Message Integrity Protection over Wireless Channel by Countering Signal Cancellation: Theory and Practice

Yantian Hou*, Ming Li*, Ruchir Chauhan*, Ryan M. Gerdes*, Kai Zeng†
houyantian@gmail.com, ming.li@usu.edu, ryan.gerdes@usu.edu, kzeng2@gmu.edu

*Utah State University, Logan, UT 84322
†George Mason University, Fairfax, Virginia 22030

ABSTRACT

Physical layer message integrity protection and authentication by countering signal-cancellation has been shown as a promising alternative to traditional pure cryptographic message authentication protocols, due to the non-necessity of neither pre-shared secrets nor secure channels. However, the security of such an approach remained an open problem due to the lack of systematic security modeling and quantitative analysis. In this paper, we first establish a novel correlated jamming framework to study the optimal signal-cancellation attacker's behavior and utility using game-theory, which precisely captures the attacker's knowledge using its correlated channel estimates in various channel environments. Besides, we design a practical physical layer message integrity protection protocol based on ON/OFF keying and Manchester coding, which provides quantitative security guarantees in the real-world. Such a guarantee is achieved by bounding the attacker's knowledge about the future channel via proactively measuring channel statistics (mimic the attacker), so as to derive a lower-bound to the defender's signal-detection probability under optimal correlated jamming attacks. We conduct extensive experiments and simulations to show the security and performance of the proposed scheme. We believe our novel threat modeling and quantitative security analysis methodology can benefit a wide range of physical layer security problems.

1. INTRODUCTION

Message integrity protection and authentication are two fundamental security services in the Internet-of-Things, given the exponential growth of wireless sensors and mobile devices [1]. Traditionally, such services have assumed the existence of pre-shared secret keys or secure channels. However, in many scenarios these premises may not be satisfied, e.g. when initial security associations need to be established among two or more constrained wireless devices. Generally, secret keys need to be distributed either via an off-line secure channel or using a public key infrastructure (PKI). But key pre-distribution may not be always feasible due to the lack of hardware interfaces and the absence of a global PKI. Some existing research proposed using out-of-band (OOB) secure auxiliary channels to build message authentication protocols without pre-shared keys [23, 3, 8, 4, 16, 17]. However, an OOB channel would require special hardware and non-trivial human interaction, while its security has been revisited [18]. In addition, whenever keys are stolen or compromised, re-keying involves significant human effort as well.

Ideally, we want to provide message integrity protection and authentication without relying on pre-shared keys or secure channels. That is, to establish the veracity of a message and its source using only wireless in-band transmissions. Čapkun et. al. [5] showed that it is possible to construct such an in-band integrity protection primitive, by preventing signal-cancellation in the wireless channel and combine it with unidirectional error detection codes. Later a few works have followed up in this direction. However, an important question remained unanswered about its security. Since the security depends on the infeasibility of signal-cancellation, work should be done to evaluate to which extent this is true, i.e. there lacks quantitative analysis of its security. Recently, Pöpper et. al. [19] demonstrated a practical relaying attack that can fully cancel the source's signal in some cases. In fact, the probability of adversarial signal-cancellation heavily depends on the wireless channel conditions. But again, so far only qualitative results are available, while no quantitative security guarantee can been provided by any of the previous designs.

Unfortunately, in general such a security guarantee is quite challenging to establish for any wireless physical-layer security mechanism. This is primarily due to a lack of systematic modeling of attacker's behaviors in this area, unlike well-known methodologies for cryptography. To do so one needs to connect knowledge from information-theory, wireless communications and practical aspects of security. In addition, the fact that wireless is an open medium makes it easy for common attacks to be launched, thus threat modeling needs to be comprehensive. A smart and strategic attacker who is knowledgeable about the wireless channel environment must be assumed. In fact, we will show that the channel state information (CSI) can be viewed as a partial secret of the legitimate communicating pairs, and we can systematically bound attacker's knowledge about this information in reality. Moreover, the attacker can possess advanced hardware and processing capabilities such as multi-antennas and directional antennas. Many other existing

ASIA CCS'15, April 14–17, 2015, Singapore..

http://dx.doi.org/10.1145/2714576.2714617.

physical layer security schemes failed to provide any security [24, 20] when the attacker is powerful as such.

In this paper, we first present a systematic threat modeling for signal cancellation attacks. In our model, the attacker exploits the intrinsic channel correlation existing in various domain(s) (e.g. temporal, spatial and frequency domains) to estimate the CSI of the legitimate communication pair with help of advanced hardware (such as directional antenna or multi-antenna). We then develop a correlated jamming and defense framework based on previous pure theoretical study in correlated jamming [12]. Our framework captures the attacker's knowledge about the legitimate communication pair's CSI using a correlation coefficient. We consider both indoor and outdoor environments in our system model. The correlated jamming and defense process is modeled as a zero-sum game, in which the attacker aims at minimizing receiver's signal detection probability, while the defender seeks to maximize this probability. Under this framework, we theoretically analyze the optimal attack/defense strategies and detection probability under correlated jamming, given any correlation coefficient.

Based on the theoretical results, we propose a practical physical-layer message integrity protection scheme that can achieve provable security guarantees. The protocol builds upon and can be regarded as an add-on to previous works using ON/OFF keying and Manchester coding. Our idea is to let the legitimate pair proactively make a worst case estimation of the attacker's knowledge of their CSI (mimic the attacker), and use that to derive a lower-bound of the signal detection probability under optimal cancellation attacks. A key challenge is how to bound the attacker's knowledge given that it can measure correlated channels at an arbitrary location, time and frequency. In our protocol, we deal with one of the strongest CSI estimation attack in temporal, frequency and spatial domains. We assume the attacker can obtain accurate CSI measurements (full knowledge) of the legitimate pair in the past, based on which it predicts the CSI in the future for optimal correlated jamming. Interestingly, by extracting the intrinsic randomness and unpredictability of the wireless channel over time, we can achieve an arbitrary goal of minimum signal detection probability by tuning the number of symbols in each ON slot.

Finally, we validate our theoretical analysis results using simulation, which shows that signal cancellation is indeed most effective when the attacker adopts the optimal jamming strategy. We also show the impact of attacker's correlation coefficient and the detection threshold to the system security level and achievable link throughput. In addition, we carry out real-world experiments and implement our scheme on USRP GNU radio devices. We found that by actively randomizing the physical channel using external disturbances, we can turn an indoor static channel into a dynamic channel which can defend against signal cancellation attack more effectively.

This paper is organized as follows. Sec. 2 presents the background and motivation to our threat modeling, followed by the system and attack models in Sec. 3. In Sec. 4, we present the game-theoretical framework to analyze the attacker/defender's strategies and their optimal utilities. Sec. 5 gives our physical layer message integrity protection scheme. Sec. 6 contains the simulation evaluation results. In Sec. 7, we present the implementation and experimental study, and discussions. Sec. 8 gives an overview of related work. Sec. 9 concludes the paper.

2. BACKGROUND AND MOTIVATION

2.1 In-band Message Integrity Protection and Authentication

A few previous works proposed in-band message integrity protection and authentication schemes without relying on pre-shared secret keys [5, 7, 10]. The common underlying idea is to combine ON/OFF keying with unidirectional error detection code. By using this coding method, bit 1 is encoded into ON_OFF slots and bit 0 is encoded into OFF_ON slots. To provide message integrity protection, a data packet is sent first using normal modulation, followed by a cryptographic hash calculated over the message which is encoded using the ON/OFF keying approach (idea is also shown in Fig. 4). The security of this approach is based on the infeasibility of signal cancellation in the wireless channel, which ensures that only unidirectional bit modification is feasible, i.e. attacker could only change OFF slot into ON slot but not in the opposite direction. Therefore any tampering with the original message will be detected (w.h.p.), and the authentication property can be derived based on integrity protection and the presence of the participating devices (authentication through presence [5]). Anti-signal-cancellation is achieved by setting the signal to be random in each ON slot, and based on the assumption that attacker could not extract any knowledge of the source signal and the channel thus it cannot cancel the signal.

However, this assumption is too strong because practical signal cancellation attack has been demonstrated [19], which uses a pair of directional antennas to relay the source signal such that the phase differs by $k\pi$ from the direct signal at the receiver. It is also referred to as correlated jamming [14]. This type of attack aims at completely cancelling out the received signal, by assuming the attacker knows the transmitted source signal x (or a correlated version of it). This is achieved in [19] by using directional antennas, such that the attacker obtains x from A in real time, and it has almost complete knowledge of the direct channel from A to B (which is a stable indoor channel). Through correlated jamming, the attacker has the potential to modify/cancel any signal in wireless channel, and the message integrity will not be protected. Thus, it is essential to investigate the possibility of signal cancellation in the real-world, so as to provide quantitative security guarantees.

