right) \geq P_{th}, \quad (7)$$

where $P_{att}^{n} = \Pr(D(p_n) = a_n)$ is the probability that the attacker successfully changes the n-th bit to its intended value, P_{att} denotes the probability of the attacker successfully changing all bits, and P_{th} is a success probability threshold that the attacker aims to achieve in a practical scenario.

3) Minimizing IDO Attack Signal's Energy: To make the injection signal's energy fall within a normal RFID reception range, the attacker should minimize its energy by carefully designing its amplitude $A_{a_n} \in \{A_L, A_H\}$, where A_L and A_H denote the injection signal's energy levels when the intended bits are 0 and 1, respectively. However, reducing both A_L and A_H may decrease the attack success probability (e.g., $A_L = A_H = 0$ means no signal injection). To find the optimal balance, we first introduce the following lemma to show the relationship between A_L, A_H and P_{att}, and then use additional theorems to simplify the problem (6) to obtain the solution of A_L and A_H.

Lemma 1. *Given a_n and b_n, let $\mu(A_{a_n}, b_n)$ be the mean of $y(t)$ in* (4) *over $t \in T_n$. Then, the power level p_n in* (2) *follows the non-central Chi-squared distribution with degree* 2 *and its cumulative distribution function is written as* $\Pr\{p_n \leq p_a | a_n, b_n\} = 1 - Q_1(\sqrt{\frac{\mu^2(A_{a_n}, b_n)}{\sigma^2}}, \sqrt{\frac{p_a}{\sigma^2}})$. *The attacker's success probability can be expressed as*

$$\begin{aligned} P_{att}^{n} = & \frac{1}{4} \sum_{i \in \{0,1\}} \left[1 - Q_1\left(\sqrt{\frac{\mu^2(A_L, i)}{\sigma^2}}, \sqrt{\frac{p_a}{\sigma^2}}\right) \right. \\ & \left. + Q_1\left(\sqrt{\frac{\mu^2(A_H, i)}{\sigma^2}}, \sqrt{\frac{p_a}{\sigma^2}}\right) \right], \end{aligned} \quad (8)$$

where $Q_1(\cdot, \cdot)$ is the Marcum Q-function of order 1, p_a is the average power of the received signal, and σ^2 is the variance of AWGN.

Proof. See Appendix A. □

Lemma 1 provides a theoretical connection from the malicious signal powers A_L and A_H to the attack success probability. Based on Lemma 1, we provide a theorem to intuitively show how A_L and A_H affect the attack success probability.

Theorem 1. *Under the condition that channel gain $|h_{AT} h_{TR} s(b_n)| < |h_{AR}|$, the following two statements hold: 1) The success probability P_{att}^{n} is monotonically decreasing with A_L. 2) The success probability P_{att}^{n} is monotonically increasing with A_H.*

Proof. Appendix B. □

The channel condition $|h_{AT} h_{TR} s(b_n)| < |h_{AR}|$ specified in Theorem 1 indicates that the reflected channel from the attacker to the tag and then to the reader should be weaker than the Line-of-Sight (LoS) channel from the attacker to the reader. This condition is reasonable in practice, because the LoS channel is generally better than the reflected one.

Theorem 1 provides an explicit guideline for the attacker to set both A_L and A_H in the injection signal. According to statement 1) in Theorem 1, we should set $A_L = 0$ during the attack, because this reduces the overall energy of the injection signal towards the goal in (6a) and increases the success probability. This indicates that to dominantly control the decoded signal based on a_n while maintaining stealthiness, the optimal approach should keep $A_L = 0$ and only increase the energy A_H to adjust the decision threshold p_a so that it overrides the tag's signal. In other words, when $a_n = 0$, the attacker should not transmit a signal over the n-bit duration. Statement 2) in Theorem 1 shows that the attack success probability is monotonically increasing with A_H. Leveraging the monotonic relationship can speed up the search for an optimal solution of A_H.

The two major results in Theorem 1 simplify (6) to a univariate optimization problem for the IDO attack as

$$\begin{aligned} & \text{Objective: } \min \ A_H, \\ & \text{Subject to: } P_{att} \geq P_{th}. \end{aligned} \quad (9)$$

The solution to the univariate optimization (9) can be efficiently found by existing optimization algorithms such as interior-point or sequential quadratic programming [28], [29] with a fast convergence rate, which are suitable for RFID scenarios.

4) Taking Into Account Reader's Transmit Power and Tag Reflection Gain: Although most RFID readers have a minimum transmit power of 0 or 10 dBm and a maximum transmit power about 30 dBm [30], [31], the exact value of the reader's transmit amplitude A is unavailable to the attacker and needed to solve (9). In addition, the reflection gains s_0 and s_1 are hardware-dependent and often proprietary information unknown to the attacker. However, $s_0 \approx 0$ holds for most tags to maintain the transmission efficiency [32] and we only need to consider the value of s_1 for the attacker to solve (9).

To address this problem of setting up A and $|s_1|$, we first analyze the relationship between A, $|s_1|$, and the attack success probability in the following theorem.

Theorem 2. *Given $A_L = 0$ and a fixed value of A_H, the probability P_{att}^{n} is a monotonically decreasing function of both A and $|s_1|$.*

Proof. See Appendix C. □

Theorem 2 indicates that the success probability P_{att}^{n} decreases as the reader's transmit amplitude A increases. Therefore, to maintain a high success probability with the uncertainty of A, the attacker should set A to be the maximum value, (e.g., 30 dBm [31]). Similarly, we set the maximum reflection gain $|s_1| = 1$ [12].

5) The Impact of Channel Gains: From the EPC Gen-2 communication process shown in Fig. 1, A_H is solved before data $\mathbf{b}$ is transmitted as all channel gains can be estimated during the Query, RN16, and ACK transmissions prior to data $\mathbf{b}$. Specifically, due to the channel reciprocity, we have

- The LoS channel from the attacker to the reader h_{AR} can be measured by receiving the ACK packet from the reader.
- The reflected channel of the attacker $h_{AT}h_{TR}$ can be estimated when the tag replies with RN16 to the reader, which is also received by the attacker.
- Since the communication between the reader and the tag is usually obstacle-free and LoS, the path loss $h_{RT}h_{TR}$ can be modeled as free-space path loss [33]. Therefore, $h_{RT}h_{TR}$ can be derived based on the distance between the reader and the tag. If the distance is unknown, we assume it to be the shortest possible distance, a wavelength, as this assumption is equivalent to using the highest transmit power.

