Wireless Networks
https://doi.org/10.1007/s11276-021-02769-x

ORIGINAL PAPER

Statistical inference attack against PHY-layer key extraction and countermeasures

Rui Zhu[1] · Tao Shu[2] · Huirong Fu[1]

Accepted: 25 August 2021

Abstract
The establishment of secure secret keys ahead of transmissions is one of the key issues in the field of information security. The security of traditional cryptographic secret key establishment mechanisms is seriously challenged by computing-intensive attacks, with the fast growth of high-performance computing. As an alternative, considerable efforts have been made to develop physical (PHY) layer security measures in recent years, such as link-signature-based (LSB) secret key extraction techniques. Those mechanisms have been believed secure, based on the fundamental assumption that wireless signals received at two locations are uncorrelated when separated by more than half a wavelength. However, this assumption does not hold in some circumstances under latest observations, rendering LSB key extraction mechanisms vulnerable to attacks. To address this problem, the formal theoretical analysis on channel correlations in both real indoor and outdoor environments is provided in this paper. Moreover, this paper proposes empirical statistical inference attacks (SIA) against LSB key extraction, whereby an adversary infers the signature of a target link. Consequently, the secret key extracted from that signature has been recovered by observing the surrounding links. In contrast to prior literature that assumes theoretical link-correlation models for the inference, our study does not make any assumption on link correlation. Instead, we employ machine learning (ML) methods for link inference based on empirically measured link signatures. We further propose a countermeasure against the SIAs, called forward-backward cooperative key extraction protocol with helpers (FBCH). In the FBCH, helpers (other trusted wireless nodes) are introduced to provide more randomness in the key extraction. Our experimental results have shown that the proposed inference methods are still quite effective even without making assumptions on link correlation. Furthermore, the effectiveness of the proposed FBCH protocol is validated by our experiment results.

Keywords PHY-layer security · Link signature · Key extraction · Channel impulse response · Channel correlation · Machine learning · Inference attack · Countermeasure

1 Introduction

Secret keys provide confidentiality and integrity in communication. The establishment of secure secret keys ahead of transmissions is one of the key issues in the field of information security. The Diffie-Hellman key agreement protocol (1976) was the first practical method for establishing a shared secret over an unsecured communication channel. The security of Diffie-Hellman protocol is based on the discrete logarithm problem, whose solution is assumed to be hard to compute. However, with the fast growth of high-performance computing, the above assumption is seriously challenged, rendering Diffie-

✉ Rui Zhu
rzhu@ieee.org

Tao Shu
tshu@auburn.edu

Huirong Fu
fu@oakland.edu

[1] Department of Computer Science and Engineering, Oakland University, Rochester, MI 48309, USA

[2] Department of Computer Science and Software Engineering, Auburn University, Auburn, AL 36849, USA

Published online: 04 September 2021

Hellman possibly vulnerable to various computation-intensive attacks.

Realizing the potential vulnerability of Diffie-Hellman, developing new key extraction mechanisms whose security does not rely on the computation hardness assumption receives a lot of attention. One solution is through PHY-layer security, which extracts symmetric secret keys from the PHY-layer channel response of the wireless link (i.e., the link signature) between the transmitter and the receiver [1–4]. The channel response or link signature is considered to be a good pick for secure key establishment because it is both reciprocal and uncorrelated. It is reciprocal because it is usually assumed that when the transmitter and receiver, Alice and Bob, make measurements on the channel state of the link between them, their measurements are symmetric (identical). On the other hand, the channel is said to be uncorrelated because it is usually assumed that the state of any other link separated by at least half of a wavelength from Alice and Bob should be independent from that of the link between Alice and Bob [5]. Based on these assumptions, it is commonly believed that common keys can be extracted by Alice and Bob based on their symmetric observation of the channel between them, while this channel is unobservable by a third party separated far enough (half a wavelength) from Alice and Bob, making the extracted keys secure and secrete.

While PHY-layer secret key extraction has been used in many applications such as encryption and authentication, recent studies have revealed that the uncorrelation assumption between separated links may not always be valid [6, 7], especially in many indoor environments where radio propagation becomes complicated due to signal reflection and multi-path. This opens the door for the statistical inference attack (SIA) against the link-signature based (LSB) key extraction, because the correlation between links may be exploited by an adversary to probabilistically infer the signature of a target link based on observations over surrounding links. In light of such a vulnerability, SIA against LSB key extraction has been analytically studied in prior work, by assuming a correlation model between neighboring links [6–8]. However, it remains to be seen, in a realistic wireless environment, without making assumptions on the link correlation model, how and to what extent SIA may undermine the security strength of LSB key extraction.

In this paper, we explore answers to the above questions. We first discuss the correlation between two wireless channels in both indoor and outdoor environments. We build two models, indoor and outdoor communication models. According to these two models, the formal theoretical analysis on the channel correlation is provided. Using theoretical analysis we find that there are still relatively strong correlations with two links separated by more than a half wavelength. Thus, it is possible for the adversary to launch correlation attacks by using the correlation information between the legitimate links and surrounding links. Then we propose inference attacks against existing key extraction schemes in three scenarios by employing ML approaches. Our proposed inference attacks study does not rely on any assumption on the link correlation model. Moreover, the applied ML approaches are able to deal with time-varying channel conditions and environments (the dataset used in our ML model is empirically measured at different times in a real environment where people may be moving around). In particular, our study roots from the utah/CIR dataset on CRAWDAD [9], which contains over 9300 measured channel traces for 1892 links in a 44-node indoor office-type wireless network. Several possible SIA scenarios are considered. To generalize the evaluation of our proposal, the statistical inference attacks are validated on another dataset, called Pozyx CIR and Range With LOS and NLOS dataset [10] (PCR dataset) as well. For each scenario, measured link signatures in each dataset are divided into two datasets: training data and test data. ML-based channel inference algorithms are developed. We start our study from establishing neural network models for inference. In this work, a two-layer feedforward network with sigmoid hidden neurons and linear output neurons are used. The network is trained with Levenberg-Marquardt backpropagation algorithm. After being trained based on the training data, these algorithms are instructed to infer the link signatures in the test dataset. The idea of our ML-based SIAs is independent from the specific key extraction implementations. For simplicity, we adopt the simple but well-known methods in [11] to generate keys from the obtained link signature. Furthermore, we utilize different ML algorithms, such as ensemble methods, support vector machine (SVM), and multivariate linear regression to launch SIA, and compare the inference performances of different ML algorithms. Our experimental results show that all these ML algorithms have approximative inference performance, and thus can effectively reduce the key search space by many orders of magnitudes compared to a brutal-force search mechanism.

In light of the above vulnerability, we propose a novel multi-link Forward-backward Cooperative Key Extraction Protocol with Helpers (FBCH) in this paper as a countermeasure to the aforementioned SIA attacks, aiming to make the LSB key extraction more secure. In particular, by introducing a set of helpers (these are legitimate nodes assisting the key extraction process), FBCH allows two communicating terminals to extract symmetric secret keys based on the combined channel impulse responses (CIRs) of several randomly selected links. This is in sharp contrast to the conventional method where the key extraction is only dependent on the particular link between the

transmitter and the receiver. Consequently, the resulting key extraction becomes less dependent on a particular fixed channel, making the aforementioned SIA attacks, which mainly target the channel between the two communicating terminals, less effective.

As a summary, the main contributions of this work are: 1) We theoretically verify the existence of correlation between neighboring links in realistic environments, 2) we suggest empirical methods to exploit the correlation to launch SIA against LSB key extraction, which do not rely on any assumption on the link correlation model, 3) we further propose a countermeasure to weaken the effects of SIA attacks and make LSB key extraction more secure, and 4) we use empirical datasets that are measured in real (varying) environments at different times to verify our findings. To the best of our knowledge, this is the first systematic and empirical study of LSB key extraction from the SIA perspective in the literature.

Part of this work has been presented previously as a conference paper in [12]. Comparing to our prior conference paper, the main differences and contributions of the journal version are summarized as follows: First, the journal version provides a formal theoretical verification for the correlation between neighboring links. Second, in this journal version, we propose a countermeasure, called FBCH protocol, to the statistical inference attacks. We also provide the theoretical verification for the security of the FBCH protocol. Moreover, it provides more simulation results and performance evaluation for the proposed FBCH countermeasure. Third, the journal version provides extended motivation, more analytical details, and more elaborations for the problem, and the formulations.

The rest of this paper is organized as follows. We review related work in Sect. 2. Section 3 presents the background of LSB key extraction schemes and defines the system model. Section IV analyzes the correlation between two wireless channels in both indoor and outdoor environments. We describe the proposed neural network based SIA attacks in Sect. 5. Section 6 evaluates the performance and effectiveness of the proposed attacks. Section 7 presents the FBCH countermeasures to the inference attacks and we conclude our work in Sect. 8.

2 Related work

The idea of exploiting wireless channel characteristics for generating secret keys has received considerable attention in recent years. A variety of physical layer characteristics-based key extraction schemes in different application scenarios have been proposed [11, 13–31]. For example, in [20], the authors propose a secret group-key extraction scheme in physical layer, where an arbitrary number of multi-antenna lens antennas (LNs) exist in a mesh topology with a multi-antenna passive eavesdropper. In [21], the authors establish the key extraction model for high dynamic wireless networks with a center node and random arrival users (e.g., roadside units (RSUs) with vehicles) for the first time. While the security of LSB key extraction relies heavily on the uncorrelation assumption of channels, the validity of this assumption has not been evaluated/verified in these works.

Theoretical analysis on the correlation among links is conducted in [6, 7, 32]. In particular, these works derive theoretical link correlation by taking into account the spatial/geometric relations among the transmitter, receiver, and signal reflectors. One ring model and the Jakes's model are employed to derive the link correlation models under various scenarios. Their main finding is that the uncorrelation-beyond-half-wavelength assumption is not always valid, and therefore it is necessary to use larger guard zones around the transmitter and the receiver for secure LSB key extraction. Unlike [6, 7, 32], our work derives theoretical link correlation in general indoor and outdoor environments with the random distributed scatterers.

With the development of a new generation wireless networks, intelligent physical layer security mechanisms based on machine learning have attracted researchers' attention. Machine learning for intelligent authentication in 5G and beyond wireless networks has been studied in [33]. Conventional cryptographic and physical layer authentication techniques are facing some challenges in complex dynamic wireless environments. In this article, Fang et al. envision new authentication approaches based on machine learning techniques by opportunistically leveraging physical layer attributes. In [34], Qiu et al. focus on secure authentication in wireless communication, and use a convolutional neural network as an intelligent authentication process to improve attack detection accuracy. Like [33, 34], our work applies the machine learning as the new tool to study the security performance of PHY-layer key extraction mechanisms.

The research topic of attacks against PHY-layer-based key extraction systems currently receives limited research input. Some researchers have reported that the current key extraction schemes are vulnerable to passive eavesdropping [8, 35–37], as well as active attacks [38–44].

For the passive attacks scenario, an experimental study on inference attack against PHY-layer key extraction is considered in [8], where the key extraction is based on the value of a received signal strength indicator (RSSI). Inferences in [8] are mainly based on simple averaging methods and the RSSI observations used for the inference are made close to the target link (ranging from 6 cm to at most 90 cm away from the target receiver). In reality, making an inference attack based on overhearing the target

channel at such a close distance may not be practical. Our work considers a significantly different inference model. In particular, rather than being a simple averaging of samples, our inference is based on machine learning models, which enables a much better inference outcome through training compared to brute force methods. This has been experimentally validated on the datasets used for our experiments. In particular, machine learning models attempt to map the patterns/relationships between input variables and target variables. Conversely, in this specific problem, simple averaging methods are not able to capture the non-linear relationship between positions of transmitter/receivers and channel impulse responses. As a result, our methods support inference based on observations made much further away from the target link, ranging from meters to over ten meters, which is of more interest in practice. Furthermore, instead of inferring RSSI, we infer the channel response, which allows faster key extraction. Steinmetzer et al. [35] also focuses on the passive attacks. It introduces a new analysis scheme that distinguishes between jammed and unjammed transmissions based on the diversity of jammed signals. Liu and Ning [36] studies the mimicry attack, where an adversary replays or forwards legitimate responses from the transmitter to the receiver. The attacker needs to meet two demanding requirement to launch a mimicry attack: first, he needs to roughly know the received symbols at the receiver from the transmitter; second, he needs to manipulate his own symbols, such that when the manipulated symbols arrive at the receiver, they are similar to the received symbols from legitimate transmitter. In fact, it is difficult for the attacker to meet these two requirements in practice. Unlike [36], our proposed SIA attacks exploit the inherent correlation between nearby links and apply ML models to infer link signatures, which do not rely on such demanding assumptions. Moreover, rather than being a theoretical study, we consider our work empirical/experimental, as we use ML algorithms to make inference based on empirically measured link signatures in a realistic environment.