2.2 Quantifying Adversary's Knowledge in Correlated Jamming

Previous results on the signal cancellation attack are qualitative [19], which show that a static environment leads to higher chance of cancellation. While some theoretical results are known in correlated jamming, the legitimate pair's CSI h is assumed to be either perfectly known by the attacker, or not known but only statistics are available. However, in practice this is often not the case. Instead, the attacker's knowledge about the channel can lie between these two extremes. And how can we quantify the attacker's capability remained as an open problem. Intuitively, the more accurate the attacker could estimate the legitimate pair's channel h, the more effective it could launch the correlated jamming attack. Therefore, we can use the correlation coefficient $r_{h\bar{g}}$ to

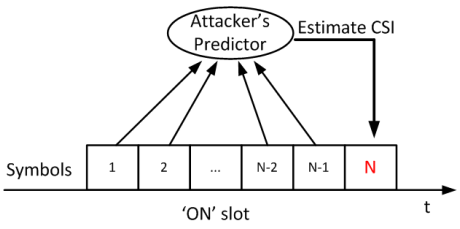


Figure 1: Example of CSI prediction attack.

quantize the correlated jammer's capability, where g denotes the attacker's estimation of h.

Generally, the attacker can exploit correlations in three domains to obtain knowledge of legitimate h: spatial domain, temporal domain, and frequency domain. In the spatial dimension, previous works [13, 11] demonstrated high correlations between channels where the receivers (or transmitters) are close to each other (typically within half wavelength). He et. al. [9] even showed that, the attacker can obtain a very accurate estimation of the legitimate pair's channel by placing multiple eavesdroppers around the legitimate receiver. The idea is to let all the eavesdroppers measure the channel simultaneously, and then combine them into a linear minimum mean square error (LMMSE) estimator. The estimation error can decrease to zero with increased number of eavesdroppers in some cases.

It has also been shown by previous works that channel is self-correlated in time domain. The correlated time scale is typically measured by the channel coherence time, which is usually several ms in dynamic environment and hundreds of ms in static environments.

Similarly, channel correlation exists in the frequency domain. The attacker can also exploit CSI measurements made in adjacent channels to derive a better estimate of the CSI in the frequency used by the legitimate pair.

In a word, the attacker could leverage channel correlation in any of the three domains and combine them. Such correlation should be considered in the threat model and design of any anti-signal-cancellation based integrity protection scheme.

In this paper, we first derive a theoretic result showing that the attacker's successful cancellation probability increases with its channel correlation with the legitimate one. However, in reality it is difficult (if not impossible) to know the attacker's capabilities in advance (e.g., location, device type, number), and it seems hopeless to upper-bound the attacker's knowledge about the legitimate channel. Fortunately, since correlated jamming is an active attack, it is only effective when the jamming signal is in the same frequency. Also, it must be timely – attacker's channel estimation needs to be done in real-time without any delay, otherwise the jamming opportunity will be missed. Therefore, even though the attacker can accurately measure the historical legitimate CSI via spatial and frequency domain correlation, it still needs to predict the CSI in the present (and future) in order to generate its correlated jamming signal (illustrated in Fig. 1). Any approach to obtain the current channel knowledge through measurements takes time, and after that the optimal jamming opportunity is already missed. That means, we can exploit the intrinsic time-domain unpredictability of the legitimate channel to prevent it from knowing the future CSI. To do so, in our scheme the legitimate TX/RX quantify the CSI's self-correlation in the time domain and use that to bound the knowledge of attacker.

3. MODEL AND ASSUMPTIONS

3.1 System Model

In our model, Alice communicates with Bob through a wireless channel. There are two types of transmission modes. In the first one (normal mode) a message is transmitted using standard modulation and data rates, such like 802.11 OFDM. The second one is called the ON/OFF keying mode, where information bits (like the hash of a normal message) are all encoded using ON/OFF keying combined with unidirectional error detection codes (e.g., Manchester coding). In each ON slot, a normal packet with random content is transmitted, while in OFF slots Alice remains silent. For this mode Bob uses energy detection to decode the received signal. Periodically (e.g., per symbol interval), Bob obtains a received signal strength (RSS) and compares it with a threshold (α). If the RSS is larger than α for N_s samples then an ON slot is detected. We assume each transmitted signal $x \in \mathbb{C}$ is arbitrary. The channel state $h \in \mathbb{C}$ between Alice and Bob is modeled under Rayleigh fading with additive white Gaussian noise n in outdoor environment, and Rician model in indoor environments.

3.2 Threat Model

The attacker's general goal is to break integrity protection, i.e., modify the message without being detected. For the normal mode, we assume the adversary can arbitrarily eavesdrop, inject, modify, replay, and block the message (standard Dolev-Yao model). For the ON/OFF keying mode, we assume a correlated jammer C who knows the exact transmitted signal x, and C's goal is to cancel out the signal received at Bob. To learn x in real-time, C can place a directional antenna closely to the legitimate transmitter A. To create and deliver a correlated jamming signal at B, C will utilize x and her "knowledge" about the CSI h from A to B. Essentially, C possesses a correlated version of h denoted as g (correlation coefficient denoted as $r \in [0, 1]$), which could be estimated from measurements (as shown in Fig. 2).

There are two types of attackers our model depending on their capabilities. We always assume the attacker cannot replace A or B, nor simply block the communication using a Faraday Cage. We do not restrict the number and type of devices the attacker may have. It can either generate its own signals or relay and process the signals from A to B.

Type I: This type of attacker relies on statistical or background information to estimate h, but makes no effort to obtain the accurate measurement of h. For example, channel propagation models can be used to derive the stable (Line-of-Sight/LoS) part of the CSI based on the distance, and large-scale fading/shadowing effects can also be predicted. However the attacker cannot derive a correlated version for the dynamic/small-scale part. This model is adopted by [19] under a stable indoor scenario, where A-B, A-C and C-B channels are all assumed constant.

Type II: This type of attacker can obtain up-to-date and correlated estimation about A to B's CSI using information from any of the three domains mentioned in previous section. For example, it could place multiple receivers close to B, and measure the channel for each transmitted symbol continu-

ously. In the worst case, it obtains the exact A-B channel for every symbol in the past and uses them to predict the future CSI.

In a word, the Type I attacker could only get the knowledge about the stable part of CSI, while the Type II attacker could also get partial knowledge of the dynamic part. We note that, the type II attack model is stronger and more general than previous works [19, 7, 5, 10], as the attacker can do real-time signal processing to generate a correlated jamming signal based on source x and the correlated CSI.

4. OPTIMAL STRATEGIES FOR SIGNAL CANCELLATION ATTACK AND DEFENSE

4.1 Game Theoretic Framework

In this section, we theoretically analyze the signal cancellation attack for one symbol in an ON slot. We model the cancellation and anti-cancellation process as a game. The attacker's goal is to transmit a signal correlated with x such that the detection probability P_d of the combined received signal is minimized at B. Therefore we define the attacker's utility function as $U_a = -P_d$. The legitimate pair's strategy is to maximize the signal detection probability and their utility function is $U_l = P_d$. Obviously, this is a zero-sum game.

For the strategy space, let the correlated jammer generate a linear signal [12, 22, 21] that is $agx + v$, in which a is a variable controlled by jammer, g is attacker's knowledge about h (an estimated or correlated version), and v is additive white Gaussian noise with variance σ_v. Thus the overall received signal will be:

$$y = (h + ag)x + n + v \tag{1}$$

W.l.o.g., we use the Rician model for A-B channel (Rayleigh model is a special case). In this model, the channel h is composed of two parts: one is the deterministic LoS component h', the other is the random Gaussian distributed fading component h''. Thus the channel is denoted by $h = h' + h''$.

We assume the attacker could estimate the LoS part precisely. The estimation g is further divided into two parts $g = g' + g''$. The attacker's strategy consists of a tuple $\mathbf{a} = [a', a'', \sigma_v]$ corresponding to each component. Its transmit power can be easily derived based on $\mathbf{a}$, g, the power of x and v, and here we assume it is not bounded. To include the attacker's power in its strategy under power constraint will be our future work. On the other hand, the defender's strategy consists of A's transmit power.