C. Practical Attack Procedures

Our previous analysis provides an input-agnostic approach for generating the malicious injection signal A_{a_n} in real-world RFID systems. Based on the actual EPC gen 2 protocol [10], the IDO attack needs to be divided into three phases.

1) **Sensing Phase**: In this phase, the attacker senses the wireless environments to gather all necessary information for generating A_H, including estimating all channel gains. This phase completes after the attacker detects the start of the ACK transmission.
2) **Attacking Phase**: The attacker uses the collected information to solve (9) and then generates the malicious signal $x_a(t)$ based on $\mathbf{a}$. After a fixed inter-frame time interval delay defined in the protocol, the attacker transmits $x_a(t)$ to the reader. The malicious bit sequence $\mathbf{a}$ is designed to pass the CRC check [10].
3) **Retry Phase**: In this phase, the attacker gets an additional chance during a communication failure between the reader and the tag. A failure to pass the CRC check will trigger the reader to launch a new inventory round. In this case, the reader will transmit a NAK command to verify the tag again [10]. The attacker, upon the reception of the NAK command, will know the previous attack fails and then launch the attack again.

IV. Experimental Evaluations

In this section, we conduct comprehensive experiments to evaluate the effectiveness of our proposed attack. We first introduce the setups, then show and analyze the results.

A. Environment Setup

The experiments were conducted in a large in-door environment that used two USRP X310s and commercial off-the-self tags working at 900 MHz. The first USRP was used to emulate the reader and the second is the IDO attacker, as an example shown in Fig. 3. We tested 10 different RFID tags with ID numbers 1-10: ID 1: Confidex Crosswave Classi UHF RFID Label; ID 2: Omni-ID Exo 750 Tag; ID 3: Vulcan RFID Embeddable Wire Tag; ID 4: Vulcan Custom Universal Mini Asset Tag; ID 5: Vulcan Custom Credential Tag; ID 6: Vulcan Arrow White Wet Inlay; ID 7: Beontag F62 Paper Tag; ID 8: Beontag Buhrer P60 Wet Inlay; ID 9: Vulcan Fire UHF White Wet Inlay; ID 10: HID SlimFlex Tag.

The default settings during our experiments were as follows: We used the tag with ID 1 as the default tag. The distance between the tag and the reader was 3 meters. The attacker was 6 meters away from the reader, which was considered as a long distance scenario. To conveniently represent A_L and A_H in experiments, we used the attacker to transmit a test signal with $A_L = 0$ and $A_H = A^*$ where A^* is an amplitude value adjusted in USRP such that the signal had a received power of -67 dBm measured at the reader, which was within the normal RFID received power range (i.e., -92 dBm to -60 dBm [34], [35]) and led to successes of most attacks. We used this signal as a reference signal and normalized all amplitudes set up in experiments by A^*. In other words, the reference signal has a setting of $A_L = 0$ and $A_H = 1$. In addition, we moved the reader, the attack, and the tag to create different distance scenarios to analyze the impact of the attack.

B. Experimental Results

1) Impact of Attack Signal: We first set up the IDO attack with different A_L and A_H values to evaluate the attack success probability in Fig. 4. We choose A_L = 0, 0.04, 0.08, and 0.12 and vary A_H from 0.2 to 1.0.

As Fig. 4 shows, the success probability rises to nearly 100% with increasing A_H for any A_L value. We observe that the probability is higher than 97% when A_H is higher than 0.8. Fixing A_H while reducing A_L improves the attack success probability. For example, given $A_H = 0.6$, the attack success probability increases from 94.3% to 97.4% when A_L is reduced from 0.12 to 0. Therefore, we should set $A_L = 0$ for the IDO attack to maximize its success probability as Theorem 1 shows.

2) Impacts of Environments: We directly change the distance between the attacker and the reader to test how it affects the success probability. In particular, we consider three different distance scenarios: the short, medium, and long scenarios with distances of 1, 3.5, and 6 meters between the reader and the attacker, respectively. We change A_H from 0.25 to 1.0. As Fig. 5 shows, the success probability is an increasing function of A_H for all three distances. The longer the distance between the attacker and the reader, the lower the probability is. For example, when $A_H = 0.44$, the probability falls to 86.6% in the long distance case; in contrast, it is 97.7% in the short distance case. Therefore, when the attacker is far away from the reader, it needs to increase its transmit power to maintain its success.

3) Impacts of Movements: Fig. 6 depicts the attack success under a movement scenario, where we move the tag around the reader and measure the attack impact during the movement. We set A_H from 0.23 to 1.0. When the tag is static, the attacker has a slightly higher success probability than the probability during the tag movement. For example, when $A_H = 0.42$, the attacker has 87.3% vs 90.8% success probabilities against static vs moving tags. This is mainly due to the reason that the received power varies during the tag movement and the power

Fig. 3: An example of RFID experiment scenarios.

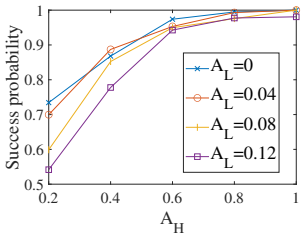

Fig. 4: Success probabilities with different amplitudes.

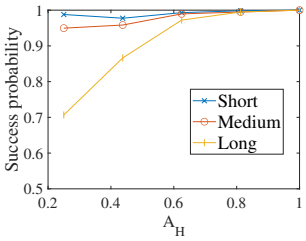

Fig. 5: Success probabilities in different distance scenarios.

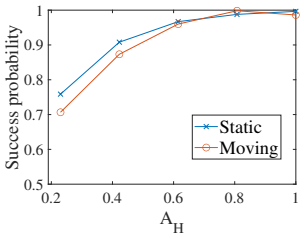

Fig. 6: Success probabilities with static vs moving tags.

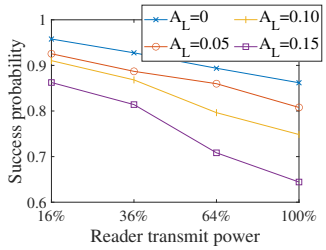

Fig. 7: Success probabilities when $A_H = 0.33$.

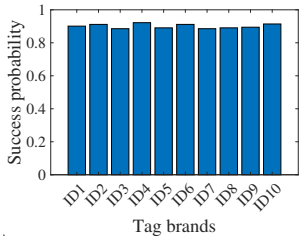

Fig. 8: Success probabilities against different tags.

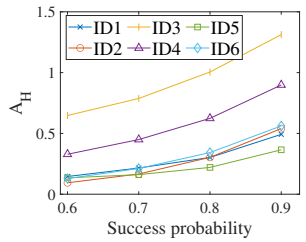

Fig. 9: Optimized A_H values for different tags.

