Analog Man-in-the-Middle Attack Against Link-Based Packet Source Identification

Yu-Chih Tung, Kang G. Shin, and Kyu-Han Kim
The University of Michigan and Hewlett-Packard Laboratories
Email: {yctung,kgshin}@umich.edu and kyu-han.kim@hpe.com

ABSTRACT

A novel attack model is proposed against the existing wireless link-based source identification, which classifies packet sources according to the physical-layer link signatures. A link signature is believed to be a more reliable indicator than an IP or MAC address for identifying packet source, as it is generally harder to modify/forge. It is therefore expected to be a future authentication against impersonation and DoS attacks. However, if an attacker is equipped with the same capability/hardware as the authenticator to process physical-layer signals, a link signature can be easily manipulated by any nearby wireless device during the training phase. Based on this finding, we propose an attack model, called the *analog man-in-the-middle* (AMITM) attack, which utilizes the latest full-duplex relay technology to inject semi-controlled link signatures into authorized packets and reproduce the injected signature in the fabricated packets. Our experimental evaluation shows that with a proper parameter setting, 90% of fabricated packets are classified as those sent from an authorized transmitter. A countermeasure against this new attack is also proposed for the authenticator to inject link-signature noise by the same attack methodology.

1. INTRODUCTION

The identification of packet sources is important in a wireless network since each packet is broadcasted over the air and can be intercepted and modified by any nearby devices. For example, one critical function of a modern wireless intrusion detection system (WIDS) is to detect packets sent from rogue wireless devices by deploying wireless sensors and analyzing spectrum usage [1]. However, unlike a wireline network which can classify a packet source by tracking its *wired* input port, it is challenging for a wireless network to figure out who sent the packet since there is no wireline constraint. Moreover, common identifiers, such as IP and MAC addresses, can be easily modified via software (e.g., by `ifconfig`). Even though the decoded packet can be authenticated later by a cryptosystem, receiving packets from an unauthorized source (in the link layer) is shown to create numerous potential threats in wireless networks. For example, injecting a fake management frame can be used to exploit a known 802.11 vulnerability [31] and replaying legitimate packets can easily cause a denial-of-service (DoS) attack [25]. These threats of injecting packets as authorized users are commonly known as identity-based attacks [36]. This problem becomes severer when devices (e.g., in sensor networks) have only limited computation resource as the cost of distributing secret keys and crypto validation is too high to be affordable [29].

One way to stop the identity-based attack is to authenticate transmitters by comparing *physical-layer fingerprints*. That is, the receiver can record the physical-layer fingerprints of authorized transmitters and reject packets sent from those with unrecognizable fingerprints. Two types of physical-layer fingerprints have been explored to distinguish transmitters: *hardware*- and *link*-based signatures. For example, modulation deviation and clock skew caused by hardware imperfection have been proven as a unique signature even when devices are built by the same manufacturer [4, 11, 13]. The link signature like received signal strength (RSS) or channel fading caused by multipath environments varies from location to location, so the difference between the attacker and the authorized transmitters can be easily identified [15, 29, 34, 35]. In this paper, we focus on the vulnerability of link-based source identification because the hardware-based signature (not coupled with the transmitter's location) has been shown to be easily estimated and reproduced [4, 14].

To attack the wireless networks protected by link-based source identification, we propose a novel attack model, called the *analog man-in-the-middle* (AMITM) attack, which can fabricate authorized link signatures in the analog signal domain. Unlike other known attacks under the assumption that the link signature of an authorized user is known (or at least partially known) so the attacker can *mimic* this link signature [8, 22, 28], AMITM injects a semi-controlled link signature in the training phase and reproduces it later. AMITM is a more practical attack since it neither relies on the presumed knowledge of authorized transmitters' signatures nor requires the receiver's cooperation.

Injection of a link signature during the training phase is the key to fabricate a packet with AMITM. AMITM exploits the latest full-duplex relay design [9] to simultaneously sniff and relay the authorized packets during the update of link signatures. (Link signatures need periodic updates due to

MobiHoc'16, July 04-08, 2016, Paderborn, Germany

DOI: http://dx.doi.org/10.1145/2942358.2942361

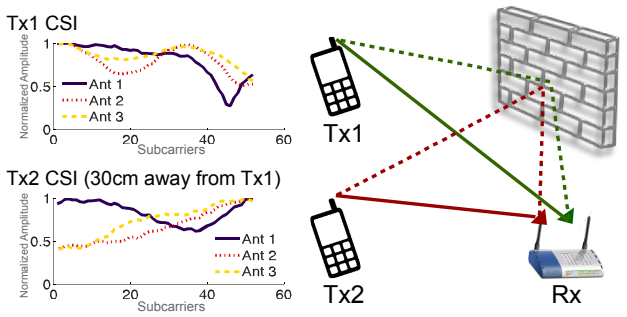


Figure 1—CSI varies with location due to multipath fading. In this example, CSI is collected by two devices only 30cm apart from each other.

environmental variations [26, 29].) AMITM is different from the jam-and-replay attack (which can be thought as a *digital* man-in-the-middle attack) because the whole packet is not decoded and retransmitted, but only the physical-layer signal is processed and relayed at the same time during the genuine packet's transmission. Thus, it is unnecessary to stop the link between the authorized users and the receiver by suspicious spectrum jamming. This property makes it hard to detect AMITM since the attacked wireless spectrum and traffic exhibit normal characteristics, but only the link signature is modified.

AMITM is not the first attack targeting the training phase of link-based source identification, but its novel full-duplex design makes it a powerful generalization of other existing attacks. For example, the attack based on forging link signatures in the training phase [18] can only hide the change of the attacker's location but cannot impersonate the authorized users (see Section 7 for details). To the best of our knowledge, AMITM is the first attack exploiting the full-duplex relay technology to inject a link signature and reproduce it in the fabricated packets.

This paper makes the following four contributions:

- The first analog man-in-the-middle attack fabricating link signatures as those sent from authorized users without knowledge of the genuine link signatures;
- Prototype implementation on software-defined radios;
- Evaluation and demonstration of AMITM's strength using real-world link traces collected from commercial devices;
- Proposal of a countermeasure to mitigate AMITM without the modification of transmitter devices.

The remainder of this paper is organized as follows. Section 2 introduces the principle of link-based source identification and our attack model. The details of AMITM attack and its analysis are provided in Section 3. Section 4 presents our prototype implementation, while Section 5 describes our evaluation setting and experimental results. Section 6 discusses countermeasures against AMITM, while Section 7 summarizes other related work. We discuss future directions and conclude the paper in Sections 8 and 9, respectively.

2. BACKGROUND AND SYSTEM MODEL

We first describe the link-based source identification and then present our proposed attack model, AMITM.

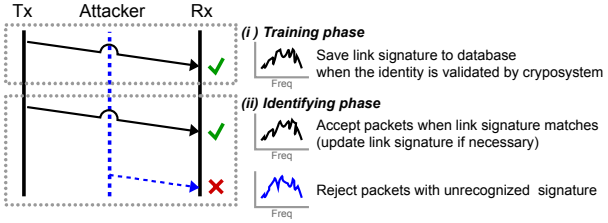


Figure 2—Link-based source identification process. If the attacker tries to send forged packets after the link signature of an authorized transmitter is remembered, the packet will be dropped immediately because of an inconsistent link signature in the forged packets.