Under this model, the received signal can be represented by:

$$y = (h' + a'g')x + (h'' + a''g'')x + n + v \tag{2}$$

4.2 Optimal Attack Strategy

Because the LoS and NLoS signal components are independent from each other, the attacker can cancel the two components separately.

4.2.1 LoS Component Strategy

As the LoS channel component h' is assumed to be precisely known, we have $g' = h'$. Therefore we can easily derive the optimal attack strategy for the LoS component:

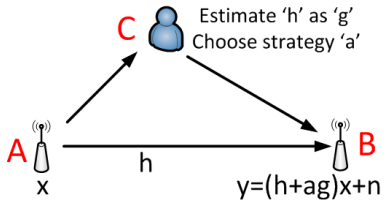


Figure 2: The system model

THEOREM 4.1. *The optimal LoS component cancellation strategy is:*

$$a' = -1 \tag{3}$$

The above indicates that, the attacker will reverse the LoS signal's phase to completely cancel it out at the receiver side.

4.2.2 NLoS Component Strategy

Given that the LoS component can be completely cancelled, we analyze the optimal attack strategy for NLoS part. We start from deriving the distribution of received power of this component under correlated jamming.

Type I attacker. For the type I attacker, the estimated channel g'' is independent from h''. Since the source signal x is randomly distributed, the power detected by receiver is $P_y = \sigma_x^2|h''|^2 + |a''g''|^2\sigma_x^2 + \sigma_n^2 + \sigma_v^2$, where $\sigma_x, \sigma_n, \sigma_v$ are the variance (power) of the source signal and noises, respectively. We can see that the variable $|h''|^2$ follows gamma distribution $\Gamma(1, 2\sigma^2)$ as $|h''|$ is Rayleigh distributed, where $\sigma^2 = \frac{1}{2}E[h''\overline{h''}]$.

THEOREM 4.2. *Given detection threshold α, the probability that a symbol within an ON slot be detected under type I attacker's correlated jamming is:*

$$P_d(\sigma^2) = e^{-\frac{\alpha - \sigma_n^2 - \sigma_v^2 - |a''g''|^2\sigma_x^2}{2\sigma_x^2\sigma^2}} \tag{4}$$

From the detection probability, we derive the optimal NLoS attack strategy:

THEOREM 4.3. *The NLoS part optimal strategy for type I attacker is:*

$$(a'' = 0, \sigma_v^2 = 0) \tag{5}$$

Due to space limit, the proof is omitted. As shown in Theorem 4.3, the best strategy for type I attacker is to not jam the NLoS part. This is because, the estimated channel g'' is not correlated with the real channel h''. Thus any non-zero signal will only add more noise at the receiver B, which increases the detection probability instead.

Type II attacker. According to the type II attacker model, the estimated channel g'' is correlated with h''. Thus in the power expression $P_y = \sigma_x^2(h'' + a''g'')^2 + \sigma_n^2 + \sigma_v^2$, the component $|h'' + a''g''|^2$ follows Gamma distribution $\Gamma(1, 2\sigma^2)$ since $(h'' + a''g'')$ is a CSCG random variable, where $\sigma^2 = \frac{1}{2}E[(h'' + a''g'')\overline{(h'' + a''g'')}]$. In addition, the part $\sigma_x^2|h'' + a''g''|^2$ also follows Gamma distribution $\Gamma(1, 2\sigma_x^2\sigma^2)$, because $\sigma_x(h'' + a''g'')$ is a CSCG random variable.

THEOREM 4.4. *Given detection threshold α, the probability that a symbol within an ON slot be detected under type II attacker's correlated jamming is:*

$$P_d(\sigma^2) = e^{-\frac{\alpha - \sigma_n^2 - \sigma_v^2}{2\sigma_x^2\sigma^2}} \tag{6}$$

According to equation 6, the detection probability is related to the estimated channel g''. Thus we will first analyze the effect of parameter σ^2 on the detection probability.

THEOREM 4.5. *The detection probability $P_d(\sigma^2)$ is a non-decreasing function with respect to σ^2.*

The proof is in Appendix. According to Theorem 4.4, the minimum detection probability is achieved when σ^2 is infinitely close to 0:

$$\lim_{\sigma^2 \to 0} P_d(\sigma^2) = 0 \tag{7}$$

The above result shows, the perfect attack precisely estimates channel h'' such that the jamming signal is exactly the opposite of the received signal from A to B, thus the original signal will be completely attenuated. However, this is an extreme case in which perfect CSI is assumed known by attacker. Some previous works are based on this extreme case [12, 14], under which the link from A to B has zero capacity. In this paper we consider a more realistic general case in which the real CSI h'' and the attacker's estimated CSI g'' is correlated with arbitrary $r_{h''\overline{g''}}$.

THEOREM 4.6. *The NLoS part's optimal correlated jamming strategy is:*

$$(a'' = -\frac{E[h''\bar{g''}]}{\sigma_g^2}, \sigma_v^2 = 0) \tag{8}$$

The proof is in Appendix. Given the optimal strategy of attacker, we can use Eq. (14) in Appendix to derive the minimum variance $\sigma^2_{min} = \frac{1}{2}\sigma_h^2(1 - |r_{h\bar{g}}|^2)$, where $|r_{h''\overline{g''}}|$ is the correlation coefficient. Substitute it into Eq. (6), we get the minimum detection probability:

$$P_d(\sigma^2_{min}) = e^{-\frac{\alpha - \sigma_n^2 - \sigma_v^2}{\sigma_x^2 \sigma_h^2 (1 - |r_{h''\overline{g''}}|^2)}} \tag{9}$$

From the analysis above, we can see that the minimum detection probability decreases with the increase of attacker's correlation coefficient $|r_{h''\overline{g''}}|$. Also, previous works that either assumed a 0 or 1 correlation coefficient are two extreme cases of our result.

4.3 Optimal Defender Strategy

Next we analyze the legitimate pair's optimal strategy. From the above, the type I attacker is only a special case of type II attacker when $r_{h''\overline{g''}} = 0$. In our model, the signal x is independent with h''. The only transmitter parameter that has influence on the final detection probability is the power σ_x^2. From Eq. (9), we can easily see that the detection probability increases when σ_x^2 increases. In reality, the transmitter's power is limited, thus it indicates that the transmitter should always choose its largest power level to defend against correlated jamming attacks.

5. INTEGRITY PROTECTION SCHEME

5.1 Design Overview

In this section, we present our integrity protection protocol. We first devise an approach to upper-bound the attacker's knowledge (correlation) under type II attack, specifically the CSI prediction attack. The idea is to extract the A-B CSI by the legitimate receiver B through channel probing, and mimic the attacker's strategy to quantify the intrinsic time-domain correlation (or unpredictability) in the channel itself, assuming perfect estimation of historical CSI by the attacker. Based on this correlation, we calculate the maximum signal cancellation probability (or minimum signal detection probability) for each symbol under correlated jamming attacks using our theoretic framework. Given a targeted security requirement (signal cancellation probability for each ON slot), we derive the transmission parameters (the number of symbols needed in each ON slot). Then the transmitter applies the parameter during its ON/OFF keying to protect message integrity, while the receiver uses energy detection to recover the source information bits. To enhance efficiency, the transmitter sends a normal message packet followed by ON/OFF keying encoding the hash of the message, which is sent in the same wireless channel such that no out-of-band communication is needed.

Input: $n_{default}$ to both A and B, P_s
Setup phase:
1. User starts receiver B and transmitter A
2. A sends an ON/OFF keying sync signal s to B using $n_{default}$
3. A sends a probing packet e to B (long known symbol sequence)
4. B measures the CSI for each received symbol
5. B mimics the attacker to calculate $r_{h''_e\overline{g''_e}}$ between g''_e and h''_e and channel variance $\sigma_{h''_e}$
6. B picks α, and calculates the minimal number of needed symbols n based on $r_{h''_e\overline{g''_e}}$ to satisfy P_s
7. B immediately notifies A the derived n via ON/OFF keying encoding
Online phase:
8. A and B update n, then A transmits each message followed by its hash using ON/OFF keying and Manchester coding

Figure 3: Overview of the Message Integrity Protection Protocol. n: amount of symbols in each ON slot; P_s: minimum guarantee of detection probability in each ON slot; α: energy detection threshold at B.

Our integrity protection protocol is divided into two phases: The first one is called setup phase, within which the channel is measured, and based on that we calculate the channel correlation coefficient and the necessary amount of symbols n in each ON slot. The second one is the online phase, during which we transmit the original message and the integrity-protected bits (hash in our case) using the parameter derived in the setup phase. Note that, to follow the optimal defender strategy, the transmitter always sends signals/packets using the highest available transmission power and antenna gain.