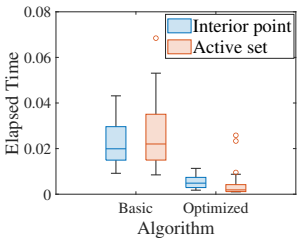

Fig. 10: Comparison of computational complexities.

decoding threshold may not be as accurate. Fig. 6 shows that the attack is still effective during the movement scenario.

4) Impacts of Reader's Transmit Powers: We also evaluate how the reader's transmit power affects the success probability given different A_L values for $A_H = 0.33$. We change the reader's transmit power from the default maximum value (i.e., 100%) to 16% with $A_L = 0$, $A_L = 0.05$, $A_L = 0.10$, and $A_L = 0.15$. Fig. 7 shows the success probability decreases as the reader's transmit power increases. For example, when $A_L = 0.10$, increasing the reader's power from 36% to 64% incurs the success probability reducing from 83.6% to 78.8%. Generally, when the reader's transmit power goes from 16% to 100%, the success probability drops from 86.0% to 59.1%. It indicates that a higher A_H value is needed to maintain the success probability. Therefore, the attacker should transmit with a higher signal amplitude as Theorem 2 shows.

5) Impacts of Tag Products: We then measure the attack success probabilities against different tag products in Fig. 8 with $A_L = 0$ and $A_H = 0.55$. We can see that generally, the attacker can achieve an average of 90.0% success probability against all 10 tags, in which the tag with ID 10 has the highest probability of 91.4%, and the one with ID 3 has the lowest probability of 88.5%.

6) Optimized Attack: Next, we calculate the optimal A_H value based on the optimized IDO attack in (9) for different tags. We set the threshold of the intended attack success probability P_{th} as 0.6, 0.7, 0.8, 0.9. As Fig. 9 shows, when we increase the threshold P_{th} (i.e., asking for more successful attacks), the solution to (9) yields a higher A_H value. For example, if we set $P_{th} = 0.7$, the attacker needs to set $A_H = 0.2$ to affect the tag with ID 6. In addition, it is also observed from Fig. 9 that multiple types of tags lead to different optimal solutions of A_H due to their hardware differences (e.g., reflection gains).

Finally, we measure the computational complexity to solve the basic formulation (6) and the optimized univariate formulation (9). We adopt two numerical optimization algorithms, interior point and active set methods [28], [29], to solve both formulations. Fig. 10 shows via box-plots that this univariate formulation (9) can substantially reduce the computational complexity. For example, when using the active set algorithm, the median of runtime to solve (6) is 0.022 seconds, but (9) only needs 0.002 seconds to solve, leading to a runtime reduction of 91%.

V. Defense Strategy

Our experimental results in Figs. 4, 5, and 6 show that the receiving power of the signal affected by the IDO attack can still fall within the same range as a normal tag-reflected signal. Existing defense methods against power overwhelming attacks, typically focused on detecting the anomaly due to high power interference (e.g., Wi-Fi [7], Bluetooth [7], and backscatter communication [6]), cannot be readily adopted in our scenario. As a result, it is necessary to create a new defense against the IDO attack in the RFID communication scenario.

A. Defense Design

Recall Theorem 1 shows that power p_n follows the non-central Chi-squared distribution with parameter $\mu^2(A_{a_n}, b_n)$. Given a particular value n and four combinations of $A_{a_n} \in \{A_L, A_H\}$ and $b_n \in \{0, 1\}$, p_n should follow one of four non-central Chi-squared distributions (with parameters $\mu^2(A_L, 0)$, $\mu^2(A_L, 1)$, $\mu^2(A_H, 0)$, and $\mu^2(A_H, 1)$). If there is no attacker, p_n should follow one of two distributions. Intuitively, this difference can be used to design a defense method to detect the

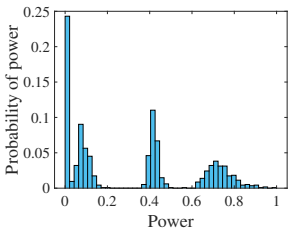

Fig. 11: Example of the signal power distribution.

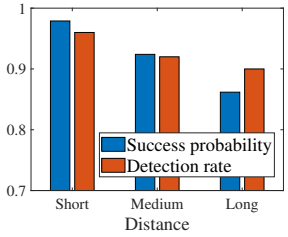

Fig. 12: Detection rates in three distance scenarios.

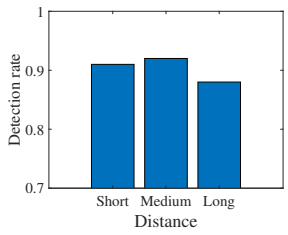

Fig. 13: Detection rates for the optimized attack.

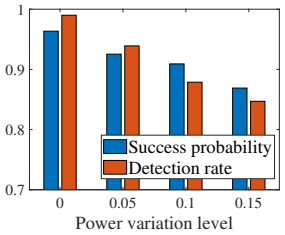

Fig. 14: Detection rates for a power-varying attack.

IDO attack. After the reader receives the signal that contains $\mathbf{b}$, it can obtain the distribution of the signal power, as an example shown in Fig. 11. Then, the reader needs a method to determine how many distributions can be recognized in the received signal and what type of distributions they are.

Based on this motivation, we create our defense mechanism, which consists of three steps: 1) initial data splitting, 2) distribution fitting, and 3) attack detection.

1) Initial data splitting: The first step splits the received power into four groups, denoted as $\mathcal{Y}_{ij}$, which is the set of received power values p_n under the conditions $a_n = i$ and $b_n = j$, where $i, j \in \{0, 1\}$. Since the optimal decoding threshold p_a is the mean of all samples, the initial splitting can be done by first dividing all samples into two halves based on the average power. Each half can then be further split based on their respective average values.

2) Distribution fitting: We use the result of distribution fitting to compare with a user-defined threshold, and determine the presence of the IDO attack.

Let P_{ij} be a random variable representing the received power under $a_n = i$ and $b_n = j$. According to Lemma 1, we know P_{ij} follows the non-central Chi-squared distribution with degree 2 and noncentrality parameter μ_{ij}^2/σ_{ij}^2, where μ_{ij} and σ_{ij}^2 are the noncentrality distance and scale of P_{ij}, respectively. Denote $f_{P_{ij}}(p_n; \mu_{ij}^2/\sigma_{ij}^2)$ as the PDF of P_{ij}. Let $\Pr\{\zeta|\mathcal{Y}_{ij}\}$ be the probability that a measured power is distributed within group $\mathcal{Y}_{ij}$ in the histogram. Then, our objective is to separate the signal into the most evident four groups via finding the optimal combination of μ_{ij} and σ_{ij} to minimize the mean squared error

$$\epsilon_{ij} = \frac{1}{|\mathcal{Y}_{ij}|} \sum_{n}^{n \in \mathcal{Y}_{ij}} \left\| \Pr\{\zeta|\mathcal{Y}_{ij}\} - \int_{\zeta} f_{P_{ij}}(p_n; \frac{\mu_{ij}^2}{\sigma_{ij}^2}) dp_n. \right\|^2 \tag{10}$$

Existing methods (e.g., [36]) can be used to solve (10).

