Privacy-Preserving Location Authentication in Wi-Fi Networks Using Fine-Grained Physical Layer Signatures

Wei Wang, *Student Member, IEEE*, Yingjie Chen, and Qian Zhang, *Fellow, IEEE*

***Abstract*—A recent measurement reveals that a large portion of the reported locations is either forged or superfluous, which raises security issues such as bogus alibis and illegal usage of restricted resources. However, most prior approaches leak users' location information or rely on external devices. To overcome these limitations, we propose *PriLA*, a privacy-preserving location authentication system that verifies users' location information based on physical layer (PHY) information available in legacy Wi-Fi preambles. The crux of PriLA is to turn detrimental features in wireless systems, namely carrier frequency offset (CFO) and multipath, into useful signatures for privacy protection and authentication. In particular, PriLA exploits CFO and channel state information (CSI) to secure wireless transmissions starting from the handshake phase between mobile users and the access point (AP), and meanwhile verify the truthfulness of users' reported locations based on users' multipath profiles. We have implemented PriLA on GNURadio/USRP platform and commercial off-the-shelf Intel 5300 NICs, and the experimental results show that PriLA achieves the authentication accuracy of 93.2% on average, while leaking merely 45.7% information in comparison with the state-of-the-art approach.**

***Index Terms*—Location authentication, location privacy, physical layer information.**

I. Introduction

DRIVEN by the proliferation of Wi-Fi hotspots in public places, location-based services (LBSs) have experienced surging development in recent years. LBSs take advantage of users' location information to provide personalized or contextual services. A typical LBS system consists of an LBS provider who offers services based on users' physical locations via trusted Access Points (APs), and mobile users who request specific service along with their own location and identity (ID) information.

Manuscript received November 6, 2014; revised March 22, 2015; accepted October 1, 2015. Date of publication October 6, 2015; date of current version February 8, 2016. The research was supported in part by grants from 973 project 2013CB329006, China NSFC under Grant 61502114, China NSFC under Grant 61173156, RGC under the contracts CERG 622613, 16212714, HKUST6/CRF/12R, and M-HKUST609/13, as well as the grant from Huawei-HKUST joint lab. The associate editor coordinating the review of this paper and approving it for publication was Majid Manteghi.

W. Wang is with the School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan, China, and also with the Fok Ying Tung Research Institute, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong (e-mail: gswwang@cse.ust.hk).

Y. Chen and Q. Zhang are with the Department of Computer Science and Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2015.2487453

Unfortunately, a recent measurement study [1] on Foursquare check-ins reports that there exists a large amount of forged and superfluous location data uploaded by mobile users. One major reason behind this phenomenon is that users concern about their location privacy and use synthetic traces to replace their true locations. The forged traces incur significant discrepancies that mislead the applications driven by the location data. What is worse, by forging location reports, malicious users can abuse services like illegally accessing restricted resources and creating bogus alibis.

To avoid location forgery, an essential step is location authentication, which verifies the truthfulness of the reported location data. An intuitive approach is to equip provider with localization capability, which, however, falls short due to the following two limitations. First, there are places such as coffee shops and stores where the number of provider-trusted APs is not enough to perform localization. Second, the growing privacy threats of sharing location information via LBS have been widely concerned [2]. Such privacy threats come from the fact that many untrusted Wi-Fi infrastructures aggressively collect the location data. Although mobile users can secure their location data via encryption, their location information is still at high risk of being leaked due to the broadcast nature of wireless medium. Adversary can easily infer the targeted user's physical location by collaboratively eavesdropping frames over the air from several sniffers (e.g., untrusted APs). Previous research [3], [4] shows that one can determine a node with meter/submeter level resolution using several receivers. Being aware of such risks, mobile users may be reluctant to use LBS applications. Note that existing location privacy preserving approaches cannot be directly integrated into location authentication systems, since hiding mobile users' location information would also prevent the LBS provider from authenticating them. Therefore, it is crucial to enable location authentication without compromising users' location privacy.

Despite growing attempts and extensive efforts, it is still challenging to facilitate privacy-preserving location authentication in Wi-Fi networks. State-of-the-art solutions either fail to consider users' privacy concerns [5]–[8] or rely on dedicated hardware or external networks for authentication [5], [9], while the capability of privacy-preserving location authentication within the LBS system is missing. External hardware or devices assisted authentication results in high start-up costs, and cannot be implemented using existing LBS infrastructures.

The target of this paper is to fill the above gaps: we argue that privacy-preserving location authentication can be realized within existing Wi-Fi-based LBS systems by exploiting physical layer (PHY) signatures in Wi-Fi preambles. To achieve this goal, this paper introduces *PriLA*, a **Pri**vacy-Preserving **L**ocation **A**uthentication system in orthogonal frequency-division multiplexing (OFDM) based Wi-Fi networks (e.g., IEEE 802.11a/g/n/ac). This system allows the LBS provider to successfully conduct authentication while and meanwhile guaranteeing location privacy preservation for all mobile users against adversaries. To this end, the following requirements should be satisfied. First, to defend against adversaries with localization capability, the users' IDs should be fully protected starting from the handshake (or association) phase. Otherwise, the adversaries can infer a user's location by analyzing the signal strength [3] or anger-of-arrival (AoA) information [4] extracted from the user's frames. Second, the provider should be able to verify users' locations even when there is not enough APs to perform localization.