5.2 The Setup Phase

First the user initiates the protocol by starting both the transmitter and receiver. Initially, the transmitting parameter $n_{default}$ will be set to a large enough default value. The default parameter can be obtained from known channel statistics or hardcoded into the devices. When the protocol starts, the devices have no much knowledge about the channel itself, thus it is reasonable to assume a worst case scenario (almost static channel). Note that the default parameter is chosen based on conservative estimation, so as to

guarantee the security for the transmission of the sync message, and parameter of n back to A from B. This may bring some time overhead due to the conservative default parameter. However it is a one-time overhead in the initialization phase, thus is not a big concern. The setup phase ends after the receiver sends back its calculated parameter n to A.

5.2.1 Synchronization

If node A wants to transmit to B, it first needs to determine the length of each slot based on the channel correlation measurement. A synchronization signal s is needed to notify the receiver to begin measurement of channel correlation. This can be done by sending a standard sync message s containing the node IDs and the initiation intent, followed by the ON/OFF keying mode encoding of its hash using $n_{default}$ as the number of symbols. By knowing there is only one legitimate transmitter within range, the receiver can detect any malicious modifications to the synchronization signal. If the attacker also tries to initiate the measurement with the receiver, two duplicate sync sequences will be detected and an alarm will be sounded at B.

5.2.2 Channel Measurement

After sending the sync sequence, the transmitter will immediately send a packet e (or several of them sent back to back) with many repeated known symbols to the receiver. The receiver will measure a CSI sequence from the received signal. The CSI h_{e_i} is computed by taking every received symbol e_i' after converting the signal to baseband (before channel equalization), and divide it by the sent symbol e_i (both complex numbers). Moreover, the receiver decodes each packet to check whether the result is the same as the known sent symbols (e.g., all '1' bits), to prevent an attacker interfering with the CSI measurement. Note the specific value of CSI measured in this step has nothing to do with the future ON/OFF keying transmission. It is only the channel statistics that we need to derive.

5.2.3 Channel Correlation Estimation

After the channel measurement, the receiver obtains $h_{e_i} = h'_{e_i} + h''_{e_i}$, $i \in [1, M]$ where M is the number of CSI samples. It then use this sample channel to mimic the attacker to obtain a upper-bound of spatial-temporal correlation. Based on our attack model, the slow-channing/LoS component h' can be cancelled by the correlated jammer. So we need to eliminate this slow-changing/LoS component from channel h. In practice, we can use a Savitzky-Golay filter to separate the slow-changing and fast-changing components in the CSI [2]. Therefore we will focus on the remaining fast-changing/NLoS part h''. We assume the attacker leverages the time domain correlation by using l historical CSI measurements $h''_{N-l}, ..., h''_{N-1}$ to predict the current channel h''_N. For example, it can use the auto-regression model which is a tool for predicting a time series of data [15]. When l=1, it reduces to the simple case when $h''_N = h''_{N-1}$. For each $i \in [l+1, M]$, the attacker can obtain a predicted CSI from the historical CSI, and we denote this predicted sequence as g''_e. Assuming the channel statistics doesn't change between the offline and online phases, we can calculate the correlation coefficient $r_{h''_e \overline{g''_e}}$ of the two sequences g''_e and h''_e (both for $i \in [l+1, M]$), and use it as an estimated upper-bound of attacker's channel correlation $r_{h'' \overline{g''}}$. Besides, the receiver also calculates the variance of the dynamic component $\sigma_{h''_e}$ and use that as an estimation of $\sigma_{h''}$ and $\sigma_{g''}$.

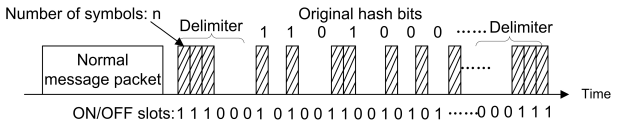


Figure 4: The messaging structure in the online phase.

5.2.4 Deriving the Minimum Symbol Amount

Given the bound of attacker's correlation coefficient, we substitute it along with others parameters (including $\sigma_{h''}$, σ_x, α) into Eq. (9) in the theoretical analysis section. Then we can derive the detection probability P_d, and the minimum necessary amount of symbols n in each ON slot:

THEOREM 5.1. *Given the required minimum detection probability in each ON slot P_s, the minimal number of symbols is:*

$$n = \lfloor log^{1-P_s}_{1-P_d} \rfloor \tag{10}$$

The proof is in Appendix. Once n is determined, the receiver B immediately notifies the transmitter A to update this parameter, by using ON/OFF keying and Manchester coding to send n to A using the default parameter $n_{default}$.

5.3 The Online Phase: Data Transmission

In the online phase, the TX and RX devices start integrity-protected data transmission using parameter n. The message structure in this phase consists of two parts (shown in Fig. 4): a normal packet containing the message m followed by a L-bit cryptographic hash $H(m)$ encoded using ON/OFF keying and Manchester coding. To ensure the receiver knows the message boundary, node A transmits an I-code delimiter "111000" both in the beginning and the end of each ON/OFF keying sequence. This delimiter is guaranteed to be different from any message bits encoded using ON/OFF keying and Manchester coding. The receiver uses energy detection to detect each ON/OFF slot, and then decode the hash string $H'(m)$, it checks whether $H'(m) = H(m)$. If so, the integrity verification is passed for m.

The above basic online protocol can be extended to different application scenarios. For example, suppose in-band communication is needed (i.e., coexist with other traffic in the same band like normal WiFi or ZigBee). To avoid interfering with other links, similar to TEP [7], we can add a Clear-To-Send (CTS) signal in front of the whole message, which reserves the channel for the period until the ON/OFF sequence ends. The two devices can exchange an authenticated Diffie-Hellman key assuming knowledge of presence of each other. If we have a dedicated channel such as a satellite link where source information is broadcasted continuously, then source authentication can also be achieved [5].

5.4 Security Analysis

Integrity Protection. First, due to the collision resistance of cryptographic hash functions, it is infeasible for the attacker to find another $m' \neq m$ such that $H(m') = H(m)$. Second, if the attacker modifies any one or more bits in the original message, approximately half of the hash bits

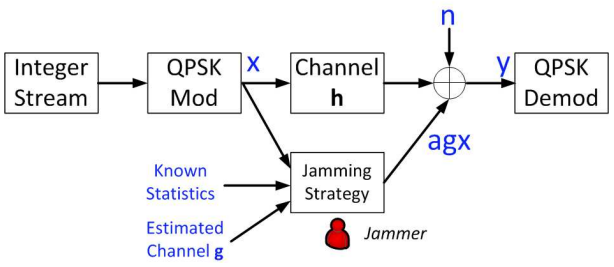


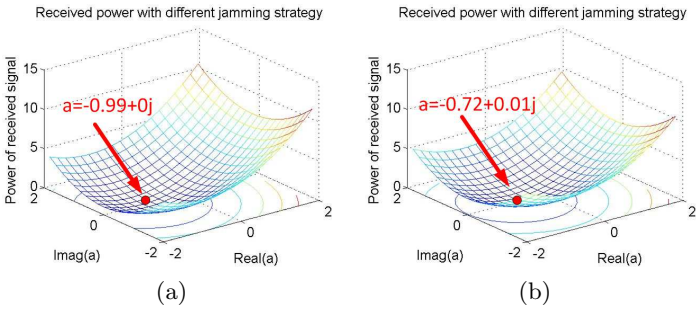
Figure 5: The simulation diagram of correlated jamming.

will flip. For each flipped bit, one ON slot needs to be converted into an OFF slot. So the probability that the attacker successfully passes verification is approximately $(1-P_s)^{L/2}$ (negligible). Thus, message integrity can be guaranteed under our attack model since we choose n to satisfy a minimum per-ON slot detection probability P_s, such that any tampering with the message m will be detected w.h.p.

Security of CSI measurement. In the setup phase, the attacker can try to interfere with the CSI measurement packet(s). If she manages to lower the perceived correlation coefficient $r_{h''_e\overline{g''_e}}$ at the receiver, the calculated n will be smaller and also the security guarantee. But if she increases the perceived correlation, it will decrease the link's throughput but increase security. We can defend against such attack in several ways. (1) Observe that to lower $r_{h''_e\overline{g''_e}}$ the attacker should generate a random jamming signal (or noise) with little correlation. It will make the received message undecodable (or a very high symbol error rate), so that the receiver will sound an alarm; whereas in normal situations the symbol error rate will be small. The attacker can also try jamming with very high power to induce capturing effect, but since the transmitter is already sending with maximum power, this will create a high RSSI at the receiver which again raises suspicion. (2) In non-mobile situations, the channel correlation statistics in different locations can be pre-gathered securely and made publicly known (such as being kept in a database). Once the receiver measures the channel and calculated $r_{h\overline{g}}$, it will compare with the preloaded data, and abort if it finds a discrepancy.

