Fast and Secure Key Generation with Channel Obfuscation in Slowly Varying Environments

Guyue Li, Haiyu Yang
Dept. of CSE, SEU
Purple Mountain Laboratories,
guyuelee, hy_yang@seu.edu.cn

Junqing Zhang
Dept. of EEE, University of Liverpool
junqing.zhang@liverpool.ac.uk

Hongbo Liu
Dept. of CS, UESTC
hongbo.liu@uestc.edu.cn.

Aiqun Hu
NCRL, SEU
aqhu@seu.edu.cn

***Abstract*—Physical-layer secret key generation has emerged as a promising solution for establishing cryptographic keys by leveraging reciprocal and time-varying wireless channels. However, existing approaches suffer from low key generation rates and vulnerabilities under various attacks in slowly varying environments. We propose a new physical-layer secret key generation approach with channel obfuscation, which improves the dynamic property of channel parameters based on random filtering and random antenna scheduling. Our approach makes one party obfuscate the channel to allow the legitimate party to obtain similar dynamic channel parameters, yet prevents a third party from inferring the obfuscation information. Our approach allows more random bits to be extracted from the obfuscated channel parameters by a joint design of the K-L transform and adaptive quantization. Results from a testbed implementation show that our approach, compared to the existing ones that we evaluate, performs the best in generating high entropy bits at a fast rate and is able to resist various attacks in slowly varying environments. Specifically, our approach can achieve a significantly faster secret bit generation rate at roughly 67 bit/pkt, and the key sequences can pass the randomness tests of the NIST test suite.**

I. Introduction

Physical-layer secret key generation (PKG) has emerged as a promising solution to enable two wireless nodes to generate shared secret keys from the observation and processing of radio channel parameters [1]. Since PKG avoids the need for key distribution by leveraging the reciprocity nature of wireless channels, it is thus appealing to complement traditional cryptographic approaches for scenarios where pre-shared keys may not exist [2].

Existing PKG approaches, however, heavily rely on the channel variation and thus suffer from low key generation rate [3] and vulnerabilities under attacks [4] in slowly varying environments. When the wireless channel varies over time, the keys can be renewed dynamically by sampling the channel parameters. The channel sampling rate is suggested to be in the order of the maximum Doppler frequency [5]. However, when users have low mobility, e.g., in a wireless sensor network, the Doppler frequency is low, and it thus requires a long time to establish long enough keys. Furthermore, the mobility of users and the dynamic change of surrounding objects, as well as the maximum Doppler frequency, are often unknown. For these reasons, there always exists inevitable and unknown temporal correlations between adjacent channel samples, resulting in a large proportion of repeated bit segments in the quantized bit sequences. Although these bits segments are scrambled through some permutation or interleaving techniques [6]–[8], the security of the key is still compromised as the permutation information is public. Another idea of introducing helper devices, e.g., relays [9], [10] and reconfigurable intelligent surface (RIS) [11] to assist secret key generation may improve the key rate and randomness, however, encounters some practical problems, such as the unavailability of trust relays and additional hardware overheads of RIS devices. Therefore, a fast and secure solution is required to facilitate the practical usage of PKG in slowly varying environments.

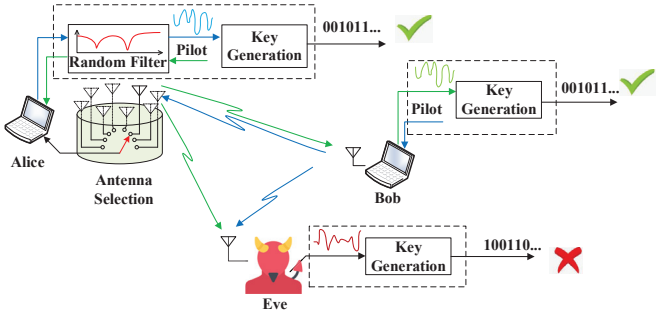


Fig. 1. The basic idea of channel obfuscation based secret key generation.

This paper designs a new physical layer key generation approach that can construct fast-changing channel parameters and thus works for slowly varying environments. We introduce a ***channel obfuscation*** technique, which makes one party obfuscate the channel in a manner that still allows the legitimate party to obtain dynamic channel parameters, yet prevents a third party from inferring the obfuscation information. We show that the Channel Obfuscation-based Secret Key Generation (CO-SKG) method achieves a fast key extraction at a high security level, even in slowly varying environments.

The basic idea underlying CO-SKG is illustrated in Fig. 1. Alice and Bob are two legitimate devices, and Eve is an eavesdropper. As a general assumption in PKG, Eve is assumed to be located at least half wavelength away from Alice and Bob, so its channel variations are assumed to be independent of Alice and Bob [1], [12]. Alice obfuscates the channel parameters by using a random filter and a random

antenna selection module. For each round in these repeated bidirectional channel probings, the pilot signal and received signal of Alice go through the filter and antenna with the same setting to obtain similar channel parameters with Bob. The settings of the antenna and filter are randomly changed to achieve fast key generation. Since these settings are known only by Alice, Eve can hardly crack the secret key.

CO-SKG builds on past work on exploiting multiple-antenna diversity to increase the bit generation rate [13]. Past work, however, typically measures the Received Signal Strength Indicator (RSSI) between each antenna pair in a round-robin way, leading to a periodic variation of RSSI in slowly varying environments. In contrast, CO-SKG measures the channel state information (CSI) between antenna pairs randomly to introduce unpredictable fluctuations in CSI values. CO-SKG also addresses the following practical challenges in using channel obfuscation for secret key generation. First, although the obfuscation information is not public, it might be speculated by a clever Eve from its prolonged channel observations. For instance, Eve may guess the order of antenna pairs by matching its current CSI with previous records. Therefore, a sophisticated design of the channel obfuscation function should be developed to prevent the attacker from inferring the obfuscation information. Second, the reciprocity of the obfuscated CSI is far from satisfactory, and adjacent CSI samples still have some auto-correlation, which may lead to long 0s and long 1s in the quantized bit sequences. Therefore, an efficient key generation approach is needed to satisfy the keys' agreement, rate, and randomness requirements.

The main contributions of this work are listed as follows:

- We propose a novel channel obfuscation approach that is conducive to fast secret bit extraction in slowly varying environments. Meanwhile, our approach prevents a third party from inferring the obfuscation information.
- We propose an effective key generation approach based on a joint design of K-L transform and adaptive quantization that would significantly improve the key agreement and randomness.
- We implemented CO-SKG using Universal Software Radio Peripheral (USRP) software defined radio (SDR) platforms and realizing antenna scheduling with an SP8T switch. Extensive experiments have been conducted, and the results demonstrate that compared with existing typical approaches, CO-SKG can provide higher key agreement, faster key generation rate, and more substantial randomness in three typical, slowly varying scenarios.
- We have verified that CO-SKG is resistant to various attacks identified as harmful to existing approaches in slowly varying environments, including the predictable channel attack and position replay attack from active attackers and effective brute-force attack and order speculation attack of passive attackers.