2.1 Link Signature

Link-based source identification utilizes the wireless channel characteristics to distinguish transmitters. For example, the received signal strength (RSS) decays proportionally to the distance between a transmitter and a receiver, so the RSS of packets received at different receivers can be used to determine the transmitter's location and its identity [6, 36]. Since each received packet only provides one RSS estimation, and the attenuation varies with the transmission power, RSS is usually considered neither accurate nor unique in identifying a packet source [23]. To improve the identification granularity further, channel state information (CSI) or channel impulse response (CIR) is used as an alternative for link signature [29, 34] due to its fast spatial de-correlation. As shown in Fig. 1, since the received signal is the combination of all delayed and attenuated copies of a sent signal, it combines constructively at certain frequencies and destructively at the others, generating a unique channel characteristic. This phenomenon is also known as frequency selective (multipath) fading. CSI-based source identification can also be extended to multiple receivers/antennas. For example, the angle-of-arrival (AoA) estimated by multiple antennas also forms a unique signature for each transmitter [35]. However, as will be introduced later, these link signatures (e.g., RSS or CSI) are vulnerable to the proposed AMITM attack since a genuine link signature can be stealthily modified by an attack node equipped with full-duplex capability.

2.2 Source Identification Process

Fig. 2 shows a general process to identify a packet source based on link signature. The link signature of an authorized packet is saved in a database after the packet is validated by a cryptosystem. When a new packet is received, a receiver will check if its link signature matches the one in the saved database. If the received link signature is different from the saved one, the received packet will be dropped immediately, thus preventing impersonation or DoS attacks. On the other hand, if the link signature matches the saved one, the packet will be decoded and this new signature can be updated in the database. However, in reality, the link signature changes over time because of environmental variations even if the location of the transmitter remains fixed. This phenomenon is more pronounced in a mobile environment [29]. Thus, for any link-based source identification, the receiver needs to update link signature periodically. By exploiting this feature, AMITM can easily and gradually inject semi-controlled signatures into authorized packets. The receiver is unable to tell if the link signature change is caused by environmental variations or the signals injected by AMITM.

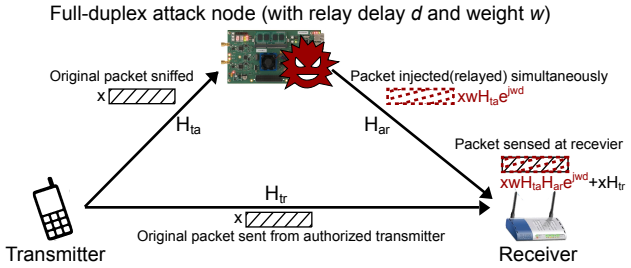


Figure 3—AMITM attack model. During the training stage, an authorized packet is sniffed by the attack node and relayed during the packet's transmutation, thus modifying the link signature estimated at the receiver.

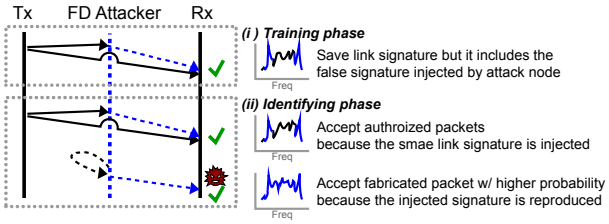


Figure 4—Attacked source identification. Since part of link signature in the authorized packet is modified by the attacker, the fabricated packet with a similar injected signature has a high probability to pass the link-based source identification.

2.3 Attack System Model

Fig. 3 shows the system model of AMITM attack. As shown in this figure, we assume the full-duplex attack node is able to sniff the genuine packet, amplify the sniffed signal by w, and relay the amplified signal with delay d. At a high level, AMITM utilizes a similar radio methodology to the well-known vehicle signal relay attack [19], but it is designed specifically to fabricate packets with the authorized link signatures. See Section 7 for the comparison of AMITM with other attacks based on manipulating phyiscal-layer signals.

The relay delay in AMITM is assumed shorter than the time of a WiFi symbol to ensure the received packet is decodable. Packets are assumed to have been modulated with OFDM, as it is the most popular 802.11 standard. In OFDM, the channel of each subcarrier is modeled by a complex coefficient and the CSI is composed by the channels over subcarriers. Let H_{tr}^i denote the channel coefficient from the transmitter to the receiver on the i-th subcarrier. Since subcarriers are orthogonal to each other, we omit the superscript i for succinct representation. Likewise, H_{ta} and H_{ar} denote the channel coefficients from the transmitter to the attack node and the channel coefficient from the attack node to the receiver, respectively. Under this setting, the received link signature at the receiver is a complex weighted combination of a genuine link signature and an injected signature.

2.4 Attack Process

AMITM injects a semi-controlled link signature during the training phase and reproduces this injected link signature in the identifying phase. As shown in Fig. 4, with the help of a full-duplex attack node, every packet sent from the authorized transmitter is injected with a relayed signal, thus making the link signature different from its original characteristics. Since every authorized packet gets the same injection, the received link signature is consistent over time. Moreover, since the processed signal is relayed within the symbol time, the whole packet is still decodable at the receiver, thus incurring minimal performance degradation during the attack. In some cases, the decoded SNR with a full-duplex relay is even $20dB$ better than the original SNR [9]. These properties make it difficult for the receiver to detect AMITM because, from the receiver's perspective, there is no difference from normal transmissions but only the received signal strength may be higher.

If a fabricated packet (i.e., the packet sent only from the attack node) is received, the packet will be classified, with a high probability, as the one sent from an authorized user because part of the received link signature matches the saved signature. The similarity of link signatures in the fabricated and authorized packets will be discussed in the following sections.

2.5 Full-duplex Attack Node

A full-duplex attack node is the most important component in AMITM, which needs to sniff, process, and relay the analog signal of authorized packets during an ongoing transmission. Otherwise, the signature would not be injected at a correct time, making the packet undecodable. There are implementation challenges for a full-duplex attack node to receive and relay a signal at the same time, but they do not restrict the realization of AMITM since the necessary functionality falls within the scope of traditional full-duplex design. For example, AMITM can directly deploy the existing full-duplex relay node introduced in [9, 12], which was originally designed to broaden WiFi transmission. Thus, in this paper, we focus on the impact of potential attacks realizable with this full-duplex relay technology, not its implementation.

Even though the existing full-duplex design can support the required functionality of attack node, the performance of AMITM depends heavily on the design parameter tuning, such as the selection of relay delay and amplification. For example, traditional full-duplex relay design enforces the sent and relayed signals combined constructively at the receiver, thus increasing the received signal strength. But AMITM achieves better performance when the sent and relayed signals received are orthogonal to each other because the purpose of AMITM is to inject and forge link signatures, rather than increasing the received SNR.