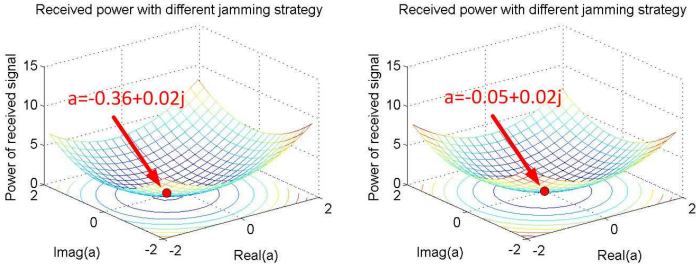
5.5 Overhead Analysis

In our protocol, the overhead is mainly brought by the setup phase. The transmissions of synchronization signals and feedback of n require $(b_s+b_n)/R_{data}+4n_{default}\cdot L\cdot\Delta_t$ time in total, in which b_s and b_n stand for the bit length of message s and n, respectively, R_{data} is the normal data rate,Δ_t is the symbol duration. The ON/OFF keying takes up the majority of transmission time. For example, assume in reality the maximum $|r_{h\overline{g}}|=0.9$ (conservative estimate), then $n_{default}=10$ would suffice. AssumeΔ_t is 1ms and let $L=256$, then we can get the time overhead is around 10 seconds. The step of sending e and calculating $r_{h''_e\overline{g''_e}}$ and $\sigma_{h''_e}$ brings only a small overhead, since typically one probing packet (lasts less than 1 second) is enough to collect a representative CSI sequence. However, we should note that this overhead is one-time. Our scheme is most suitable to a static scenario where the channel statistic doesn't change overtime. If we want to apply it to mobile settings, we may need to re-do the setup phase and update the parameter n periodically. However, for applications such as key establishment, only a few messages need to be exchanged, so we can still assume the channel does not change even in the mobile cases. Fully handling mobile settings will be part of our future work.



(a) (b)

Figure 6: The curved surface denotes average powers of received signal under different attacking strategies. The red dot is the optimal strategy derived by our theoretical analysis. The left figure shows the case when channel correlation coefficient $|r_{h\overline{g}}|=1$. The right figure shows the case of $|r_{h\overline{g}}|=0.7$.



(a) (b)

Figure 7: The left figure shows the case when channel correlation coefficient $|r_{h\overline{g}}|=0.3$. The right figure shows the case of $|r_{h\overline{g}}|=0$.

6. SIMULATION EVALUATION

In this section we use Matlab simulation to validate the correctness of our correlated jamming analysis. Then we evaluate the throughput of our scheme.

6.1 Effectiveness of Correlated Jamming

To analyze the effectiveness of correlated jamming attack, we study the received signal power in the presence of correlated jamming. We choose NLoS Rayleigh fading channel in this simulation. We generate two CSI sequences with a given correlation coefficient $r_{h\overline{g}}$ to simulate the legitimate channel and attacker's estimation. For convenience, we assume A's transmission power is $0dB$, and we normalize the channel gain to be 1. The SNR at the receiver side is assumed to be $25dB$. The signal is modulated using QPSK. We assume the attacker knows $r_{h\overline{g}}$ and the estimated channel's variance σ_g^2, thus can calculate its best jamming strategy a. The simulation result is shown in Fig. 5.

6.1.1 *Optimal Jamming Strategy*

We first validate the theoretical analysis of optimal jamming strategy. We test several cases, each with a different

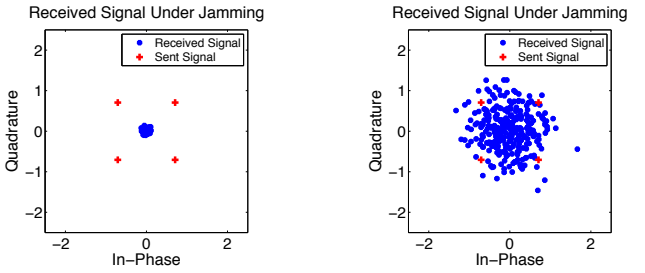


Figure 8: The received signal's constellation under optimal correlated jamming. Left: channel correlation coefficient $|r_{h\bar{g}}| = 1$. Right: $|r_{h\bar{g}}| = 0.7$.

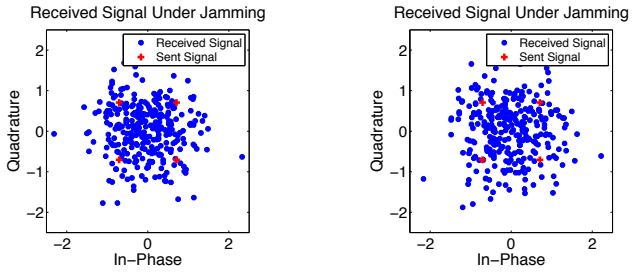


Figure 9: Left: channel correlation coefficient $|r_{h\bar{g}}| = 0.3$. Right: $|r_{h\bar{g}}| = 0$.

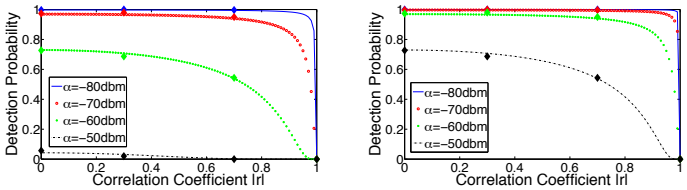


Figure 10: The energy detection probability in each ON slot under different detection thresholds α. Subfigure (a) and (b) uses $0dbm$ and $10dbm$ for transmission power respectively.

correlation coefficient $r_{h\bar{g}}$. The simulation results are shown in Figs. 6 and 7. The curved surface illustrates the average power of received signal under correlated jamming with different attacker strategies. The red dot in the center is the optimal jamming strategy derived from theoretical analysis in Section IV. We can see that the power of received signal always achieves the minimum when the jammer applies the optimal strategy, thus confirming the correctness of our theoretical analysis.

From Figs. 6 and 7, we have another insight. When $|r_{h\bar{g}}|$ decreases from 1 to 0, the amplitude of attacker's jamming coefficient a also decreases from $|a| \approx 1$ to $|a| \approx 0$. This phenomenon corresponds to our intuition: the higher the accuracy of jammer's estimation about h, the more confidence it has, then the more energy it uses to cancel out the signal. When the correlation coefficient $|r_{h\bar{g}}|$ is small, the jammer chooses to be quiet, because extra jamming signal is more likely to increase the received signal's power, instead of weakening it.

6.1.2 *Jamming Effectiveness*

Next we evaluate the impact of optimal jamming strategy to the received signal's power under various channel correlation coefficients. We assume random symbols are sent. The results are shown in Figs. 8 and 9. We can see that when the correlation coefficient is 1, all points concentrate around the original point, i.e. the received signal's average power is close to 0 (the transmitted signal x is almost completely cancelled out). The remaining power is caused by noise. When the channel correlation coefficient decreases, the amplitude (and power) of received signal increases. Specifically, when $|r_{h\bar{g}}| = 0.7$ and 0.3, average powers are 0.46 and 0.87. When $|r_{h\bar{g}}|$ drops to 0, the average power is nearly 1, which means the signal is not canceled at all.

6.2 Signal Detection Probability

Next we will quantize the effectiveness of the defense mechanism by studying its detection probability. The background noise is chosen to be $-95dBm$. We choose the frequency as $2.4Ghz$. The transmission power is set to be 0 or $10dBm$. The distance between legitimate sender and receiver is $5m$. The channels are generated following Rayleigh distribution, together with path-loss model with path-loss exponent as 3.

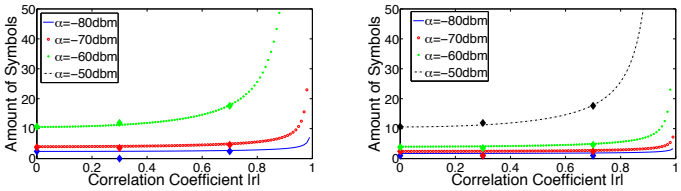
The detection probability results are shown in Fig. 10. We set the detection threshold α at different values. The continuous lines are generated from theoretical calculation, while the big diamond dots are derived from simulation. We can see that our simulation results are consistent with the theoretical results. We can also observe that: 1) A higher correlation coefficient will yield a lower detection probability. 2) The higher threshold we choose, the lower the detection probability is. 3) Higher transmitting power can increase the detection probability.