3) Attack detection: Given the optimal fitting error ϵ_{ij}^* in minimizing (10) and the corresponding parameters μ_{ij}^* and σ_{ij}^*, we detect the presence of an attack if the following two conditions hold.

1) The normalized fitting error for all sets is below a threshold ϵ_{th}, i.e., $\epsilon_{ij}^*/\mathbb{E}^2(p_n) < \epsilon_{th}$.
2) Due to the short period of an RFID signal, all p_n values should have similar variances due to the same impact of AWGN. Thus, the Index of Dispersion among all σ_{ij}, defined as the variance of σ_{ij} over the mean of σ_{ij}, should be less than a threshold σ_{th}.

B. Experimental Results

Next, we use experiments with default settings adopted from Section IV to evaluate the effectiveness of the proposed attack detection method. This detection method is tuned to have a low false alarm rate to avoid undermining the normal communication between the tag and the reader. We use the detection rate, which is defined as the probability that the detection method can successfully detect the presence of the attack from a reader's signal.

We first consider an injection attack signal with fixed $A_L = 0$ and $A_H = 0.43$ in three distance scenarios similar to Section IV-B (i.e., short: 1 meter, medium: 3.5 meters and long: 6 meters). In Fig. 12, we measure the detection rate along with the attacker's success probability in the case of no detection. We can see that in the short distance scenario, the attack can be quite successful with a probability of 97.9%, but can also be detected by our method with a rate of 96.2%. Fig. 12 also shows that a lower success attack with a longer distance leads to a lower detection rate, because the attack's signal received at the reader becomes weaker. Although more difficult to detect, it has a smaller success probability.

Fig. 13 shows the performance of the detection method in the face of the optimized attack, which optimizes A_H such that $P_{att} \geq P_{th} = 90\%$ based on (9). We can find that in the same short, medium and long distance scenarios, the detection rate is generally about 90%. For example, the detection rate is 92.0% in the medium distance scenario.

The intuition of the defense is based on the recognition of four distributions. We also study whether an attack, which is aware of the defense, can affect the detection by varying its power. In particular, we test how varying A_H can impact the detection rate. We set $A_L = 0$ and vary $A_H = 0.75\beta$ where β follows a uniform distribution within the range $[1-\gamma, 1+\gamma]$. We set γ to be 0, 0.05, 0.1, 0.15 to control the power variation level. Fig. 14 shows that by increasing γ, the success probability drops from 96.4% to 86.9% due to the varying nature of the signal from attacker. At the same time, the detection rate decreases from 99.1% to 84.7%. We note that with a detection rate of 84.7%, the attack has a probability of 1 - 84.7% = 15.3% to evade the detection. With the attack's success probability itself being 86.9%, the probability

of overall attack success (i.e., the attack successfully evading the detection and injecting its intended bits into the reader) is $15.3\% \times 86.9\% = 13.29\%$, which shows the detection can effectively limit the IDO attack's impact.

VI. Related Work

Eavesdropping attacks. Eavesdropping attacks that passively steal information were extensively considered in [3], [4], [37], [38]. For example, a detection algorithm is proposed in [4] to detect eavesdropping by leveraging the magnetic coupling between the antennas of RFID and the attacker. Random physical stimulation and subsequent behavior analysis are used to detect whether a tag is being attacked.

Relay and tag cloning attacks. A tag is vulnerable to relay attacks that maliciously send the reflected signal by the tag to another place [5], [9], [18]–[25]. For example, the work in [23] developed a man-in-the-middle attack with a high success ratio on today's contactless payment methods over NFC communication.

Jamming attacks. An attacker can send the jamming signal at the same RFID frequency with a higher power than the RFID signal [6]–[8], [26]. For example, the work in [6] proposed an anti-jamming attack strategy using deep reinforcement learning to learn the jamming behaviors, and adjust the packet rate and energy consumption of the transmitter.

Injection Attacks. Existing studies tried to inject signals into various communication systems, though they cannot be directly adapted to the RFID case. The work in [39], [40] proposed injection attacks targeting the frequency-hopped spread-spectrum communication that is vulnerable to the packet capture phenomena, but it is not applicable to the ASK modulation-based RFID [10]. Studies in [41], [42] exploited unique procedures of IEEE 802.11 and Glossy protocols, but such procedures can not be found in EPC gen 2 [10]. In addition, [43], [44] proposed to inject signals to devices using the pre-knowledge of circuit hardware design, but it can be difficult to know the circuit in practice.

Authentication and Fingerprinting. Encryption and authentication techniques have been proposed to protect RFID [18], [19], [21], [22], [24], [25], [45]. The work in [18] develops an authentication protocol to encrypt the communication between reader and tag using physically unclonable functions. This unpredictable function is based on the minor hardware difference among tags, hence it is challenging to compromise the encryption. The work of [21] found that the power of tag signal varies over frequency and every tag has a unique reflection coefficient, which can be used as the fingerprint of tags. However, those defense methods may not be reliable to the temperature and supply voltage fluctuation [46].

VII. Conclusion

In this paper, we proposed an IDO attack, which is able to inject a malicious signal into a tag's reflection signal at a reader to arbitrarily modify the tag's data. We provided mathematical modeling to formulate the attack as an optimization problem to balance the attack success probability and the stealthiness in terms of a low energy level for the injection signal in practical scenarios. We also proposed a countermeasure design that can detect the presence of the IDO attack. Experimental results demonstrated that the proposed detection achieved high detection rates in evaluation scenarios.

Acknowledgement: The work at USF was supported in part by NSF Grants 2044516 and 2316719. The work at OU was supported in part by NSF Grant 2316720.