To overcome the above predicaments, PriLA exploits carrier frequency offset (CFO) and multipath, which can be obtained via Wi-Fi preambles. In communication systems, CFO and multipath are considered to be detrimental, while PriLA leverages them for authentication and privacy-preservation. PriLA takes advantage of the channel reciprocity property, and uses CFO together with channel state information (CSI) to generate CFO patterns that are exclusively known by the transmission pair. In particular, to defend against adversaries with localization capability, PriLA uses CFO pattern to secure users' IDs starting from the handshake (or association) phase. As such, the adversaries cannot link a frame to a certain user, or infer the presence of a user, and thus fail to localize a user via localization. To enable authentication without performing localization, PriLA leverages users' multipath profiles, which can be extracted from CSI using multiple antennas. In addition, the multipath profiles are reliable as it is determined by the environment's physical layout and hard to forge. As reported in [10], users in proximity share similar multipath profiles. Thus, the LBS provider can verify the reported location information through comparing users' multipath profiles.

The main contributions of this paper are summarized as follows. First, We propose a detailed design for privacy-preserving location authentication in Wi-Fi networks without assistance from extra hardware or networks. Second, we leverage the CFO and CSI information to secure the transmissions between users and the provider. The proposed security technique leaks merely 45.7% information compared to the state-of-the-art approach. Third, we prototype PriLA using GNURadio/USRP testbed [11] and off-the-shelf Intel 5300 NICs [12].

II. System Model

Fig. 1 depicts a typical LBS system architecture, which consists of an LBS provider, mobile users, and adversaries. In an LBS system, a mobile user requests service from the LBS provider by reporting the user's location information with its ID to the trusted AP, which connects to the LBS servers via a secured backhaul. In this paper, we assume that a user's ID is its MAC address, or its ID can be inferred from its MAC address.

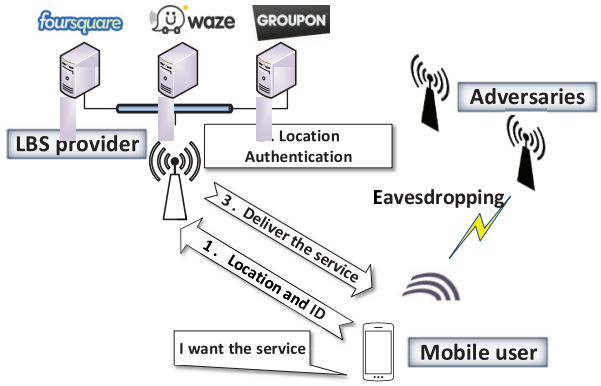


Fig. 1. System architecture of location-based service in Wi-Fi networks.

As assumed in many existing location privacy preservation proposals [13]–[15], the mobile user only reports coarse location information to preserve privacy. Based on the frames sent by users, the LBS provider checks the truthfulness of the location information. Only when the reported information is confirmed to be true, the LBS provider delivers the service to the mobile user via downlink transmission from the trusted AP.

Hardware Impairments. In a typical wireless communication system [16], the signal to be transmitted is upconverted to a high frequency carrier prior to transmission. The receiver is expected to tune its frequency to the same carrier frequency for downconverting the signal to baseband, prior to demodulation. However, due to impairments of RF chipsets, the carrier frequency of the receiver is impossible to be exactly same as the carrier frequency of the transmitter. Hence in practice, the received baseband signal, instead of being centered at DC (0 Hz), will be centered at a frequency offset Δf, where

$$\Delta f = fc_{TX} - fc_{RX}, \tag{1}$$

The representation of received baseband signal is (ignoring the noise)

$$r(t) = x(t) * e^{\frac{j2\pi\Delta f t}{F_s}}, \tag{2}$$

where $x(t)$ denotes the transmitted signal, $r(t)$ the received signal, and F_s the sampling frequency. In the single carrier case, this equation can be further interpreted as

$$r(t) = A(t)e^{j\theta(t)} * e^{\frac{j2\pi\Delta f t}{F_s}} = A(t) * e^{j(\theta(t)+\frac{2\pi\Delta f t}{F_s})}, \tag{3}$$

where $A(t)$ and $\theta(t)$ are the magnitude and phase components of the received signal respectively. It is obvious that the frequency offset will cause the received symbol suffering from phase rotation depending of the sampling time t and the amount of Δf. In multiple carrier modulation like OFDM system, the impact of CFO becomes more complicated. Large CFO not only causes phase offset in received symbol, but also introduces amplitude reduction of desired subcarrier, which will largely degrade the decoding signal-to-noise ratio (SNR).

Threat Model. We consider adversaries who are interested in tracking the location of mobile users. We assume that an adversary is computationally unconstrained, and can eavesdrop and analyze all frames over the air in Wi-Fi networks. An adversary can be either an external node or a compromised mobile

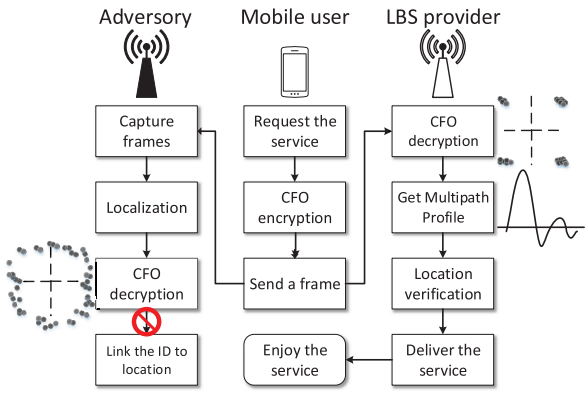


Fig. 2. The flowchart of PriLA.

user. Adversaries are assumed to be equipped with multiple antennas, and there may be multiple adversaries that collude together to locate mobile users using existing localization techniques (e.g., CSI-based [3] or AoA-based [4] approaches). To this end, adversaries first identify the mobile user's frames, and then use CSI or AoA information of the frames obtained at multiple eavesdroppers to determine the user's location. We do not consider active adversaries that perform active channel jamming, mobile worm attacks, or other denial-of-service (DoS) attacks, as these attacks cannot be used to compromise user's location privacy.