6.3 Throughput Analysis

Next we analyze the throughput of our scheme. If we only consider using the ON/OFF keying mode to carry data, given the parameter n and security requirement P_s, we can derive the maximum link throughput between A and B: $c = \frac{1}{2\lfloor log_{1-P_d}^{1-P_s} \rfloor \cdot \Delta t}$ (ignore decoding errors). If we consider both normal mode and the hash ON/OFF encoding, the maximum throughput will be $c' = \frac{L_{data}}{T_{data}+2L\cdot\lfloor log_{1-P_d}^{1-P_s} \rfloor \cdot \Delta t}$, where L_{data} and T_{data} are the bit length and transmission time of a normal data packet, respectively, while L is hash length. We can see that the higher the per-symbol detection probability P_d, the higher the throughput.

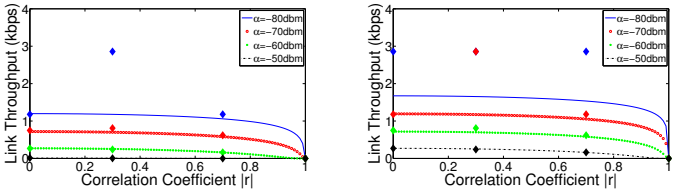
For simplicity, we evaluate the ON/OFF keying mode only. Let the symbol duration be $0.35ms$. The security requirement for successfully detecting each ON slot is set to be $P_s = 0.999999$. All other parameters are the same as in the previous sub-section.

The minimal number of symbols under the given security requirement is shown in Fig.11, and the corresponding maximum link throughput is shown in Fig.12. We have several observations. 1) As the correlation coefficient $|r|$ increases, the energy detection probability in each ON slot decreases, which leads to an increasing number of needed symbols and a decreasing link throughput. This corresponds to our intuition that the more accurate the jammer's estimation about channel h, the more symbols we need in order to achieve



(a) (b)

Figure 11: The minimum necessary amount of symbols under different detection thresholds α, given $P_s = 0.999999$. Subfigure (a) and (b) use $0dbm$ and $10dbm$ transmission power respectively. The black curve in subfigure (a) is not shown as it is above 300.



(a) (b)

Figure 12: The link throughput under different detection thresholds α, given $P_s = 0.999999$ detection probability. The length of each symbol is $0.35ms$. Subfigure (a) and (b) are with $0dbm$ and $10dbm$ transmitting power respectively.

the same security requirement. 2) With the increase of detection threshold α, the detection probability decreases, and the necessary amount of symbols increases. We note that, in reality the detection threshold is set based on the noise level. The higher the noise level, the higher the threshold should we use, which can decrease the false positive rate for OFF slots. The tradeoff is that, this will decrease the true positive probability (for ON slots) and also the link throughput eventually.

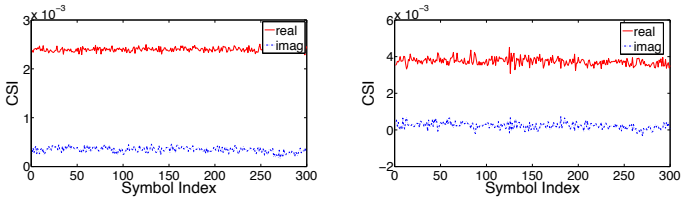
Note that, in Fig. 12, some of the result points for low detection thresholds seem far away from our theoretical calculation. This is because the high detection probability leads to enlarged error in the minimal number of symbols in our discrete simulation, which could be remedied by choosing longer symbol sequences in simulation.

7. EXPERIMENT AND IMPLEMENTATION

In this section, we first carry out experiments to measure the CSI of real indoor environments as a case study, and analyze the temporal domain channel correlation to show how it affects the security of our scheme. Then we implement our ON-OFF keying scheme on USRP devices to show the performance.

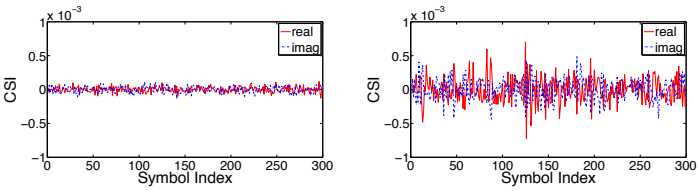
7.1 Experiment Setup

We setup two USRP GNU radio N210 devices with SBX daughter boards on a table in an indoor lab, the distance between them is about $0.5m$. The transmitting power is 20 dbm, and each symbol duration is $0.35ms$. We tested two scenarios: 1) static channel without external disturbance; 2) dynamic channel by using automatic random disturbance. In our case we attach several aluminum foil strips on the transmitting antenna, and use an electric fan to blow air towards it so the foil vibrates randomly. The second experiment's layout is shown in Fig. 17 (a).



(a) (b)

Figure 13: Subfigure (a) shows the channel CSI measured between transmitter and receiver in static scenario. (b) shows the artificial dynamic channel scenario.



(a) (b)

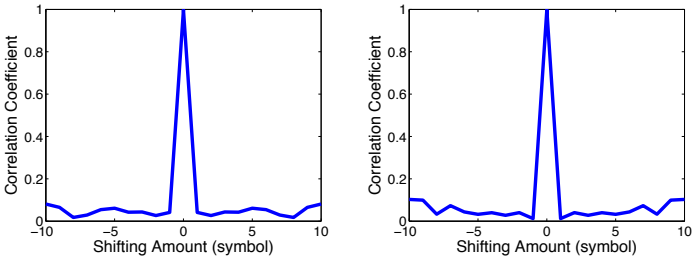
Figure 14: Subfigure (a) shows the NLoS CSI component extracted between transmitter and receiver in static case. (b) shows artificial dynamic channel scenario.

We implement an OFDM transmitter and receiver on the USRP devices. The transmitter sends 5 packets in a row with known symbols in the 2.4GHz band with bandwidth set to 500kHz. In order to obtain the true physical channel state, we connect the two USRPs with a MIMO cable to synchronize their clocks to eliminate the impact of frequency and phase offset. In reality, if the devices are far apart and no cable is available, we can use accurate external clocks such as GPS clocks to synchronize TX/RX. The receiver extracts the frequency domain CSI for each symbol in one OFDM subcarrier from baseband before equalization, and we analyze the CSI sequence on the computer using Matlab. Because each OFDM subcarrier is narrow-band (7.8kHz), the frequency domain CSI is equivalent to time domain CSI. In the first scenario, since the devices are close to each other, the LoS part is strong and thus the channel is very stable. Therefore it is an ideal case for the attacker.

7.2 CSI Randomness and Correlation

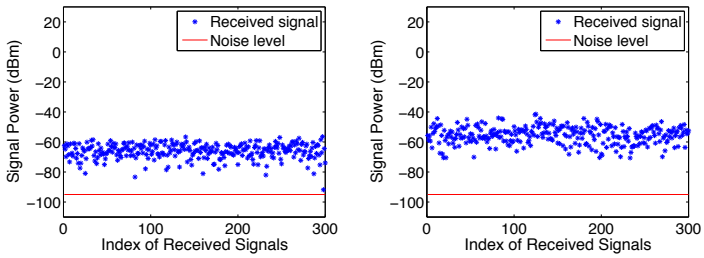
To mimic the correlated jamming attacker, after obtaining the CSI h, we first eliminate its static component and obtain the remaining dynamic component. We use the Savitzky-Golay filter to separate the slow-changing (mainly caused by LoS) and fast-changing components. To generate the attacker's estimated CSI sequence g, we use autoregression with the exact previous two CSI samples to predict the next sample. Then we analyze the correlation between h and g and compute the minimum necessary number of symbols, which is used to derive the maximum secure link throughput.

The experiment result for scenario 1) is shown in Fig. 13 (a). We can see that the CSI between the transmitter and receiver is quite stable, due to the static channel condition. However, there is still a small fast-changing and random component (caused by noise and other factors). The fast-



(a) (b)

Figure 15: Subfigure (a) shows the auto-correlation coefficient of the NLoS CSI sequence in static scenario. (b) shows the artificial dynamic channel scenario.