For the active attack scenario, in [40], Zhou et al. assume that Eve is active and can send attack signals to minimize the key extraction rate of the current key extraction scheme. A formal active adversary model which takes into account an adversary's knowledge/control of the wireless channel is presented in [41]. In [42], Jin et al. propose a new form of highly threatening active attack, named signal injection attack. The attacker can inject the similar signals to both two terminals to manipulate the channel measurements and compromise a portion of the key. PHY-UIR as a countermeasure to the signal injection attack is proposed in this work. In PHY-UIR, both two terminals introduce randomness into the channel probing frames. Thus, the random series that are used to extract secret keys are the combination of randomness in the fading channel and the ones introduced by users. Then the composed series are uncorrelated to the injected signals. As a result, the attacker is not able to compromise the composed secret keys. Later, in [43], Hu et al. propose a new kind of key manipulating attack which PHY-UIR can not prevent, called session hijacking attack. The attacker hijacks the key agreement by injecting high power signals and force legitimate devices running PHY-UIR protocol with the attacker. In such way, the attacker and device generate the same key. Recently, researchers study a jamming attack on received signal phase-based key extraction system in [44]. The jammer is an active attacker that tries to make a disturbance in the key derivation procedure and changes the phase of the received signal by transmitting an adversary signal. In contrast to these kinds of active attacks, our proposed SIA attacks are passive attacks, which do not try to affect the process of key extraction, but try to silently infer the key bits that the legitimate users obtain.

In [45], authors investigate and compare the secret-key capacity based on the sampling of the entire complex channel state information (CSI) and the received signal strength (RSS). The fact that the eavesdropper's observations might be correlated has been taken into account in this work. In the conclusion, the authors find the RSS-based secret-key generation is heavily penalized as compared to CSI-based systems. In our work, the study of statistical inference attacks are based on the CSI-based secret-key generation systems. Our results show that the CSI-based systems still have security vulnerability under statistical inference attacks, even though the eavesdropper's observations are not correlated.

Unlike the attacks on physical-layer key extraction between two ends, in [46], Harshan et al. consider the attacks on group key extraction systems. They address insider attacks from the legitimate participants of the wireless network during the key extraction process. Instead of addressing conspicuous attacks such as switching-off communication, injecting noise, or denying consensus on group keys, they introduce stealth attacks that can go undetected against state-of-the-art GSK schemes. On a different track other than security, channel inference/estimation has been extensively studied for efficient radio resource management in wireless networks, e.g., for MIMO systems [1, 47]. However, such inference/estimation is made only for the channel between the target transmitter and the receiver, rather than for channels beside the target link.

3 Model description

3.1 Multi-path effect and link signature

In wireless communications, radio signals generally reach the receiving antenna by two or more paths due to reflection, diffraction, and scattering, which is called multipath propagation. Since different paths have different distances between transmitter and receiver, a receiver usually receives multiple copies of the transmitted signal at different time. Different copies have different attenuations due to the different path losses. The received signal is the sum of these delayed signal copies.

A radio channel consists of multiple paths from a transmitter to a receiver, and each path of the channel has a response (e.g., distortion and attenuation) to the multipath component traveling on it, which is called a component response. The superposition of all component responses is the *channel impulse response* (CIR). Since the multi-path effects between different pairs of nodes, as well as channel impulse responses, are usually different, a channel impulse response between two nodes is also called a *link signature*.

3.2 Key extraction from link signature

Once channel impulse responses have been estimated, the process of key extraction is rather straightforward. First, channel impulse responses should be quantized for secret key extraction since they are continuous random variables. There are several kinds of mechanisms to quantize link signatures. In this paper, for completeness purposes, we are taking as an example the uniform quantization, which was adopted by one of the seminal works on link-signature-based key extraction [3]. Other quantization methods also work for the proposed attacks. First, we normalize each CIR with its maximum element value to obtain vectors of discrete decimals. Next, the resulting discrete decimals are multiplied with 32 and then are rounded to the nearest integers. In this way, we obtain vectors of integers in the range of [0,31], which are the quantization results of continuous channel impulse responses in integer representation. Then, the vectors of integers are converted to their binary presentation. Lastly, N-bit binary string is cut out from the whole string as the initial N-bit secret key. Figure 1 shows the framework to extract key bits.

However, the straightforward quantization mechanism usually is not sufficient. Due to the variation of real environment and hardware differences between two measurement devices, it does not guarantee that the pairwise measurements from two communication ends (Alice and Bob) are identical. In this case, the sequences of bit keys extracted will not be identical. Therefore, we should employ an error reconciliation mechanism to solve this problem. We still adopt the relative simple error reconciliation mechanism in [10] as an example. For example, we can apply challenge-response verification protocol. Let K_A, K_B are the bit keys extracted by Alice and Bob, respectively, and ϕ is a random number that Alice picked. To launch the verification protocol, Alice encrypts ϕ by her secret key K_A, and sends Bob $E_{K_A}(\phi)$ and Bob responds with $E_{K_B}(\phi+1)$. If Alice gets $\phi+1$ after she decrypts Bob's message, she can conclude that Bob obtains the correct key. Bob can do likewise. Otherwise, Alice and Bob will extract bit keys from new measurements, and continue to launch error reconciliation processes until they obtain the same keys.

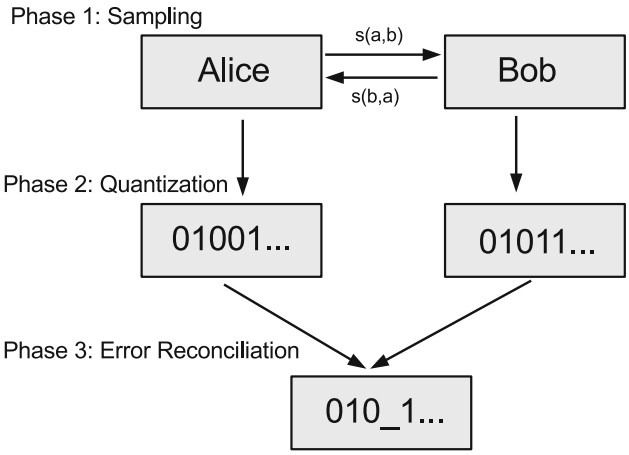


Fig. 1 Framework to extract key bits from link signature

4 Theoretical analysis

The existing LSB key extraction schemes have been believed secure, based on the fundamental assumption that wireless signals received at two locations are uncorrelated, when they were separated by more than half a wavelength. However, some of the latest work has observed that this assumption does not hold in some circumstances [6, 7]. If this assumption does not hold, the adversary can infer the target link signature based on his measurements of the correlated channels. To investigate the potential vulnerability of existing LSB key extraction system, in this section, we provide a formal theoretical verification for the existence of correlation between neighboring links in both indoor and outdoor environments. Based on the results of our theoretical analysis, we will further propose the methods of inference attacks to LSB key extraction systems in the next section.

There are various classical channel models, e.g., one ring model and Jakes's model, serving for the correlation

analysis between two adjacent channels [6, 7, 32]. However, there still exists insufficiency of these classical models. In practice, the locations of eavesdroppers/attackers are in general unknown, we cannot assume that they are located on a certain circle, as specified by the one-ring and Jake's model. Instead, a more reasonable model is to assume that these eavesdroppers could be located at arbitrary locations within a certain area, which is the feature of our newly proposed model. In contrast to these classical channel models, we further popularize a model for Rayleigh fading and let the scatterers be randomly distributed in an area A. Our analysis relies on the assumption that the two legitimate communication ends are fixed and the physical environment does not vary dramatically with the time. The important notations used in our analysis are defined in Table 1:

4.1 Channel correlation in outdoor environment

First of all, the wireless communication in an outdoor cellular system is considered, where a mobile user equipped with one antenna communicates with a base station (BS), as shown in Fig. 2. To simplify the presentation, but without loss of generality, we only consider the downlink (i.e., BS transmitting to the user) in the following analysis. Such consideration is representative, as realistic cellular communication is usually dominated by downlink traffic. Moreover, since the distance between the BS and the mobile user is much greater than the distance between each antenna element of BS, we will treat the BS as one node in the following discussion. We assume that there is no scatterer around the BS, since the BS antennas are typically installed at a high place. Attack nodes (AN) are placed by the adversary in any potential location. Moreover, we assume that scatterers are randomly distributed around the mobile user, inside a circular area A around the user's receiving antenna and the attack node, as illustrated in Fig. 2. In Fig. 2, the radius of area A is R, and the center of A is the midpoint of the user and attack node. We further assume that the number of scatterers is large, and the mobile user and attack node receive the signal from surrounding scatterers, so the line-of-sight (LOS) paths are not considered in our discussion. Our derivation is inspired by the channel correlation analysis in classical MIMO system, first proposed in [48].

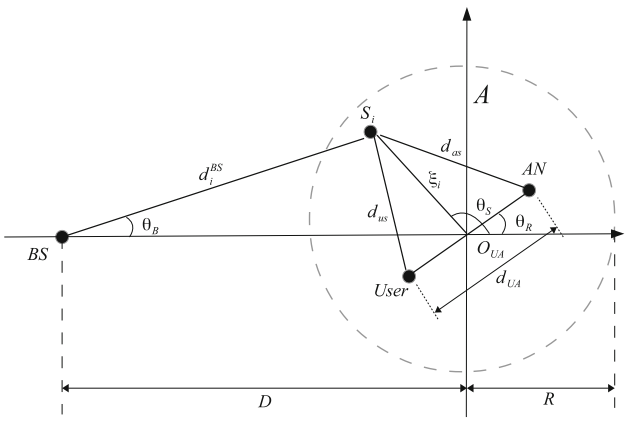


Fig. 2 Geometrical configuration of the channel model in outdoor environment

We start our discussion by defining the correlation coefficient between two neighboring links BS→User and BS→ AN as follows:

Definition IV.1 The normalized cross correlation coefficient between the neighboring links *BS* to *User* and *BS* to *AN* is expressed as:

$$\rho_{bu,ba} = \frac{E\{h_{bu} \cdot h_{ba}^*\}}{\sqrt{\Omega_{bu} \cdot \Omega_{ba}}}, \tag{1}$$

where $*$ is the complex conjugate, h_{bu} and h_{ba} denote the received signal at the user and the attack node, respectively, Ω_{bu} and Ω_{ba} are the received link power at the user and the attack node, respectively.