II. Related Works

There have been ongoing research efforts on fast key generation from wireless channels in slowly varying environments. First, diversity techniques, e.g., orthogonal frequency division multiplexing (OFDM) and multiple input multiple output (MIMO), have been exploited to improve the key rate by extracting more bits from one channel sample [14]–[16]. For example, Liu *et al*. [17] use channel responses from multiple OFDM subcarriers to provide fine-grained channel information, and Zeng *et al*. [13] use multiple-antenna diversity to increases the bit generation rate by more than four times over single-antenna systems. Despite these efforts, these protocols still require dynamic environments to generate keys with high entropy continuously. To address this issue, opportunistic beamforming [15], [18], [19] is exploited in multiple antenna systems. However, its performance has only been evaluated through theoretical analysis and simulations.

Second, secret key generation with helper devices such as relays has also been proposed [9], [10], [18]. Although this relay-based solution may increase the key rate, it is affected by design challenges. The key rate depends on the relay movement, but it is impractical to persuade a relay to move all the time. Moreover, relays are not always available and trustable [20], which limits the practical usage of this solution. Recently, Paul *et al* [11] has reported promising experimental results on integrating a passive reconfigurable intelligent surface (RIS) device to help in achieving a higher secret bit generation rate in static environments. Although RIS device does not have a trustworthy problem, it adds additional hardware overhead to the key generation system.

Another solution is to apply permutation [7], [8] or interleaving technique [6], [21] to increase the randomness of the weak key generated in the static environments. The channel samples are rearranged according to a pseudo-random sequence generated by specific algorithms. However, for legitimate users to agree on the same key, the pseudo-random sequence of interleaving is shared on the public channel, which means that an attacker also knows the information. The attacker can guess the original weak key and then interleave it for correctness verification. Hence, from the view of key cracking, it is futile to interleave the CSI samples.

Our work differs from existing works in the following significant ways. First, we propose a channel obfuscation method to allow Alice and Bob to obtain similar dynamic channel parameters, yet prevents Eve from inferring the obfuscation information. Second, we further increase the secret bit rate and randomness by a joint design of K-L transform and adaptive quantization. Third, we perform extensive real-world measurements in various slowly varying environments and settings to determine the effectiveness and security of the proposed CO-SKG approach.

III. Problem Formulation and the CO-SKG Scheme

A. Problem Formulation

As shown in Fig. 1, Alice (A) and Bob (B) intend to generate a shared secret key from their reciprocal CSI over N OFDM subcarriers. Eve (E) attempts to find the key based on its channel observations and information transmitted over the public channel. Alice is equipped with M antennas, and Bob

has a single antenna. They share M spatial channels, which are collected one after another, by switching the working antenna of Alice. Therefore, the system works in a single-input single-output (SISO) mode, and Eve is assumed to be equipped with a single antenna.

Alice and Bob complete K rounds of channel probing to obtain their CSI samples. For the k-th round, the signal received on subcarrier n by user $u \in \{A, B, E\}$ can be expressed as

$$Y_u(k,n) = H_u(m_k,n)S(k,n) + Z_u(k,n), \tag{1}$$

where $H_u(m_k,n)$ represents the spatial channel of the m_k-th antenna pair in the k-th round, $S(k,n)$ represents the known probe signal, and $Z_u(k,n)$ is the noise. In the existing work of [13], the antennas are used in a round-robin way, which implies that

$$m_k = (k+c)_M, \tag{2}$$

where $(\cdot)_M$ represents the modulo M operation and c is a constant integer with the range of $\{1, 2, \cdots, M-1\}$. Based on the received signal, the estimated CSI at user u is given by

$$\hat{H}_u(k,n) = H_u(m_k,n) + Z'_u(k,n), \tag{3}$$

where $Z'_u(k,n)$ represents the equivalent noise after estimating $H_u(m_k,n)$. According to the principle of channel reciprocity, when Alice and Bob probe the channel within the channel coherence time, the estimated CSI values, i.e., $\hat{H}_A(k,n)$ and $\hat{H}_B(k,n)$ should be highly correlated, which allows Alice and Bob to extract the same secret bit sequence from their CSI estimates, respectively. The secret bit sequence is continually renewed in the following key generation time rounds.

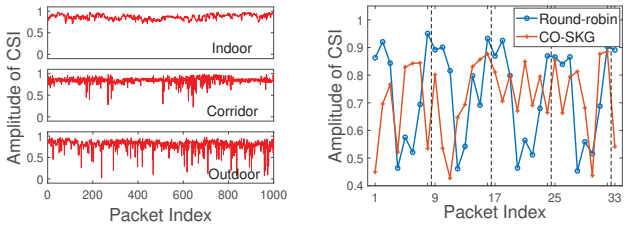
However, the CSI for each antenna pair may have little fluctuation in slowly varying environments. When antennas are scheduled in a fixed order periodically, we have

$$\hat{H}_u(k+M,n) \approx \hat{H}_u(k,n), \tag{4}$$

which indicates that the collected CSI samples may change regularly with time. To verify it, we conducted experiments under three slowly varying scenarios, i.e., indoor, corridor, and outdoor. Two USRP N210 devices operating in the 2.535 GHz channel are deployed as Alice and Bob, which remain stationary during the experiments. We extracted CSI samples from packets over 1000 time rounds, each with a time interval of 3 seconds. Fig. 2(a) shows that for a single antenna case, the amplitude of CSI on one subcarrier has slight variation during the 1000 rounds. Next, Alice is connect to an SP8T switch to realize the function of antenna selection. When the antenna is used in a round-robin way, as predicted in (4), the blue curve in Fig 2(b) exhibits significant regularity with a period of 8. These periodic changes may result in a large proportion of repeated bit segments in the quantization results.

B. The CO-SKG Scheme

1) Preliminary Study: One way to realize channel obfuscation is by randomly changing the index of the employed antenna in each packet. In this way, the order of CSI samples



(a) Single antenna (b) Multiple antennas

Fig. 2. The amplitude of CSI collected in slowly varying environments.

varies from one probing round to another. As a result, the CSI samples will have more fluctuation, largely avoiding the periodic repetition. We also experiment with randomly changing antennas on the USRP platform, and the amplitude of CSI sequences is shown as the red line in Fig. 2(b). As expected, the CSI estimated in different rounds has significant variation, and the previous regularity has disappeared.

2) Practical Challenges: The above preliminary study provides encouraging results on the feasibility of improving channel fluctuation with channel obfuscation in slowly varying environments. However, there are still two main challenges to integrate channel obfuscation with secret key generation securely and efficiently.

How to hide the obfuscation information? A desired channel obfuscation approach should obfuscate the channel without leaking the obfuscation information. However, when channel fluctuation is simply caused by antenna switching, the obfuscation information might be known by Eve. In a slowly varying environment, the CSI of Eve also varies little when Alice selects the same transmit antenna. This fact helps Eve to derive the relationship of the used antenna pair by matching the current CSI with previous CSI records.

