FREE: A Fast and Robust Key Extraction Mechanism via Inaudible Acoustic Signal*

Youjing Lu
Shanghai Jiao Tong University
luyoujing@sjtu.edu.cn

Fan Wu
Shanghai Jiao Tong University
fwu@cs.sjtu.edu.cn

Shaojie Tang
University of Texas at Dallas
tangshaojie@gmail.com

Linghe Kong
Shanghai Jiao Tong University
linghe.kong@sjtu.edu.cn

Guihai Chen
Shanghai Jiao Tong University
gchen@cs.sjtu.edu.cn

ABSTRACT

To build a secure wireless networking system, it is essential that the cryptographic key is known only to the two (or more) communicating parties. Existing key extraction schemes put the devices into the physical proximity, and utilize the common inherent randomness between the devices to agree on a secret key, but they often rely on custom devices and have low bit rate. In this paper, we seek a key extraction approach that only leverages off-the-shelf mobile devices, while achieving significantly higher key generation efficiency. The core idea of our approach is to exploit fast varying inaudible acoustic channel for generating enough randomness and wireless parallel communication for exchanging reconciliation information to improve the key generation rate. We have carefully studied and validated the feasibility of our approach through both theoretical analysis and a variety of measurements. We implement our approach on different mobile devices and conduct extensive experiments in different real scenarios. The experiment results show that our approach achieves high efficiency and satisfactory robustness. Compared with the state of art methods, our approach improves the key generation rate by 38.46% and reduces the bit mismatch ratio by 42.34%.

CCS CONCEPTS

• **Security and privacy → Mobile and wireless security**; • **Networks → Mobile and wireless security**; *Security protocols.*

KEYWORDS

Key extraction, Inaudible acoustic signal, Channel estimation

*F. Wu is the corresponding author.

†This work was supported in part by the National Key R&D Program of China 2018YFB1004703, in part by China NSF grant 61672348 and 61672353, in part by Supported by the Open Project Program of the State Key Laboratory of Mathematical Engineering and Advanced Computing 2018A09, and in part by Alibaba Group through Alibaba Innovation Research Program. The opinions,fi ndings, conclusions, and recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agencies or the government.

Mobihoc '19, July 2–5, 2019, Catania, Italy

ACM ISBN 978-1-4503-6764-6/19/07…$15.00
https://doi.org/10.1145/3323679.3326529

ACM Reference Format:
Youjing Lu, Fan Wu, Shaojie Tang, Linghe Kong, and Guihai Chen. 2019. FREE: A Fast and Robust Key Extraction Mechanism via Inaudible Acoustic Signal. In *The Twentieth ACM International Symposium on Mobile Ad Hoc Networking and Computing (Mobihoc '19), July 2–5, 2019, Catania, Italy*. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3323679.3326529

1 INTRODUCTION

Nowadays, with the emerging of various mobile devices, communication security becomes increasingly important and challenging. Different from wired network, wireless network is often built in a mobile and dynamic environment, where there are no preexisting infrastructures that support centralized cryptographic key generation and distribution. This poses a big challenge when it comes to generating and sharing cryptographic key among mobile devices in a secure manner.

Public key system and secret key exchange algorithms, such as Diffie-Hellman (DH) protocol, RSA algorithm, etc., can generate cryptographic key between two entities. Some works also rely on public key system to generate and exchange cryptographic key [27, 33]. However, all of them require the two entities to have some common priori knowledge about the modulus and base [6], and this is hard to achieve especially when two entities meet each other for thefi rst time.

Currently, the most common approach for extracting the cryptographic key is by using the common inherent randomness between the entities. These efforts can be classified into two categories. The first category of approaches exploits the reciprocity of wireless channel and extracts the channel randomness to generate secret key. For example, radio-telepathy extracts secret key from unauthenticated wireless channel [17]. Wang et al. exploited channel phase randomness to generate secret key [30]. The second category of approaches utilizes the randomness of environment sensing to extract secret key. Their main idea is to put the devices into physical proximity to get similar environment sensing data to generate common secret key. For example, Bichler et al. exploited acceleration data of shaking process to generate secret key [3]. MAGIK utilizes the dynamic geomagneticfi eld sensing data to extract secret key [23]. However, existing key extraction approaches often rely on custom devices and have low key generation rate. For example, the average key generation rate of radio-telepathy is only 1 bit per second [17]. An enhanced approach has been proposed in [21] with key generation rate $10 \sim 20$ bits per second. It needs several minutes to generate a 512-bit key. Although some approaches can achieve

higher key generation efficiency, they often require custom devices, i.e., Intel 5300 Network Interface Card (NIC) [8], or Atheros AR 9380 NIC, and a laptop that is compatible with above NICs [31, 32]. Therefore, there is a lack of appropriate key extraction methods that have high key generation efficiency and can be implemented on off-the-shelf mobile devices.