Table 1 Important notations

S_i	The ith scatterer
θ_S	The angle of arrival (AOA) of the wave traveling from the scatterer toward the mobile user
R	Radius of scatterer area
Ω	Received link power
$f_S(S)$	Probability density function (PDF) of scatterers in an area
d_{UA}	Distance between the mobile user and attack node
d_{RA}	Distance between the legitimate receiver and attack node
d_{as}	Distance between the attack node and the scatterer
d_{us}	Distance between the mobile user and the scatterer
d_{RA}	Distance between the legitimate receiver and the scatterer
λ	Wavelength

The received signal h_{bu} between the mobile user and the BS is given by [49] (the LOS component is neglected),

$$h_{bu} = \lim_{N\to\infty} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} g_i (d_i^{BS} \cdot \xi_i / D)^{-n/2} \cdot exp\{j\psi_i - j\frac{2\pi}{\lambda}(d_i^{BS} + d_{as})\}. \tag{2}$$

where N is the number of scatterers; g_i is the amplitude of the wave scattered by the ith scatterer; ψ_i is the phase shift introduced by the ith scatterer, respectively; d_i^{BS} and d_{as} are the distances shown in Fig. 2; λ denotes the wavelength. The term $(d_i^{BS} \cdot \xi_i / D)^{-n/2}$ accounts for the power loss relative to the distance D between the user and the BS with path loss exponent n. The total received power Ω_{bu} of this link is expressed as

$$\Omega_{bu} = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} E\{g_i^2\} (d_i^{BS} \cdot \xi_i / D)^{-n} \tag{3}$$

We assume all links have equal received power, i.e., $\Omega_{bu} = \Omega_{ba} = \Omega$. By substituting Eqs. (2) and (3) into Eq. (1), the normalized correlation coefficient between the neighboring links BS→ User and BS→AN can be derived as follows,

$$\rho_{bu,ba} = \frac{1}{\Omega} \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} E\{g_i^2\} (d_i^{BS} \cdot \xi_i / D)^{-n} \cdot exp\{-j\frac{2\pi}{\lambda}(d_{as} - d_{us})\}. \tag{4}$$

In particular, we assume the scatterers are independently distributed according to some 2-D probability density function (PDF) $f_S(S)$ on the circular area A as shown in Fig. 2. When N becomes large, the diffuse power scattered by the ith scatterer has quite small contribution out of the total Ω, which is proportional to $E\{g_i^2\}/N$. This is equal to the infinitesimal power coming from the different area d_S with probability $f_S(S)$, i.e., $E\{g_i^2\}/N = f_S(S)d_S$. Therefore, Eq. (4) can be rewritten in the following integral form:

$$\rho_{bu,ba} = \frac{1}{\Omega} \int_A (d_i^{BS} \cdot \xi_i / D)^{-n} \cdot exp\{-j\frac{2\pi}{\lambda}(d_{as} - d_{us})\} f_S(S) dS, \tag{5}$$

Since we assume scatterers are randomly distributed inside the circular area A around the user and attack node, Eq. (5) can be written as

$$\rho_{bu,ba} = \frac{1}{\Omega} \int_0^R \int_{-\pi}^{\pi} (d_i^{BS} \cdot \xi_i / D)^{-n} \cdot exp\left\{ -j\frac{2\pi}{\lambda} \cdot (d_{as} - d_{us}) \right\} f(\theta_S, \xi_i)\, d(\theta_S) d(\xi_i). \tag{6}$$

where the $f(\theta_S, \xi_i)$ is the PDF of the locations of scatterers relative to the user and attack node with θ_S and distance ξ_i and d_i^{BS}, θ_S is the angle of arrival (AOA) of the wave traveling from the scatterer toward the mobile user.

According to the laws of cosine and sine in [50], we get

$$\begin{aligned} d_{as}^2 &= d_{UA}^2/4 + \xi_i^2 - d_{UA} \cdot \xi_i \cdot cos(\theta_S - \theta_R) \\ d_{us}^2 &= d_{UA}^2/4 + \xi_i^2 + d_{UA} \cdot \xi_i \cdot cos(\theta_S - \theta_R) \end{aligned} \tag{7}$$

and

$$\frac{D}{sin(\theta_S - \theta_B)} = \frac{\xi_i}{sin(\theta_B)} = \frac{d_i^{BS}}{sin(\theta_S)}, \tag{8}$$

where the d_{UA} is the distance between the mobile user and the attack node.

In the outdoor environment, the assumption of $D \gg R \gg d_{UA}$ is realistic. Therefore, the difference of path lengths can be approximated as

$$\begin{aligned} -(d_{as} - d_{us}) &\approx d_{UA} \cdot cos(\theta_S - \theta_R) \\ d_i^{BS} &\approx D \end{aligned} \tag{9}$$

Substituting the arguments in Eq. (6) with Eqs. (7), (8) and (9) yields

$$\rho_{bu,ba} = \frac{1}{\Omega} \int_0^R \int_{-\pi}^{\pi} (\xi_i)^{-n} exp\{-j\frac{2\pi}{\lambda} \cdot d_{UA} \cdot cos(\theta_S - \theta_R)\} f(\theta_S, \xi_i)\, d(\theta_S) d(\xi_i). \tag{10}$$

Since the scatterers are uniformly distributed inside the circular ring, we can use the PDF as $f(\theta_S, \xi_i) = 1/2\pi R$.

Then, we obtain

$$\begin{aligned} \rho_{bu,ba} &= \frac{1}{\Omega} \int_0^R \int_{-\pi}^{\pi} (\xi_i)^{-n} exp\{-j\frac{2\pi}{\lambda} \cdot d_{UA} \cdot cos(\theta_S - \theta_R)\} \frac{1}{2\pi R}\, d(\theta_S) d(\xi_i) \\ &= \frac{1}{\Omega} \int_0^R \int_{-\pi}^{\pi} \frac{(\xi_i)^{-n}}{2\pi R} exp\{-j2\pi \frac{d_{UA}}{\lambda} * cos(\theta_S - \theta_R)\}\, d(\theta_S) d(\xi_i) \end{aligned} \tag{11}$$

We plot the correlation coefficient $\rho_{bu,ba}$ as a function of the ratio d_{UA}/λ in Fig. 3, in which we assume the distance D is fixed. From this numerical result, it can be observed that there exists correlation, even if the distance d_{UA} between the mobile user and the attack node is greater than half wavelength. For instance, $\rho_{bu,ba} = 0.21$ when d_{UA} is equal to 5 wavelength, and $\theta_R = 1$. Moreover, the angle θ_R affects the correlation coefficient $\rho_{tr,ta}$ dramatically in this model.

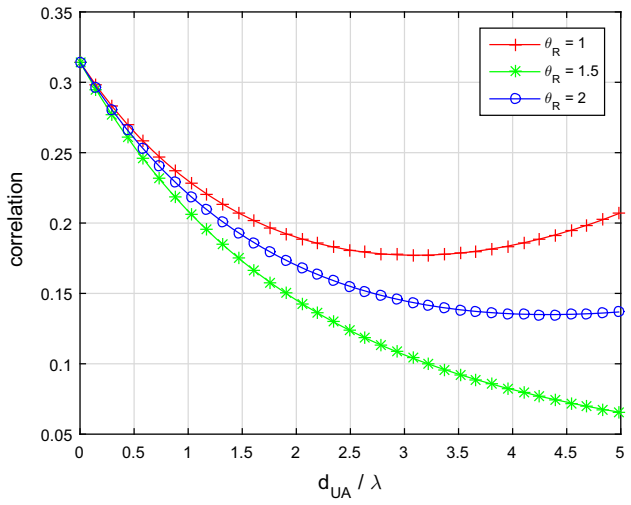


Fig. 3 Channel correlation coefficient $\rho_{bu,ba}$ versus distance ratio d_{UA}/λ

4.2 Channel correlation in indoor environment

In indoor environments, two communication ends, the transmitter (Tx) and the receiver (Rx), generally are not far from each other and surrounded by scatterers nearby. In this case, we build a model, in which a big scatterer-ring area A encloses both the transmitter and the receivers (including attack nodes), as depicted in Fig. 4. In this circular area A, the radius is R, and the center O is the midpoint of the Tx and O_{RA}, where O_{RA} is the midpoint of the Rx and the attack node (AN). In Fig. 4, we follow most of the notations as those defined in Fig. 2. Comparing to the outdoor model as shown in Fig. 2, in the indoor model, we substitute the notations "Tx" and "Rx" for "BS" and "User", respectively. Moreover, to facilitate the derivation, we substitute d_i^{TS}, d_{RA}, d_{rs} and θ_T for d_i^{BS}, d_{UA}, d_{us} and θ_B, respectively, and introduce the new variable γ to denote the angle of the scatterer in the polar coordinate system, as illustrated in Fig. 4.

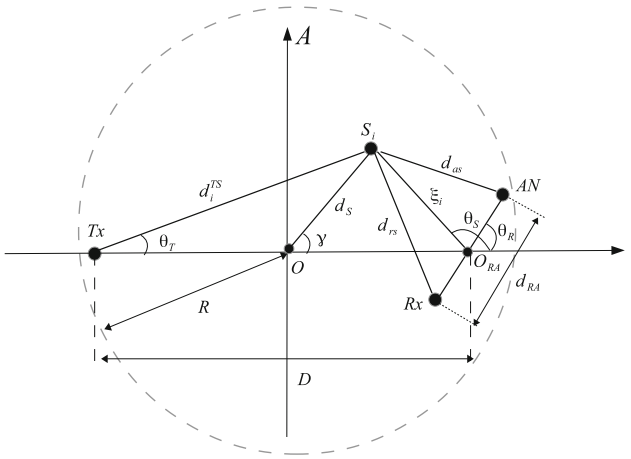


Fig. 4 Geometrical configuration of the channel model in indoor environment

In Fig. 4, the correlation coefficient between two neighboring links Tx→Rx and Tx→AN is defined as follows:

Definition IV.2 The normalized cross correlation coefficient between the neighboring links *Tx* to *Rx* and *Tx* to *AN* is expressed as:

$$\rho_{tr,ta} = \frac{E\{h_{tr} \cdot h_{ta}^*\}}{\sqrt{\Omega_{tr} \cdot \Omega_{ta}}}, \tag{12}$$

where $*$ is the complex conjugate, h_{tr} and h_{ta} denote the received signal at the receiver and the attack node, respectively, Ω_{tr} and Ω_{ta} are the received link power at the receiver and the attack node, respectively

To discuss the correlation between the neighboring links Tx→Rx and Tx→AN in this model, we still use Eq. (5) to yield the correlation coefficient function. The Eq. (5) is rearranged in the indoor model as follows:

$$\begin{aligned}\rho_{tr,ta} = &\frac{1}{\Omega}\int_A (d_i^{TS} \cdot \xi_i / D)^{-n} \\ &\cdot exp\left\{-j\frac{2\pi}{\lambda}(d_{as} - d_{rs})\right\} f_S(S) dS,\end{aligned} \tag{13}$$

where Ω is the received link power (equal power for all radio links is assumed).

According to the laws of cosine and sine in [50], we have

$$\begin{aligned}(d_i^{TS})^2 &= d_S^2 + D^2/4 + D \cdot d_S cos(\gamma) \\ \xi_i^2 &= d_S^2 + D^2/4 - D \cdot d_S cos(\gamma)\end{aligned} \tag{14}$$

and

$$\begin{aligned}sin(\theta_T) &= d_S sin(\gamma)/d_i^{TS} \\ &= d_S sin(\gamma)/\sqrt{d_S^2 + D^2/4 + D \cdot d_S cos(\gamma)}\end{aligned} \tag{15}$$

$$\begin{aligned}cos(\theta_T) &= (d_S cos(\gamma) + D/2)/d_i^{TS} \\ &= (d_S cos(\gamma) + D/2)/ \\ &\quad \sqrt{d_S^2 + D^2/4 + D \cdot d_S cos(\gamma)}\end{aligned} \tag{16}$$

$$\begin{aligned}sin(\theta_S) &= d_S sin(\gamma)/\xi_i \\ &= d_S sin(\gamma)/\sqrt{d_S^2 + D^2/4 - R \cdot d_S cos(\gamma)}\end{aligned} \tag{17}$$

$$cos(\theta_S) = (d_S cos(\gamma) - D/2)/\xi_i = (d_S cos(\gamma) - D/2)/\sqrt{d_S^2 + D^2/4 - R \cdot d_S cos(\gamma)} \quad (18)$$

By doing the substitution and rearrangements, we obtain the approximation of $-(d_{as} - d_{rs})$ as

$$-(d_{as} - d_{rs}) \approx \frac{d_{RA} \cdot (d_S cos(\gamma - \theta_R) - D/2 \cdot cos(\theta_R))}{\sqrt{d_S^2 + D^2/4 - D \cdot d_S cos(\gamma)}} \quad (19)$$

Since the scatterers are uniformly distributed inside the circular ring, we can use the PDF as $f_S(S) = f(d_S, \gamma) = \frac{1}{2\pi R}$. Substituting the arguments in Eq. (13) yields the correlation coefficient $\rho_{tr,ta}$ between neighboring links Tx→Rx and Tx→AN as

$$\rho_{tr,ta} = \frac{D^n}{\Omega} \int_0^R \int_{-\pi}^{\pi} ((d_S^2 + D^2/4)^2 - D^2 d_S^2 cos^2\gamma)^{-n/2} \frac{1}{2\pi R} \cdot exp\left\{ \frac{-j2\pi \cdot d_{RA}(d_S cos(\gamma - \theta_R) - D/2 \cdot cos(\theta_R))}{\lambda \sqrt{d_S^2 + D^2/4 - D \cdot d_S cos(\gamma)}} \right\} d_\gamma d(d_S) s \quad (20)$$

The correlation coefficient $\rho_{tr,ta}$ as a function of the ratio d_{RA}/λ is plotted in Fig. 5. It can be observed that there still exists correlation, even if the distance d_{RA} between the legitimate receiver and the attack node is greater than half wavelength. For instance, the correlation coefficient $\rho_{tr,ta} = 0.26$ when d_{RA} is equal to 5 wavelength, and $\theta_R = 1.2$. In contrast to the outdoor model, the correlation is a decreasing function of the angle θ_R, for a certain distance ratio d_{RA}/λ. Moreover, comparing to the outdoor model, the correlation coefficient $\rho_{tr,ta}$ is more sensitive to θ_R.