Appendix A
Proof of Lemma 1

As the channel remains constant over T_n, $n \in [1, N]$, $y(t)$ in (4) within T_n follows a complex Gaussian distribution with a non-zero mean $\mu(A_{a_n}, b_n)$ due to AWGN $w(t)$. Given (5), power p_n follows the non-central Chi-squared distribution with degree 2 [47]. Therefore, we have

$$\begin{aligned}
&\mu(A_{a_n},b_n)=|h_{AR}A_{a_n}+h_{AT}h_{TR}s(b_n)A_{a_n}+h_{RT}h_{TR}s(b_n)A| \\
&= ((h_{AR}A_{a_n} + h_{AT}h_{TR}s(b_n)A_{a_n} + h_{RT}h_{TR}s(b_n)A) \\
&\times (h^*_{AR}A_{a_n} + h^*_{AT}h^*_{TR}s(b_n)A_{a_n} + h^*_{RT}h^*_{TR}s(b_n)A))^{\frac{1}{2}}, \\
&p_a = 2\sigma^2 + \sum_{i=0}^{1}(\mu^2(A_L, i)P(a_n = 0)P(b_n = i) \\
&+ \mu^2(A_H, i)P(a_n = 1)P(b_n = i)). \qquad (11)
\end{aligned}$$

Appendix B
Proof of Theorem 1

We introduce Lemmas 2 to 4 to prepare our proof.

Lemma 2. $Q_1(\sqrt{\frac{\mu^2(A_{a_n},b_n)}{\sigma^2}}, \sqrt{\frac{p_a}{\sigma^2}})$ is increasing with A_{a_n} if the following inequality holds:

$$\begin{aligned}
&I_1(\frac{\sqrt{\mu^2(A_{a_n}, b_n)p_a}}{\sigma^2})/I_0(\frac{\sqrt{\mu^2(A_{a_n}, b_n)p_a}}{\sigma^2}) \\
&> \frac{\partial\sqrt{\frac{p_a}{\sigma^2}}}{\partial A_{a_n}}/\frac{\partial\sqrt{\frac{\mu^2(A_{a_n},b_n)}{\sigma^2}}}{\partial A_{a_n}}. \qquad (12)
\end{aligned}$$

Proof. Based on total derivative, $Q_1(\sqrt{\frac{\mu^2(A_{a_n},b_n)}{\sigma^2}}, \sqrt{\frac{p_a}{\sigma^2}})$ is increasing with A_H if the following inequality holds:

$$\begin{aligned}
&\frac{\partial Q_1(\sqrt{\frac{\mu^2(A_{a_n},b_n)}{\sigma^2}}, \sqrt{\frac{p_a}{\sigma^2}})}{\partial\sqrt{\frac{\mu^2(A_{a_n},b_n)}{\sigma^2}}}\frac{\partial\sqrt{\frac{\mu^2(A_{a_n},b_n)}{\sigma^2}}}{\partial A_{a_n}} \\
&> -\frac{\partial Q_1(\sqrt{\frac{\mu^2(A_{a_n},b_n)}{\sigma^2}}, \sqrt{\frac{p_a}{\sigma^2}})}{\partial\sqrt{\frac{p_a}{\sigma^2}}}\frac{\partial\sqrt{\frac{p_a}{\sigma^2}}}{\partial A_{a_n}}. \qquad (13)
\end{aligned}$$

According to [48], for any $A_{a_n} \geq 0$ and $b_n \in \{0, 1\}$, the following relations hold:

$$\begin{aligned}
&\frac{\partial Q_1(\sqrt{\frac{\mu^2(A_{a_n},b_n)}{\sigma^2}}, \sqrt{\frac{p_a}{\sigma^2}})}{\partial\sqrt{\frac{\mu^2(A_{a_n},b_n)}{\sigma^2}}} \geq 0 \geq \frac{\partial Q_1(\sqrt{\frac{\mu^2(A_{a_n},b_n)}{\sigma^2}}, \sqrt{\frac{p_a}{\sigma^2}})}{\partial\sqrt{\frac{p_a}{\sigma^2}}}, \\
&\frac{\partial Q_1(\sqrt{\frac{\mu^2(A_{a_n},b_n)}{\sigma^2}}, \sqrt{\frac{p_a}{\sigma^2}})}{\partial\sqrt{\frac{\mu^2(A_{a_n},b_n)}{\sigma^2}}}/-\frac{\partial Q_1(\sqrt{\frac{\mu^2(A_{a_n},b_n)}{\sigma^2}}, \sqrt{\frac{p_a}{\sigma^2}})}{\partial\sqrt{\frac{p_a}{\sigma^2}}} \\
&= I_1(\frac{\sqrt{\mu^2(A_{a_n}, b_n)p_a}}{\sigma^2})/I_0(\frac{\sqrt{\mu^2(A_{a_n}, b_n)p_a}}{\sigma^2}) < 1, \qquad (14)
\end{aligned}$$

where I_1 and I_0 are modified Bessel functions of the first kind with orders 1 and 0, respectively, which leads to (12). □

Lemma 3. Inequality (12) holds when $b_n = 1$.

Proof. The left-hand side (LHS) of (12), $I_1(\cdot)/I_0(\cdot)$, increases from 0 to 1 in $[0, \infty)$ with a_n increasing. When $\sqrt{\mu^2(A_{a_n}, b_n)p_a}/\sigma^2$ is small, it is not likely to achieve (12). With a_n increases, $\sqrt{\mu^2(A_{a_n}, b_n)p_a}/\sigma^2$ also increases. Then, (12) can hold because the right-hand side (RHS) of (12) could be much smaller than 1. Given (11), the RHS of (12) is calculated as:

$$\text{RHS} = \frac{\partial\sqrt{p_a}}{\partial A_{a_n}} / \frac{\partial \mu(A_{a_n}, b_n)}{\partial A_{a_n}} = \frac{1}{4\sqrt{p_a}} (\mu(A_{a_n}, 1) \frac{\partial \mu(A_{a_n}, 1)}{\partial A_{a_n}} + \mu(A_{a_n}, 0) \frac{\partial \mu(A_{a_n}, 0)}{\partial A_{a_n}}) / \frac{\partial \mu(A_{a_n}, b_n)}{\partial A_{a_n}}. \quad (15)$$

When $b_n = 1$, as $|s_0| \ll |s_1|$ [32], it can be shown that $\frac{\partial \mu(A_{a_n}, 0)}{\partial A_{a_n}} / \frac{\partial \mu(A_{a_n}, 1)}{\partial A_{a_n}} < \frac{\mu(A_{a_n}, 1)}{\mu(A_{a_n}, 0)}$, then RHS $< \frac{1}{2\sqrt{p_a}} \mu(A_{a_n}, 1) \ll 1$ always holds for (15). We approximate $\frac{I_1(\cdot)}{I_0(\cdot)} \approx 1$ because of the following reason: $\frac{\mu^2(A_{a_n}, 1)}{2\sigma^2}$ and $\frac{p_a}{2\sigma^2}$ are the signal-to-noise ratios, which are usually higher than 10 dB in real-world-scenarios [49]. Thus, $\frac{\sqrt{\mu^2(A_{a_n}, 1)p_a}}{\sigma^2} > 20$. $\frac{I_1(20)}{I_0(20)} = 0.97$ and this ratio is approximately 1 for larger $\frac{\sqrt{\mu^2(A_{a_n}, 1)p_a}}{\sigma^2}$ with the relative error lower than 3% [50]. Thus, (12) holds, which means that all bits can be manipulated when $b_n = 1$. □

Lemma 4. Inequality (12) holds when $b_n = 0$ and channel gain $|h_{AT} h_{TR} s(b_n)|$ is smaller than $|h_{AR}|$.