III. Design Overview

The crux of PriLA is to facilitate the LBS provider to authenticate users' location by exploiting multipath profiles while preserving mobile users' location privacy by encrypting the location reports using fine-grained PHY information. The LBS server and a mobile user follow the protocol described in Fig. 2. First, the mobile user requests the service by exchanging handshake frames with the provider's AP. Then, both the mobile user and the provider extracts CSI and CFO information from the preambles to generate a secret key, which is used to encrypt the following frames sent by the user. After receiving the encrypted frames, the provider decrypts the frames using the CSI and CFO information obtained in the handshake frames, and then extracts the user's ID (MAC address) and location information from the frames. Afterwards, the provider uses the CSI obtained from the user's frames to construct a multipath profile, which is compared with multipath profiles stored at the provider for location authentication. After verifying the truthfulness of the reported location, the provider delivers the service to the user.

IV. CFO Encryption Using Wi-Fi Preamble

A. Exploiting PHY Signatures

Recall that a Wi-Fi receiver always suffers from the CFO when downconverting the signal to baseband due to the hardware impairments. Specifically, CFO not only results in a loss in SNR, but also creates inter carrier interference (ICI), which can severely degrade the receiver's decoding performance. Inspired by this observation, we propose the CFO encryption technique that leverages this inherent feature of CFO to thwart adversaries from obtaining users' locations. The basic intuition behind CFO encryption is to inject an intended CFO pattern to each frame sent by the user. As the CFO pattern is injected at PHY, both the header and data are protected. Without the knowledge of the CFO pattern, the adversaries are unable to obtain the user's ID or decode the frame.

Moreover, a wireless channel is reciprocal, and cannot be estimated by a node whose distance with the sender/receiver is larger than half the wavelength of the transmitted signal [19]. Such a property of wireless channel can be leveraged to generate secret CFO patterns that are privately shared between the sender and the receiver.

However, to leverage the above PHY signatures for location privacy preservation, we have the following challenges. First, to fully protect a user's location privacy, adversaries should learn nothing about the user's ID or location from the first frame (i.e., handshake frame) that a user sends to the provider. However, existing PHY security techniques [20]–[22] are primarily designed to secure the data transmission of subsequent frames after the handshake frame, which leaves the header of the handshake frame exposed to the adversaries. As such, the adversaries can extract the user's MAC address and localize the user based on the CSI estimation. Hence, a special designed protocol is required to protect the handshake frame. Second, the CFO pattern encoding should be as robust as possible to ensure effective communications between users and the provider, while at the same time as efficient as possible so that we can generate as many secret bits as possible in each frame to minimize the number of handshakes. However, due to RF impairments and local interference, the estimated CSI values at the sender and the receiver are not exactly the same. Directly extracting bits from CSI (e.g., threshold-based approach [21]) can be very efficient, while it is not robust to local interference or hardware impairments. Thus, a careful investigation into the impact of local interference and hardware impairments should be conducted to devise a robust and efficient coding scheme.

B. Securing The Handshake

To prevent the adversaries localizing the user at handshake phase, the MAC address of each frame sent by the user should be kept as private information between the user and the provider. To this end, we leverage the CFO signature to secure the handshake protocol. As discussed earlier, the CFO of a link is unique and can be obtained using existing Wi-Fi preamble. Such appealing features make CFO ideal for user identification at handshake phase. In particular, the user sends a frame to the provider's AP to request service. As depicted in Fig. 3, the request frame sets the transmitter address as "NULL" to hide the user's MAC address from adversaries. The provider extracts the CFO and CSI information from the preamble, and maintains a mapping $CFO_u \rightarrow CSI_u$ for a user u. Then, the provider returns an acknowledge frame (ACK) to the user. The user extracts the CSI information from the ACK, and uses the CSI information to encrypt the subsequent frames. The provider first extracts the CFO, and then finds the matched CSI information based on previously logged mappings. Due to reciprocity of a wireless link, the CFO and CSI information obtained by

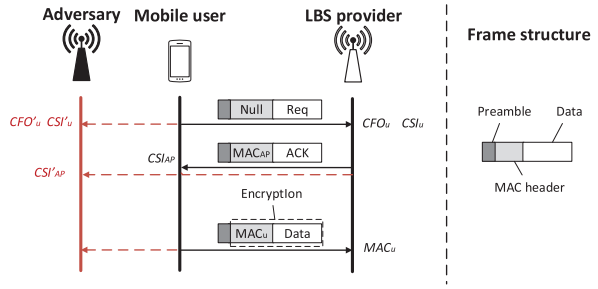


Fig. 3. Secure handshake protocol.

the user and the provider is (theoretically) identical. Therefore, the provider can use the CFO and CSI information obtained at the AP side to decrypt the frame. On the other hand, even if the adversaries can eavesdrop all frames sent by the user and the provider, they cannot acquire the MAC address of the user. Through eavesdropping, the adversaries can estimate the CFO and CSI information of the adversary-user link and the adversary-provider link. However, the adversaries are not able to use such information to decrypt the subsequent frames sent by the user, as they cannot infer the CSI information of the user-provider link based on CSI information of other links. Without the knowledge of a frame's ID, the adversaries cannot infer a user's presence, or link a user to a certain location. Hence, the user's location privacy is fully preserved according to the notion of privacy [23], [24].