How to produce secret keys effectively? Converting the obfuscated CSI samples into consistent secret keys is also challenging. Since the equivalent obfuscation channel changes rapidly, smoothing algorithm [22], usually used to improve the similarity of channel parameters in existing works, does not apply to the obfuscated CSI samples. When the inconsistent rate of the raw key is high, the information reconciliation imposes a large burden to correct these errors, which can significantly affect the efficiency of key generation. In addition, there is still inevitable auto-correlation between these CSI samples, which leads to long 0 and long 1 in the quantized results.

To cope with these challenges, we propose a CO-SKG approach, which is divided into two parts, i.e., **channel obfuscation** and **effective key generation**, as shown in Fig. 3. Alice and Bob obtain the channel estimates during channel obfuscation by sending pilot signals to each other in turn. Different from existing works, Alice combines the technologies of random filtering and antenna scheduling to fluctuate the CSI samples. After channel obfuscation, Alice and Bob convert their channel estimates into secret keys through effective key generation. We exploit K-L transform and adaptive quantization to obtain raw key bits with high similarity and strong

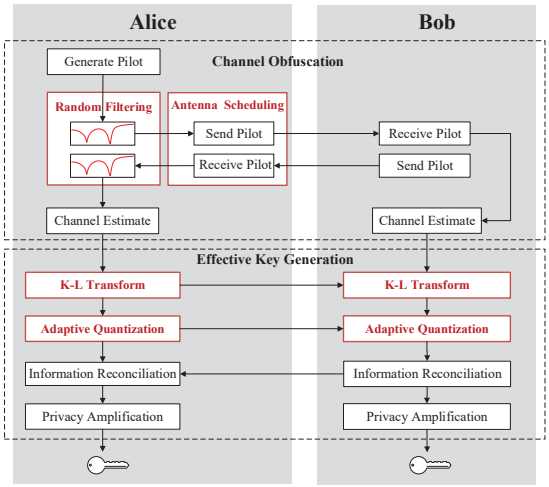


Fig. 3. The flow of CO-SKG.

randomness. Then, Alice and Bob obtain consistent keys from their raw key bits through information reconciliation and privacy amplification. Details of the protocol design of channel obfuscation and effective key generation are elaborated in Section IV and Section V, respectively.

C. Attack Model

We consider four types of attacks, which exploit the vulnerability of key generation in slowly varying environments. In the first two types of attacks, Eve intends to infer the key of the current round, while in the last two types of attacks, Eve aims at cracking the renewed key with an assumption that the current key has been known.

Predictable Channel Attack [23]: Eve tries to cause desired or predictable changes in the channel measurements by controlling the movement of intermediate objects between Alice and Bob.

Position Replay Attack [15]: In stationary environments, the channel parameters might be identical at the same position even when the measurements are made at different times. In this case, Eve records the position of Bob and moves to it after Bob leaving to acquire similar channel measurements.

Effective Brute-force Attack: Eve predicts the proportion of repeated bit segments between successive key generation processes, according to the variation of its channel observations. Hence, the time to recovering a renewed key can be drastically shortened as Eve only needs to try a small proportion of bits that have been changed.

Order Speculation Attack: In the PKG approaches based on permutation, the fluctuation of CSI samples is caused by the randomness of the sample order. A clever Eve will speculate the order from its channel observations and the information transmitted over the public channel.

Besides, Eve is assumed to be more interested in the extracted keys between Alice and Bob, but not disrupting their key establishment process. Therefore, we also assume that Eve can neither jam the communication channel between Alice and Bob nor modify the information exchanged between them.

IV. Channel Obfuscation

A. Design Philosophy

The desired CSI $\hat{H}_u(k,n)$ should meet the following three requirements.

- Randomness: $\hat{H}_A(k,n)$ should change over k in an unpredictable way. As a counterexample, periodic variable m_k in (2) should be avoided, because it obfuscates the CSI samples with regular changes that can be predicted.
- Reciprocity: Channel obfuscation should not affect the reciprocity principle, i.e., Alice and Bob are able to obtain similar CSI samples as $\hat{H}_A(k,n) \approx \hat{H}_B(k,n)$.
- Security: The obfuscation information should not been leaked from $\hat{H}_E(k,n)$ to Eve, otherwise the security of $\hat{H}_A(k,n)$ is compromised.

Accordingly, we design a time-varying function, $\mathbf{G}(k,n)$, which is given by

$$\mathbf{G}(k,n) = \alpha(k,n)\mathbf{g}_k, \tag{5}$$

where $\mathbf{g}_k \in \mathbb{R}^{M\times 1}$ is the antenna selecting vector, in which the m_k-th element is 1 and others are 0, and $\alpha(k,n)$ is the weighting coefficient of the selected channel. Denoting the channel vector as $\mathbf{H}_u(n) = [H_u(1,n), H_u(2,n), \cdots, H_u(M,n)]$, the CSI estimated by user u at the k-th round is

$$\begin{aligned}\hat{H}_u(k,n) &= \mathbf{H}_u(n)\mathbf{G}(k,n) + Z'_u(k,n)\\ &= \alpha(k,n)H_u(m_k,n) + Z'_u(k,n).\end{aligned} \tag{6}$$

We incorporate random variables $\alpha(k,n)$ and m_k into function $\mathbf{G}(k,n)$ to obtain the desired CSI. The necessity of their combination is expounded through two cases as follows.

Case 1: When $\alpha(k,n) = 1$ and m_k is an integer uniformly distributed variable, $m_k \sim U[1,M]$, the CSI is obfuscated by randomizing index m_k of M spacial channels. The obfuscated CSI can meet the requirements of randomness and reciprocity. However, its variation range is limited mainly by the number of antennas M. When Alice uses the same antenna m_k for channel probing round of k and k', Eve is able to observe $\hat{H}_E(k,n) \approx \hat{H}_E(k',n)$. With the increase of channel probing time K, the CSI differences between different antenna pairs will become more distinguishable.

Case 2: When $m_k = (k)_M$ and $\alpha(k,n)$ is a complex random variable, the CSI is obfuscated by multiplying the periodic CSI with a random coefficient $\alpha(k,n)$ that is known only by Alice. However, when only random coefficient is exploited, Eve is possible to speculate the obfuscation information of the coefficient according to the known periodic index of m_k. To clarify the speculation process, we omit the noise term and put the CSI with the same m_k into a vector. The CSI vectors of Alice and Eve are respectively given by

$$\vec{H}_{A,n} = [\alpha(k,n), \alpha(k+M,n), \cdots]\, H_A\left((k)_M, n\right), \tag{7}$$

$$\vec{H}_{E,n} = [\alpha(k,n), \alpha(k+M,n), \cdots]\, H_E\left((k)_M, n\right). \tag{8}$$

It is worth noting that Eve can hardly know the exact value of $\vec{H}_{A,n}$, due to the unknown information of $H_A((k)_M, n)$ and $H_E((k)_M, n)$, but it can observe the obfuscation information of coefficients. For example, the normalized amplitude of $\vec{H}_{A,n}$ and $\vec{H}_{E,n}$ are the same.