3. AMITM ATTACK ANALYSIS

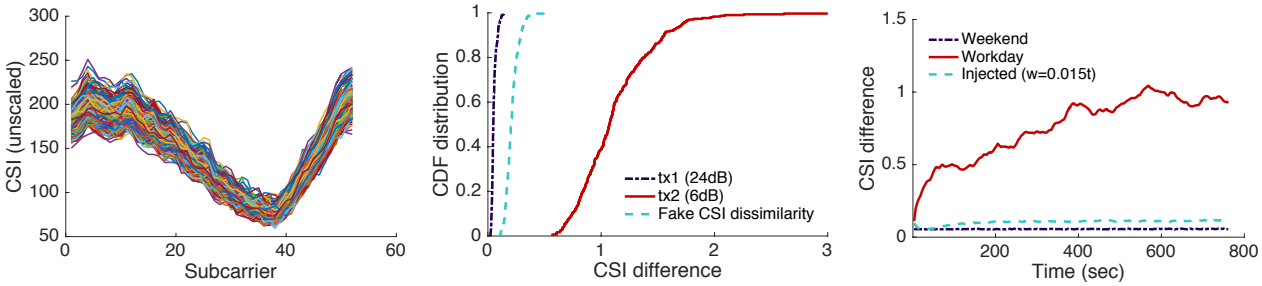
We analyze the effectiveness of AMITM based on the system model in Fig. 3. We will discuss the signature-injection process, packet-fabrication process, and the performance under different parameter settings.

3.1 Link-Signature Injection

Suppose the payload of the original packet sent from an authorized user is x and its frequency-domain channel response at the receiver is H_{tr}, then the received signal, y_r, at the receiver is:

$$y_r = H_{tr}x + n_r, \tag{1}$$

where $n_r \sim \mathcal{CN}(0, \sigma_r)$ is a complex Gaussian noise with zero mean and σ_r variance. In the case of 802.11 probing process, the sent signal x is set to a known training sequence so that the receiver can easily derive the channel response,



(a) Unscaled CSI changes over packets (b) CSI changes over locations (c) CSI changes over time

Figure 5—Practical limitations of link-based source identification. Link signature varies with (i) packets due to ADC and sampling imperfection, (ii) time because of environmental changes, and (iii) locations due to multipath fading. Accommodating the attack setting to ensure link-signature changes to be not greater than these natural variations makes it difficult to detect AMITM because it is unable to know if an observed link-signature change is caused by an injected signal.

H_{tr}, by dividing the received signal y_r by this known x to get $\hat{H}_r = y_r/x$. The sent signal is also received at the AMITM attack node as $y_a = H_{ta}x + n_a$, where H_{ta} is the channel response from the transmitter to the attacker, and n_a is the zero-mean receiving noise at the attacker with variance σ_a. When this sniffed signal is relayed by a full-duplex attack node with delay d and amplification w, the relayed signal received at the receiver is:

$$y_{relayed} = e^{-j2\pi fd} wH_{ar}(H_{ta}x + n_a) + n_r, \quad (2)$$

where $e^{-j2\pi fd}$ is the phase change caused by the relay delay d at the transmission frequency f. This equation is valid only when the delay is less than the WiFi symbol time. In 802.11a/g/n, this symbol time is about 400ns, which is shown to be sufficient for full-duplex to relay a signal within the symbol time, thus causing no inter-symbol interference [9]. Since the signal is relayed within the WiFi symbol time, the received signal at the receiver is a combination of genuine and relayed signals as $y_{injected} = y_r + y_{relayed}$. Once the receiver uses this polluted signal to derive the channel response by dividing the received signal by x (known). The estimated channel response under AMITM attack is:

$$\hat{H}_{injected} = H_{tr} + e^{-j2\pi fd} wH_{ar}(H_{ta} + n'_a) + n'_r, \quad (3)$$

where n'_r and n'_a are the received noises following the same distribution as n_r and n_a since the WiFi probing sequence is modulated with BPSK. As shown in this equation, part of the estimated channel response at the receiver is controlled by the injected link signature, $e^{-j2\pi fd} wH_{ar}H_{ta}$. We call this a *semi-controlled injected signature* because H_{ar} and H_{ta} are the genuine channel responses, which are characterized by the surrounding environment. However, as also shown in this equation, the relay amplification and the received phase difference, $we^{-j2\pi fd}$, are under the attacker's control by setting different w and d. For example, the attacker may select a large w which makes the injected signature dominate the genuine signature, but a very large w might make AMITM easily detectable. Thus, a proper parameter tuning helps fabricate packets with the authorized link signatures and avoid being identified by the receiver, as discussed in the following sections.

3.2 Reproduction of Injected Signatures

Since the purpose of AMITM is to fabricate packets so that link-based source identification may classify them as sent from authorized users, the fabricated packets should have link signatures similar to $\hat{H}_{injected}$. However, it is impossible for the attacker to perfectly reproduce $\hat{H}_{injected}$ as H_{tr} (i.e., the genuine CSI) is only known to the receiver. Nevertheless, the attacker can still try to reproduce the injected signature as much as possible. For example, suppose the attacker wants to fabricate a packet with payload x and the sent signal is z. Instead of naively sending $z = x$ which makes the received link signature equal to H_{ar} (different from $H_{injected}$), the attacker can transform the fabricated signal z to $wH_{ta}x$ and add an additional delay d after the data transmission starts. The received signal of this fabricated packet at the receiver is:

$$y_{fake} = H_{ar}z + n_r = e^{-j2\pi fd} wH_{ar}H_{ta}x + n_r. \quad (4)$$

The estimated channel response of this fake packet at the receiver then becomes $\hat{H}_{fake} = e^{-j2\pi fd} wH_{ar}H_{ta} + n'_r$, which is identical to the part of the injected signature as shown in Eq. (3). Note that the attacker is able to reproduce this signature since s/he can estimate H_{ta} using the same channel probing process, so this coefficient can be intentionally multiplied to x before the fake signal is sent. Moreover, even though H_{ar} is unknown to the attacker, sending a packet from the attacker to the receiver will impose this channel response naturally because both the injected signal and the fabricated packet are sent from the same device.

3.3 Attack Analysis

As shown above, the success in fabricating packets depends on the similarity between the link signature of the fabricated packet and the packet injected during the training phase. Note that only the channel amplitude, $|\hat{H}|$, is usually used as a link signature since the signal phase is noisy and intractable [23, 26]. Thus, the dissimilarity between fabricated and injected packets can be estimated as $|\hat{H}_{fake}| - |\hat{H}_{injected}|$. If the injected signal is perfectly reproduced in the fabricated packet (i.e., no full-duplex self-interference), the difference between fabricated and injected packets depends strongly on H_{tr}, which is the genuine link signature not under the control of and unknown to AMITM. However, in the real world, there are 3 practical limitations in link-based identification, thus making it hard to detect this minor signature difference.