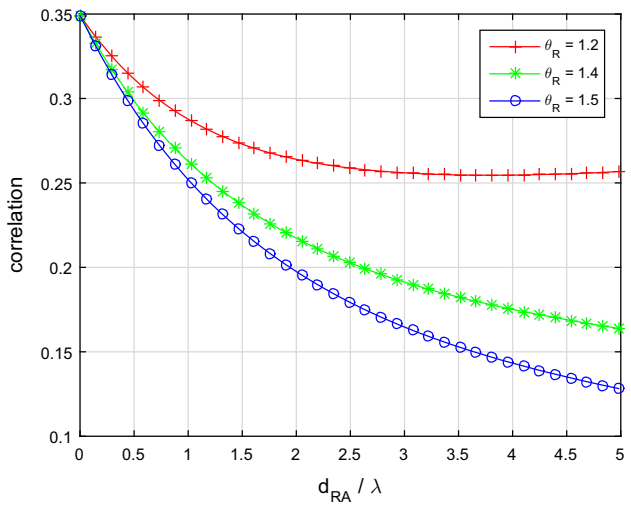


Fig. 5 Channel correlation coefficient $\rho_{tr,ta}$ versus distance ratio d_{RA}/λ

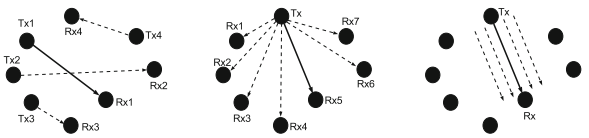
In conclusion, the wiretap channel (i.e., Tx→AN) and the legitimate channel (i.e., Tx→Rx) still have relatively strong correlation, even if the attack node is far away from the transmitter and receiver (here the "far away" means that the distance between the attack node and receiver is greater than $\lambda/2$). Therefore, the adversary can infer the secret key bits based on his measurements of channel states, which makes the existing key extraction system vulnerable. In the following section, we propose a class of attacks, called statistical inference attacks (SIAs), to reveal the vulnerability of the LSB key extraction system.

5 Statistical inference attack

In the previous section, we have theoretically verified the existence of correlation between neighboring links. The remaining issue is how this correlation may be exploited to infer secret keys in practice. In this section, we apply several ML-based algorithms to propose statistical inference attacks (SIAs) to infer keys by exploiting this correlation. To launch SIAs, we assume the physical environment does not vary dramatically with the time, e.g, the relative speed between the Tx and Rx is low, $\Delta V \approx 0$, and the number of high speed obstacles can be neglected. However, the environments should remain changed over time. For instance, different people may be moving around in an office. Depending on the information available for the training of the ML model, we consider the following three scenarios for SIAs:

(a) Inference based on links disjoint from the target link (i.e., links of different transmitters and receivers from the target link)
(b) Inference based on links sharing the same transmitter as the target link
(c) Inference based on historical signatures of the target link

These scenarios are illustrated in Fig. 6, respectively, and are elaborated as follows.



(a) Case I: Using links of different tx and rx locations. **(b)** Case II: Using links of the same tx locations. **(c)** Case III: Using signatures of the same link, but measured at different times.

Fig. 6 Attacks in three scenarios, solid lines denote the target links, dotted lines denote the existing links

5.1 Case I: SIA based on disjoint links

This case is pertinent to the scenario where the adversary has prior knowledge regarding the area in which the target (i.e., the transmitter and receiver of the target link) will appear, but does not know the exact location of the target until the target appears. In this case, the adversary may perform a site survey in the area before the target appears. During the site survey, the adversary may collect sample signatures of links throughout the area at different times, and then use these samples to train a model that represents the signature of an arbitrary link in the area as a function of the transmitter and receiver locations of the link. The adversary divides the sample dataset into training and testing sets for cross-validation. Since the environments are varying during the site survey, the sample link signatures are not identical. Later in the online inference phase, the adversary can observe the location of the target, and supply this information to the trained model to infer the signature of the target link. SIA to links in a mobile ad hoc network is a typical example of this scenario.

We apply the classical topology of ML model for the above link signature inference is as follows:

Input: $S(L_T(i), L_S(i))$, $L_T(i)$, $L_S(i)$,
Output: $S(L_T(t), L_S(t))$,

where

- i = the index of surveyed links,
- t = the index of the target link,
- $L_T(i)$ = the locations of transmitters,
- $L_S(i)$ = the locations of receivers,
- $S(L_T(i), L_S(i))$ = link signatures on the links $(L_T(i), L_S(i))$,
- $S(L_T(t), L_S(t))$ = link signature on the target link.

5.2 Case II: SIA based on links sharing the same transmitter

This case applies to the scenario where the adversary does not know the exact location of the target until the target appears, but has prior knowledge of the area in which the target will appear, and the communication in this area is through a centralized access point such as a base station or an access point (AP). Typical examples of such scenarios include cellular networks and wireless local area networks (WLAN). In this case, the adversary can also survey the area before the target appears, during which he collects sample signatures of various downlinks throughout the area. These sample signatures are then used to train a downlink signature model of the area, which represents the signature of an arbitrary link as a function of the receiver's location. In the online inference phase, the adversary observes the location of the target (the receiver), and supply this information to the trained model to infer the signature of the target downlink.

The topology of the ML model in this case is given as follows:

Input: $S(L_S(i))$, $L_S(i)$,
Output: $S(L_S(t))$,

where

- i = the index of surveyed downlinks,
- t = the index of the target downlink,
- $L_S(i)$ = the receiver location of the ith surveyed downlink,
- $S(L_S(i))$ = link signature of the ith surveyed downlink,
- $S(L_S(t))$ = link signature of the target downlink.

5.3 Case III: SIA based on historical signatures of the target link

In this case, we consider the temporal variation of the signatures of the same link. This case applies to the scenario where the adversary has prior knowledge of the exact location of the target. Such information can be obtained by the adversary by peeking into the location privacy of the target. This is especially true if the target's activity or schedule follows a regular rule.

To infer the signature of the target link at a given time, the adversary may first measure the signatures of the target link at different times. The sample signatures are then used to train a model that represents the link signature as a function of time. In the online inference phase, the adversary simply feeds the desired time into the trained model to make inference on the signature of the target link at that time.

The topology of the ML model for this case is given as follows:

Input: $S(t_i)$, t_i,
Output: $S(t_c)$,

where

- i = the index of time,
- t_i = time t_i,
- $S(t_i)$ = link signature at time t_i,
- $S(t_c)$ = link signature at the target moment.

5.4 Overview of statistical inference attacks

Figure 7 shows the overview of statistical inference attacks. Specifically, the adversary starts the statistical inference attacks by collecting sample signatures during the site survey phase. Once the data are collected, the

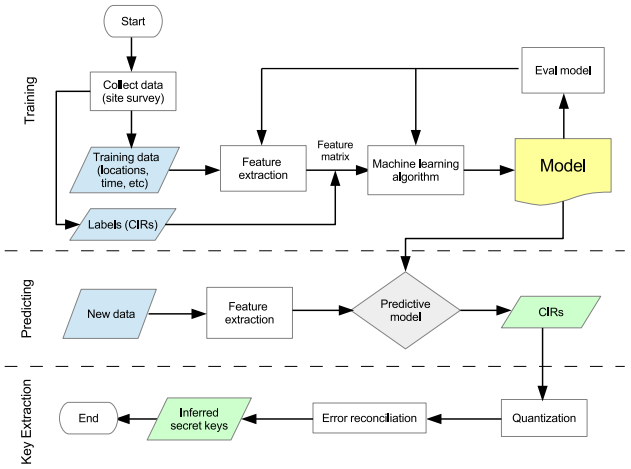


Fig. 7 Flowchart of the statistical inference attacks

adversary labels all the collected data, and feeds the data to selected machine learning algorithms to train the model. The input training data is transformed into a feature matrix, which contains a number of features that are descriptive of the sample link signatures. In this work, the neural networks, support vector machines, linear regression, and ensemble methods are used as learning algorithms. In the following, the adversary evaluates the accuracy of the learned function by using a test set that is separate from the training set, and tunes the model parameters to improve the training of the model. Once the training and parameters are optimized, the adversary uses this predictive model to infer the signature/CIR on the target link. In the final inferred key extraction phase, since the estimated channel impulse responses are continuous variables, they should be quantized to sequences of bits. Due to the variation in the real environment and hardware differences between two measurement devices, it does not guarantee that the pairwise measurements from two communication ends are identical. In this case, the sequences of bit keys extracted will not be identical. Therefore, the error reconciliation (ER) mechanism should be employed to obtain the inferred secret key, and complete the inference attack. There are kinds of existing quantization and ER mechanisms that the adversary can use, which are beyond the scope of this work.

6 Experiment verification

In this section, we evaluate the effectiveness of the above statistical inference attacks based on the utah/CIR dataset from CRAWDAD [9] and PCR dataset from the IEEE dataport [10].

6.1 Dataset

The experiment is based on the utah/CIR dataset on CRAWDAD [9]. The CRAWDAD is a widely used archive for sharing wireless network data resources across the research community. This archive stores wireless trace data, from real networks under real conditions, from many contributing locations. There are 125 different datasets in CRAWDAD so far, and the utah/CIR dataset is the one we used in this work. In the utah/CIR dataset, over 9300 link signatures are recorded in a 44-node wireless network, which are measured in an indoor environment with obstacles and scatters. By moving the transmitter and receiver between node locations 1 - 44, it gives the number of transmitter and receiver permutations counted as $44 * 43 = 1892$. At each permutation of transmitter and receiver, 5 link signatures are measured over a period of about 30 seconds. In this dataset, each link signature is recorded as a 50-component vector. The campaign measures the relative static channels, while two or three people were typically walking in the measurement environment. The measurements are completed over the course of eight days, and as a result, the samples vary with time.

To generalize the evaluation of our proposal, the statistical inference attacks are validated on another dataset, called Pozyx CIR and Range With LOS and NLOS dataset (PCR dataset) [10] as well. This dataset includes UWB range measurements performed with Pozyx devices. The measurements were collected between two tags placed at several distances and in two different conditions: with Line of Sight (LOS) and Non-Line of Sight (NLOS). The measurements include the range estimated by the Pozyx tag, the actual distance between devices, the timestamp of each measurement and the values corresponding to the samples of the Channel Impulse Response (CIR) after each transmission.