C. CFO Encoding

A key technique that ensures the secure handshake phase is to use the channel reciprocity for encryption. Ideally, the estimated CSIs are identical at both ends of a link. However, there are discrepancies caused by hardware imperfection and local interference. In practice, the CSI estimation at one side are normally a shifted, enlarged, or shrinked version of the CSI estimated by the other side. To alleviate the impact of those discrepancies, we leverage the fact that the shifting, enlarging, or shrinking mainly affect the amplitude of CSI rather than its curve shape [20]. To save most of the information stored in the CSI curve, we employ a two-layer differential coding scheme.

Two-layer differential coding (TLDC). The core idea of the proposed coding scheme is to extract the first and second order derivatives of the curve simultaneously, and map the derivatives into secret bits. Both the user and the provider independently executes Algorithm 1 to encode the CSI curve, which can be represented as a vector consisting of CSI values in all subcarriers. The CSI vector is divided into multiple buckets of equal length. Then, we map each bucket to a predefined pattern for encoding. In particular, we define four curve patterns, i.e., descending trend with decreasing gradient, descending trend with increasing slope, ascending trend with decreasing slope, and ascending trend with increasing slope. Those curve patterns can be determined using the first and second order derivatives. Such derivative-based encoding can resist the impact of shifting. To alleviate the zooming effect caused by hardware imperfection, we set the gradient of each predefined pattern according to the CSI variance of its own

Algorithm 1. Two-Layer Differential Coding (TLDC)

Input: CSI vector $[c_1, \ldots, c_n]$; bucket size L
Output: Secret key $\mathbf{k}$
I. Initialization
Initialize $\mathbf{k}$ as an empty vector: $\mathbf{k} \leftarrow [\ ]$;
Compute the differential CSI vector $[d_1, \ldots, d_{n-1}]$, where $d_i = c_{i+1} - c_i, \forall i = 1, \ldots, n$;
Put the differential CSI values into different buckets one by one following the rule: $d_i \rightarrow$ the $\lceil i/L \rceil$th bucket;
Find the maximal and minimal differential CSI values d_{max}, d_{min};
Generate four shape pattern vectors $\mathbf{v_{00}} = [\frac{d_{min}}{n}, \ldots, \frac{id_{min}}{n}, \ldots, \frac{Ld_{min}}{n}]$, $\mathbf{v_{01}} = [\frac{Ld_{min}}{n}, \ldots, \frac{id_{min}}{n}, \ldots, \frac{d_{min}}{n}]$, $\mathbf{v_{10}} = [\frac{d_{max}}{n}, \ldots, \frac{id_{max}}{n}, \ldots, \frac{Ld_{min}}{n}]$, $\mathbf{v_{11}} = [\frac{Ld_{max}}{n}, \ldots, \frac{id_{max}}{n}, \ldots, \frac{d_{max}}{n}]$;
II. Key Generation
for each bucket **do**
  Compute Fréchet distances between the bucket and the four shape pattern vectors;
  Find the vector $\mathbf{v_i}$ with the smallest distance;
  Add the corresponding bits to $\mathbf{k}$: $\mathbf{k} = [\mathbf{k}, \mathbf{i}]$;
end for
Return $\mathbf{k}$;

Algorithm 2. CFO Injection

Input: Secret key $\mathbf{k}$; inherent CFO Δf; CFO injection range $[f_l, f_u]$; number of symbols in the frame S
Output: Encrypted frame;
Generate a vector of CFOs of length $\lfloor \frac{2n}{ML} \rfloor$ by multiply each M bits of $\mathbf{k}$ with Δf;
Hash each generated CFO to $[f_l, f_u]$;
for $i \leftarrow 1$ to S **do**
  Compute the index j of CFO used for injection: $j = i \mod \lfloor \frac{2n}{ML} \rfloor$;
  Inject the jth CFO value to the ith symbol;
end for

received signals. Specifically, the gradient of the ascending and descending trends is specified to be $\frac{d_{max}}{n}$ and $\frac{d_{min}}{n}$, respectively, where d_{max}, d_{min} are the maximal and the minimal elements in the differential CSI vector, respectively. As a bucket may not perfectly match one of the predefined patterns, we map each bucket to the most similar pattern, and generate secret bits using the indices of the mapped pattern, as described in Algorithm 1. The similarity between a bucket and a curve pattern is measured using the discrete Fréchet distance [25], which is defined to be the minimum length of a leash required to connect two spots who follow two separate paths.

CFO Injection. After generating a secret key using the CSI curve, we leverage the secret key $\mathbf{k}$ to encode a CFO pattern for encryption. Algorithm 2 summarizes the CFO injection process. The CFO pattern is determined by the multiplication results of the inherent CFO Δf and the private key $\mathbf{k}$.

Concretely, we first multiply each M bits of $\mathbf{k}$ with Δf to generate $\lfloor \frac{2n}{ML} \rfloor$ CFOs. Then, we hash each generated CFO to a predefined CFO injection range $[f_l, f_u]$. As such, we derive a sequence containing $\lfloor \frac{2n}{ML} \rfloor$ hashed CFOs. Finally, we inject the jth CFO into the ith symbol, where $j = i \mod \lfloor \frac{2n}{ML} \rfloor$. The mobile user will repeat these three processes until the end of the frame. Since $\mathbf{k}$ is the private message merely shared between the communication pair, the adversaries have no way to guess the generated CFO pattern.

V. Multipath-Based Location Authentication

A. Design Rationale

One might think using localization techniques to verify the location of a user. However, this is infeasible in practice, as the scenarios of adopting user-reported location are when the provider or infrastructure is unable to identify the user's location by themselves. This is because existing localization techniques either require multi-AP cooperation or modifications of existing infrastructure. To solve this predicament, the target of the proposed location authentication is to verify the location of a user based on the information that is already available in existing Wi-Fi infrastructure.