First, the received signal amplitude varies from packet to packet even if the packets are sent from the same transmitter due to the sampling imperfection and the automatic gain control. Thus, normalizing link signatures is necessary to extract a consistent signature at each receiver. For ex-

ample, as the unscaled CSI shown in Fig. 5(a), even CSI from different packets seems following the same distribution over subcarriers, the amplitude envelope of certain packets is 1.5x higher than the other. Thus, adjusting the amplitude of estimated CSI to a fixed range is shown to improve the accuracy of source identification [23, 26, 29]. This normalization also implies that injecting a signal with proper relay amplification w will not cause a suspicious signature change but will actually reduce the relative signature difference (after normalization), thus removing the signature dissimilarity between the fabricated and the genuine packets.

Second, even for packets sent from the same transmitter, link signatures are not exactly the same for different packets due to the receive noise caused by hardware imperfections and environmental changes. Thus, link-based source identification generally relies on thresholding the signature difference between a new received packet and the previously authorized packet for classification. Some systems set a static threshold [34, 29] while the others adjust the threshold dynamically [23, 26]. However, irrespective of the threshold selection algorithm used, it is necessary to set the threshold to be larger than the maximum link variation at the authorized transmitter. Since WiFi is designed to accommodate different link qualities (i.e., $5dB$ to $30dB$), this thresholding process provides room for AMITM to fabricate a valid link signature even when there exists a link-signature difference between fabricated and injected packets. For example, Fig 5(b) shows the link-signature difference inside two transmitters, which is measured by Euclidean distance of the normalized CSI between a newly arrived packet and the previously received packet. In this example, the threshold should be set to be greater than 2 in order to prevent the rejection of the genuine packet from $tx2$. By leveraging this property, once the link-signature difference between fabricated and injected packets is smaller than the link variation at the authorized transmitter, the receiver cannot tell if the link-signature change is caused by AMITM or the estimation errors (after normalization).

Last, the link signature changes over time even if the location of transmitter is fixed, thus providing an attack surface to inject packets even when the genuine signature is recorded. Note if AMITM is deployed before the authorized user appears, a relayed packet with a fixed large w is adequate to keep AMITM stealthy because the receiver cannot tell if the large received channel gain is caused by the relayed signal or the genuine signal itself. However, if the genuine link signature H_{tr} is already known to the receiver, injecting a signature with large w will make the receiver discover AMITM easily because the new (injected) link signature is significantly different from the link signature of previously authorized packets. In such a case, gradual injection of link signatures with an increasing value of w reduces the possibility of AMITM being discovered because the receiver doesn't know if the link signature change is caused by an injected signature or environmental change. For example, Fig. 5(c) plots the CSI difference between the packet received at time t versus the packet received at time 0. In a static environment without any environmental change (e.g., in an office during the weekend), the CSI difference remained similar to each other even after 10 minutes. However, if the trace is collected during working days (with people moving around the test location), the CSI difference increases over time, thus requiring a periodic update of the genuine link signature. By exploiting this property, injecting packets with $w(t) = 0.015t$ helps hide AMITM because the link variations over time stay in a normal range. This one-time bootstrap process takes only 4 minutes to inject a signature with necessarily large w to make 90% of fabricated packets classified as those sent from an authorized transmitter.

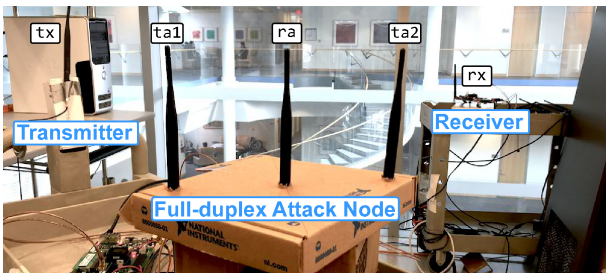
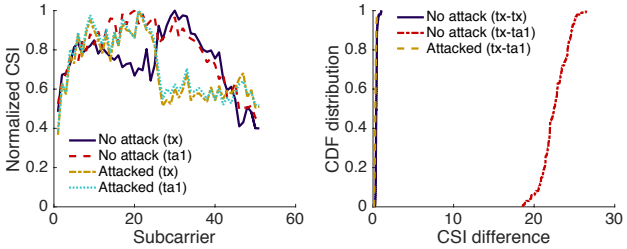


Figure 6—Prototype setup. A prototype is built with WARP software-defined radios in which the analog cancellation is achieved by proper antenna placements.

4. PROTOTYPE IMPLEMENTATION

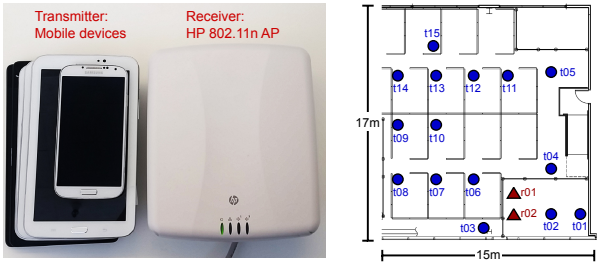
To validate the feasibility of AMITM attack, we implemented a full-duplex prototype on WARP software-defined radios [24]. WARP can control physical-layer wireless signals by its Matlab/FPGA library. The main challenge in implementing a full-duplex attack node is the cancellation of self-interference due to the relayed signals [5, 9, 10]. Instead of utilizing customized hardware, such as the analog circuit used in *FastForward* [9], we chose to build a simplified full-duplex testbed which cancels the relayed signal at the attack node by a proper antenna placement [5]. As shown in Fig. 6, in our prototype, the full-duplex attack node is equipped with 3 antennas where antennas `ta1` and `ta2` are used to relay the signal received at antenna `ra`. The relayed signals sent from `ta1` and `ta2` have a similar delay profile and attenuation at `ra` because `ta1` and `ta2` are placed half-wavelength (12.5cm) away from `ra`. Based on this placement, adding the relayed signals sent from `ta2` with a phase shift π cancels the interference caused by the relayed signal at `ra`. Note that there is an initial random phase shift between each transmitter chain in WARP [35], so it is necessary to make a one-time calibration before relaying the signal. Assuming the difference of this initial phase shift between the two transmitting antennas is ρ, the signal sent from the second transmitting antenna is compensated with a phase shift $\pi - \rho$ for the proper signal cancellation. After this analog cancellation, a standard digital cancellation that further nullifies the received baseband signal [21] is later applied.

The ms-level delay in WARP is larger than the WiFi symbol time, so it cannot relay signals in real time. In our current prototype, we use a similar method as introduced in [16] in which each packet is transmitted twice. The first packet is sent only from the transmitter `tx` and the received signal at `ra` is recorded and relayed later. In the second transmission, both the transmitter and the attack node send the original/relayed signals to the target receiver `rx`. The collected trace is later analyzed by the source-identification process as discussed earlier. This prototype might not be used as a commercial product but it suffices to demonstrate the capability of AMITM. In our current implementation, the sys-



(a) Example of CSI (b) CSI difference

Figure 7—Prototype validation. CSI difference between the attacker and the authorized transmitter is significantly reduced with AMITM enabled.