6.2 SIA results and analysis

We first evaluate how accurately the proposed neural network can infer the signature of a target link using utah/CIR dataset. To this end, we randomly pick a link (a transmitter-receiver pair) from utah/CIR dataset as the target link. The five signatures of the target link are used as ground truth for testing. Training data is selected from the remaining links in utah/CIR dataset in the following way. For case-I SIA, we use all remaining links, in total $9300 - 43 * 2 * 5 = 8870$ signatures, as training data. For case-II SIA, we use all the 43 links that share the same transmitter of the target link but have a different receiver as the training data. So in total $43 * 5 = 215$ link signatures are included in the training dataset for case II. For case III SIA, we randomly

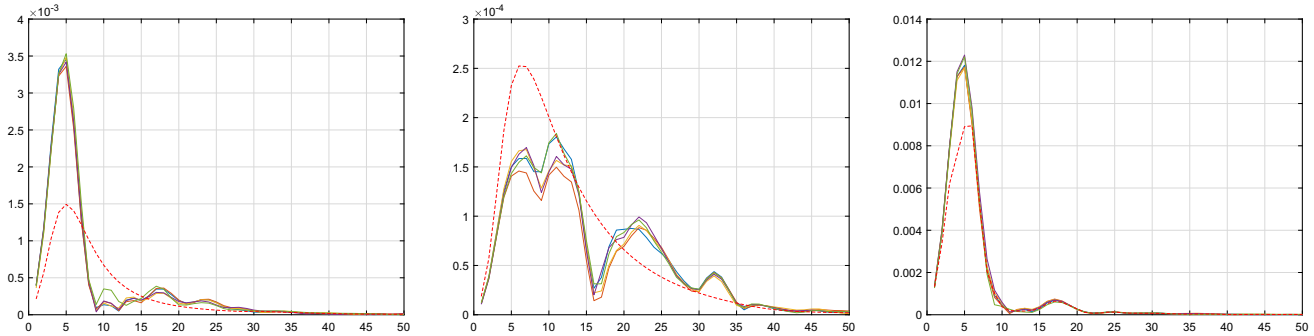
pick 4 signatures of the target link and use them as training data, and the remaining link signature is used for testing. In a nutshell, in this experiment we use all relevant data for training to avoid the complicated issue of training data selection. As a baseline of the performance, our goal here is to see how well the neural network can do without discriminating the available training data. The optimization of the inference, e.g., through training data filtering, is studied shortly. The important experiment parameters/hyperparameters are highlighted in Table 2. In this table, the data splitting strategy, key parameters/hyperparameters of neural network, support vector machine (SVM), and linear regression models, which are used in the experiments, are listed.

Different target links were inferred in our SIA experiments. Figure 8 plots a typical case for the comparison between the inferred signature and the ground truth version for the target link $1 \rightarrow 4$ under the three SIA cases, respectively. The inferences for other target links present similar trends, and thus are omitted here due to space limit. Three observations can be made on Fig. 8. First, in all three cases there exists significant similarity between the inferred signature and the ground truth, and the trends of curves match quite well. This observation implies that there are indeed correlations between neighboring wireless links, even when their separation is farther than half a wavelength, and these correlations are harnessed by neural network models in the experiments for inference. Second, the inferred signature presents different accuracy in the three cases. This is not surprising, because the inferences are based on different amount of knowledge about the target. In particular, training data in case III is the closest to what is being inferred, and therefore the inference accuracy in that case is the highest among the three cases. Third, the inferred signature is much smoother than the truth version. The current neural network models cannot capture enough high frequency details in the correlation to make a better inference. This observation suggests that the neural network models we are using may not be the optimal ones and there is room for improvement from a ML's point of view.

We study the impact of inference accuracy on the security strength of LSB key extraction as shown in Fig. 9. In this experiment, the goal of the adversary is to figure out, or guess, the secret key extracted from the true link signature, based on the inferred version of the signature. To simplify the experiment, but without loss of generality, we take the quantization method in [10] as an example to illustrate the idea of the key extraction. In particular, we assume that each time-series point on the true link signature is represented as a 5-bit binary number according to the quantization scheme described in Section III.B (32-interval quantization). So in total a 250-bit binary string can be extracted from the 50-point true link signature. Because of the limitation of our computation resources, to save time, we pick a short 75-bit string as the true key in our experiment. To guess the true key, the adversary uses trial and error, starting from the 75 bits quantization of the inferred signature (5 bits per point, $15 * 5$-bit binaries in total). In each round of trial and error, the adversary explores the key search space by incrementing or decrementing by one to one of the $15 * 5$-bit binaries, where the exploration is sequential over the $15 * 5$-bit binaries. Under such an inference attack, the security strength of the LSB key extraction can be measured by the average number of trials needed to find the true key. Equivalently, this metric can be normalized on a per-point basis, i.e., measured by the average number of guesses required to find the true 5-bit quantization for a point on the link signature.

Table 2 Important Parameters in Experiments

Dataset	Training dataset	70%
	Validation dataset	15%
	Testing dataset	15%
Neural Networks	Number of hidden layers	2
	Number of hidden neurons per layer	64
	Initial learning rate	1.0
	L2 regularization λ	1.0
	Activation function	Sigmoid
SVM	Kernel	RBF
	C parameter	1.0
	Gamma	auto
Multivar Linear Regress	Alpha	1.0
	Fit intercept	True
	Normalize	True
	N jobs	1
	Max iteration	1000



(a) Case I **(b)** Case II **(c)** Case III

Fig. 8 Link signature inference accuracy. Note that x-axis denotes index of points, y-axis denotes the amplitude; the solid lines represent the true link signatures, and dotted line represents our inferred link signature

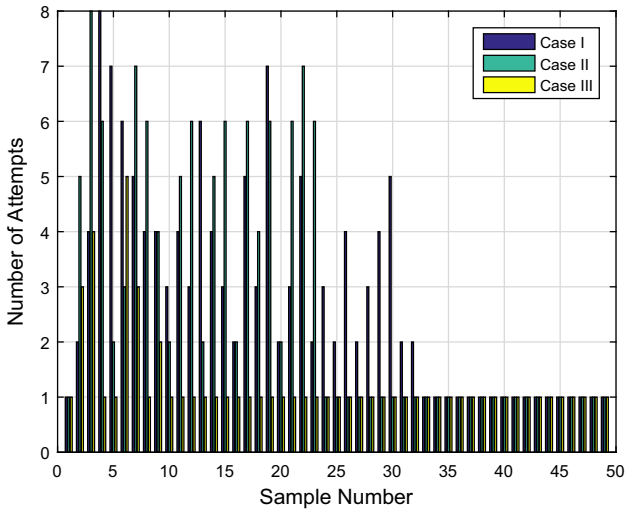


Fig. 9 Average number of guesses needed to find the 50 points in a link signature (neural network)

Figure 9 plots the average required number of guesses for each of the whole 50 points on a link signature under the three SIA cases. The average is based on 26 target links randomly selected from utah/CIR dataset. From this figure, it can be observed that for cases I and II on average at most 8 guesses are enough to find the true quantization of a point using the neural network. In contrast, to find a 5-bit quantization, a brutal force search algorithm needs on average 16 guesses. Therefore, for the 75-bit key in this experiment, the key search space of the brutal force algorithm is 16^{15}, or 2^{60}. Using the proposed neural network, the key search space is at most 8^{15}, or 2^{45}: a reduction of 2^{15} compared to the brutal force search! On a computer with a 4-core Intel CPU (2.0 GHz CPU clock speed), it takes about 4,000 hours to find the correct secret key by brutal force algorithm, however, it takes only about 1 hour and 13 minutes to find it by our proposed mechanism. Note that this is the upper bound (worst case) key search space for the case I and case II SIAs, because the number of required guesses per point is much smaller than 8 for the points on the tail of the link signature. For example, only one guess is needed for points after index 33. Furthermore, it can also be observed that SIA in case III is able to figure out the true key much more efficiently, as 90% points can be found in just one guess in that case.

To obtain a statistical view about the strength of the proposed SIAs, Fig. 10 compares the CDF (cumulative distribution function) of the number of guesses needed per point under various SIA cases. The CDFs in the figure are calculated based on the same 26 target links as in Fig. 9. This figure shows that statistically the relative strength of SIAs are case III > case II > case I. For example, case III can find 90% points in just one round, while 56% points are found in one round in case II, and only 38% are found in one round in case I. This trend is aligned with the inference accuracies as observed in Fig. 8.

Furthermore, we study the optimization of the inference accuracy through training data selection in Fig. 11.

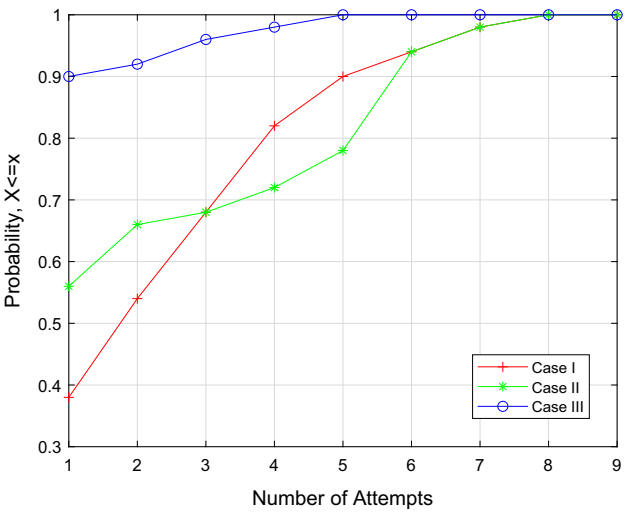
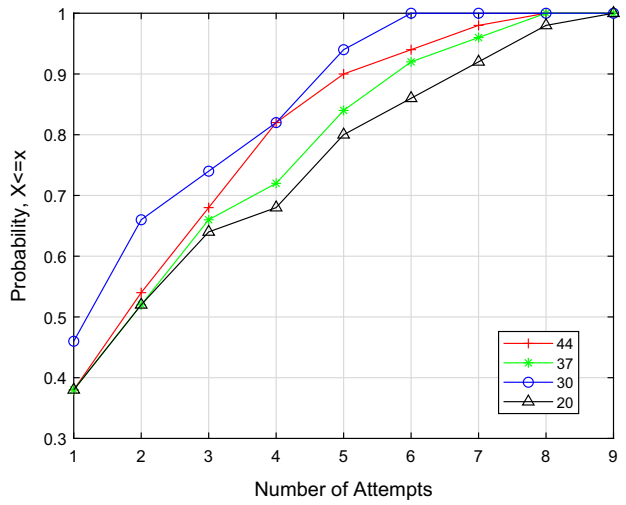


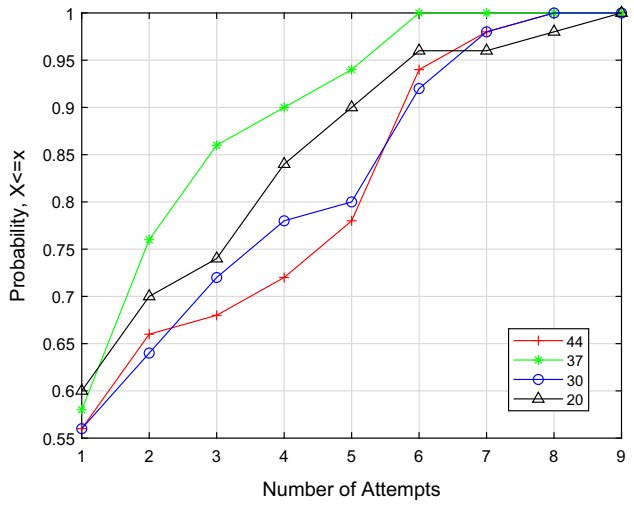
Fig. 10 Comparison of inference performance in different cases by using neural network

Selecting the right training data is usually vital to ensure a good performance of a neural network due to the well-known over-training issue. Because the inference in case III is very accurate, here we only focus on the optimization of cases I and II. For each case, we pick k nearest links to the target link, and use their signatures ($5 * k$ in total) to train the neural network. Such a treatment is based on the rationale that a closer link to the target should possess higher correlation, and thus can provide better training effects. So now the training data selection is converted to deciding the optimal size of the training dataset (i.e., the k). In our experiment we vary the value of k (ranging from 20 to 44) and evaluate the strength of the resulting attacks in terms of the CDF of the number of guesses needed to find a point on the link signature.