In this paper, we seek a key generation approach that only needs off-the-shelf mobile devices, while achieving significantly higher key generation efficiency. The core idea of our approach is to leverage the randomness of acoustic channel to extract secret key. Our study is motivated by the following observations fromfi eld tests. First, off-the-shelf mobile devices are equipped with microphone and speaker, which can be used to transmit and receive acoustic signal. Second, users can shake their mobile devices to vary the acoustic channels to extract more randomness, thus improving the key generation rate. Third, the key generation rate can be further improved by enabling parallel communication, i.e., users can communicate in wireless channel to reconciliate common secret key, while communicating in acoustic channel for extracting randomness. Besides, we choose inaudible frequency bands for not disturbing others.

However, it is highly non-trivial to realize this idea. In particular, we are facing three challenges: thefi rst challenge is to identify an effective way to grab similar acoustic channel randomness between mobile users; the second one is to quantize the acoustic randomness into bit stream; the last one is to reconciliate a common secret key from two similar bit streams in a secure manner.

To address these challenges, we propose a Fast and Robust key Extraction mEchanism, named FREE. Wefi rst study the feasibility of utilizing acoustic channel randomness for key extraction from the perspectives of theory and experiment. Fortunately, the acoustic channel is proved to have the significant properties of temporal variation, channel reciprocity, and spatial decorrelation, and thus it is a great medium to establish secret key. Then, we use the transmitted inaudible acoustic signal to estimate the common channels, and grab the channel randomness. To quantize the acoustic channel randomness, we use adaptive secret bit generation method to quantize a channel tap into a signal bit or multiple bits. To generate identical bit stream, we design a protocol for the two entities to reconciliate the mismatched bits.

To evaluate the performance of FREE, we build a FREE prototype on different pairs of mobile devices, and conduct extensive experiments in different scenarios. The experiment results validate the effectiveness and efficiency of FREE.

We now summarize the main contributions of this paper.

- We consider using the randomness of inaudible acoustic channel for key extraction, and demonstrate that it is a great medium to establish secret key. Our approach utilizes the validated properties of acoustic channel to defend against eavesdropping, approaching, and repeating imitation attacks.
- We successfully implement the acoustic channel based key extraction approach on the commodity mobile devices. We use parallel communication to improve the key generation rate, which is significantly higher than existing solutions [10, 14, 17, 23, 31, 32], e.g., the state of art approach can generate 180 ∼ 260 bits/sec at most [23, 32], while our approach can generate 320 ∼ 360 bits/sec average.

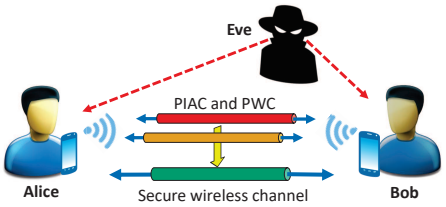


Figure 1: System model. Alice and Bob estimate the public inaudible acoustic channel (PIAC) to get the randomness, they exchange reconciliation information through public wireless channel (PWC), and extract identical cryptographic key from the randomness to establish a secure wireless channel. A passive adversary Eve, can eavesdrop all acoustic and wireless signals transmitted through PIAC and PWC, but cannot extract the same cryptographic key.

The rest of this paper is organized as follow. Section 2 presents the system model and attack model. Section 3 studies the feasibility of using acoustic channel randomness for key extraction. Section 4 details the design of FREE. Section 5 analyzes the security of FREE. Section 6 evaluates the performance of FREE in real-word experiments. Section 7 reviews the related work. Section 8 concludes this paper.

2 SYSTEM OVERVIEW

In this section, we present the overview of our system model and attack model.