Notably, the function $\mathbf{G}(k,n)$ solely relies on random variable m_k or $\alpha(k,n)$ will leak the obfuscation information. Fortunately, we find that the index m_k and coefficient $\alpha(k,n)$ have mutual remedying parameters in hiding the obfuscation information. First, $\alpha(k,n)$ prevents Eve from knowing the obfuscation information of m_k, as channel estimates of different time rounds are not the same due to the change of $\alpha(k,n)$. Second, m_k also prevents Eve from knowing the obfuscation information of $\alpha(k,n)$, as Eve can hardly collect CSI with the same antenna to find the obfuscation information of coefficients. Therefore, by combining random index m_k and random coefficient $\alpha(k,n)$, the desired CSI can be achieved through the channel obfuscation function in (6).

B. Design Protocol

To implement the function $\mathbf{G}(k,n)$ in reality, we propose a channel obfuscation protocol by adopting the technologies of *random antenna scheduling* and *random filtering*. The former randomizes the index m_k by randomly selecting the employed antenna of Alice for each channel probing. The latter randomizes the value of $\alpha(k,n)$ by adding a time-varying finite impulse response (FIR) filter with L_f taps to process the pilot signal and received signal at Alice. The tap coefficients of the FIR filter are denoted by $\vec{a} = [a_1, a_2, \cdots, a_{L_f}]$, where $a_i \sim \mathcal{CN}(0,1)$ is a random variable following the complex Gaussian distribution with zero mean and unit variance.

The pilot signal of Alice is cyclic convolved with the impulse response of the FIR filter as well as the channel. Since the cyclic convolution in the time domain is equivalent to the multiplication in the frequency domain, the value of $\alpha(k,n)$ in the vector $\alpha(k) = [\alpha(k,1), \alpha(k,2), \cdots, \alpha(k,N)]$ is

$$\alpha(k,n) = \sum_{i=1}^{L_f} a_i e^{-j2\pi \frac{ni}{N}}. \tag{9}$$

The detailed channel obfuscation for the k-th time round of key generation is summarized in Algorithm 1.

Thanks to the reciprocity principle, the CSI vectors received by Alice and Bob, i.e., $\vec{H}_A(\mathrm{k})$ and $\vec{H}_B(\mathrm{k})$, are proportional to each other. In the following rounds of channel probing, the process will be repeated with changing filter coefficients and antenna index. After K time rounds, Alice and Bob obtain their CSI matrices of $\hat{\mathbf{H}}_A = \left[\vec{H}_A(1), \vec{H}_A(2), \cdots, \vec{H}_A(K)\right]$ and $\hat{\mathbf{H}}_B = \left[\vec{H}_B(1), \vec{H}_B(2), \cdots, \vec{H}_B(K)\right]$, respectively. Then, they use these observations to generate the secret keys.

V. Effective Key Generation

Alice and Bob will generate secret keys from the collected CSI matrices $\hat{\mathbf{H}}_A$ and $\hat{\mathbf{H}}_B$, each of which has N subcarriers and K samples.

Algorithm 1 Channel obfuscation.

Input: The pilot signal $S = [S(1), S(2), \cdots, S(N)]$.
Output: The CSI vectors $\vec{H}_A(\mathrm{k})$ and $\vec{H}_B(\mathrm{k})$.

Alice:
for $i := 1$ to L_f **do**
  $a_i \leftarrow$ random variable following $\mathcal{CN}(0,1)$.
end for
Generate $\alpha(k)$ with $\{a_1, a_2, \cdots, a_{L_f}\}$ according to (9).
$S_A \leftarrow$ element-wise multiple S and $\alpha(k)$.
$m_k \leftarrow$ random integer following $U[1, M]$.
Sends S_A to Bob via the m_k-th antenna.
Bob:
Receives the signal $Y_B(k,n)$.
Estimates CSI by $\hat{H}_B(k,n) = \frac{Y_B(k,n)}{S(n)}$ and obtains $\vec{H}_B(\mathrm{k}) = \left[\hat{H}_B(k,1), \hat{H}_B(k,2), \cdots, \hat{H}_B(k,N)\right]^T$.
Sends S to Alice.
Alice:
Receives signal $Y_A(m_k, n)$ on m_k-th antenna .
$Y'_A(k,n) \leftarrow$ multiple $Y_A(k,n)$ and $\alpha(k,n)$.
Estimates CSI by $\hat{H}_A(k,n) = \frac{Y'_A(k,n)}{S(n)}$ and obtains $\vec{H}_A(\mathrm{k}) = \left[\hat{H}_A(k,1), \hat{H}_A(k,2), \cdots, \hat{H}_A(k,N)\right]^T$.

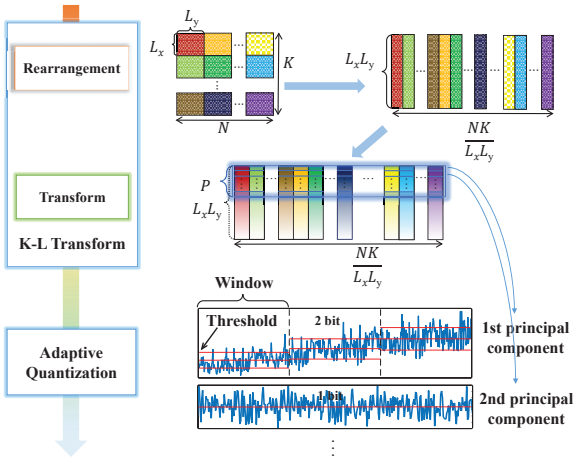


Fig. 4. Process of K-L transform and adaptive quantization.

A. K-L Transform

We apply the K-L transform to CSI samples for reducing the residual correlation between them. The process has two steps, i.e., CSI rearrangement and transform, as shown in Fig. 4.

Rearrangement: As shown in Fig. 4, elements in $\hat{\mathbf{H}}_A$ and $\hat{\mathbf{H}}_B$ are segmented into multiple blocks, with $L_x \times L_y$ CSI values in each block. There are $\frac{N}{L_x}$ and $\frac{K}{L_y}$ blocks in the direction of subcarrier and samples, respectively. The CSI values in different blocks are assumed to be uncorrelated, so we aim at removing the correlation between CSI values in the same block. Therefore, L_x and L_y are designed according to the assumption. The CSI values in the same block are vectorized into a column. Then, $\hat{\mathbf{H}}_A$ and $\hat{\mathbf{H}}_B$ are rearranged

as $\dot{\mathbf{H}}_A$ and $\dot{\mathbf{H}}_B$, each has $L_x \times L_y$ rows and $\frac{N}{L_x} \times \frac{K}{L_y}$ columns.

Transform: Based on $\dot{\mathbf{H}}_A$, Alice calculates the channel covariance matrix, $\mathbf{R}_A$, and performs the eigenvalue decomposition of $\mathbf{R}_A$ to obtain the eigenvector matrix $\mathbf{U}_A$ and eigenvalue matrix Λ_A. Only a small part of principal components are suitable for key generation, and others are severely corrupted by noise. So we construct the transform matrix $\mathbf{V}_A$ by selecting the first P column vectors in $\mathbf{U}_A$, which correspond to η percentage of energy. Alice sends $\mathbf{V}_A$ to Bob and they complete the transform by multiplying $\dot{\mathbf{H}}_A$ and $\dot{\mathbf{H}}_B$ with $\mathbf{V}_A$ as $\ddot{\mathbf{H}}_A = \mathbf{V}_A\dot{\mathbf{H}}_A, \ddot{\mathbf{H}}_B = \mathbf{V}_A\dot{\mathbf{H}}_B$, respectively. The transmission of eigenvector over an insecure public channel can cause information leakage and the information leakage ratio is $\eta_1 \approx 1/(\frac{N}{L_x} \times \frac{K}{L_y})$ as analyzed in [24].