Figure 11 plots the CDFs under various training sizes. It shows that the security strength of the SIA in both cases is sensitive to the size of the training data. For example, the 75-percentiles in case II may range from 2 to 5 under various training data sizes, corresponding to a factor of $(5/2)^{50}$ difference in size of the key search space! This observation suggests the necessity of optimizing the training dataset in order to improve the inference accuracy, and hence enforce the attack strength, of the SIAs. Figure 11 also suggests that the inference accuracy of the neural network is a non-monotonic function of the training data size, and there seems to be an optimal training data size in each case that maximizes the inference accuracy. For example, the optimal training data sizes are 30 and 37 for case I and case II, respectively.



(a) Case I



(b) Case II

Fig. 11 SIA strength vs. training data size (neural network)

To test the inference performances of different ML algorithms, we utilize more ML algorithms, such as ensemble methods, support vector machine (SVM), and multivariate linear regression to launch SIAs in case I. Figure 12 plots the CDFs under different ML inference algorithms. The CDFs in this figure are also calculated based on the same 26 target links as shown in Fig. 9, and the training data size is 44. This figures shows that more than 50% points can be found in just one round by applying multivar linear regress method. In comparison, less than 40% points can be found in one round by applying neural network. However, all these ML algorithms can successfully guess the truth value of each point within 10 attempts. Table 3 shows that statistically SVM has the highest inference accuracy, since in average, it just need 2.9 attempts to reach truth value of each point. However, when training data size becomes greater, it will not be efficient enough to launch SIA by using SVM. In this case, SVM will spend costly computation and memory resources. In addition, the adversary has to spend a lot of time selecting an optimal kernel and adjusting the parameters in SVM.

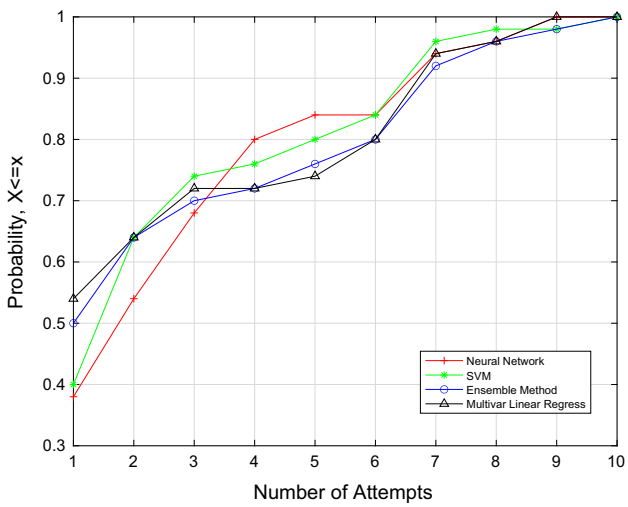


Fig. 12 Comparison between different ML inference algorithms in CDF representation

Table 3 Summary of statistics for accuracy

	Mean	Standard deviation (SD)
Neural network	3.04	2.3997
SVM	2.90	2.4021
Ensemble method	3.02	2.7018
Mutivar linear regress	2.94	2.6564

Note that the listed number of attempts and SD represents for each point

To generalize the evaluation of our proposal, the statistical inference attacks are validated on the PCR dataset as well. Figure 13 shows the secret key inference accuracies of the proposed algorithms over this new dataset. In particular, in Fig. 13, it can be observed that more than 40% of the points can be found in just one round by applying these models. Moreover, all these ML algorithms can successfully guess the truth value of each point within 10 attempts. Comparing this to Fig. 12, the secret key inference accuracies of our proposed SIAs are pretty similar on these two different datasets.

From this figure, we can also observe that the trends are aligned with the inference accuracies as observed in Fig. 12. These observations validate that the effectiveness of our proposed SIAs do not depend on a specific dataset.

In our experiment, we apply several general ML algorithms to launch SIAs. How to improve the inference algorithms, and analytically decide the optimal training data size, so that the adversary can construct the best site survey strategy to maximize its attack strength, remain questions to be explored in our future study.

7 Countermeasure

In this section, we develop a novel LSB key extraction scheme to defend against the statistical inference attacks.

7.1 Forward-backward cooperative key extraction protocol with helpers (FBCH)

In conventional LSB key extraction scheme, only the link between legitimate transmitter and receiver is measured to obtain CIRs as the random series. Based on the nature of channel correlation, the adversary can effectively utilize the channel information of surrounding links to infer CIR of the target link. Our experiments in Section VI demonstrate the search space of the secret key has been significantly shrunk and the inference attacks are feasible. To overcome the weakness of existing scheme, we propose a novel LSB key extraction protocol, called forward-backward cooperative key extraction protocol with helpers (FBCH). In FBCH, helpers participate in key extraction process to construct several channels, and pass the CIR information between transmitter and receiver by manipulating their transmission power. And we assume the physical environment is quasi-static as well. Algorithm 1 describes the procedures of FBCH.

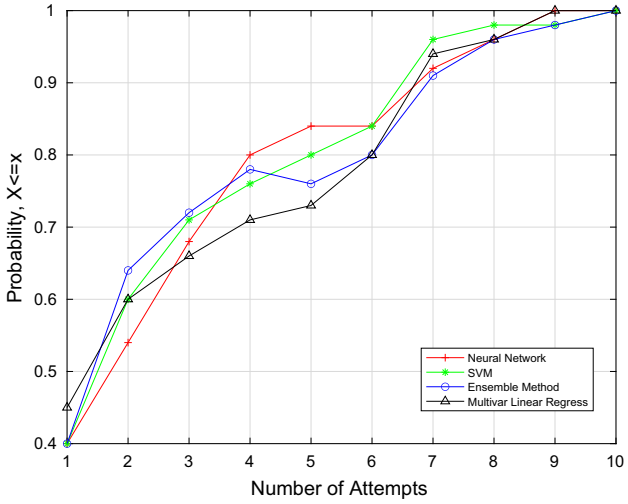


Fig. 13 Comparison between different ML inference algorithms in CDF representation on the PCR dataset

Algorithm 1 FBCH Protocol

Step 1: The Tx randomly picks N helpers H_i from all of the M available relays in its transmission range.
Step 2: The Tx broadcasts training symbols under the standard transmission power P_T, each helper H_i receives signal from the Tx and measures the channel Tx $\rightarrow H_i$ to obtain the CIR $h_{tx,i}$.
Step 3: Each helper H_i broadcasts a training symbol under the standard transmission power P_H, the Tx and Rx receive signals from H_i and measure channels $H_i \rightarrow$ Tx and $H_i \rightarrow$ Rx to obtain the CIRs $h_{tx,i}$, and $h_{i,rx}$, respectively.
Step 4: The Rx broadcasts a training symbol under the standard transmission power P_R, each helper H_i receives signal from the Rx and measures the channel Rx $\rightarrow H_i$ to obtain the CIR $h_{i,rx}$.
Step 5: Each helper H_i manipulates its transmission power to P'_H and transmits training symbol to the Tx under the transmission power P'_H, where $P'_H = \frac{P_H * h_{i,rx}}{h_{tx,i}}$; the Tx measures the channel $H_i \rightarrow$ Tx to obtain the CIR $h'_{tx,i}$.
Step 6: Each helper H_i manipulates its transmission power again to P''_H and transmits training symbol to the Rx under the transmission power P''_H, where $P''_H = \frac{P_H * h_{tx,i}}{h_{i,rx}}$; the Rx measures the channel $H_i \rightarrow$ Rx to obtain the CIR $h'_{i,rx}$.
Step 7: Tamper detection : the Tx and Rx calculate summation $\sum(h_{tx,i} + h_{i,rx})$. If the summations at two ends are not equal to each other, they drop the CIRs and go back to Step 1.
Step 8: The Tx and Rx utilize the summation $\sum(h_{tx,i} + h_{i,rx})$ as random series to extract secret keys.

We now detail our key extraction protocol FBCH, which consists of the following steps:

Let a transmitter Tx and a receiver Rx be the two parties that wish to extract a key; when Tx wants to establish a secret key with Rx, the Tx first determines an integer N as the number of helpers, and randomly picks N helpers H_i from all of the available relays in its transmission range. The number of available relays is M. Then the Tx broadcasts training symbols $\mathbf{x}_T$ under the standard transmission power P_T, the received signal at each helper H_i from the Tx is given by

$$\mathbf{y}_{i,tx} = P_T h_{tx,i}\mathbf{x}_T + \mathbf{N} \tag{21}$$

where $\mathbf{N}$ is the additive Gaussian white noise. Thus, each H_i can measure the channel Tx $\rightarrow H_i$ and obtain the CIR $h_{tx,i}$.

Then each helper H_i broadcasts a training symbol $\mathbf{x}_i$ under the standard transmission power P_H, the Tx and Rx receive signals from H_i, the received signals at the Tx and Rx are given by

$$\mathbf{y}_{tx,i} = P_H h_{tx,i}\mathbf{x}_i + \mathbf{N} \tag{22}$$

and

$$\mathbf{y}_{rx,i} = P_H h_{i,rx}\mathbf{x}_i + \mathbf{N} \tag{23}$$

, respectively. The Tx and Rx measure channels $H_i \rightarrow$ Tx and $H_i \rightarrow$ Rx to obtain the CIRs $h_{tx,i}$, and $h_{i,rx}$, respectively.

In the next step, the Rx broadcasts a training symbol $\mathbf{x}_R$ under the standard transmission power P_R, each helper H_i receives signal from the Rx, the received signals at each helper H_i is given by

$$\mathbf{y}_{i,rx} = P_R h_{i,rx}\mathbf{x}_R + \mathbf{N} \tag{24}$$

Therefore, each helper H_i can measure the channel Rx $\rightarrow$ H_i and obtain the CIR $h_{i,rx}$.

By using power control technology, each helper H_i manipulates its transmission power to P'_H and transmits the same training symbols $\mathbf{x}_i$ to the Tx again, where $P'_H = \frac{P_H h_{i,rx}}{h_{tx,i}}$; the received signal $\mathbf{y}'_{tx,i}$ at the Tx is given by

$$\begin{aligned}\mathbf{y}'_{tx,i} &= P'_H h_{tx,i}\mathbf{x}_i + \mathbf{N} \\ &= \frac{P_H h_{i,rx}}{h_{tx,i}} h_{tx,i}\mathbf{x}_i + \mathbf{N} \\ &= P_H h_{i,rx}\mathbf{x}_i + \mathbf{N}\end{aligned} \tag{25}$$

When the Tx measures the channel $H_i \rightarrow$ Tx again, the CIR $h'_{tx,i}$ that the Tx obtains is given by

$$h'_{tx,i} \triangleq \frac{y'_{tx,i}}{P_H} = h_{i,rx} \tag{26}$$

Again, each helper H_i manipulates its transmission power to P''_H and transmits the same training symbols $\mathbf{x}_i$ to the Rx, where $P''_H = \frac{P_H h_{tx,i}}{h_{i,rx}}$; the received signal $\mathbf{y}'_{rx,i}$ at the Rx is given by

$$\begin{aligned}\mathbf{y}'_{rx,i} &= P''_H h_{i,rx}\mathbf{x}_i + \mathbf{N} \\ &= \frac{P_H h_{tx,i}}{h_{i,rx}} h_{i,rx}\mathbf{x}_i + \mathbf{N} \\ &= P_H h_{tx,i}\mathbf{x}_i + \mathbf{N}\end{aligned} \tag{27}$$

When the Rx measures the channel $H_i \rightarrow$ Rx again, the CIR $h'_{i,rx}$ that the Rx obtains is given by

$$h'_{i,rx} \triangleq \frac{y'_{rx,i}}{P_H} = h_{tx,i} \tag{28}$$

Since the Tx and Rx obtain the CIRs $h_{tx,i}$ and $h_{i,rx}$ in step 3, respectively, they have the agreement that

$$\begin{aligned}h_{tx,i} + h'_{tx,i} &= h_{i,rx} + h'_{i,rx} \\ &= h_{tx,i} + h_{i,rx}\end{aligned} \tag{29}$$

Therefore, the Tx and the Rx are able to use the summation $\sum_{i=1}^{N}(h_{tx,i} + h_{i,rx})$ as random series to extract secret key bits, where N is the number of helpers. We should note that since the physical environments are quasi-static, the coherence time of the corresponding channels is relative large. The procedures of FBCH can be completed within a very short time window (several millisecond). During such short time window, the CIRs, e.g, $h_{tx,i}$ and $h_{i,rx}$, do not change.