2.1 System model

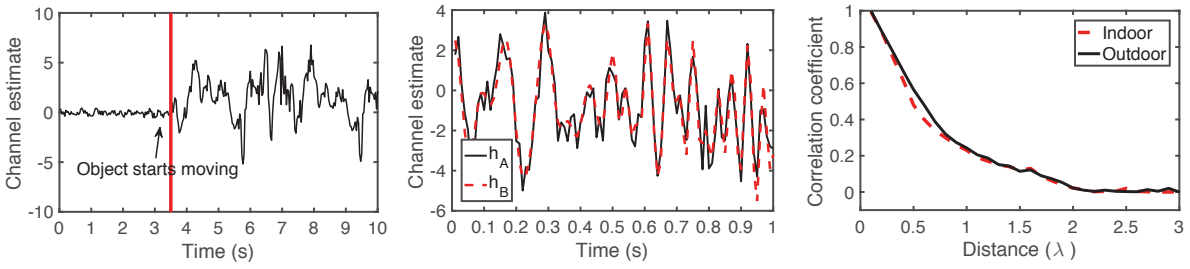
We illustrate our system model in Figure 1. There are two legitimate mobile users, Alice and Bob, which are located in physical proximity. To prevent the passive adversary Eve from eavesdropping their communication, they need a common secret key to establish a secure and authenticate channel between each other.

Alice and Bob are equipped with off-the-shelf mobile device, such as a consumer-grade mobile phone. They extract the randomness from the public inaudible acoustic channel (PIAC) between them, and use public wireless channel (PWC) to exchange some message to reconciliate a common secret key from the randomness, as shown in Figure 1.

Our goal is to utilize the randomness of inaudible acoustic channel to extract secret key, while achieving higher key generation rate and lower bit mismatch ratio.

2.2 Attack model

Next, we introduce possible attacks from a passive adversary Eve. Eve can overhear all signals transmitted through public inaudible acoustic channel and public wireless channel. Eve also can estimate his own acoustic channel and extract the channel randomness. He knows the key extraction algorithm with the parameters setting. We assume that Eve is not too close to either Alice or Bob, i.e., they are separated by at least 5 cm. We also assume that Eve's goal is to intercept the cryptographic key instead of jamming their communications. If Eve jams the communication between Alice and Bob (e.g., transmitting high-power acoustic signals), it is blocking the key extraction between Alice and Bob. When the acoustic channel



(a) Temporal variation. (b) Channel reciprocity. (c) Spatial decorrelation.

Figure 2: The properties of acoustic channel: temporal variation, channel reciprocity, and spatial decorrelation.

communication is blocked, neither Alice nor Bob can extract the secret key. Thus, the attacker cannot get the key either.

In particular, we mainly consider the following three kinds of attacks.

- Eavesdropping attack: Eve can eavesdrop all signals transmitted through public inaudible acoustic channel and public wireless channel between Alice and Bob, and he knows everything about the key extraction algorithm. Thus, he can analyze the captured acoustic and wireless signals to guess the secret key.
- Approaching attack: To get similar channel estimates, Eve can approach to legitimate users to receive similar acoustic signals from other side. Eve intents to exploit the similarities to generate the same cryptographic key as Alice or Bob.
- Repeating imitation attack: After Alice and Bobfi nished key extraction and left the site, Evefi nds a partner Dave to imitate the key extraction process conducted by Alice and Bob, i.e., they estimate the inaudible acoustic channel to get randomness, and then try to extract the same secret key as Alice or Bob.

3 FEASIBILITY STUDIES

In this section, we study the feasibility of using acoustic channel information to generate secret key. Wefi rst validate three important properties of acoustic channel, including temporal variation, channel reciprocity, and spatial decorrelation. These three properties together serve as the basis of our approach. (The frequency band of tested acoustic signals ranges from 0 Hz to 22k Hz. 22k Hz is maximum audio frequency which can be played by ordinary mobile devices. We have tested the channel estimates using different devices in four kinds of scenarios. The details of channel estimation and the experiment settings can be found in Section 4.2 and Section 6.1, respectively.)

Temporal variation: Wefi rst observe that due to multipath propagation, acoustic signals could reach the receiving antenna by two or more paths. Causes of multipath propagation include various obstacles, static or mobile, act as reflectors to the signals, change the reflection, refraction, and scattering of the channel paths. Besides, the movements of transmitter and receiver, e.g., the shaking of the mobile device equipped with microphone and speaker, also change the original channel paths.