Our observation is that the signals emitted by nearby users propagate along closer paths in indoor environments where there are multiple reflectors and scatters. Moreover, the multipath profile is hard to forge as it is determined by the environment's physical layout. With these two merits, the LBS provider can determine which areas the mobile user belongs to, while such coarse-grain information is enough to help authenticate but not comprises user's location privacy. The remaining questions are how to obtain multipath profiles and how to exploit these information to conduct authentication.

B. Multipath Profile Matching

Antenna array can be used to construct the multipath profile based on arrival angle of received signal [10]. The basic idea is to measure the power of different paths coming different directions by steering antenna beam across 180°. After acquiring the multiple profile of a user, the LBS provider needs to compare it with the existing multipath profiles of users who have already been authenticated. However, even in the same zone, two points only apart from few meters will not hold the exactly same profiles due to the channel noise and the spatial gap. Hence, simple correlation between two profiles does not work. To address this issue, we observe that although two profiles may experience scale variation and misalignment, the underlying shapes remain stable. We leverage Dynamic Time Warping (DTW) [26] to cope with the impact of local shifts. In our design, the core idea is trying to extract the similarity between two misaligned profiles. DTW tries to find a path that minimizes the overall cost of the continuous mapping pairs. The cost of each mapping pair is defined to be the Euclidean distance between two points. To find the shortest path between two multipath profiles, DTW looks for a path starting from the bottom left cell to the top right cell, and computes intuitive distance between two curves. The cost of the path between two series is normalized by the length of the path. If the similarity is still very low after DTW calculation, we treat these two profiles coming from users at different locations.

C. Countermeasures For User Collusion

The above discussion focuses on the case of a single dishonest user. Now we consider the threat of user collusion, that is, multiple dishonest users collaboratively report bogus locations. In particular, multiple co-located users may collude to report the same bogus location. As such, the provider may consider that their location reports are consistent with their multipath profiles. In line with a common practice in collusion-resistance protocols [9], [27], [28], we make an assumption that the number of dishonest or collusive users is no more than a fraction of the total number of users, which is referred to as the threshold. As such, the AP can leverage the non-collusive users to verify a user's location by comparing their multipath profiles. Specifically, when a user reports its location loc_u to the provider, the provider compares the multipath profiles of the user with that of multiple users whose reported location is within a certain range from loc_u. The number of users in the comparisons is set to exceed the collusion threshold. If their multipath profiles are similar enough, the user location is proved to be true. Otherwise, the user fails to pass the location authentication.

VI. System Implementation

PriLA can be realized in the existing OFDM PHY with no change in hardware. We have implemented the prototype of CFO encryption atop the OFDM structure of GNURadio/USRP platform. We implement the entire CFO encryption design specified in Section IV directly in the USRP Hardware Drive (UHD). All the PHY parameters conform to PHY layer convergence procedure (PLCP) format of IEEE 802.11. We use DELL Optiplex 9010 with Intel i3 Dual-core processor and 4GB memory for the testbed setup. Nodes in our experiments are equipped with RFX2450 daughterboards as RF frontend, which is configured to operate in the 2.4–2.5 GHz range. The frame synchronization and channel equalization algorithms are implemented according to IEEE 802.11a/n. Due to hardware limitations of USRP, we turn to Intel 5300 NIC for multipath profile construction.

VII. Experimental Evaluation

The layout of the experimental environments is sketched in Fig. 4, where Lab1 has 4 desks and Lab2 consists of 36 cubics. We conduct experiments on different days during work hours. There were 4 and 36 students in Lab1 and Lab2, respectively, and most of them sat in front of their desks, while only a few students were walking during experiments. Such movements cause certain levels of mismatch in channel reciprocity, but the impact on the performance of PriLA is small, as shown in our results in the following subsections.

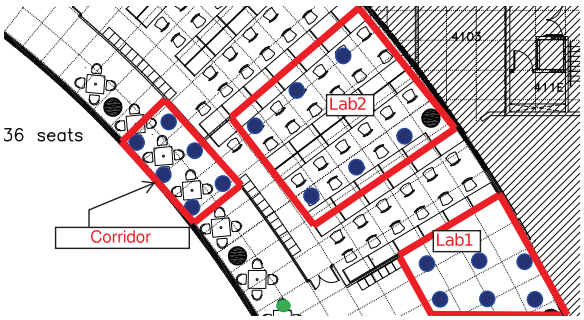


Fig. 4. Testbed layout with three zones, lab1, lab2 and corridor. The green spot is the position where the LBS provider is placed. The blues spots are the position where the mobile users are placed.

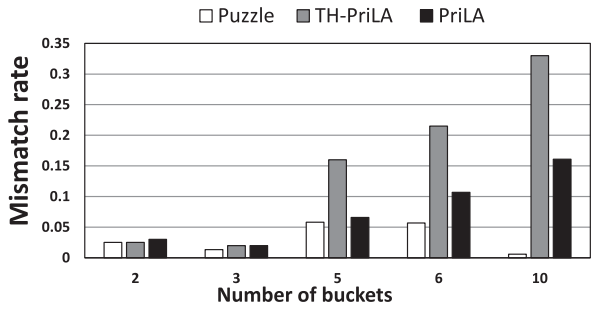


Fig. 5. Mismatch rate under different numbers of buckets.

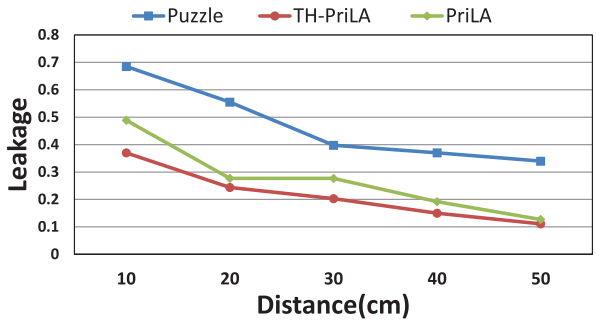


Fig. 6. Information leakage to the adversary with different distances.