7.2 Security analysis

In this subsection, we discuss the security of the proposed secret key extraction protocol FBCH under attack models of SIAs. In practice, the helpers may be compromised by the adversary and not every helper is trustworthy. So we discuss the security of FBCH in two scenarios: un-trusted helpers and trusted helpers.

7.2.1 Un-trusted helpers

In this scenario, we assume that some of helpers can be compromised by the adversary, so that not every helper is trustworthy. The un-trusted helpers can act as passive bad nodes or active bad nodes.

For passive bad nodes, they receive and transfer CIRs as normal. However, they may reveal the corresponding CIRs to the adversary. In fact, this leakage of partial information is not threatening at all, because the adversary cannot obtain key bits from the partial information and it is difficult for him to compromise all of the N helpers. As we described in Algorithm 1, the Tx and the Rx use the

summation of CIRs from N helpers, i.e., $\sum_{i=1}^{N}(h_{tx,i}+h_{i,rx})$, as random series to extract secret key bits. Therefore, from the partial CIR information, which is obtained from compromised helpers, the adversary cannot get the summation $\sum_{i=1}^{N}(h_{tx,i}+h_{i,rx})$, unless he compromises all of the N helpers. However, in the first step of Algorithm 1, the Tx randomly picks N helpers from M available relays in its transmission range. When M grows large, the possibility that the compromised nodes cover all of the N helpers becomes very small.

For active bad nodes, they will deny to pass the CIRs between the two legitimate communication ends or tamper with the CIR information before its transmission. In this case, it is easy for the Tx and the Rx to detect the falsifications, because the summations $\sum_{i=1}^{N}(h_{tx,i}+h_{i,rx})$ will not be consistent at the two ends. As such, they will simply drop the obtained random series, and re-pick N helpers to perform FBCH protocol again.

7.2.2 Trusted helpers

In this scenario, we assume all helpers are trustworthy. According to this assumption, first of all, in the FBCH protocol, the Tx and Rx use the summation $\sum_{i=1}^{N}(h_{tx,i}+h_{i,rx})$ to extract key bits (N is the number of helpers), the number N is secret to the passive adversary, it is hard for the passive adversary to get this summation of CIRs for key extraction. Moreover, by introducing helper nodes, the FBCH protocol hides the relevance between one link signature and the corresponding two locations of Tx and Rx. As a result, in the operational phase of SIAs, the ML models are not able to provide the proper group link signatures, since the adversary has no knowledge about the construction of new channels during its site survey phase. Likewise, it is hard for the adversary to infer the key bits in case II and case III of SIAs by applying ML methods.

Second, we analyze the security strength of FBCH from the spatial randomness perspective. By launching SIAs, the adversary has the ability to infer the link signature of any channel between any two locations. If there exists at least one attack node that is in close proximity to each helper node (here the "close" means that the distance between attack node and helper is smaller than $\lambda/2$), the strong correlation will lead to the accurate inference of link signature on each legitimate link. Therefore, the close distance between attack nodes and helpers will threaten the security of FBCH protocol. We will study the probability that an attack node is placed very close to a legitimate node by lucky coincidence.

Since the distributions of attack nodes and legitimate nodes (i.e., Tx, Rx and helpers) are fully random, the number of attack nodes that are close to each legitimate node follows the Poisson distribution. Therefore, we can apply the spatial Poisson point process to analyze this probability as follows:

The Poisson point process is defined in the plane $\mathbf{R^2}$. And we consider a circular area $B_i \subset \mathbf{R^2}$ ($i = 1...N$, and N is the number of legitimate nodes) for one legitimate node, which takes the legitimate node as the center and r as the radius. We treat the attack nodes as points in the plane $\mathbf{R^2}$. The number of points of a point process X existing in this area B_i is a random variable, denoted by $X(B_i)$. The points belong to a Poisson process with parameter $\lambda > 0$, then the probability of k points existing in B_i is given by

$$P\{X(B_i)=k\} = \frac{(\lambda|B_i|)^k}{k!}\cdot exp(-\lambda|B_i|), \tag{30}$$

where $|B_i|$ denotes the area of B_i, and $|B_i| = \pi r^2$; λ is the density of points in the plane $\mathbf{R^2}$.

Given by Eq. (30), the probability that at least one point existing in B_i is given by

$$\begin{aligned} 1 - P\{X(B_i)=0\} &= 1 - exp(-\lambda|B_i|) \\ &= 1 - exp(-\lambda\pi r^2) \end{aligned} \tag{31}$$

To launch SIAs successfully, there should be at least one attack node existing in B_i for every legitimate node. Then the total probability P that at least one attack node existing in B_i for each legitimate node is given by

$$\begin{aligned} P &= \prod_{i=1}^{N}(1 - exp(-\lambda|B_i|)) \\ &= (1 - exp(-\lambda\pi r^2))^N, \end{aligned} \tag{32}$$

where N is the number of legitimate nodes.

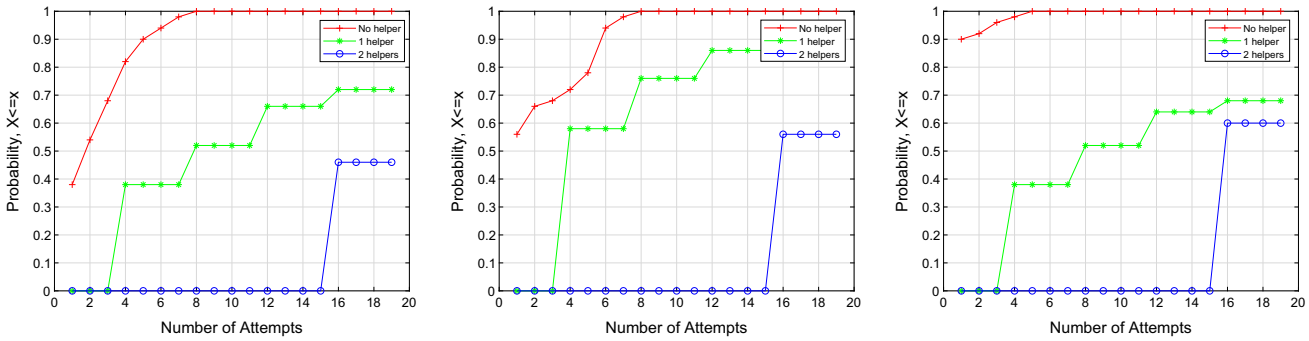
It can be observed from Eq. (32) that when the number of helpers N becomes large or the density of attack nodes λ becomes small, the probability P becomes small.

In particular, if N gets very large, we obtain

$$P = \lim_{N\to\infty}(1 - exp(-\lambda\pi r^2))^N = 0 \tag{33}$$

Given by Eq. (33), we can observe that when we pick large enough number of helpers (i.e., $N \to \infty$), the probability that there exists at least one attack node in B_i for each legitimate node is 0, which implies that there is no possible for the adversary to launch SIAs successfully.

Obviously, FBCH will exponentially increase the overhead of communication in the number of helpers during the key extraction process. Nevertheless, it is an effective countermeasure solution to defend against the SIAs.



(a) Case I **(b)** Case II **(c)** Case III

Fig. 14 SIA strength vs. number of helpers (SVM)

7.3 Numerical results

To illustrate the security strength of FBCH, in this subsection we study the performance of statistical inference attacks (SIAs), which we proposed in previous sections, to the new key extraction scheme. Our experiments are based on the same setup as in Section VI-B: we use the same dataset and launch SIAs in 3 cases. As pointed out in Section VI, support vector machine (SVM) has the highest inference accuracy in the previous experiments. To make a fair comparison, we only use SVM as the ML inference method to launch SIAs.

In this experiment, we first pick N nodes from utah/CIR dataset as the legitimate Tx, Rx and helpers, respectively. Furthermore, we randomly pick several links as the transmitter-helpers (Tx-H) links and helpers-receiver (H-Rx) links. Then we use the summations of CIRs of these several links, $h + h' + ...$, as ground truth to extract secret keys. Since in each case of SIAs, the adversary has no knowledge about the helper' selection, he has to infer CIRs on all links according to each potential helper (the location of each helper is known to the adversary). For case I SIA, the adversary uses all remaining links in utah/CIR dataset to infer the CIRs of potential Tx-H and H-Rx links and calculate the summation of these CIRs. Then he uses this summation to guess secret keys, as we mentioned in Section III-B. Likewise, for case II SIA, the adversary attempts to infer each potential Tx-H and H-Rx link using links that share the same receiver but have a different transmitter. And for case III SIA, the adversary randomly picks helpers and infers potential Tx-H and H-Rx links (the locations of Tx, Rx, and each helper are known).

To obtain the statistical view about the security strength of the proposed protocol FBCH, we study the percentage of bits in a secret key string that can be inferred, as a function of the number of attempts. Figure 14 compares the CDF (cumulative distribution function) of the number of guesses needed per point under 3 SIA cases. In each attack case, we vary the number of helpers and use different size of training data to measure the security strength of the proposed scheme. This figure shows that the adversary needs more attempts to find the true key, and FBCH can exponentially amplify the the adversary's search space. For example, in case I SIA, the adversary can find 94% points in just 6 rounds when Tx and Rx use conventional LSB key extraction scheme. On the contrary, 38% points are found in 6 rounds when there is 1 helper, while 0 point is found in 6 rounds when there are 2 helpers. In particular, FBCH can significantly increase the adversary's search space and effectively prevent the SIA in Case III. For example, when Tx and Rx use conventional LSB key extraction scheme, the adversary can find 90% true bits in secret key in just one round, while they cannot find any true bit in 14 rounds when two or more helpers are involved.

8 Conclusion and future work

In summary, the formal theoretical analysis in channel correlations have been done relying on both outdoor and indoor models. Following the machine learning (ML) framework, we have studied empirical statistical inference attacks against LSB key extraction. Different from prior analytical work that assumes a link-correlation model, our study is based on from empirically measured channel data and does not rely on any assumption on the link correlation. We applied several ML-based methods to launch SIA against LSB key extraction under various scenarios, and evaluated the effectiveness of these attacks based on the utah/CIR dataset. Our finding has verified the existence of correlation between neighboring links in realistic environments, and also showed that such correlation can be practically exploited by ML algorithms to undermine the security strength of PHY-layer security measures. Upon investigation, we proposed a countermeasure against the statistical inference attacks called forward-backward

cooperative key extraction protocol with helpers (FBCH). Our experiments verify that FBCH is more robust under the statistical inference attacks.

In future work, how to improve the inference algorithms, and analytically decide the optimal training data size, so that the adversary can construct the best site survey strategy to maximize its attack strength, remain questions to be explored. Moreover, it would be interesting to incorporate mmWave systems, and study the security and efficiency of key extraction protocols in mmWave ultra-dense networks or hybrid networks. It would be another interesting topic to implement the new robust cooperative key extraction protocol on real mobile devices. To the best of our knowledge, there is no secure and efficient LS-based secret key extraction tool on real mobile devices (e.g., smart phone, tablet, etc).

Acknowledgements This research work is partially supported by the National Science Foundation under Grants CNS-1837034, CNS-1745254, CNS-2006998, CNS-1460897 and DGE-1623713. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors would like to thank Mr. James F. Huber for proofreading and editing the language of this paper.