We can formulate the signal propagation as below:

$$y(t) = \sum_{i=1}^{M} a_i x(t-\tau_i) = \sum_{i=1}^{M} a_i e^{-j2\pi f_c \tau_i} s(t-\tau_i) = h(t) * x(t). \quad (1)$$

In the above formula, we assume that the received acoustic signal $y(t)$ is received from M paths respectively. The signal transmitted along path i has amplitude a_i and delay τ_i, which are determined by reflectors and the signal travel distance. $x(t)$ and $s(t)$ are the transmitted passband and baseband signals at time t, respectively, f_c is the center frequency of passband, and $h(t)$ is the channel impulse response (CIR). $h(t) = \sum_{i=1}^{M} a_i e^{-j2\pi f_c \tau_i} \delta(t-\tau_i)$, where $\delta(t)$ is Dirac's delta function [20].

We aim to use the received signals to estimate the time-varying acoustic channel to grab enough randomness. The channel estimation from the received baseband symbol is a discrete output of $h(t)$ sampled every T_s interval [29], which is

$$h[n] = \sum_{i=1}^{M} a_i e^{-j2\pi f_c \tau_i} \delta(t-\tau_i) \mathrm{sinc}(n - \tau_i W), \quad (2)$$

where $\mathrm{sinc}(t) = \frac{\sin(\pi t)}{\pi t}$. Generally, CIR is regarded as a discrete-timefi lter in Linear Time-Invariant (LTI) system, and $h[n]$ is called the n-th channel tap. Same as time-varying acoustic channel, the calculated channel estimate $h(t)$ is also time-varying.

In addition to the theoretical analysis, we also verify the temporal variation of acoustic channel in experiments. We conducted extensive experiments in corridors and our labs. We estimated the acoustic channels between several different pairs of mobiles and averaged thefi rst 10 channel taps of their estimates, as shown in Figure 2 (a). It illustrates the variation of acoustic channel in a short time. When a reflector, such as a person, starts moving, the acoustic channel changes instantly and obviously. Thus, the temporal variation of acoustic channel offers enough randomness for key extraction.

Channel reciprocity: We next show that at the same carrier frequency, the multipath and fading of Alice → Bob direction are same as the Bob → Alice direction in the same link in a short period. Channel reciprocity is a fundamental property of signal wave propagation [24], and it is the basis of using acoustic channel randomness to generate common secret key. More specifically, we use symbol $\mathbf{h}$ to denote the channel parameter and assign it the channel impulse response $h(t)$. To get the channel parameter $\mathbf{h}$, Alice and Bob calculate the channel estimates $\mathbf{h_A}$ and $\mathbf{h_B}$, respectively. Theoretically, $\mathbf{h_A}$ and $\mathbf{h_B}$ are highly correlated.

We validate the channel reciprocity property in both indoor and outdoor environments. Figure 2 (b) shows that the channel estimates $\mathbf{h_A}$ and $\mathbf{h_B}$ have significant correlations. Due to page limit, we only plot the channel estimation outdoor but omit the channel estimation indoor. The indoor experiment also shows the same result.

Spatial decorrelation: We further observe that when the adversary Eve is more than half of wavelength away from legitimate users Alice and Bob, their mutlipath and fading are uncorrelated. This property guarantees the security of legitimate users' key extraction. It has been shown that both large-scale fading and small-scale fading contribute to channel variation [7]. The small-scale fading is dominant in our experiment, due to the short travel distance of the receiver and short time duration. In small-scale fading, the signals decorrelate over distance of approximately one half-wavelength [7].

We also validate the spatial decorrelation property of acoustic channel indoor and outdoor in our experiments. Figure 2 (c) shows the correlation between Eve's and legitimate user's estimates when the distance varies from 0 to 3 wavelengths. The carrier frequency is 20k Hz, and the wavelength λ is 1.7 cm. We use Pearson correlation coefficient to measure the correlation between their channel estimation. We observe that both indoor and outdoor correlation coefficients are lower than 0.2 when the distance is greater than one wavelength. Therefore, Eve cannot receive similar acoustic signal when he is located 5 cm away from legitimate users.