(a) Experiment Devices (b) Topology

Figure 8—Experimental setup. Two commercial APs and 15 handheld devices are deployed for collecting link signatures at different locations.

tem can provide about $23dB$ cancellation of self-interference. This cancellation is less effective than [5] because our testing environment is full of multipath reflections as shown in Fig. 6. Given a scenario that the SNR of the original packet is lower than $20dB$, our current implementation can relay the signal properly with minimal impact on the signal received at `rx`. Improvement of this prototype such as utilizing other self-interference cancellation technology is part of our future work.

Fig. 7(a) shows an example of CSI received at `ra`, where the difference between CSI of the transmitter and the attacker can be seen clearly and utilized easily for source identification [29, 34]. However, when AMITM is enabled, this dissimilarity is reduced significantly due to the injected signals. This phenomenon can also be seen in Fig. 7(b), which shows the CSI difference of 100 transmissions in an hour. For example, setting the identification threshold to 10 in this case can differentiate the transmitter from the attacker with a 100% probability, but it fails to distinguish packet sources when AMITM is enabled. These results are consistent with our system model as described in Section. 3.

As introduced earlier, the attack performance varies with scenarios and parameter settings. In the validation of our current prototype, the parameter w is set to 1 and the distances from the transmitter to the attacker and the receiver are about 1m and 4m, respectively. The testing environment is full of multipaths but there is no significant environment change over time. To understand the characteristics of AMITM in real-life scenarios, an extensive evaluation based on the CSI collected via commercial devices will be presented next.

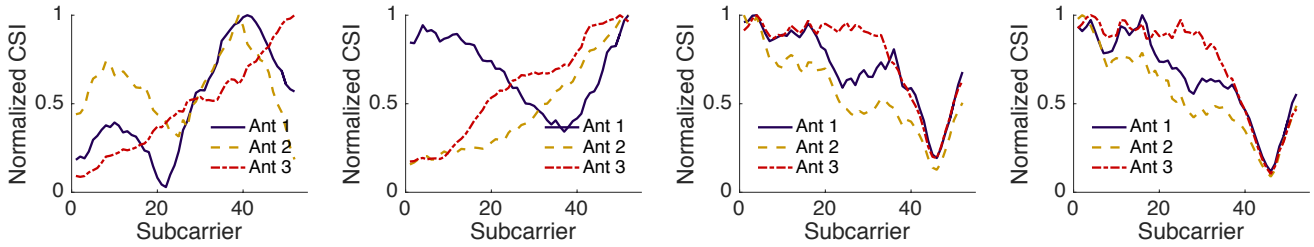
5. TRACE-DRIVEN EVALUATION

To evaluate the performance of AMITM, we simulated our attack model based on real-world CSI traces, comparing the performance of AMITM under different parameter/topology settings. The CSI information is acquired by the standard 802.11n probing process from HP E-series AP. The AP's driver is modified to send the 802.11 probing request periodically to each transmitter (i.e., tablets and smartphones) and report the estimated CSI to our server. CSI of each transmitter is recorded about every 1∼7 seconds for 2 hours. Experiments are done in the office environment as shown in Fig. 8. Both weekend (i.e., no environment change) and workday traces are collected. To the best of our knowledge, this is the first to evaluate link-based source identification through uncompressed 52-subcarrier CSI acquired from commercial APs.

After acquiring CSI from a commercial AP, link-based source identification and AMITM are simulated using Matlab. We implemented the identification algorithm based on [23, 26, 29] for the purpose of illustrating the effectiveness of AMITM. We compared the normalized CSI using Euclidean distance to the 5 previously received packets from each transmitter. If the shortest distance to previous CSI is less than a defined threshold thr, the source of current received packet is classified as sent from the transmitter which has the most similar CSI signature. Note that this scheme is just an example of link-based source identification, but it is general enough to show the effectiveness of AMITM. For example, the RSS-based authentication introduced in [36] follows a similar classification process, and it is also vulnerable to the proposed attack because the link signature is polluted stealthily in the training phase. AMITM attacks on different link-based identification schemes are discussed in Section 8.

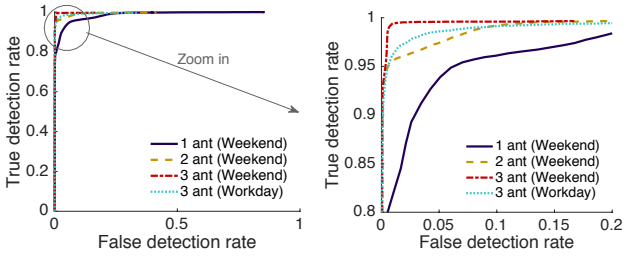
During the evaluation of AMITM, one transmuter/receiver pair is selected to simulate the full-duplex attack node. That is, the injected signal, $H_{ta}x$, is the signal received at this selected receiver and it is relayed by the selected transmitter to the AP as introduced in Section 3. The full-duplex parameter setting follows the results shown in [9, 10], where the relay delay is set to achieve the selected phase difference while satisfying the hardware limitations. Although AMITM is not yet fully implemented in real devices, the required capability of the attack node is just identical to any other full-duplex relay design, such as $FastForward$ [9], but for a different purpose. This evaluation result can be regarded as the optimal performance of AMITM since we didn't simulate the self-interference cancellation considered in our prototype. Building a large-scale testbed of AMITM with commercial devices is part of our future work.

Fig. 9 shows an evaluation example of the normalized CSI with and without AMITM. As shown in this figure, when a signal is injected as discussed earlier, the CSI difference between the attacker and the authorized transmitter is decreased significantly. As shown in Fig. 7, this result is consistent with our prototype validation. The only difference is that there are three CSIs for each transmission because the router we used is equipped with three receiving antennas, which provide higher capability to differentiate transmitters by their different link signatures. Nevertheless, once the CSI difference is smaller than the receiving noise by injecting a proper link signature, the fabricated packet has a high probability of being classified as the one sent from the authorized transmitter.



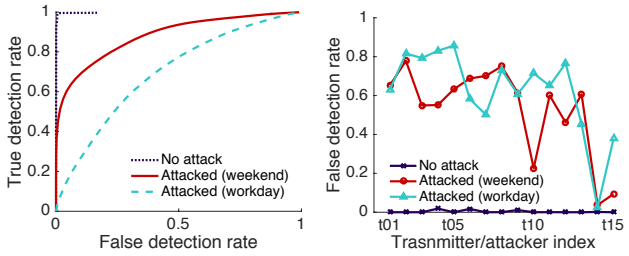
(a) Original Transmitter CSI (b) Original Attacker CSI (c) Injected Transmitter CSI (d) Reproduced Attacker CSI

Figure 9—Example of AMITM attack. Before the link signature is injected, the difference of CSI between the attacker and the victim transmitter is pronounced. Once AMITM is enabled, a fabricated packet sent from the attacker has a similar link signature as the injected packet.



(a) ROC curve w/o attack (b) Zoomed in ROC

Figure 10—Result of CSI-based identification without attack. CSI-based identification in static environments can achieve 95% TD rate with only 0.4% false detection.



(a) ROC curve w/ AMITM (b) Attacks at locations

Figure 11—Result of AMITM attack. The attacker located near the receiver has a higher probability to reproduce a link signature similar to the one injected before.