B. Adaptive Quantization

The transformed CSI values in $\ddot{\mathbf{H}}_A$ and $\ddot{\mathbf{H}}_B$ are converted into bit sequences, $\mathbf{q}_A$ and $\mathbf{q}_B$, respectively, through an adaptive quantization algorithm that can adjust the *quantization level* and *thresholds* dynamically. These bit sequences are referred to as raw keys.

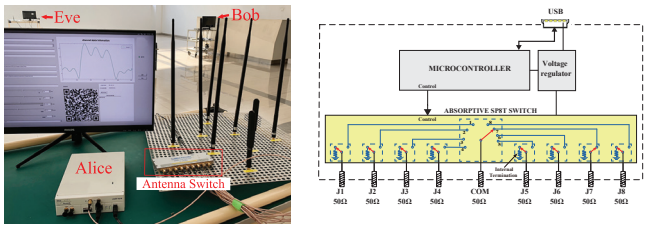
Since $\ddot{\mathbf{H}}_A$ and $\ddot{\mathbf{H}}_B$ contain multiple principal components with different signal-to-noise ratios (SNRs), a fixed quantization level does not apply to all components. Thus, we employ flexible quantization levels in the quantization algorithm. The levels L_p are decided by the desired disagreement ratio of raw keys and the corresponding SNRs.

For each component, the threshold should be adaptive to the variation of the values to avoid long 0s and long 1s. Thus, we add windows to the CSI sequence and quantize them in each window. The trade-off between the randomness and the disagreement ratio of the raw key is taken into account in the design of the window length, L_w. In the implementation, L_w is set according to the variance of the CSI values in the window. In each window, we apply multiple-level quantization with a guard band. The threshold is set based on the probability of the CSI values. Finally, the bit sequences, $\mathbf{q}_A$ and $\mathbf{q}_B$, are the combination of quantization results of each column in $\ddot{\mathbf{H}}_A$ and $\ddot{\mathbf{H}}_B$, respectively.

C. Information Reconciliation and Privacy Amplification

To obtain identical bit sequences, we apply a BCH code to correct the disagreements in raw keys according to their disagreement ratio. Bob calculates the syndrome of the $\mathbf{q}_B$ and sends it to Alice over the public channel. Alice corrects $\mathbf{q}_A$ according to the received syndrome. Transmitting the syndrome of BCH transmitted over public channels leaks information to eavesdroppers. The leakage rate in information reconciliation is $\eta_2 = L_s/L_q$ where L_q and L_s are the lengths of the raw key and the syndrome, respectively.

Privacy amplification allows legitimate users to distill a shorter but almost completely secret key from a common random variable about which Eve has partial information. We use Message-Digest Algorithm 5 (MD5) for privacy amplification. MD5 is a widely used hash function which maps data of arbitrary size to data of 128 bits. Considering the information leakage in the phases of K-L transform and information reconciliation, Alice and Bob should at least generate $L_{req} = \lceil \frac{128}{(1-\eta_1)(1-\eta_2)} \rceil$ bits common random sequence, where $\lceil \cdot \rceil$ represents ceiling operation.



(a) Hardware testbed (b) Antenna switch

Fig. 5. The experimental platform.

VI. Performance Evaluation

A. Experimental Setup

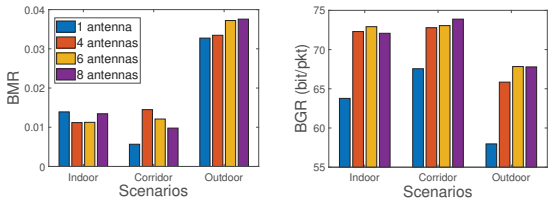
To evaluate the performance of the proposed CO-SKG approach in real environments, we built a hardware testbed using USRP SDR platform. Three USRP N210 SDR platforms embedded with the CBX daughterboards [25] were used as Alice, Bob, and Eve, as shown in Fig. 5(a). The experiments were carried out at the 2.535 GHz channel. To enable the function of antenna scheduling in the CO-SKG approach, Alice was connected with an SP8T switch [26] as shown in Fig. 5(b). Alice sent instructions to the microcontroller through software and controlled the SP8T switch to connect COM to one of J1 to J8. We carried out extensive experiments under three slowly varying scenarios, i.e., indoor, corridor, and outdoor. Three USRP devices remained stationary, but there might be people or vehicles moving around. We collected at least 1000 CSI vectors, each containing 512 values on these OFDM subcarriers [27].

We adopted three evaluation metrics: bit mismatch rate (BMR), bit generation rate (BGR), and random tests. The BMR is defined as the number of different bits between Alice and Bob divided by the total number of bits of the raw key, which is calculated using the raw bits before information reconciliation. The BGR is defined as the number of generated raw key bits per packet (bit/pkt). We used the National Institute of Standards and Technology (NIST) test suite [28] to evaluate the randomness of our generated key sequences.

B. Optimal Design Parameters in the CO-SKG Method

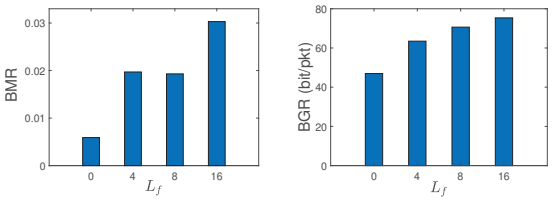
To improve the performance of CO-SKG , we conducted extensive experiments to find the optimal parameters in the processes of channel obfuscation and effective key generation.

1) Channel Obfuscation: Since antenna number and filter length are two critical parameters in the channel obfuscation, we evaluate the impacts of them on the performance. Fig. 6 compares the performance of BMR and BGR under different number of antennas, M. The impact of M on BMR varies for different scenarios, while for all three scenarios, the BGR grows with M. In particular, compared with the single antenna



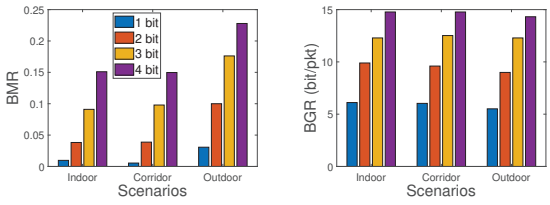
(a) BMR (b) BGR

Fig. 6. Performance under different number of antennas.



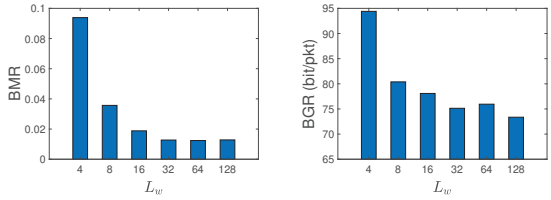
(a) BMR (b) BGR

Fig. 7. Performance under different filter length in an indoor scenario.