4 DESIGN OF FREE

In this section, we present the architecture and design details of FREE.

4.1 Design rationale

We illustrate the architecture of FREE in Figure 3. The architecture is mainly divided into four stages: acoustic channel estimation, quantization, reconciliation, and privacy amplification.

The above three properties make it possible to use acoustic channel randomness to extract secret key. Temporal variation of acoustic channel offers enough randomness for key extraction. Channel reciprocity is the basis of Alice and Bob having similar channel randomness. Spatial decorrelation makes users can resist against above passive attacks. We next address the following two challenges: 1. how to extract common channel randomness; 2. how to generate a common secret key from the channel randomness.

To extract common channel randomness, Alice and Bobfirst need to estimate the acoustic channels. There are many methods for channel estimation. Given the implementation on mobile devices, we choose Least-Square (LS) method [22], which only requires low computation overhead. Alice and Bob transmit acoustic signals to each other, and then use the received signals from each other to estimate the acoustic channel and get the channel randomness.

To generate a common secret key from the channel randomness, Alice and Bobfi rst need to quantize the channel randomness into bit stream. To improve the key generation rate, we choose adaptive secret bit generation method to quantize a channel tap into a signal bit or multiple bits [10]. After quantization, Alice and Bob get similar bit streams with some mismatched bits. To eliminate the mismatched bits, we design a protocol for Alice and Bob to reconciliate an identical bit stream.

In the above process, the interactions between Alice and Bob may leak some information about the secret bit stream. To eliminate such leakage, Alice and Bob perform the privacy amplification on their own bit stream, respectively. Finally, both Alice and Bob acquire a common secret key to establish a secure wireless channel between them.

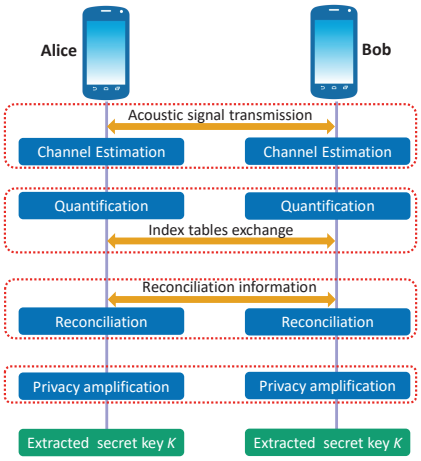


Figure 3: The architecture of FREE.

4.2 Design details

4.2.1 Acoustic channel estimation. We propose an inaudible acoustic signal transmission scheme to estimate the acoustic channel, referring to [34]. Since most adults can only hear the sound on the frequency band lower than 18k Hz and we do not want to disturb others, we choose to use frequency band from 18k Hz to 22k Hz, with the bandwidth of 4k Hz, which can be captured by microphones embedded on general smartphones and tablets.

Next, we use single-carrier to estimate the acoustic channel in time domain. To enhance the accuracy of channel estimation, we choose not to use multi-carriers, e.g., multi-carriers technique surrenders the channel estimate result in frequency domain.

Then we present the transmitter design and receiver design.

Transmitter design: To estimate the acoustic channel, transmitter transmits a known training sequence, which is indicated as $P = \{p_1, p_2, ..., p_L\}$, where L is the sequence length. The training sequence can be any random bit stream. We adopt the channel estimation method in [22] to choose a 26-bit Global System for Mobile Communication (GSM) training sequence, which is well known to outperform on synchronization and widely used in channel estimation. Then the training sequence P is modulated to the symbols of Gaussian Filtered Minimum Shift Keying (GMSK), which maps bits 0 and 1 to baseband symbols -1 and 1, respectively.

Figure 4 (a) depicts the system diagram of transmitter in the inaudible signal transmission. Wefi rst upsample the baseband symbol at a rate of $\frac{f_s}{B}$, where f_s and B represent sampling rate and bandwidth, respectively. The purpose of upsampling is to smooth discontinuity, by zero padding and low-passfi ltering [20]. Let f_c represent the center frequency of passband. We transform the signal frequency into the baseband signal: $x(t) = \sqrt{2}\cos(2\pi f_c t)p(t)$, where $p(t)$ and $x(t)$ represent upsampled baseband and passband signals, respectively.