A. *Performance of CFO Encryption*

We evaluate the performance of secure handshake protocol using three USRP2 nodes. One acts as a mobile user. Unless otherwise stated, the other two are both placed 5 meters away from the mobile users, acting as the LBS provider and the adversary, respectively. The adversary merely acts as a passive eavesdropper that aims to decode the user's frames for localization purpose.

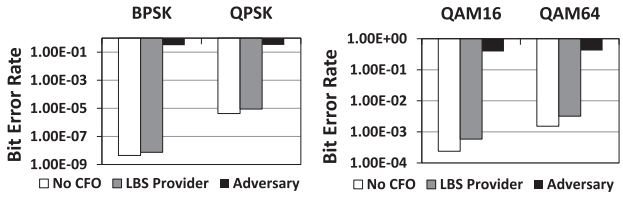
Evaluation metrics. We use three metrics, i.e., *mismatch rate*, *entropy*, and *leakage*, to evaluate the performance of CFO encryption. Mismatch rate is defined to be ratio of mismatched bits between the secret keys independently generated by the user and the provider. Mismatch rate measures the robustness of the encryption scheme. Entropy is the average amount of information contained in a message. For a random variable X, its entropy is defined to be $H(X) = -\sum_{i=1}^{n} \Pr[x_i] \log_2 \Pr[x_i]$, where $\Pr[x_i]$ is the probability of X's possible value x_i. Here, we use entropy to measure the uncertainty of the generated secret bits. In our evaluation, we compute the entropy of the curve patterns used for encryption. The probability of each curve pattern is computed by counting the its frequency in repeated experiments. The secret bits with higher entropy contain more information, and are harder for the adversary to infer. Leakage measures the amount of information learned by the adversary. In our evaluation, leakage is defined to be the ratio of matched bits between the sender (the user or provider) and the adversary. An encryption scheme with lower leakage is more secure.

Baselines. To evaluate the performance gain of the proposed CFO encryption, we compare it with two baselines. The first baseline is Puzzle [20], which the only-known secret key generation scheme that extracts bits from the curve shape of a channel's frequency response. *Puzzle* generates bits by mapping each segment of the power spectral density to ascending, descending, or steady shapes. For fair comparison, Puzzle is modified to use the same secure handshake protocol as used in PriLA. Another baseline named *TH-PriLA* adopts all the same techniques used in PriLA, except that TH-PriLA uses threshold-based approach to map each bucket to one of the four predefined shapes.

To verify the robustness of the CFO encryption, we measure the mismatch rate of different schemes in Fig. 5. PriLA achieves comparable mismatch rate with Puzzle when the number of buckets is no more than 5, while the mismatch rate of PriLA is significantly higher than that of Puzzle when the number of buckets grows to 10. The reason is that when the number of buckets is large, the entropy of Puzzle is quit small, implying low uncertainty in the bits generated by Puzzle. Hence, the security level of Puzzle in the case of large number of buckets is low. Moreover, PriLA outperforms TH-PriLA in mismatch rate while enjoying comparable entropy, which implies that TLDC is more robust than the threshold-based approach.

To validate the security level provided by PriLA, we conduct experiments where the user and the provider are placed at a fixed distant (5 m) while the adversary is placed at various distances apart from the sender. The number of buckets is fixed to be 5. As shown in Fig. 6, more information is leaked to the adversary with smaller distance. This is quit intuitive as nearer adversary shares more similar multipath profiles and channel responses. Besides, both PriLA and TH-PriLA leak less information compared to Puzzle in all cases demonstrated. On average, PriLA leaks merely 45.7% information compared to Puzzle. It is easy to explain as the bits generated by both PriLA and TH-PriLA is more evenly distributed than that of Puzzle. It is worth noting that in practice the distance between the adversary and the sender is very likely to be much larger than 50 cm, in which case the amount of information leakage is even smaller.

We further evaluate the BER performance of PriLA after the secure handshake phase. In this experiment, the secret key is already obtained by the user and the provider, who decodes the frames using the secret key. The user continuously sends CFO-encrypted 1000-Byte frames back-to-back to the provider. To demonstrate that the CFO encryption incurs neglectable impact on decoding performance, we also measure the BER of the frames without CFO injection and treat it as



(a) PSK (b) QAM

Fig. 7. BER performance of the provider and the adversary in PriLA.

TABLE I
THE ACCURACY OF MULTIPATH-BASED AUTHENTICATION

Zone	Lab1	Lab2	Corrider
Lab.1	91.7%	6.2%	2.1%
Lab.2	7.7%	83.4%	8.9%
Corrider	7.0%	17.2%	75.8%

the normal decoding benchmark. Fig. 7 reveals that the frame decoding performance of the provider is very closed to that of the benchmark, which implies that the CFO encryption has little impact on the decoding performance. The slight difference in the decoding performance is caused by CFO mismatch measured by the user and the provider. Meanwhile, the BER performance of the adversary is significantly poorer, reaching to a level (more than 0.3) that is not unacceptable for frame decoding. To sum up, we claim that CFO encryption can prevent the attack from the adversary while not comprising frame decoding performance of the provider.