Proof. When $b_n = 0$, we have

$$\begin{aligned} &\frac{\partial \mu(A_{a_n}, 1)}{\partial A_{a_n}} / \frac{\partial \mu(A_{a_n}, 0)}{\partial A_{a_n}} = \frac{\mu(A_{a_n}, 0)}{\mu(A_{a_n}, 1)} (2(h_{AR} + h_{AT} h_{TR} s_1) \\ &(h_{AR}^* + h_{AT}^* h_{TR}^* s_1) A_{a_n} + ((h_{AR} + h_{AT} h_{TR} s_1) \\ &h_{RT}^* h_{TR}^* + (h_{AR}^* + h_{AT}^* h_{TR}^* s_1) h_{RT} h_{TR}) A s_1) \\ &/ (2(h_{AR} + h_{AT} h_{TR} s_0))(h_{AR}^* + h_{AT}^* h_{TR}^* s_0) A_{a_n} \\ &+ ((h_{AR} + h_{AT} h_{TR} s_0) h_{RT}^* h_{TR}^* \\ &+ (h_{AR}^* + h_{AT}^* h_{TR}^* s_0) h_{RT} h_{TR}) A s_0), \end{aligned} \quad (16)$$

which shows that RHS (15) may exceed 1 if $|h_{AT} h_{TR} s(b_n)|$ is large.Then, the bit can not be manipulated by the attacker when $b_n = 0$. If $|h_{AT} h_{TR} s(b_n)|$ is smaller than h_{AR}, the RHS of (15) is still smaller than 1. We discuss two cases for a_n.

First, we discuss a larger value for A_{a_n}. Suppose $\frac{\sqrt{\mu^2(A_{a_n}, 0)p_a}}{\sigma^2} > 20$, we can use the same proof in Lemma 3. Since $\frac{I_1(\cdot)}{I_0(\cdot)} \approx 1$ and the RHS is less than 1, (12) holds.

Second, we discuss a smaller value for a_n. Since $\frac{I_1(\cdot)}{I_0(\cdot)}$ is monotonically increasing and convex [51], we can relax it within the bound of 20 as $\frac{I_1(\cdot)}{I_0(\cdot)} > \frac{I_1(20)}{20 I_0(20)} \frac{\sqrt{\mu^2(A_{a_n}, 0)p_a}}{\sigma^2} = 0.0485 \frac{\sqrt{\mu^2(A_{a_n}, 0)p_a}}{\sigma^2}$. Based on (16) and (15), (12) holds under the assumption that $|h_{AT} h_{TR} s(b_n)| < |h_{AR}|$. □

Proof of Theorem 1: Now we use Lemmas 2 to 4 to prove the monotonicity of equation (8). First we prove the monotonicity over A_H. Given A_L and n-th bit of y_n, we know p_a is increasing with A_H, and $Q_1(\sqrt{\frac{\mu^2(A_L, b_n)}{\sigma^2}}, \sqrt{\frac{p_a}{\sigma^2}})$ is deceasing with p_a. In (8), $\Pr\{p_n \le p_a | a_n = 0, b_n = i\}$ for $b_n = 0$ and 1 is increasing and approaches 100% when A_H is increasing. Since Lemmas 2 to 4 prove that $Q_1(\sqrt{\frac{\mu^2(A_{a_n}, b_n)}{\sigma^2}}, \sqrt{\frac{p_a}{\sigma^2}})$ is increasing with A_{a_n}, $\Pr\{p_n > p_a | a_n = 1, b_n = i\}$ is increasing with A_H. Therefore, (8) is increasing with A_H. Second, given A_H and n-th bit of y_n, we can adopt a similar way to prove that (8) is decreasing with A_L. □

Appendix C
Proof of Theorem 2

We use the following equation to analyze the effect of A:

$$\frac{Q_1(\sqrt{\frac{\mu^2(A_{a_n}, b_n)}{\sigma^2}}, \sqrt{\frac{p_a}{\sigma^2}})}{\partial A} = \frac{\partial Q_1(\sqrt{\frac{\mu^2(A_{a_n}, b_n)}{\sigma^2}}, \sqrt{\frac{p_a}{\sigma^2}})}{\partial \sqrt{\frac{\mu^2(A_{a_n}, b_n)}{\sigma^2}}} \cdot \frac{\partial \sqrt{\frac{\mu^2(A_{a_n}, b_n)}{\sigma^2}}}{\partial A} + \frac{\partial Q_1(\sqrt{\frac{\mu^2(A_{a_n}, b_n)}{\sigma^2}}, \sqrt{\frac{p_a}{\sigma^2}})}{\partial \sqrt{\frac{p_a}{\sigma^2}}} \frac{\partial \sqrt{\frac{p_a}{\sigma^2}}}{\partial A}. \quad (17)$$

When $b_n = 0$ holds, we have $\frac{\partial \sqrt{\mu^2(A_{a_n}, 0)}}{\partial A} = 0$ because $|s_0| \approx 0$ in (11) [32]. The inequality (14) shows that (17) ≤ 0. Since Marcum Q function is decreasing with A, Lemma 1 shows that $\Pr\{p_n \le p_a | a_n = 0, b_n = 0\}$ and $\Pr\{p_n > p_a | a_n = 1, b_n = 0\}$ are monotonically increasing and decreasing with A, respectively.