To remove the noise outside transmission band, wefi lter the signal $x(t)$ with passband from $f_c - \frac{B}{2}$ to $f_c + \frac{B}{2}$ Hz. Then, the processed signal is transmitted by the speaker. Since the training sequence isfi xed, the generated signal is alsofi xed. To reduce the computation overhead, we save the generated signal as a Waveform

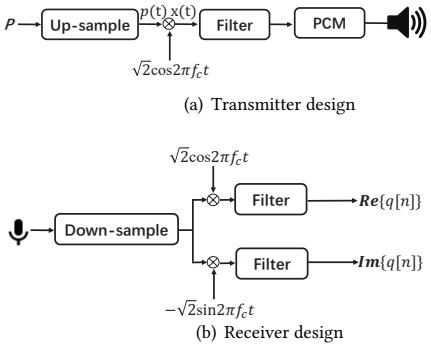


Figure 4: Design of transmitter and receiver.

Audio (WAV)fi le with a format of 16-bit Pulse Coded Modulation (PCM), which can be played by most of mobile devices.

To avoid inter-frame interference, transmitter cannot continuously transmit training sequence with no gap. We insert zeros at the end of training sequence, as shown in Figure 5. We consider a training sequence as a frame. The gap between training sequences must be long enough to prevent a frame interfering the previous frame. It has been shown that 24 zeros are long enough to avoid inter-frame interference in our experiments. Finally, we refer to the training sequence with zeros as a new frame, consisting of 50 bits. Since the baseband symbol interval is $T_s = \frac{1}{B} = 0.25$ ms, each frame lasts 12.5 ms.

Receiver design: Figure 4 (b) demonstrates the processing of received signal in receiver's side. The received signal $y[n]$ from microphone is converted into baseband symbol $q[n]$ as follows:

$$\begin{aligned} q[n] &= \sqrt{2}\cos(2\pi f_c t)y(t) - j\sqrt{2}\sin(2\pi f_c t)y(t) \\ &= \sqrt{2}e^{-j2\pi f_c t}y(t), \end{aligned} \tag{3}$$

where t is the time, and k-th baseband symbol is sampled, that is, $t = k \times T_s$, where T_s is the corresponding baseband symbol interval. We multiply $y(t)$ with $\sqrt{2}cos(2\pi f_c t)$ and $-sin(2\pi f_c t)$, and get the real and imaginary parts of received signal, respectively. Then both of real and imaginary parts are processed by low-passfi ltering and down-sampling.

To detect the arrival of a frame, the receiver uses energy detection and cross-correlation for received signal after baseband symbol conversion. We use energy detection to roughly determine the starting point of a frame: we set a threshold δ to measure whether the magnitude of three consecutive symbols is the starting point of a frame. The threshold δ is set as 0.005 in our study, and the setting relies on the microphone and the volume of the speaker. Then, we use cross-correlation method tofi nd precise starting point of the frame.

Since mobile devices are often resource constrained, we choose Least-Square (LS) method [22], which only requires low computation overhead. In LS channel estimation, wefi rst need to determine a reference length X and a guard length Y, and $X + Y$ is the length L of training sequence. The guard length Y determines the number of channel taps that we can estimate. To balance the number of channel taps and the estimation quality, we choose $X = 16$ and $Y = 10$ in our study. Readers canfi nd more details in [22].

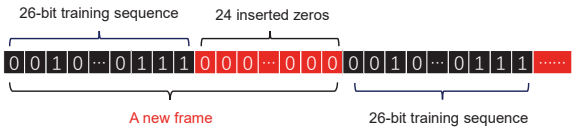


Figure 5: Training sequence with inserted zeros.

As mentioned above, the training sequence is denoted as $P = \{p_1, p_2, ..., p_i, ..., p_L\}$, $p_i \in \{-1, +1\}$. Then the corresponding circulant matrix $\mathbf{M} \in \mathbb{R}^{X \times Y}$ is formed as:

$$\mathbf{M} = \begin{bmatrix} m_Y & m_{Y-1} & \cdots & m_1 & m_0 \\ m_{Y+1} & m_Y & \cdots & m_2 & m_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ m_{Y+X-1} & m_{Y+X-2} & \cdots & m_X & m_{X-1} \end{bmatrix}. \tag{4}$$

Let $y = \{y_1, y_2, ..., y_i, ..., y_{X+Y}\}$ represent the received training sequence. Then we get the acoustic channel estimate as follows:

$$\hat{h} = (\mathbf{M}^H\mathbf{M})^{-1}\mathbf{M}^H y_Y, \tag{5}$$

where $\mathbf{M}^H$ and $\mathbf{M}^{-1}$ indicate the Hermitian and inverse matrices of $\mathbf{M}$, respectively. $y_Y = \{y_{Y+1}, y_{Y+2}, ..., y_{Y+X}\}$, which offers the randomness source for key generation.