B. Performance of Multipath-Based Location Authentication

To validate the feasibility of multipath profile based location authentication, the key metric is the accuracy that the LBS provider succeed to identify which zone the mobile user belongs to. Hence, we conduct trace-driven experiment in a real-world environment. As shown in Fig. 4, we divide the test floor into three zones, two labs and one corridor. The LBS provider is emulated by one fixed laptop, which is assembled with a three antennas Intel 5300 NIC. Mobile users are emulated by one TP_Link router, sharing a 2.4 GHz channel with 20 MHz bandwidth. The fixed laptop continuously pings to the TP_LINK router deployed in each zone. We repeat this measurement by placing the router at six different position in each zone. After trace collection, we process them offline. Multipath profiles can be constructed based on each CSI feedback frame received by three antennas. We divide the profiles data in half, one as the users that need to be authenticated, the other as already-authenticated users.

We first assume that there is no user collusion and compare a user's multipath profile with three already-authenticated multipath profiles, each of which locates in one zone. Table I shows the matching accuracy in different zones. The results report that the matching accuracy is relatively higher in Lab1 whereas worse in corridor. One reason behind is that the physical layout is much consistent in Lab1 where all mobile users experience similar multipath effect. However, corridor is a free space environment where reflection is much less. The other reason is that we only use three antennas to construct the multipath profile,

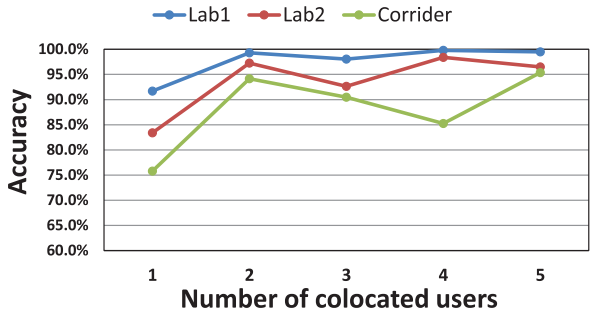


Fig. 8. The accuracy of multipath-based authentication under user collusion.

which offer limited multipath features like the number of peak and valley. We believe that the performance will be better if the more antennas are equipped. Then, we evaluate the authentication performance under user collusion in Fig. 8. We set the collusion threshold to be 50%. PriLA authenticates a user's location only when its multipath profile matches with the multipath profiles of over 50% co-located users. The results show that the authentication performance is still high under different numbers of co-located users.

VIII. Related Work

Several recent research works are presented to enable location authentication using wireless infrastructures or signals. Lenders et al. [5] utilize dedicated measuring hardware to generate unforgeable location proofs for user-generated content. Saroiu et al. [6] present a set of applications that require location authentication to enable their core functionality, and leverage the physical proximity between a transmission pair to verify a user's location. Brassil et al. [8] try to detect the location of mobile user through monitoring traffic signatures of voice call. These studies do not consider users location privacy. A fairly recent work [9] propose a location proof update system with privacy protection. Different from [9], PriLA extracts PHY signatures for privacy protection and does not require external device assistance.

PHY information has been exploited to facilitate the security and privacy mechanisms in wireless networks. The CFO encryption technique proposed in this paper follows on the heels of several recent efforts [20]–[22] that use channel reciprocity for encryption. Premnath et al. [21] study the received signal strength (RSS) variations on the wireless channel between the two devices, and propose an environment adaptive secret key generation scheme using the temporal variations. Liu et al. [22] take one step further by enabling secure group communications using RSS-generated secret keys. Different from these proposals, Qiao et al. [20] extract secret bits from the shape of frames' power spectral density to provide more robust encoding. PriLA differs from these proposals in two aspects. First, these proposals focus on data encryption after secret key extraction, while PriLA ensures secure handshake even when the secret key has not been generated. Second, PriLA exploits more fine-grained shape information in CSI curve to develop a coding scheme with higher entropy. Other PHY information has also been investigated to enable

different functionalities in wireless networks. Multipath profiles are leveraged to assist RFID positioning in non line-of-sight environments [10]. Differently, PriLA leverages multipath profiles to facilitate privacy-preserving location authentication in Wi-Fi networks.

IX. Conclusion

This paper presents PriLA, a privacy-preserving location authentication framework in Wi-Fi networks. PriLA extracts the inherent CFO and CSI signatures from legacy Wi-Fi preambles to verify users' locations without compromising their privacy. We have prototyped PriLA to demonstrate its feasibility and merits. PriLA is a clean-slate design that is transparent to upper layer protocols, and can be integrated into OFDM-based Wi-Fi devices without hardware modifications. With those features, we believe that PriLA can be easily applied to existing LBS systems with a slight upgrade.

References

[1] Z. Zhang, L. Zhou, and X. Zhao, "On the validity of geosocial mobility traces," in *Proc. ACM HotNets*, 2013, pp. 1–7.

[2] W. Wang and Q. Zhang, "Toward long-term quality of protection in mobile networks: A context-aware perspective," *IEEE Wireless Commun.*, vol. 22, no. 4, pp. 34–40, Aug. 2015.

[3] K. Wu, J. Xiao, Y. Yi, D. Chen, X. Luo, and L. M. Ni, "CSI-based indoor localization," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 7, pp. 1300–1309, Jul. 2013.

[4] J. Xiong and K. Jamieson, "Arraytrack: A fine-grained indoor location system," in *Proc. USENIX Symp. Netw. Syst. Des. Implement. (NSDI)*, 2013, pp. 71–84.

[5] V. Lenders, E. Koukoumidis, P. Zhang, and M. Martonosi, "Location-based trust for mobile user-generated content: Applications, challenges and implementations," in *Proc. ACM HotMobile*, 2008, pp. 60–64.

[6] S. Saroiu and A. Wolman, "Enabling new mobile applications with location proofs," in *Proc. ACM HotMobile*, 2009, p. 3.