5.1 Evaluation Metric

To generalize our evaluation for different classification algorithms, we choose to plot the ROC curve of true detection rate (TD) and false detection rate (FD) under different threshold settings. We define the TD rate as the probability of the transmitter being correctly classified based on link signature while FD rate as the probability of classifying the packet sent from the other device as the attacked transmitter. In other words, FD rate represents the attack success rate that the fabricated packet is classified as sent from the authorized transmitter.

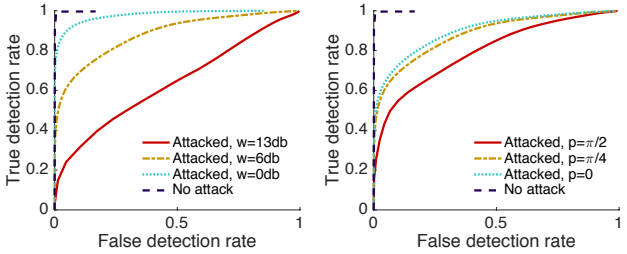
5.2 Source Identification without AMITM

We first evaluate the characteristics of link-based identification without AMITM attack. The ROC curve of distinguishing 15 clients under 2 APs is plotted in Fig. 10. When all three antennas at the receiver are used to estimate link signature, existing link-based source identification algorithms can achieve higher than 95% TD rate with only 0.4% FD rate in a steady environment (i.e., during weekends). The same test was also repeated during workdays where people move around the test location. Even in this case, link signature can still provide 93% TD rate with the same FD rate. This high TD and low FD rates indicate that link-based source identification can easily stop identity-based attacks because fabricated packets are dropped immediately once their link signatures do not match the saved database. For example, an attack utilizing known WEP/WPA vulnerability [31] to inject fake management frames will need about 200x more efforts to make fabricated packets acceptable to link-based source identification. Note that the performance of this identification can be improved further by using more antennas (i.e., more uniqueness of link signature). Our results match the performance reported in previous studies [23, 26, 29], showing the real-world benefit of link-based source identification.

5.3 AMITM Result

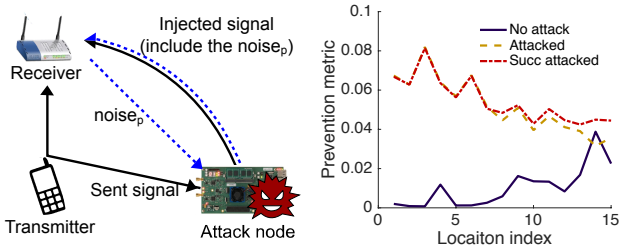
Fig. 11 shows the result based on the same weekend dataset when AMITM is set to inject signal with $w = 6dB$ and phase difference $p = \pi/4$. As shown in this figure, if the threshold is set to 2.9 for ensuring 95% authorized packets can be correctly classified, fabricated packets can also have 61% acceptance, thanks to their similar fabricated link signatures. This rate is significantly higher than the 0.4% FD rate without AMITM. On the other hand, if *thr* is reduced for ensuring 99.5% rejection of packets fabricated by AMITM, 55% of authorized packets will also be dropped, thus incurring a significant overhead. This result demonstrates the effectiveness of AMITM, making the receiver unable to differentiate packet sources by the link signature, because it is unclear if the link signature variation is caused by environmental changes or injected signals. This phenomenon is more severe in a workday dataset since the link signature variation of genuine packets increases when the environment changes.

Fig. 11(b) further decomposes the overall FD rate based on the attack mounted at different locations to `r01`. In this result, *thr* is set for ensuring 95% TD rate. As shown in this figure, for attackers located in a nearby area with a strong received gain at the receiver, such as `t02`, the attacker can have a higher probability of injecting signatures effectively because its sent signal is naturally stronger than other transmitters. In contrast, an attack from far locations such as `t14` only has limited chance to inject the link signature successfully. This result demonstrates that if the placement of attacker is carefully deployed, only $6dB$ antenna gain can make 80% of fabricated packets as sent from the authorized transmitter.



(a) ROC ($p = \pi/4$) (b) ROC ($w = 6dB$)

Figure 12—Result of attack performance with different settings. Higher the power coefficient w or the larger phase offset between original transmission and injected transmission, the better attack performance (i.e., a larger FD rate for the same TD rate).



(a) Prevention Model (b) Prevention Metric

Figure 13—Prevention result at different transmitter locations. Artificial noise has less effect on far-away transmitters but those transmitters also get less chance for mounting successful attacks.

5.4 Attack Parameter Characteristics

Changing the attack parameter w and p can achieve different performance of AMITM attack. As derived in Section 3, the success probability of AMITM attack increases when w is larger and phase difference p is close to $\pi/2$ because it helps the injected signal dominate the received link signature. However, in real-world, w cannot be increased infinitely due to hardware limitation. Thus, in our evaluation, we chose w ranging from 0 to $13dB$ because it is possible for the attacker to buy a high-gain antenna like [2] which provides about $0 \sim 13dB$ higher gain than authorized transmitter. Note handheld devices usually use low-gain (e.g., less than $2dBi$) antenna to reduce battery usage. Moreover, as shown in the previous section, if the attack node is placed appropriately, the attacker's location can provide $10dB$ higher gain than the attacked device; in this case, only a $0 \sim 6dB$ antenna gain is necessary for AMITM to successfully inject and reproduce link signatures. The current setting of AMITM is practical in showing its effectiveness. Note that an AP is unable to identify AMITM by only monitoring received signal strength because WiFi is designed to operate with a wide operation range. For example, `t01` shown in Fig. 8 naturally has $15dB$ higher gain than `t15` even without an attack. Thus, it is unable to tell whether or not the received signal is relayed and amplified.

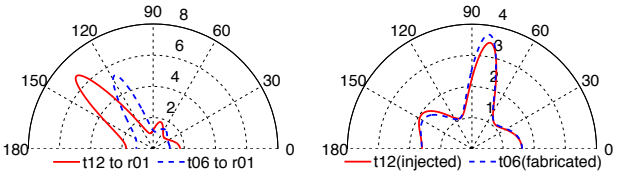
We plot the ROC curve for different w and p settings in Fig. 12. AMITM is shown to achieve the best performance when $w = 13dB$ and $p = \pi/2$, matching the derivation in Section 3. With the setting of $w = 13dB$, the receiver is unable to distinguish packets sent from the attacker and the authorized transmitter, thus making a more than 90% of packets falsely classified when TD rate is set to 95%. On the other hand, if p is tuned from $\pi/4$ to $\pi/2$, FD rate with $6dB$ antenna gain can be increased from 61% to 76%. Tuning p is relatively harder than tuning w because the attack node doesn't get cooperation from the AP. However, it is possible for the attacker to search for the best p by changing it and monitoring the attack rate.