References

1. Kyritsi, P., Cox, D., Valenzuela, R., & Wolniansky, P. (2003). Correlation analysis based on mimo channel measurements in an indoor environment. *IEEE Journal on Selected Areas in Communications, 21*(5), 713–720.
2. Patwari, N., & Kasera, S. K. (2007). Robust location distinction using temporal link signatures. In *Proceedings of the ACM International Conference on Mobile Computing and Networking (MobiCom)*, 2007, pp. 111–122.
3. Zhang, J., Kasera, S. K., & Patwari, N. (2010). Mobility assisted secret key generation using wireless link signatures. In *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, 2010, pp. 1–5.
4. Wilson, R., Tse, D., & Scholtz, R. A. (2007). Channel identification: Secret sharing using reciprocity in ultrawideband channels. *IEEE Transactions on Information Forensics and Security, 2*(3), 364–375.
5. Patwari, N., Hero, A., Perkins, M., Correal, N., & O'Dea, R. (2003). Relative location estimation in wireless sensor networks. *IEEE Transactions on Signal Processing, 51*(8), 2137–2148.
6. He, X., Dai, H., Shen, W., & Ning, P. (2013). Is link signature dependable for wireless security?" In *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, 2013, pp. 200–204.
7. He, X., Dai, H., Shen, W., Ning, P., & Dutta, R. (2016). Toward proper guard zones for link signature. *IEEE Transactions on Wireless Communications, 15*(3), 2104–2117.
8. Edman, M., Kiayias, A., & Yener, B. (2011). On passive inference attacks against physical-layer key extraction? In *Proceedings of the Fourth European Workshop on System Security (EUROSEC)*, 2011, pp. 1–6.
9. Patwari, N. (2007). CRAWDAD dataset utah/cir (v. 2007-09-10). [Online]. Available: https://crawdad.org/utah/CIR/20070910
10. Barral, V. (2020). Pozyx cir and range with los and nlos. [Online]. Available: https://doi.org/10.21227/sr92-6s06.
11. Mathur, S., Trappe, W., Mandayam, N., Ye, C., & Reznik, A. (2008). Radio-telepathy: Extracting a secret key from an unauthenticated wireless channel. In *Proceedings of the ACM International Conference on Mobile Computing and Networking (MobiCom)*, pp. 128–139.
12. Zhu, R., Shu, T., & Fu, H. (2017). Empirical statistical inference attack against phy-layer key extraction in real environments. In *Proceedings of the IEEE Military Communications Conference (MILCOM)*, pp. 46–51.
13. Liu, Y., Draper, S. C., & Sayeed, A. M. (2012). Exploiting channel diversity in secret key generation from multipath fading randomness. *IEEE Transactions on Information Forensics and Security, 7*(5), 1484–1497.
14. Chen, K., Natarajan, B. B., & Shattil, S. (2015). Secret key generation rate with power allocation in relay-based lte-a networks. *IEEE Transactions on Information Forensics and Security, 10*(11), 2424–2434.
15. Im, S., Choi, J., & Ha, J. (2015). Secret key agreement for massive mimo systems with two-way training under pilot contamination attack. In *IEEE GLOBECOM Workshops*, pp. 1–6.
16. Zeng, K. (2015). Physical layer key generation in wireless networks: Challenges and opportunities. *IEEE Communications Magazine, 53*(6), 33–39.
17. Truyen Thai, C. D., Lee, J., & Quek, T. Q. S. (2015). Secret group key generation in physical layer for mesh topology. In *Proceedings of the IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6.
18. Moara-Nkwe, K., Shi, Q., Lee, G. M., & Eiza, M. H. (2018). A novel physical layer secure key generation and refreshment scheme for wireless sensor networks. *IEEE Access, 6,* 11374–11387.
19. Shimizu, T., Iwai, H., & Sasaoka, H. (2011). Physical-layer secret key agreement in two-way wireless relaying systems. *IEEE Transactions on Information Forensics and Security, 6*(3), 650–660.
20. Thai, C. D. T., Lee, J., Prakash, J., & Quek, T. Q. S. (2019). Secret group-key generation at physical layer for multi-antenna mesh topology. *IEEE Transactions on Information Forensics and Security, 14*(1), 18–33.
21. Jin, R., Du, X., Zeng, K., Huang, L., Xiao, L., & Xu, J. (2017). Delay analysis of physical-layer key generation in dynamic roadside-to-vehicle networks. *IEEE Transactions on Vehicular Technology, 66*(3), 2526–2535.
22. Fang, H., Wang, X., & Hanzo, L. (2019). Learning-aided physical layer authentication as an intelligent process. *IEEE Transactions on Communications, 67*(3), 2260–2273.
23. Kong, Y., Lyu, B., Chen, F., & Yang, Z. (2018). The security network coding system with physical layer key generation in two-way relay networks. *IEEE Access, 6,* 40673–40681.
24. Fang, H., Xu, L., Zou, Y., Wang, X., & Choo, K.-K.R. (2018). Three-stage stackelberg game for defending against full-duplex active eavesdropping attacks in cooperative communication. *IEEE Transactions on Vehicular Technology, 67*(11), 10788–10799.
25. Li, G., Hu, A., Sun, C., & Zhang, J. (2018). Constructing reciprocal channel coefficients for secret key generation in fdd systems. *IEEE Communications Letters, 22*(12), 2487–2490.
26. Zhao, H., Zhang, Y., Huang, X., & Xiang, Y. (2019). An adaptive secret key establishment scheme in smart home environments. In *Proceedings of the IEEE International Conference on Communications (ICC)*, pp. 1–6.
27. Hajomer, A. A. E., Zhang, L., Yang, X., & Hu, W. (2020). Post-processing protocol for physical-layer key generation and

distribution in fiber networks. *IEEE Photonics Technology Letters, 32*(15), 901–904.
28. Henkel, W., Turjman, A. M., Kim, H., & Qanadilo, H. K. H. (2020). Common randomness for physical-layer key generation in power-line transmission. In *Proceedings of the IEEE International Conference on Communications (ICC)*, pp. 1–6.
29. Aldaghri, N., & Mahdavifar, H. (2020). Physical layer secret key generation in static environments. *IEEE Transactions on Information Forensics and Security, 15*, 2692–2705.
30. Ribouh, S., Phan, K., Malawade, A. V., Elhillali, Y., Rivenq, A., & Faruque, M. A. A. (2020). Channel state information-based cryptographic key generation for intelligent transportation systems. *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–12.
31. Lin, R., Xu, L., Fang, H., & Huang, C. (2020). Efficient physical layer key generation technique in wireless communications. *EURASIP Journal on Wireless Communications and Networking, 2020*,
32. Jakes, W. C., & Cox, D. C. (Eds.). (1994). *Microwave Mobile Communications*. New York: Wiley.
33. Fang, H., Wang, X., & Tomasin, S. (2019). Machine learning for intelligent authentication in 5g and beyond wireless networks. *IEEE Wireless Communications, 26*(5), 55–61.
34. Qiu, X., Dai, J., & Hayes, M. (2020). A learning approach for physical layer authentication using adaptive neural network. *IEEE Access, 8,* 26139–26149.
35. Steinmetzer, D., Schulz, M., & M. Hollick, (2015). Lockpicking physical layer key exchange: Weak adversary models invite the thief. In *Proceedings of the ACM Conference on Security & Privacy in Wireless and Mobile Networks (WiSec)*, pp. 1–11.
36. Liu, Y., & Ning, P. (2012). Enhanced wireless channel authentication using time-synched link signature. In *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, pp. 2636–2640.
37. Dautov, R., & Tsouri, G. R. (2019). Effects of passive negative correlation attack on sensors utilizing physical key extraction in indoor wireless body area networks. *IEEE Sensors Letters, 3*(7), 1–4.
38. Zafer, M., Agrawal, D., & Srivatsa, M. (2012). Limitations of generating a secret key using wireless fading under active adversary. *IEEE/ACM Transactions on Networking, 20*(5), 1440–1451.
39. Law, Y. W., Palaniswami, M., Hoesel, L. V., Doumen, J., Hartel, P., & Havinga, P. (2009). Energy-efficient link-layer jamming attacks against wireless sensor network mac protocols. *ACM Transactions on Sensor Networks, 5*(1), 1–38.
40. Zhou, H., Huie, L. M., & Lai, L. (2014). Secret key generation in the two-way relay channel with active attackers. *IEEE Transactions on Information Forensics and Security, 9*(3), 476–488.
41. Clark, M. (2012). Robust wireless channel based secret key extraction. In *Proceedings of the IEEE Military Communications Conference (MILCOM)*, pp. 1–6.
42. Jin, R., & Zeng, K. (2015). Physical layer key agreement under signal injection attacks. In *Proceedings of the IEEE Conference on Communications and Network Security (CNS)*, pp. 254–262.
43. Hu, Q., & Hancke, G. P. (2017). A session hijacking attack on physical layer key generation agreement. In *Proceedings of the IEEE International Conference on Industrial Technology (ICIT)*, pp. 1418–1423.
44. MirhoseiniNejad, S. M., Rahmanpour, A., & Razavizadeh, S. M. (2018). Phase jamming attack: A practical attack on physical layer-based key derivation. In *Proceedings of the International ISC Conference on Information Security and Cryptology (ISCISC)*, pp. 1–4.
45. Rottenberg, F., Nguyen, T.-H., Dricot, J.-M., Horlin, F., & Louveaux, J. (2021). Csi-based versus rss-based secret-key generation under correlated eavesdropping. *IEEE Transactions on Communications, 69*(3), 1868–1881.
46. Harshan, J., Chang, S.-Y., & Hu, Y.-C. (2017). Insider-attacks on physical-layer group secret-key generation in wireless networks. In *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6.
47. Malmirchegini, M., & Mostofi, Y. (2012). On the spatial predictability of communication channels. *IEEE Transactions on Wireless Communications, 11*(3), 964–978.
48. Shiu, D.-S., Foschini, G., Gans, M., & Kahn, J. (2000). Fading correlation and its effect on the capacity of multielement antenna systems. *IEEE Transactions on Communications, 48*(3), 502–513.
49. Abdi, A., & Kaveh, M. (2002). A space-time correlation model for multielement antenna systems in mobile fading channels. *IEEE Journal on Selected Areas in Communications, 20*(3), 550–560.
50. Chen, P.-Y., & Li, H.-J. (2007). Modeling and applications of space-time correlation for mimo fading signals. *IEEE Transactions on Vehicular Technology, 56*(4), 1580–1590.

Rui Zhu received the Ph.D. degree from Oakland University, USA, in 2019. He received the B.S. and M.S. degree from Beijing Jiaotong University, China, and Valparaiso University, USA, respectively. He is currently a software development and robotics engineer at Verizon Location Technology. Previously, he worked as a machine learning research scientist in the Glycomics Center at the University of New Hampshire. His research interests include PHY-layer security in wireless network and millimeter-wave communication security, machine learning, autonomous mobile robots (AMRs), multi-access edge computing (MEC), and digital surface modeling (DSM).

Tao Shu received the B.S. and M.S. degrees in electronic engineering from the South China University of Technology, Guangzhou, China, in 1996 and 1999, respectively, the Ph.D. degree in communication and information systems from Tsinghua University, Beijing, China, in 2003, and the Ph.D. degree in electrical and computer engineering from The University of Arizona, Tucson, AZ, USA, in 2010. He is currently an Associate Professor in the Department of Computer Science and Software Engineering at Auburn University, Auburn, AL. His research aims at addressing the security, privacy, and performance issues in wireless networking systems, with strong emphasis on system architecture, protocol design, and performance modeling and optimization.

Huirong Fu (M'01) received the Ph.D. degree from Nanyang Technological University, Singapore, in 2000. She is currently a Professor with the Department of Computer Science and Engineering, Oakland University, Rochester, MI, USA, where she joined as an Assistant Professor in 2005. Previously, she was an Assistant Professor with North Dakota State University, Fargo, ND, USA, for three years, and she was a Postdoctoral Research Associate with Rice University, Houston, TX, USA, for two years. As a Lead Professor and the Principal Investigator for several projects funded by the National Science Foundation, she has been actively conducting research in the areas of networks, security, and privacy.