[7] M. Talasila, R. Curtmola, and C. Borcea, "Link: Location verification through immediate neighbors knowledge," in *Proc. Mobile Ubiquitous Syst. Comput. Netw. Serv.*, 2012, pp. 210–223.

[8] J. Brassil, P. Manadhata, and R. Netravali, "Traffic signature-based mobile device location authentication," *IEEE Trans. Mobile Comput.*, vol. 13, no. 9, pp. 2156–2169, Sep. 2014.

[9] Z. Zhu and G. Cao, "Toward privacy preserving and collusion resistance in a location proof updating system," *IEEE Trans. Mobile Comput.*, vol. 12, no. 1, pp. 51–64, Jan. 2013.

[10] J. Wang and D. Katabi, "Dude, where's my card?: RFID positioning that works with multipath and non-line of sight," in *Proc. ACM SIGCOMM*, 2013, pp. 51–62.

[11] *USRP N210* [Online]. Available: https://www.ettus.com/product/details/UN210-KIT.

[12] *Intel Ultimate N Wi-Fi Link 5300: Product Brief* [Online]. Available: http://www.intel.com/content/www/us/en/wireless-products/ultimate-n-wifi-link-5300-brief.html

[13] K. G. Shin, X. Ju, Z. Chen, and X. Hu, "Privacy protection for users of location-based services," *IEEE Wireless Commun.*, vol. 19, no. 1, pp. 30–39, Feb. 2012.

[14] C.-Y. Chow, M. F. Mokbel, and T. He, "A privacy-preserving location monitoring system for wireless sensor networks," *IEEE Trans. Mobile Comput.*, vol. 10, no. 1, pp. 94–107, Jan. 2011.

[15] X.-Y. Li and T. Jung, "Search me if you can: Privacy-preserving location query service," in *Proc. IEEE INFOCOM*, 2013, pp. 2760–2768.

[16] J. Terry and J. Heiskala, *OFDM Wireless LANs: A Theoretical and Practical Guide*. Indianapolis, IN, USA: Sams, 2002.

[17] *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 5: Enhancements for Higher Throughput*, IEEE Std 802.11n, 2009, pp. 1–565.

[18] J. Fang *et al.*, "Fine-grained channel access in wireless LAN," *IEEE/ACM Trans. Netw.*, vol. 21, no. 3, pp. 772–787, Jun. 2013.

[19] T. S. Rappaport *et al.*, *Wireless Communications: Principles and Practice*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1996, vol. 2.

[20] Y. Qiao, K. Srinivasan, and A. Arora, "Puzzle: A shape-based secret sharing approach by exploiting channel reciprocity in frequency domain," in *Proc. USENIX Symp. Netw. Syst. Des. Implement. (NSDI)*, 2014, pp. 1–14.

[21] S. N. Premnath *et al.*, "Secret key extraction from wireless signal strength in real environments," *IEEE Trans. Mobile Comput.*, vol. 12, no. 5, pp. 917–930, May 2013.

[22] H. Liu, J. Yang, Y. Wang, Y. Chen, and C. E. Koksal, "Group secret key generation via received signal strength: Protocols, achievable rates, and implementation," *IEEE Trans. Mobile Comput.*, vol. 13, no. 12, pp. 2820–2835, Dec. 2014.

[23] L. Sweeney, "k-Anonymity: A model for protecting privacy," *Int. J. Uncertainty Fuzziness Knowl. Based Syst.*, vol. 10, no. 5, pp. 557–570, 2002.

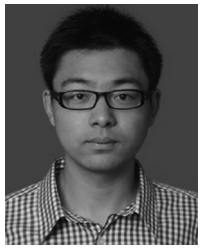
[24] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," *Theory Cryptogr.*, vol. 3876, pp. 265–284, 2006.

[25] T. Eiter and H. Mannila, "Computing discrete Fréchet distance," Tech. Rep. CD-TR 94/64, Information Systems Department, Technical University of Vienna, 1994.

[26] S. Salvador and P. Chan, "Toward accurate dynamic time warping in linear time and space," *Int. Data Anal.*, vol. 11, no. 5, pp. 561–580, 2007.

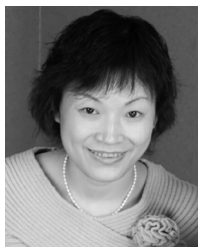
[27] S. Goldwasser, "Multi party computations: Past and present," in *Proc. ACM Symp. Principle Distrib. Comput.*, 1997, pp. 1–6.

[28] K. Suzuki and M. Yokoo, "Secure combinatorial auctions by dynamic programming with polynomial secret sharing," in *Financial Cryptography*. New York, NY, USA: Springer, 2003, pp. 44–56.



Wei Wang (S'10) received the Bachelor's degree in electronics and information engineering from Huazhong University of Science and Technology, Hubei, China, in 2010, and the Ph.D. degree in computer science and engineering from Hong Kong University of Science and Technology (HKUST), Clear Water Bay, Hong Kong. He is currently a Research Assistant Professor with Fok Ying Tung Graduate School, HKUST.

Yingjie Chen received the M.Phil. degree in computer science from Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, in 2012. He is currently a Research Assistant with Hong Kong University of Science and Technology. His research interests include PHY and MAC layer design in Wi-Fi network, and mobile computing.



Qian Zhang (M'00–SM'04–F'12) received the B.S., M.S., and Ph.D. degrees from Wuhan University, China, all in computer science, in 1994, 1996, and 1999, respectively. She joined Hong Kong University of Science and Technology in September 2005, where she is a Full Professor with the Department of Computer Science and Engineering. Before that, she was with Microsoft Research Asia, Beijing, China, in July 1999, where she was the Research Manager with the Wireless and Networking Group.