6. COUNTERMEASURES

We have shown that AMITM can breach the existing link-based source identification once the attacker has the capability to inject proper link signatures. As shown in Section 5, this attack can be avoided naturally if the authorized transmitters are equipped with better hardware than the attackers (i.e., more antennas or a higher transmission gain). However, this is not a proper solution since it turns the link-based source identification into an arms race between the transmitters and the attackers, undoing the benefit to secure wireless systems by physical security. It can also stop AMITM by shuffling/changing the location/value of 802.11 probing sequence, as suggested in [27]. Changing this process helps prevent attackers from evaluating H_{ta}, which is essential to reproduce the injected signature as shown in Section 3. However, this method requires major modifications on both the transmitter and the receiver, thus making it impractical.

Instead of modifying the hardware of transmitter devices or changing the existing WiFi protocol, we propose a novel protection mechanism based on the same principle of signal injections to make H_{ta} unknown to the attacker. Fig. 13(a) illustrates our idea that requires only the modification at the receiver (i.e., access point). As shown in this figure, the receiver uses a spare antenna to add a random white noise, $noise_p$, during the probing process. This way, only the receiver can estimate CSI correctly because the sent noise pattern needs to be known and removed before estimating CSI [21]. A similar method was also proposed in [3, 20] for stopping eavesdropping but we use it in preventing the estimation of H_{ta}. Note the noise sent from the receiver's additional antenna need not be synchronized with transmitters since it is only used by the receiver (synced with the same internal clock) to estimate CSI.

Since H_{ta} is estimated incorrectly, if there is an attack node relaying the authorized packet, the transmitted noise will also be magnified and then forwarded. The receiver is unable to remove this changed noise from the relayed/injected packet, thus knowing the existence of AMITM by monitoring the received noise pattern. For example, if the average CSI difference (including the relayed noise) in the last 5 packets is used as a prevention metric, the result of this metric for attacked and not attacked datasets is plotted in Fig. 13(b). When the attack node is located near the receiver (e.g., `t01`∼`t05`), the difference of this prevention metric between the attacked and not attacked datasets is significant because the attacker receives and relays the strong noise sent from the receiver if it is located nearby. When the attacker is located farther away, such as `t14`, the receiver would not be able to identify the attack because the attacker receives less noise when it is farther away. However, as shown in Fig. 11(b), the attacker also gets less chance to reproduce the injected link signature at these far-away locations. If we exclude the cases where AMITM is unable



(a) Angle of Arrival (AoA) (b) Attacked AoA

Figure 14—Result of signal Angle of Arrival (AoA). AoA difference between fabricated and genuine packets is reduced when the underlying channel characteristic is injected by AMITM.

to reproduce the injected signature (assuming the TD rate is set to 95%), this prevention metric gets less confused to identify AMITM. Based on our test with the same dataset used in Section 5, the current setting of our countermeasure can identify 93% AMITM with only 7% false positives. The performance of this method can be further improved by sending a well-designed artificial noise so the relayed noise can be easily identified. In future, we would like to conduct a more in-depth study of this countermeasure. We will discuss other potential methods to stop AMITM in Section 8.

7. RELATED WORK

Physical security, also known as non-cryptographic security, is a way to provide system security by utilizing physical-layer or hardware characteristics. It offers an additional security layer with minimal overhead (i.e., no cost of key change and encryption/decryption). For example, it can hide confidential messages by adding artificial noise [3, 20] or identify wireless transmitters by hardware imperfection [4, 11, 13] and wireless link properties [29, 34, 35]. Since security is enforced by low-layer information, it is commonly believed to be able to stop traditional attacks efficiently and hard to be exploited at software layer.

Nevertheless, with the advance of radio technology, physical-layer characteristics are shown to be forged and replayed if the attacker knows the genuine physical feature at the receiver. For example, hardware imperfection, such as the modulation errors and the clock skews, can be easily estimated and forged by any nearby rogue devices since the packet (and also its hardware characteristics) is broadcasted over the air [4, 14]. Existing studies also show that once the link signature (e.g., CSI or RSS) is known, rogue devices can easily forge those channel characteristics even if the attacker is located at the different locations [8, 22, 28]. However, in reality, a genuine link signature is hard for attackers to obtain unless they get the AP's cooperation or an additional receiver is placed at a half wavelength (about 5cm) away from the AP. Unlike these studies, AMITM removes this stringent assumption and extends the attack scheme by injecting fake link signatures in the training phase.

The closest to AMITM are [17, 18, 26], all of which inject controlled signals in the training phase but for different purposes. For example, Simon *et al.* [17] proposed specialized jamming to fool the key-generation process based on channel fluctuation. Song *et al.* [18] discussed the possibility for a user to fabricate his own channel response for hiding his location change, which is unable to impersonate an authorized transmitter as shown in this paper. Hongbo *et al.* [26] discussed the same vulnerability of polluting link signature, but the attacker in their scheme was designed only to send a burst of unprocessed packets during the training phase and expect the AP to use packets sent from the attacker as the authentication reference. This scheme can be easily identified by either checking link signature distribution or monitoring the unusual traffic pattern. In contrast, AMITM injects link-based signatures stealthily and reproduces them with a high probability. An attack methodology to relay analog signals was also mentioned in [27] without its evaluation. Note that this work focuses on the practical issues of link-based source identification. Specifically, in AMITM, the attackers pretend to be legitimate transmitters by injecting link signatures. Other attacks which instead breach the confidentiality provided by physical security can be found in [30, 32, 33].

8. DISCUSSION

We have shown the potential vulnerability of link-based source identification. The proposed attack model is practical and will likely and readily be exploited in the near future. For example, the ongoing research of commercial software-defined radios [7] will make AMITM possible only by modifying the installed software.

There are several proposals of link-based source identification, but none of them has been standardized. We have demonstrated the concept of AMITM against a general scheme of CSI-based identification. However, it doesn't restrict the capability of AMITM on other systems. For example, the received angle of arrival (AoA) estimated by MUSIC algorithm represents the direction of sent signal and it can be used to identify the packet source [35]. An example of this technique is shown in Fig. 14, the packets sent from transmitters `t12` and `t06` can be easily distinguished by their AoA since the packets come from different directions. However, once the signal is injected as introduced in this paper, the receiver is unable to distinguish packet source as shown in Fig. 14(b) because the underling channel property is modified by the injected signal.

We have evaluated AMITM by building a prototype based on WARP boards and collecting the real-world CSI traces from commercial devices. Due to the hardware limitation of WARP, our current implementation is unable to operate in real time. Implementing AMITM attack in real time is helpful to characterize it further, which is part of our ongoing work. Note that there have been several real-time implementations of full-duplex relay [9, 12], which is an active area of wireless communications. AMITM attack can be implemented by utilizing/modifying any of these existing systems.

During our experiments, we found some artifacts at the estimated link signature when CSI is fabricated. For example, the side lope of estimated AoA is slightly inflated as shown in Fig. 14(b) because the injected CSI (i.e., $H_{ta}H_{at}$) violates the assumption of MUSIC algorithm. Another example is that the channel correction among antennas of fabricated packets is higher than a normal packet because the relayed channel signature, H_{ta}, is the same at all receiving antennas. This phenomena can also be found in Fig. 9 where the injected CSI at three antennas look similar to each other. Utilizing these artifacts will help us identify AMITM with less overhead, which is also part of our future work.