(a) (b)

Figure 16: Subfigure (a) shows the received power level under optimal correlated jamming attack in static scenario. (b) shows the artificial dynamic channel scenario.

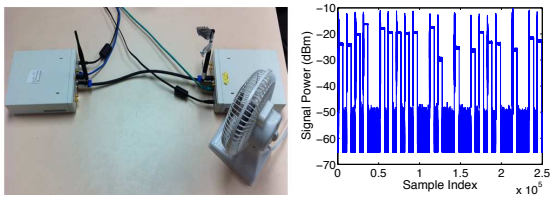
changing CSI component is shown in Fig. 14 (a). We can see that this part's randomness is high. To verify this we calculate this sequence's auto-correlation coefficient and show the result in Fig. 15. The correlation coefficient drops below 0.1 after one symbol's duration, which is about 0.35 ms.

On the other hand, in the artificial dynamic case shown in Fig. 14 (b) it has a higher variance than in the static case. The average gain of the NLoS component is $-84dB$ for static case and $-72dB$ for dynamic case, which means it is barely large enough compared with background noise level in static scenario. However, in the artificial dynamic scenario, the received power level is about $12dB$ higher, which will result in higher detection probability, thus higher secure link throughput. We then plot the received signals' power levels after removing the LoS component in Fig. 16. We can see that the received signals' power levels are both higher than the noise level; but if we use a higher detection threshold, the static case will need much more number of symbols to defend against signal cancellation. This shows by actively randomizing the channel, we can achieve better integrity protection under the same throughput, or achieve a higher secure link throughput under the same security guarantee.

We can calculate ON/OFF keying parameters for both scenarios. After the CSI statistics estimation, we derived $r_{h''_e \overline{g''_e}} < 0.1$ for both scenarios. Along with $\sigma^2_{h''}$ and detection threshold $\alpha = -90dbm$, we derive the necessary amount for each slot as $n = 2.2$ and $n = 1.5$ for the static and dynamic scenarios respectively; In reality we choose $n = 3$ and $n = 2$, respectively. The theoretical maximum link throughput of the ON/OFF mode can be derived as $1.29kbps$ and $1.86kbps$, respectively. More comparison results of the two scenarios under different thresholds are shown in Table 1.

Threshold	-90dBm	-80dBm	-70dBm	-60dBm
Static n	2.2	3.8	9.7	217
Dynamic n	1.5	2.1	3.4	7.5
Static R(kbps)	1.3	0.7	0.3	0.01
Dynamic R(kbps)	1.8	1.3	0.8	0.4

Table 1: The comparison of symbol number in ON slots (n) and ON/OFF mode throughput under different detection thresholds and channel conditions.



(a) (b)

Figure 17: Subfigure (a) shows the experiment layout; (b) illustrates the receiving power of ON/OFF key energy detection method with 3 symbols in a ON slot. The sampling rate is $500ksps$.

7.3 Protocol Implementation

We have implemented a basic prototype of our integrity-protection scheme in USRPs. By now our prototype can achieve three main functions: synchronization, channel measurement, and ON/OFF keying encoded hash transmission. For synchronization and hash transmission, they are all based on ON/OFF keying encoding and energy detection decoding, which is the key part in the implementation.

The transmitting gain is $30dB$ and receiving gain is $15dB$. The noise floor is about $-50dBm$. We use the static scenario's parameter $n = 3$ as an example and the received power sequence is shown in Fig. 17 (b). The receiver counts the number of high power samples in each slot and compare it with the threshold. We note that there is an implementation issue: the ON and OFF slots are not exactly the same in length and there is a lag after each ON slot (in Fig. 17 (b)). This is because when the USRP switches from ON to OFF, there is always a time delay induced by hardware. This delay could be tolerated when the ON slot is long. However for the short slots, it will prevent the receiver from decoding an OFF slot between two ON slots. To remedy it we make the OFF slots longer, beyond the end of the lag. The throughput will be decreased, however as the original ON slot is short, so such a loss of throughput is tolerable. If we enlarge the OFF slot and make each ON/OFF to be 22 ms, the derived practical throughput is $45bps$ for ON/OFF keying mode. After adjusting the length of OFF slot, we achieve zero bit error with random sequences of 10^5 bits long. If the message transmitted in the normal mode is 4096 Bytes long at $500kbps$ and hash length is 256 bits, this yields an overall throughput of $5.75kbps$.

7.4 Discussion

Although indoor channel is typically stable, we use active channel disturbance to make it a more dynamic channel to better defend against signal cancellation attacks. In outdoor scenarios, the slow-changing/LoS component is typically weak and the channel is dynamic due to fading, so it will be harder for the attacker to cancel out the signal. Eval-

uating our scheme in outdoor scenarios will be part of the future work.

In our protocol, we only considered the type of attack which uses estimated future CSI (spatial and temporal domain) to carry out jamming. This type of attack includes the one proposed in [19] as a special case. In the future we will consider another type of attack, especially spatial correlated jamming where the jammer itself is located near to the TX or RX and one of its own channels is correlated with legitimate CSI h. The attacker in this case does not need to measure and estimate h, but needs to do additional processing of signal x and relay it. However, since it does not know h and thus the correlation coefficient, it is hard to achieve optimal cancellation. To defend it we can impose a restricted region around the TX/RX to prevent the attacker from being too close, to upper-bound its spatial correlation. With our channel changing approach, such correlation should decrease compared with static channel case.

8. RELATED WORK

Physical Layer Message Integrity Protection and Authentication: The integrity code (I-code) proposed by Čapkun et. al. in [5] protects the message integrity transmitted in an insecure wireless channel. It provides message tamper-evidence using unidirectional error detection code, based on the infeasibility of signal cancellation. Similarly, Tamper-Evident Pairing (TEP) proposed by Gollakota et. al [7] is an in-band device pairing protocol for 802.11 devices, which protects message integrity by embedding cryptographic authentication information (e.g., a hash) into the physical signals. Both of the above works assume that wireless signal cancellation is infeasible. However, Pöpper et. al. demonstrated a practical cancellation attack using a pair of directional antennas in [19], under quasi-static channel conditions. Recently, Hou et. al. proposed Chorus [10], which extends the idea of in-band message authentication to a group of devices, and adapts the uncoordinated frequency hopping (UFH) mechanism to defend against correlated signal cancellation. However, it assumes the attacker does not possess advanced real-time processing capabilities and is located at some distance away from the legitimate pair. In summary, a quantitative security guarantee is still lacking for anti-signal-cancellation based message integrity protection schemes. In this paper, our proposed defense approach can be applied to any protocol with the same core idea.

Correlated Jamming: Médard and Goldsmith [14] first studied the capacity of wireless channels under correlated jamming. This work considers two cases: first, the jammer has complete knowledge of the transmitted signal, and second, the jammer could only obtain partial knowledge of the transmitted signal through eavesdropping. The channel is always assumed to be constant and known by the jammer. It is shown that the channel capacity decreases to zero. Kashyap et. al. [12] also studied channel capacity under correlated jamming, and they expand to MIMO case and assume the CSI is totally random (attacker only knows the statistics). Shafiee and Ulukus [22, 21] continued the research of correlated jamming and expanded to the multi-user scenario. All these works try to maximize the link capacity in the information theory sense, while our work considers a more practical goal - the energy detection probability which is the key to integrity protection. In addition, in our work we consider a different model by assuming the attacker can obtain an arbitrarily correlated CSI, which is more practical. We study the relationship between link throughput and the correlation coefficient between the attacker's estimated channel and the real channel. Recently, Chang et. al. proposed a countermeasure for signal-cancellation based jamming [6], by introducing redundancy frequency offsets at the transmitter-side. However, it can only prevent a weak form of correlated jamming attack where the attacker generates its own signal but cannot prevent the relaying attacks.

9. CONCLUSION

In this work, we studied the security of physical layer message integrity protection scheme. We established a correlated jamming framework to model the attacker's behavior. We quantitatively analyzed the security guarantee for physical layer message integrity protection protocol under correlated jamming attack with arbitrary CSI-estimation correlation coefficient. Based on the analysis we proposed a physical layer message integrity protection protocol which achieves any given security requirement. Extensive experiments and simulations results are shown to verify the correctness of our optimal jamming strategy analysis and to evaluate the achievable throughput under different security requirements. In the future, we will extend our scheme to the case of MIMO.

10. ACKNOWLEDGMENTS

We would like to thank the anonymous reviewers for helpful comments. We thank Dan Shan for helpful discussions on USRP implementation. This work was supported in part by the US National Science Foundation under grants CNS-1410000, CNS-1350655 and CNS-1502584.