(a) BMR (b) BGR

Fig. 9. Performance of the impact of quantization levels of first component.



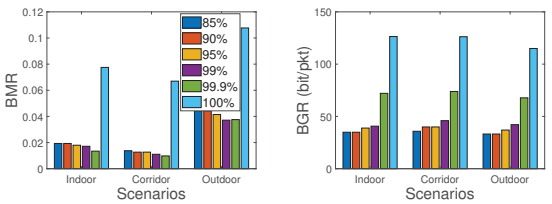
(a) BMR (b) BGR

Fig. 10. Performance under different window length in an indoor scenario.

case, the BGRs of other antenna numbers have increased by at least 5 bit/pkt. It is also observed that the BGR levels off when the antenna number becomes large, e.g., $M = 8$. In this case, random filtering has already provided enough randomness, and thus no need to deploy more antennas. According to the result, we choose the antenna number to be 8 in the design of random antenna scheduling.

Fig. 7 shows the impact of filter length, L_f, on the performance of BMR and BGR. In the experiment, L_f is set as 0, 4, 8 and 16, where $L_f = 0$ represents the case that random filtering is not in use. It is observed that when L_f increases from 0 to 16, both BMR and BGR increase. When random filtering is not used, BMR is below 0.01. When $L_f = 4$, BMR rises rapidly to 0.019, and after having a slightly decrease at $L_f = 8$, continuously increases to 0.03 at $L_f = 16$. For BGR, when random filtering is not used, BGR is 46.9 bit/pkt. When $L_f = 4$, BGR also rises rapidly to 63.4 bit/pkt, and then slowly increases to 76.3 bit/pkt at $L_f = 16$. To make a trade-off between BMR and BGR, we choose the filter length $L_f = 8$ in the design of the random filtering.

2) Effective Key Generation: We study the performance of secret key under different design parameters in the steps of K-L transform and adaptive quantization. Fig. 8 compares the performance of BMR and BGR with different proportions of principal components used for key generation. We specify the percentage of the principal components, η, to choose the first P principal components. We examine the BMR and BGR by changing η from 85% to 100% and find that BGR increases from around 30 bit/pkt to over 100 bit/pkt with the increase of η. The BMR decreases when η increases from 85% to 99.9%, but increases rapidly over 0.06 at $\eta = 100\%$. It is because the rest components are severely corrupted by noise. Therefore, we set the percentage threshold η as 99.9% in the K-L transform.



(a) BMR (b) BGR

Fig. 8. Performance under the impact of proportion of principal components.

Next, we perform multi-bit quantization on the first principal component, as its energy is much higher than others. We measure the BMR and BGR from 1 bit to 4 bits quantization of the first principal component. As shown in Fig. 9, both BGR and BMR grow as the quantization level increases. Since some bits are discarded in the guard band during quantization, the BGR is not a strict double relationship as the quantization level increases. Therefore, we need to balance the impact on BMR and BGR when choosing an appropriate quantization level. According to the experimental results, we set the quantization level for the first principal component as 2 bits.

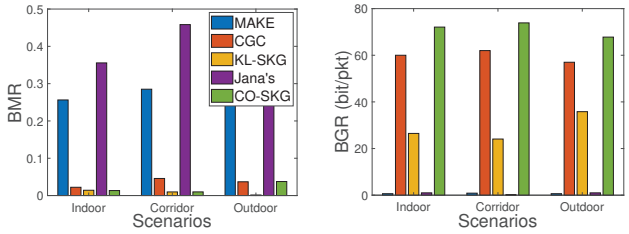
Finally, Fig. 10 shows the impact of window length, L_w, on the key generation performance. We find that BMR decreases from 0.094 to 0.012 as L_w increases from 4 to 64, and then increases to 0.138 when $L_w = 128$. BGR decreases from 94.4 bit/pkt to 71.5 bit/pkt as L_w increases from 4 to 128. To make a trade-off between these metrics, we set the window length $L_w = 64$. Table I summarizes above optimal parameters used in the CO-SKG approach.

C. Comparison with Existing Typical Key Generation Methods

Fig. 11 compares the BMR and BGR of the proposed CO-SKG approach with four existing typical key generation approaches, including MAKE [13], CGC [17], KL-SKG [24] and Jana's [23]. From Fig. 11(a), MAKE and Jana's have higher BMR, as they use RSSI for key generation, which has

TABLE I
OPTIMAL PARAMETERS IN THE CO-SKG APPROACH

Parameters	Values
Number of antennas of Alice, M	8
Filter length, L_f	8
Energy Percentage of selected principal components, η	99.9%
Quantization level of first principal component	2 bit
Window length, L_w	64



(a) BMR (b) BGR

Fig. 11. Performance comparison of different key generation approaches.

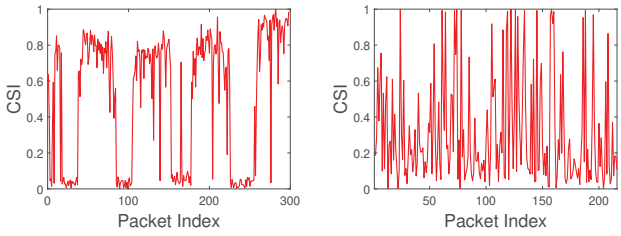
TABLE II
NIST STATISTICAL TEST RESULTS

Test	CO-SKG	KL-SKG	CGC	Jana's	MAKE
Freq.	0.4106	0.0709	1	0.7113	0.0240
Block Freq.	0.1108	0	0.0011	0.5035	0.4371
Runs.	0.4195	0	0.9500	0	0.1323
Longest of 1's.	0.1059	0	0	0	0.0080
Serial.	0.2118	0	0	0	0
	0.3169	0	0	0	0.0001
Approx. Entropy.	0.0468	0	0	0	0
Cumusum	0.5014	0.4433	0.4994	0.5035	0.0204

slight fluctuation, mainly dominated by the noise in slowly varying environments. CGC, KL-SKG, and CO-SKG extract secret key from CSI, which fluctuates over multiple subcarriers and thus have better performance of the key agreement in the slowly varying environments. Since KL-SKG and CO-SKG use the K-L transform to reduce noise, they have lower BMRs than CGC.

Fig. 11(b) shows that CO-SKG provides higher BGR than other approaches, achieving the rate of 72.08, 73.88, and 67.79 bit/pkt for scenarios of the indoor, corridor, and outdoor, respectively. It is because that CO-SKG exploits both random antenna scheduling and random filtering to increase the CSI fluctuation, so more components can be extracted in K-L transform. Since MAKE and Jana's use RSSI as the channel parameters, they obtain only one sample for each round, so their BGRs are significantly lower than those of CSI-based approaches, i.e., CGC, KL-SKG, and CO-SKG.

In Table II, we evaluate the randomness of the raw key sequences before privacy amplification using the widely adopted NIST random test suite. When the P-value test result is greater than the threshold, usually chosen as 0.01, the sequence passes the test. The key sequences are extracted in an indoor environment, and its length is set as 1024 bits. Table II shows



(a) Original (b) CO-SKG

Fig. 12. The amplitude of CSI under predictable channel attack in an indoor scenario.

that the proposed CO-SKG approach can pass all of these seven tests, while other approaches can only pass parts of them.