When $b_n = 1$ holds, we approximate $\frac{I_1(\cdot)}{I_0(\cdot)} \approx 1$ because $\frac{\sqrt{\mu^2(A_{a_n}, 1)p_a}}{\sigma^2} > 20$, as we have shown in Lemma 3. Given $\frac{I_1(\cdot)}{I_0(\cdot)} = 1$ and (11), the inequality (14) shows (17) ≥ 0 if the following inequality holds:

$$4\sqrt{p_a} \ge (\mu(A_H, 1) \frac{\partial \mu(A_H, 1)}{\partial A} + \mu(0, 1) \frac{\partial \mu(0, 1)}{\partial A}) / \frac{\partial \mu(A_{a_n}, 1)}{\partial A}.$$

The amplitude $\mu(A_{a_n}, 1)$ can be separated into the components from the attacker and the reader. Thus, $\frac{\partial \mu(A_{a_n}, 1)}{\partial A}$ is basically not affected by A_{a_n}. Then, $\frac{\partial \mu(A_H, 1)}{\partial A} \approx \frac{\partial \mu(0, 1)}{\partial A}$, and the above inequality is proved. Therefore, we obtain that (17) ≥ 0. $\Pr\{p_n < p_a | a_n = 0, b_n = 1\}$ is monotonically decreasing with A, and $\Pr\{p_n > p_a | a_n = 1, b_n = 1\}$ is monotonically increasing with A.

In a real-world scenario, the reader always transmits with enough power for A to guarantee that the communication has a low bit error rate. Hence, both $\Pr\{p_n < p_a | a_n = 0, b_n = 0\}$ and $\Pr\{p_n > p_a | a_n = 1, b_n = 1\}$ shall be nearly 100%. They remain to be 100% even if A is increasing. However, both $\Pr\{p_n > p_a | a_n = 1, b_n = 0\}$ and $\Pr\{p_n < p_a | a_n = 0, b_n = 1\}$ decrease with A. Therefore, P_{att} is deceasing with A.

Now consider the effect of $s(b_n)$. The probability $\Pr\{p_n > p_a | a_n = 1, b_n = 0\}$ is not affected because $|s_0| \approx 0$ [32]. Since a larger $|s_1|$ is equivalent to increasing A for $\Pr\{p_n < p_a | a_n = 0, b_n = 1\}$, we can finally conclude that P_{att} is deceasing with $|s_1|$. □

REFERENCES

[1] C. Boyer and S. Roy, "Backscatter communication and RFID: Coding, energy, and MIMO analysis." *IEEE Trans. Commun.*, 2013.

[2] J. C. Chen, C.-H. Cheng, and P. B. Huang, "Supply chain management with lean production and rfid application: A case study," *Expert Systems with applications*, vol. 40, no. 9, pp. 3389–3397, 2013.

[3] C. Wang, L. Xie, Y. Lin, W. Wang, Y. Chen, Y. Bu, K. Zhang, and S. Lu, "Thru-the-wall eavesdropping on loudspeakers via rfid by capturing sub-mm level vibration," *ACM IMWUT*, vol. 5, no. 4, pp. 1–25, 2021.

[4] Y. Shoukry, P. Martin, Y. Yona, S. Diggavi, and M. Srivastava, "PyCRA: Physical challenge-response authentication for active sensors under spoofing attacks," in *ACM CCS*, 2015, pp. 1004–1015.

[5] Y. Sun, S. Kumar, S. He, J. Chen, and Z. Shi, "You foot the bill! attacking NFC with passive relays," *IEEE Internet of Things Journal*, vol. 8, no. 2, pp. 1197–1210, 2020.

[6] N. Van Huynh, D. N. Nguyen, D. T. Hoang, and E. Dutkiewicz, "Jam me if you can: defeating jammer with deep dueling neural network architecture and ambient backscattering augmented communications," *IEEE J. Selected Areas in Communications*, vol. 37, 2019.

[7] H. Pirayesh and H. Zeng, "Jamming attacks and anti-jamming strategies in wireless networks: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, 2022.

[8] J. H. Sarker and A. M. Nahhas, "Mobile RFID system in the presence of denial-of-service attacking signals," *IEEE Transactions on Automation Science and Engineering*, vol. 14, no. 2, pp. 955–967, 2016.

[9] K. Bu, M. Weng, Y. Zheng, B. Xiao, and X. Liu, "You can clone but you cannot hide: A survey of clone prevention and detection for RFID," *IEEE Communications Surveys & Tutorials*, vol. 19, 2017.

[10] "EPC Radio-Frequency Identity Protocols Generation-2 UHF RFID," 2015.

[11] V. Liu, A. Parks, V. Talla, S. Gollakota, D. Wetherall, and J. R. Smith, "Ambient backscatter: Wireless communication out of thin air," *ACM SIGCOMM Computer Communication Review*, vol. 43, 2013.

[12] J. D. Griffin and G. D. Durgin, "Gains for RF tags using multiple antennas," *IEEE Transactions on Antennas and Propagation*, vol. 56, no. 2, pp. 563–570, 2008.

[13] A. Boaventura, J. Santos, A. Oliveira, and N. B. Carvalho, "Perfect isolation: Dealing with self-jamming in passive RFID systems," *IEEE Microwave Magazine*, vol. 17, no. 11, pp. 20–39, 2016.

[14] F. Pfeiffer, K. Finkenzeller, and E. Biebl, "Theoretical limits of ISO/IEC 14443 type A RFID eavesdropping attacks," in *Smart SysTech*, 2012.

[15] G. P. Hancke, "Practical eavesdropping and skimming attacks on high-frequency RFID tokens," *Journal of Computer Security*, vol. 19, no. 2, pp. 259–288, 2011.

[16] H. Hassanieh, J. Wang, D. Katabi, and T. Kohno, "Securing RFIDs by randomizing the modulation and channel," in *USENIX NSDI*, 2015.

[17] D. Dobrykh, D. Filonov, A. Slobozhanyuk, and P. Ginzburg, "Hardware RFID security for preventing far-field attacks," *IEEE Trans. on Antennas and Propagation*, vol. 70, no. 3, pp. 2199–2204, 2021.

[18] A. Rullo, C. Felicetti, M. Vatalaro, R. De Rose, M. Lanuzza, F. Crupi, and D. Sacca, "PUF-based authentication-oriented architecture for identification tags," *IEEE Trans. Dependable and Secure Computing*, 2024.

[19] Y. Yang, J. Cao, Z. An, Y. Wang, P. Hu, and G. Zhang, "NFChain: A practical fingerprinting scheme for NFC tag authentication," in *IEEE INFOCOM*, 2023, pp. 1–10.

[20] A.-I. Radu, T. Chothia, C. J. Newton, I. Boureanu, and L. Chen, "Practical EMV relay protection," in *IEEE S&P*, 2022, pp. 1737–1756.

[21] J. Li, A. Li, D. Han, Y. Zhang, T. Li, and Y. Zhang, "RCID: Fingerprinting passive RFID tags via wideband backscatter," in *INFOCOM*, 2022, pp. 700–709.