11. REFERENCES

[1] Top 50 internet of things applications - ranking. `http://www.libelium.com/top_50_iot_sensor_applications_ranking/`.

[2] S. Ali, V. Sivaraman, and D. Ostry. Secret key generation rate vs. reconciliation cost using wireless channel characteristics in body area networks. In *Embedded and Ubiquitous Computing (EUC), 2010 IEEE/IFIP 8th International Conference on*, pages 644–650, Dec 2010.

[3] D. Balfanz, D. K. Smetters, P. Stewart, and H. C. Wong. Talking to strangers: authentication in ad-hoc wireless networks. In *NDSS '02*, 2002.

[4] M. Cagalj, S. Capkun, and J.-P. Hubaux. Key agreement in peer-to-peer wireless networks. *Proceedings of the IEEE*, 94(2):467–478, Feb. 2006.

[5] S. Capkun, M. Cagalj, R. Rengaswamy, I. Tsigkogiannis, J.-P. Hubaux, and M. Srivastava. Integrity codes: Message integrity protection and authentication over insecure channels. *Dependable and Secure Computing, IEEE Transactions on*, 5(4):208–223, Oct 2008.

[6] S.-Y. Chang, Y.-C. Hu, J. Chiang, and S.-Y. Chang. Redundancy offset narrow spectrum: countermeasure for signal-cancellation based jamming. In *Proceedings of the 11th ACM international symposium on Mobility management and wireless access*, pages 51–58. ACM, 2013.

[7] S. Gollakota, N. Ahmed, N. Zeldovich, and D. Katabi. Secure in-band wireless pairing. In *USENIX security symposium*, 2011.

[8] M. T. Goodrich, M. Sirivianos, J. Solis, G. Tsudik, and E. Uzun. Loud and clear: Human-verifiable authentication based on audio. In *In IEEE ICDCS 2006*, page 10, 2006.

[9] X. He, H. Dai, W. Shen, and P. Ning. Is link signature dependable for wireless security? In *INFOCOM, 2013 Proceedings IEEE*, pages 200–204, April 2013.

[10] Y. Hou, M. Li, and J. D. Guttman. Chorus: Scalable in-band trust establishment for multiple constrained devices over the insecure wireless channel. In *Proceedings of the Sixth ACM Conference on Security and Privacy in Wireless and Mobile Networks*, WiSec '13, pages 167–178, New York, NY, USA, 2013. ACM.

[11] P. Kafle, A. Intarapanich, A. Sesay, J. McRory, and R. Davies. Spatial correlation and capacity measurements for wideband mimo channels in indoor office environment. *Wireless Communications, IEEE Transactions on*, 7(5):1560–1571, May 2008.

[12] A. Kashyap, T. Basar, and R. Srikant. Correlated jamming on mimo gaussian fading channels. *Information Theory, IEEE Transactions on*, 50(9):2119–2123, Sept 2004.

[13] P. Kyritsi, D. Cox, R. Valenzuela, and P. Wolniansky. Correlation analysis based on mimo channel measurements in an indoor environment. *Selected Areas in Communications, IEEE Journal on*, 21(5):713–720, June 2003.

[14] A. G. M. MÃl'dard. Capacity of correlated jamming channels. In *Allerton Conference on Communications, Computing and Control*, 1997.

[15] H. Madsen. *Time Series Analysis: Forecasting and Control*, volume 77. CRC Press, 2008.

[16] J. M. McCune, A. Perrig, and M. K. Reiter. Seeing-is-believing: Using camera phones for human-verifiable authentication. In *IEEE S & P*, pages 110–124, 2005.

[17] L. Nguyen and A. Roscoe. Authentication protocols based on low-bandwidth unspoofable channels: a comparative survey. *Journal of Computer Security*, 19(1):139–201, 2011.

[18] T. Perkovic, M. Cagalj, T. Mastelic, N. Saxena, and D. Begusic. Secure initialization of multiple constrained wireless devices for an unaided user. *Mobile Computing, IEEE Transactions on*, 11(2):337–351, Feb 2012.

[19] C. Pöpper, N. O. Tippenhauer, B. Danev, and S. Capkun. Investigation of signal and message manipulations on the wireless channel. ESORICS'11, pages 40–59, 2011.

[20] M. Schulz, A. Loch, and M. Hollick. Practical known-plaintext attacks against physical layer security in wireless mimo systems. *Proceedings of the Network and Distributed System Security Symposium, NDSS 2014*, 2014.

[21] S. Shafiee and S. Ulukus. Capacity of multiple access channels with correlated jamming. In *Military Communications Conference, 2005. MILCOM 2005. IEEE*, pages 218–224 Vol. 1, Oct 2005.

[22] S. Shafiee and S. Ulukus. Mutual information games in multiuser channels with correlated jamming. *Information Theory, IEEE Transactions on*, 55(10):4598–4607, Oct 2009.

[23] F. Stajano and R. J. Anderson. The resurrecting duckling: Security issues for ad-hoc wireless networks. In *IWSP '00*, pages 172–194, 2000.

[24] N. Tippenhauer, L. Malisa, A. Ranganathan, and S. Capkun. On limitations of friendly jamming for confidentiality. In *Security and Privacy (SP), 2013 IEEE Symposium on*, pages 160–173, May 2013.

12. APPENDIX

12.1 Proof of Theorem 4.5

PROOF. We will show the monotonicity w.r.t $k = \sigma^2$. According to Leibniz integral rule we have:

$$P'_d(k) = \int_{\alpha-\sigma_n^2-\sigma_v^2}^{\infty} \frac{d}{dk} \frac{e^{-\frac{r}{2k\sigma_x^2}}}{2k\sigma_x^2} \, \mathrm{d}r \tag{11}$$

$$= \int_{\alpha-\sigma_n^2-\sigma_v^2}^{\infty} \frac{(r-2k\sigma_x^2)\cdot e^{-\frac{r}{2k\sigma_x^2}}}{4k^3\sigma_x^4} \, \mathrm{d}r \tag{12}$$

$$= \frac{-r}{2k^2\sigma_x^2 e^{\frac{r}{2k\sigma_x^2}}}\Big|_{r=\alpha-\sigma_n^2-\sigma_v^2}^{r=\infty} \tag{13}$$

It is easy to see from equation 13 that $P'_d(k) > 0, \forall k > 0$, which means the smaller $k = \sigma^2$ brings smaller detection probability. □

12.2 Proof of Theorem 4.6

PROOF. We start from expression of the variance of random variable $\beta = h + ag$. The variance σ^2 is denoted as:

$$\begin{aligned}
2\sigma^2 &= E|h+ag|^2, \\
&= E[(h+ag)\overline{(h+ag)}], \\
&= E[(h+ag)(\overline{h}+\overline{g}\overline{a})], \\
&= \sigma_h^2 + E[h\overline{(ag)}] + E[(ag)\overline{h}] + |a|^2\sigma_g^2, \\
&= \sigma_h^2 + |a|^2\sigma_g^2 + \overline{a}E[h\overline{g}] + aE[\overline{(h\overline{g})}], \\
&= \sigma_h^2 + |a|^2\sigma_g^2 + 2Re\{a\}Re\{E[h\overline{g}]\} + 2Im\{a\}Im\{E[h\overline{g}]\}, \\
&= \sigma_h^2 + (Re\{a\}^2 + Im\{a\}^2)\sigma_g^2 + 2Re\{a\}Re\{E[h\overline{g}]\} \\
&\quad + 2Im\{a\}Im\{E[h\overline{g}]\}
\end{aligned} \tag{14}$$

From above we can see that σ^2 is a convex function w.r.t $Re\{a\}$ and $Im\{a\}$. By setting partial derivatives to zero, We can get $Re\{a\}* = -\frac{Re\{E[h\bar{g}]\}}{\sigma_g^2}$, $Im\{a\}* = -\frac{Im\{E[h\bar{g}]\}}{\sigma_g^2}$.

The proof of $\sigma_v^2 = 0$ is similar to that of type I attacker, thus it is not presented here. □

12.3 Proof of Theorem 5.1

PROOF. The probability of signal $'ON'$ being not detected at one symbol is $(1-p_d)$, thus consider n continuous symbols, the probability of $'ON'$ being not detected within all n symbols is $(1-p_d)^n$. By making it lower than the non-secure probability $1-p_s$, we can derive the necessary amount of symbols:

$$n = \lfloor log_{1-p_d}^{1-p_s} \rfloor \tag{15}$$

□