The above experimental results have verified that the proposed CO-SKG can provide secret keys with higher key agreement, faster key generation rate, and stronger randomness in slowly varying environments than existing approaches.

VII. SECURITY EVALUATION

In this section, we evaluate the security of the proposed CO-SKG approach against the four attacks described in Sec. III-C.

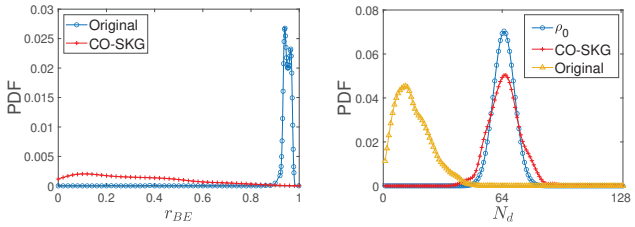
A. Preventing Predictable Channel Attack

CO-SKG prevents Eve from causing the predictable changes in the channel measurements by controlling the movements of some intermediate objects. According to (6), the variation of $\hat{H}_u(k,n)$ depends on the physical channel, as well as the selected antenna and the filter coefficients. Therefore, the pattern of variation of $\hat{H}_u(k,n)$ does not follow the movements of Eve, even in a slowly varying environment.

We evaluate the performance in a corridor environment, where Alice and Bob are placed on the two sides of the corridor with a door in the middle. The predictable channel attack is implemented by opening and closing the door periodically over 2 minutes. Fig. 12 compares the amplitude of CSI sequences on one subcarrier of CO-SKG against the original one without channel obfuscation. The pattern of variation in Fig. 12(a) is highly related to the opening or closing state of the door, while the relevance is significantly reduced in Fig. 12(b). In particular, when the door is closed, the channel becomes non-line-of-sight (NLoS), and thus the CSI amplitude in Fig. 12(a) is significantly less than that in the line-of-sight (LoS) case in which the door is opening. The CSI amplitude in Fig. 12(b) has a significant fluctuation under both cases, which indicates that CO-SKG can resist the predictable channel attack.

B. Preventing Position Replay Attack

Since $\hat{H}_u(k,n)$ is not determined solely upon the locations of Alice and Bob, CO-SKG can prevent Eve from obtaining the same key via location replay attacks. When Eve moves to the same position at another time round of k', it will obtain $\hat{H}_E(k',n)$, which is independent from $\hat{H}_B(k,n)$, as $\alpha_{k,n}$ and m' change independently over k.



(a) r_{BE} (b) N_d

Fig. 13. The probability distribution of r_{BE} and N_d in an indoor scenario.

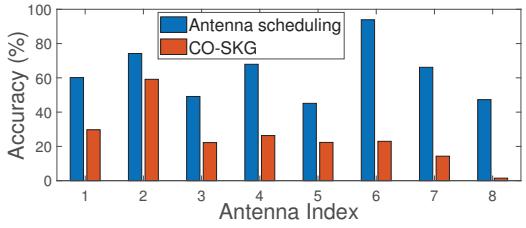


Fig. 14. Speculation accuracy of Eve in an indoor scenario.

We implemented the position replay attacks in an indoor scenario. After Alice and Bob communicate for a while, Bob moves to a new location, and Eve quickly moves to Bob's original position and attempts to obtain a similar CSI as that of Bob. Fig. 13(a) reports the probability distribution of the correlation coefficient over all subcarriers, r_{BE}. It is observed that the value of r_{BE} is concentrated at the range from 0.9 to 1 for the original approach, while its value is largely reduced to below 0.5 for the CO-SKG approach, which indicates that CO-SKG can effectively defend against the position replay attack in a slowly varying environment.

C. Preventing Effective Brute-Force Attack

CO-SKG prevents the Eve from shortening the time-complexity through the effective brute-force attack, which is described in Section III-C. By exploiting channel obfuscation, CSI samples will be changed irregularly and thus reduce the proportion of repeated bit segments in $\mathbf{q}_B$. We divide $\mathbf{q}_B$ into multiple groups, each of which contains 128 bits. The number of different bits between adjacent groups, N_d, is defined as $N_d = ||\mathbf{q}_B^{j+1} - \mathbf{q}_B^{j}||_1$, where j is the index of group and $||\cdot||_1$ represents the norm-1 function. Theoretically, when the raw keys in adjacent groups are independent, the probability distribution function of N_d should be $\rho_0 = C_{128}^{N_d}\left(\frac{1}{2}\right)^{128}$. Fig. 13(b) compares the probability distribution of N_d that is calculated from the raw keys generated by the original approach, the CO-SKG approach, and the theoretical function of ρ_0. It is observed that the value of N_d in the original approach is concentrated around 18, which deviates largely from the theoretical value ρ_0. In this case, Eve is able to infer the raw key in adjacent groups by the effective brute-force attack. Conversely, the probability distribution of N_d for the CO-SKG approach is close to that of the ρ_0, which indicates the effectiveness of CO-SKG in reducing the proportion of repeated bit segments in raw keys. In this way, CO-SKG can resist an effective brute-force attack.

D. Preventing Order Speculation Attack

By exploiting two degrees of freedom, i.e., m_k and $\alpha(k,n)$, CO-SKG prevents Eve from speculating the obfuscation information. We evaluate the resistance of CO-SKG under this attack in comparison with the approach solely using random m_k, which is referred to as the antenna scheduling approach. Eve implemented the order speculation attack with 1000 rounds of experimental data collected in an indoor environment. In each round, Eve obtained a CSI vector with 512 CSI values. Eve intends to speculate the index of the used antenna in the current time around by matching its CSI vector with those obtained from previous rounds. The K-means algorithm [29] is exploited by Eve to partition its unlabelled CSI vectors into eight distinct groupings, corresponding to $M = 8$ antennas. The speculation is defined as accurate when the output index of the K-means algorithm is equal to the index of antennas in use. Fig. 14 compares the speculation accuracy of Eve on different antennas, and the results show that Eve can speculate the antenna index with high accuracy in the antenna scheduling approach, while the accuracy is largely reduced in the proposed CO-SKG approach.

VIII. Conclusion

In this paper, we proposed a key generation protocol, CO-SKG, that combines channel obfuscation and effective key generation to provide fast and secure key generation in slowly varying environments. CO-SKG exploits two degrees of freedom, i.e., antenna index and filter coefficients, to obfuscate channel parameters and hide the obfuscation information. In addition, the joint design of K-L transform and adaptive quantization improves key agreement and randomness effectively. We implemented CO-SKG using USRP SDR platforms and realizing antenna scheduling with an SP8T switch. Extensive experiments were conducted, and the results demonstrate that compared with existing typical approaches, CO-SKG can provide higher key agreement, faster key generation rate, and stronger randomness in three slowly varying scenarios. Finally, experimental results have verified that our protocol achieves a high-security level against various attacks, including the predictable channel attack and position replay attack from active attackers and effective brute-force attack and order speculation attack of passive attackers. The authors have provided public access to their data at [27].