9. CONCLUSION

We have proposed and evaluated AMITM, a novel attack model targeting link-based source identification. It exploits the nature of shared wireless medium to inject semi-controlled link signatures in the training phase and reproduces a similar link signature in fabricated packets. Our evaluation based on commercial APs shows AMITM can make 90% more fabricated packets classified as sent from authorized transmitter. A countermeasure against AMITM is also proposed without modifying the transmitter's hardware and existing wireless protocols.

10. REFERENCES

[1] A Closer Look at Wireless Intrusion Detection : How to Benefit from a Hybrid Deployment Model Introduction. Aruba White Paper.

[2] TP-Link 15dBi Omni-directional Antenna. http://www.tp-link.com/en/products/details/cat-5063_TL-ANT2415D.html.

[3] N. Anand, S.-J. Lee, and E. Knightly. Strobe: Actively securing wireless communications using zero-forcing beamforming. In *Proc. of IEEE INFOCOM '12*, pages 720–728.

[4] C. Arackaparambil, S. Bratus, A. Shubina, and D. Kotz. On the reliability of wireless fingerprinting using clock skews. In *Proc. of ACM WiSec '10*, pages 169–174.

[5] E. Aryafar, M. A. Khojastepour, K. Sundaresan, S. Rangarajan, and M. Chiang. Midu: Enabling mimo full duplex. In *Proc. of ACM Mobicom '12*, pages 257–268.

[6] P. Bahl and V. Padmanabhan. Radar: an in-building RF-based user location and tracking system. In *Proc. of IEEE INFOCOM '00*, pages 775–784 vol.2.

[7] M. Bansal, J. Mehlman, S. Katti, and P. Levis. Openradio: A programmable wireless dataplane. In *Proc. of ACM HotSDN '12*, pages 109–114.

[8] P. Baracca, N. Laurenti, and S. Tomasin. Physical layer authentication over mimo fading wiretap channels. *Wireless Communications, IEEE Transactions on*, 11:2564–2573, 2012.

[9] D. Bharadia and S. Katti. Fastforward: Fast and constructive full duplex relays. In *Proc. of ACM SIGCOMM '14*, pages 199–210.

[10] D. Bharadia, E. McMilin, and S. Katti. Full duplex radios. In *Proc. of ACM SIGCOMM '13*, pages 375–386.

[11] V. Brik, S. Banerjee, M. Gruteser, and S. Oh. Wireless device identification with radiometric signatures. In *Proc. of ACM MobiCom '08*, pages 116–127.

[12] B. Chen, Y. Qiao, O. Zhang, and K. Srinivasan. Airexpress: Enabling seamless in-band wireless multi-hop transmission. In *Proc. of ACM MobiCom '15*, pages 566–577.

[13] B. Danev and S. Capkun. Transient-based identification of wireless sensor nodes. In *Proc. of IEEE IPSN '09*, pages 25–36.

[14] B. Danev, H. Luecken, S. Capkun, and K. El Defrawy. Attacks on physical-layer identification. In *Proc. of ACM WiSec '10*, pages 89–98.

[15] M. Demirbas and Y. Song. An rssi-based scheme for sybil attack detection in wireless sensor networks. In *Proc. of WoWMoM '06*, pages 5 pp.–570.

[16] M. Duarte, C. Dick, and A. Sabharwal. Experiment-driven characterization of full-duplex wireless systems. *Wireless Communications, IEEE Transactions on*, 11(12):4296–4307, 2012.

[17] S. Eberz, M. Strohmeier, M. Wilhelm, and I. Martinovic. A practical man-in-the-middle attack on signal-based key generation protocols. In *ESORICS '12*, volume 7459, pages 235–252. Springer.

[18] S. Fang, Y. Liu, W. Shen, and H. Zhu. Where are you from?: Confusing location distinction using virtual multipath camouflage. In *Proc. of ACM MobiCom '14*, pages 225–236.

[19] A. Francillon, B. Danev, S. Capkun, S. Capkun, and S. Capkun. Relay attacks on passive keyless entry and start systems in modern cars. In *NDSS '11*.

[20] S. Goel and R. Negi. Guaranteeing secrecy using artificial noise. *IEEE Transactions on Wireless Communications*, 7(6):2180–2189, 2008.

[21] S. Gollakota, S. D. Perli, and D. Katabi. Interference alignment and cancellation. In *Proc. of the ACM SIGCOMM 2009*, pages 159–170.

[22] X. He, H. Dai, W. Shen, and P. Ning. Is link signature dependable for wireless security? In *Proc. of IEEE INFOCOM '13*, pages 200–204.

[23] Z. Jiang, J. Zhao, X.-Y. Li, J. Han, and W. Xi. Rejecting the attack: Source authentication for wi-fi management frames using csi information. In *Proc. of IEEE INFOCOM '13*, pages 2544–2552.

[24] A. Khattab, J. Camp, C. Hunter, P. Murphy, A. Sabharwal, and E. W. Knightly. Warp: A flexible platform for clean-slate wireless medium access protocol design. *SIGMOBILE Mob. Comput. Commun. Rev.*, 12:56–58, 2008.

[25] C. Liu and J. Yu. Rogue access point based dos attacks against 802.11 wlans. In *Proc. of IEEE AICT '08*, pages 271–276.

[26] H. Liu, Y. Wang, J. Liu, J. Yang, and Y. Chen. Practical user authentication leveraging channel state information (csi). In *Proc. of ASIA CCS '14*, pages 389–400.

[27] Y. Liu and P. Ning. Enhanced wireless channel authentication using time-synched link signature. In *Proc. of IEEE INFOCOM '12*, pages 2636–2640.

[28] Y. Liu and P. Ning. Poster: Mimicry attacks against wireless link signature. In *Proc. of ACM CCS '11*, pages 801–804.

[29] N. Patwari and S. K. Kasera. Robust location distinction using temporal link signatures. In *Proc. of ACM MobiCom '07*, pages 111–122.

[30] M. Schulz, A. Loch, and M. Hollick. Practical known-plaintext attacks against physical layer security in wireless mimo systems. In *Proc. of NDSS '14*.

[31] E. Tews and M. Beck. Practical attacks against wep and wpa. In *Proc. of ACM WiSec '09*, pages 79–86.

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

[33] Y.-C. Tung, S. Han, D. Chen, and K. G. Shin. Vulnerability and protection of channel state information in multiuser mimo networks. In *Proc. of ACM CCS '14*, pages 775–786, 2014.

[34] L. Xiao, L. Greenstein, N. B. Mandayam, and W. Trappe. Using the physical layer for wireless authentication in time-variant channels. *Wireless Communications, IEEE Transactions on*, 7:2571–2579, 2008.

[35] J. Xiong and K. Jamieson. Securearray: Improving wifi security with fine-grained physical-layer information. In *Proc. of ACM MobiCom '13*, pages 441–452.

[36] K. Zeng, K. Govindan, D. Wu, and P. Mohapatra. Identity-based attack detection in mobile wireless networks. In *Proc. of IEEE INFOCOM, '11*, pages 1880–1888.