[22] K. Joo, W. Choi, and D. H. Lee, "Experimental analyses of RF fingerprint technique for securing keyless entry system in modern cars," in *USENIX NDSS*, 2020.

[23] S. Akter, S. Chellappan, T. Chakraborty, T. A. Khan, A. Rahman, and A. A. Al Islam, "Man-in-the-middle attack on contactless payment over NFC communications: design, implementation, experiments and detection," *IEEE Trans. Dependable and Secure Computing*, vol. 18, no. 6, pp. 3012–3023, 2020.

[24] J. Li, C. Wang, A. Li, D. Han, Y. Zhang, J. Zuo, R. Zhang, L. Xie, and Y. Zhang, "RF-Rhythm: Secure and usable two-factor RFID authentication," in *IEEE INFOCOM*, 2020, pp. 2194–2203.

[25] M. Chen, S. Chen, and Y. Fang, "Lightweight anonymous authentication protocols for RFID systems," *IEEE/ACM Trans. Networking*, 2017.

[26] L. Avanco, A. E. Guelfi, E. Pontes, A. Silva, S. T. Kofuji, and F. Zhou, "An effective intrusion detection approach for jamming attacks on RFID systems," in *EURFID*, 2015, pp. 73–80.

[27] G. K. Balachandran and R. E. Barnett, "A passive uhf rfid demodulator with rf overvoltage protection and automatic weighted threshold adjustment," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 57, no. 9, pp. 2291–2300, 2010.

[28] A. Nemirovski, "Interior point polynomial time methods in convex programming," *Lecture notes*, vol. 42, no. 16, pp. 3215–3224, 2004.

[29] P. E. Gill and E. Wong, "Sequential quadratic programming methods," in *Mixed integer nonlinear programming*. Springer, 2011, pp. 147–224.

[30] M. S. Trotter, J. D. Griffin, and G. D. Durgin, "Power-optimized waveforms for improving the range and reliability of RFID systems," in *IEEE RFID*, 2009, pp. 80–87.

[31] A. Ghahremani, V. D. Rezaei, and M. S. Bakhtiar, "A UHF-RFID transceiver with a blocker-canceller feedback and +30 dBm output power," *IEEE Trans. Circuits and Systems I*, vol. 60, 2013.

[32] A. Bletsas, A. G. Dimitriou, and J. N. Sahalos, "Improving backscatter radio tag efficiency," *IEEE Transactions on Microwave Theory and Techniques*, vol. 58, no. 6, pp. 1502–1509, 2010.

[33] J. Garcia-Alfaro, J. Herrera-Joancomartí, and J. Melià-Seguí, "Security and privacy concerns about the RFID layer of EPC Gen2 networks," *Advanced Research in Data Privacy*, pp. 303–324, 2015.

[34] R. Chen, S. Yang, R. V. Penty, and M. Crisp, "UHF RFID reader sensitivity requirements due to poor tag matching," in *RFID-TA*, 2022.

[35] I. Kwon, Y. Eo, H. Bang, K. Choi, S. Jeon, S. Jung, D. Lee, and H. Lee, "A single-chip CMOS transceiver for UHF mobile RFID reader," *IEEE J. Solid-state Circuits*, vol. 43, no. 3, pp. 729–738, 2008.

[36] D. Cousineau, S. Brown, and A. Heathcote, "Fitting distributions using maximum likelihood: Methods and packages," *Behavior Research Methods, Instruments, & Computers*, vol. 36, no. 4, pp. 742–756, 2004.

[37] F. Huo, P. Mitran, and G. Gong, "Analysis and validation of active eavesdropping attacks in passive FHSS RFID systems," *IEEE Trans. Information Forensics and Security*, vol. 11, no. 7, pp. 1528–1541, 2016.

[38] B.-Q. Zhao, H.-M. Wang, and J.-C. Jiang, "Safeguarding backscatter RFID communication against proactive eavesdropping," in *IEEE ICC*, 2020.

[39] R. Jin and K. Zeng, "Physical layer key agreement under signal injection attacks," in *IEEE CNS*, 2015, pp. 254–262.

[40] J. Lee, W. Kim, S.-J. Lee, D. Jo, J. Ryu, T. Kwon, and Y. Choi, "An experimental study on the capture effect in 802.11 a networks," in *ACM WiNTECH*, 2007, pp. 19–26.

[41] W. Kim, S. Kim, and H. Lim, "Malicious data frame injection attack without seizing association in IEEE 802.11 wireless LANs," *IEEE Access*, vol. 9, pp. 16 649–16 660, 2021.

[42] K. C. Hewage, S. Raza, and T. Voigt, "Protecting glossy-based wireless networks from packet injection attacks," in *IEEE MASS*, 2017.

[43] D. F. Kune, J. Backes, S. S. Clark, D. Kramer, M. Reynolds, K. Fu, Y. Kim, and W. Xu, "Ghost talk: Mitigating EMI signal injection attacks against analog sensors," in *IEEE S&P*, 2013.

[44] S. Kaji, M. Kinugawa, D. Fujimoto, and Y.-i. Hayashi, "Data injection attack against electronic devices with locally weakened immunity using a hardware trojan," *IEEE Transactions on Electromagnetic Compatibility*, vol. 61, no. 4, pp. 1115–1121, 2018.

[45] P. Gope, J. Lee, and T. Q. Quek, "Lightweight and practical anonymous authentication protocol for RFID systems using physically unclonable functions," *IEEE Trans. Information Forensics and Security*, vol. 13, no. 11, pp. 2831–2843, 2018.

[46] M. Kaur, R. Rashidzadeh, and R. Muscedere, "Reliability of physical unclonable function under temperature and supply voltage variations," in *IEEE MWSCAS*, 2018, pp. 1008–1011.

[47] D. J. Maširević, "On new formulas for the cumulative distribution function of the noncentral chi-square distribution," *Mediterranean Journal of Mathematics*, vol. 2, no. 14, pp. 1–13, 2017.

[48] W. Pratt, "Partial differentials of Marcum's Q function," *Proceedings of the IEEE*, vol. 56, no. 7, pp. 1220–1221, 1968.

[49] Z. Blažević, P. Šolić, M. Škiljo, M. Stella, Č. Stefanović, P. Popovski, G. Fr *et al.*, "Signal-to-noise ratio measurements and statistical characterization in Gen2 RFID," in *SpliTech*, 2017.

[50] B. B. Geelen, "Accurate solution for the modified bessel function of the first kind," *Advances in Engineering Software*, 1995.

[51] Z.-H. Yang, S. Zheng *et al.*, "Monotonicity and convexity of the ratios of the first kind modified bessel functions and applications," 2017.