IX. Acknowledgment

We thank Dr. Jiabao Yu, Mr. Yanjun Ding, Dr. Dong Wang and Dr. Linning Peng, from the Purple Mountain Laboratories, for their help with the USRP SDR platform. This work was supported by the National Natural Science Foundation of China under Grants 6217011510, 61801115 and 61941115, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20211160.

REFERENCES

[1] J. Zhang, G. Li, A. Marshall, A. Hu, and L. Hanzo, "A new frontier for IoT security emerging from three decades of key generation relying on wireless channels," *IEEE Access*, vol. 8, pp. 138 406–138 446, 2020.

[2] G. Li, C. Sun, J. Zhang, E. Jorswieck, B. Xiao, and A. Hu, "Physical layer key generation in 5G and beyond wireless communications: Challenges and opportunities," *Entropy*, vol. 21, p. 497, 2019.

[3] Suwadi, Wirawan, and M. Yuliana, "Performance evaluation of secret key generation system for static and dynamic condition," in *IEEE ICC*, Batam, Indonesia, Dec. 2020, pp. 423–428.

[4] S. N. Premnath, S. Jana, J. Croft, P. L. Gowda, M. Clark, S. K. Kasera, N. Patwari, and S. V. Krishnamurthy, "Secret key extraction from wireless signal strength in real environments," *IEEE Trans. Mobile Comput.*, vol. 12, no. 5, pp. 917–930, 2013.

[5] S. Mathur, W. Trappe, N. Mandayam, C. Ye, and A. Reznik, "Radio-telepathy: extracting a secret key from an unauthenticated wireless channel," in *Proc. ACM MobiCom*, San Francisco, California, USA, Sep. 2008, pp. 128–139.

[6] W. Xi, M. Duan, X. Bai, K. Zhao, L. Mo, and J. Zhao, "Keep: Secure and efficient communication for distributed iot devices," *IEEE Internet Things J.*, pp. 1–1, 2020.

[7] W. Xi, C. Qian, J. Han, K. Zhao, S. Zhong, X.-Y. Li, and J. Zhao, "Instant and robust authentication and key agreement among mobile devices," in *Proc. ACM CCS*, New York, NY, USA, Oct. 2016, pp. 616–627.

[8] H. Liu, Y. Wang, Y. Ren, and Y. Chen, "Bipartite graph matching based secret key generation," in *Proc. IEEE INFOCOM*, Virtual Conference, May. 2021, pp. 1–9.

[9] N. Aldaghri and H. Mahdavifar, "Physical layer secret key generation in static environments," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 2692–2705, Feb. 2020.

[10] H. Zhou, L. M. Huie, and L. Lai, "Secret key generation in the two-way relay channel with active attackers," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 3, pp. 476–488, 2014.

[11] P. Staat, H. Elders-Boll, M. Heinrichs, R. Kronberger, C. Zenger, and C. Paar, "Intelligent reflecting surface-assisted wireless key generation for low-entropy environments," *ArXiv*, vol. abs/2010.06613, 2020.

[12] G. Li, C. Sun, E. A. Jorswieck, and et al., "Sum secret key rate maximization for TDD multi-user massive MIMO wireless networks," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 968–982, 2021.

[13] K. Zeng, D. Wu, A. Chan, and P. Mohapatra, "Exploiting multiple-antenna diversity for shared secret key generation in wireless networks," in *Proc. IEEE INFOCOM*, San Diego, CA, USA, Mar. 2010, pp. 1–9.

[14] D. Qin and D. Zhi, "Exploiting multi-antenna non-reciprocal channels for shared secret key generation," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 12, pp. 2693–2705, Dec. 2016.

[15] M. G. Madiseh, S. W. Neville, and M. L. McGuire, "Applying beamforming to address temporal correlation in wireless channel characterization-based secret key generation," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 4, pp. 1278–1287, 2012.

[16] Q. Wang, H. Su, K. Ren, and K. Kim, "Fast and scalable secret key generation exploiting channel phase randomness in wireless networks," in *Proc. IEEE INFOCOM*, Shanghai, China, Apr. 2011, pp. 1422–1430.

[17] H. Liu, Y. Wang, J. Yang, and Y. Chen, "Fast and practical secret key extraction by exploiting channel response," in *Proc. IEEE INFOCOM*, Turin, Italy, Apr. 2013, pp. 3048–3056.

[18] P. Huang and X. Wang, "Fast secret key generation in static wireless networks: A virtual channel approach," in *Proc. IEEE INFOCOM*, Turin, Italy, Apr. 2013, pp. 2292–2300.

[19] Y. Huang, L. Jin, H. Wei, Z. Zhong, and S. Zhang, "Fast secret key generation based on dynamic private pilot from static wireless channels," *China Communications*, vol. 15, no. 11, pp. 171–183, 2018.

[20] R. Guillaume, S. Ludwig, A. Muller, and A. Czylwik, "Secret key generation from static channels with untrusted relays," in *Proc. IEEE WiMob*, Abu Dhabi, United Arab Emirates, Oct. 2015, pp. 1–8.

[21] H. M. Furqan, J. M. Hamamreh, and H. Arslan, "New physical layer key generation dimensions: Subcarrier indices/positions-based key generation," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 59–63, 2021.

[22] X. Zhu, F. Xu, E. Novak, C. C. Tan, Q. Li, and G. Chen, "Using wireless link dynamics to extract a secret key in vehicular scenarios," *IEEE Trans. Mobile Comput.*, vol. 16, no. 7, pp. 2065–2078, 2017.

[23] S. Jana, S. N. Premnath, M. Clark, S. K. Kasera, N. Patwari, and S. V. Krishnamurthy, "On the effectiveness of secret key extraction from wireless signal strength in real environments," in *Proc. ACM MobiCom*, Beijing, China, Sep. 2009, pp. 321–332.

[24] G. Li, A. Hu, J. Zhang, and et al., "High-agreement uncorrelated secret key generation based on principal component analysis preprocessing," *IEEE Trans. Commun.*, vol. 66, no. 7, pp. 3022–3034, 2018.

[25] E. research, *USRP N200/N210 NETWORKED SERIES*, Mountain View, CA, 2012. [Online]. Available: http://www.ettus.com/

[26] Mini-Circuits, *Solid state USB RF SP8T Switch: USB-1SP8T-63H*, Brooklyn, NY, 2012. [Online]. Available: https://www.minicircuits.com

[27] G. Li, H. Yang, J. Zhang, H. Liu, and A. Hu, "Datasets for fast and secure key generation with channel obfuscation in slowly varying environments." Zenodo, Dec. 2021. [Online]. Available: https://doi.org/10.5281/zenodo.5795180

[28] A. Rukhin, J. Soto, J. Nechvatal, M. Smid, and E. Barker, "A statistical test suite for random and pseudorandom number generators for cryptographic applications," DTIC Document, Tech. Rep., 2001.

[29] J. Tou and R. Gonzalez, *Pattern Recognition Principles*. Addison-Wesley, 1977.