From Equation (5), we can see that the most computationally expensive part is the computation of $(\mathbf{M}^H\mathbf{M})^{-1}\mathbf{M}$ which is a matrix-to-vector multiplication, and its computation complexity is $O(X \times Y)$. Given that both of X and Y are constants, the computation overhead is very low for mobile devices.

To generate enough randomness for key extraction, Alice and Bob need to continuously transmit modulated acoustic signal to each other. After i-th acoustic communication accomplished, both Alice and Bob get channel estimates, which are denoted by $\hat{H^i}_A$ and $\hat{H^i}_B$, respectively.

4.2.2Quantization. Once Alice and Bob get the acoustic channel estimates, they need to quantize these channel estimates to bit streams. To improve secret key generation rate, we use Adaptive Secret Bit Generation (ASBG) method to quantize a channel tap into a signal bit or multiple bits [10].

To illustrate the single bit quantization method formally, we suppose that Alice and Bob get a channel tap sequence, denoted by $\hat{H}_A = \left\{\hat{h}_A[1], \hat{h}_A[2], ..., \hat{h}_A[l]\right\}$ and $\hat{H}_B = \left\{\hat{h}_B[1], \hat{h}_B[2], ..., \hat{h}_B[l]\right\}$, respectively, where l is the length of channel estimates. The process of single bit quantization method is illustrated as follows:

• Alice divides $\hat{H}_A = \left\{\hat{h}_A[1], \hat{h}_A[2], ..., \hat{h}_A[l]\right\}$ into small blocks of size *block_size*, which is an adjustable parameter. Bob also performs the same operations.

• For each block, they calculate two adaptive thresholds $q+$ and $q-$, $q+ = mean + \alpha * \sigma$ and $q- = mean - \alpha * \sigma$, where $\alpha > 0$, *mean* is the mean of the magnitude of the estimates in a block, and σ is the standard deviation.

• Alice compares their channel estimates to the two thresholds, $q+$ and $q-$. If channel estimate $\hat{h}_A[i] > q+$, then $\hat{h}_A[i]$ is recorded as 1; if $\hat{h}_A[i] < q-$, then $\hat{h}_A[i]$ is recorded as 0; when $\hat{h}_A[i]$ lies in between $q+$ and $q-$, then $\hat{h}_A[i]$ is discarded, and the index i is recorded in an index table T_A. Bob performs the same operations on $\hat{H}_B$, and generates index table T_B.

• After the above operations accomplished, Alice and Bob exchange their index table, T_A and T_B. They only keep the channel estimates that are not discarded by either of them. Finally, they obtain the bit stream S_A and S_B, respectively.

In single bit quantization, the adaptive thresholds are calculated for each block separately. We alsofi nd the optimal *block_size* to divide channel estimates.

We can also quantize channel estimates to multiple bits. The process of multiple bits quantization is as follow:
• Alicefi nds the minimum and maximum of $\hat{H}_A$ to calculate the $Range_A$, $Range_A = \max(\hat{H}_A) - \min(\hat{H}_A)$. Bobfinds $Range_B$.
• Determine N, which is the number of bits for quantizing a channel estimate N, must satisfy $N < \lfloor log_2 Range_A \rfloor$ and $N < \lfloor log_2 range_B \rfloor$.
• Divide *Range* into $M = 2^N$ intervals, and choose N bits assignment for each of M intervals. To reduce the mismatch ratio, we use Gray code to encode them.
• For each channel estimate, Alice and Bob extract N bits, according to their location in M intervals. Finally, they obtain the bit stream S_A and S_B, respectively.

4.2.3 Reconciliation. Alice and Bob aim to generate identical bit stream to extract the same secret key K. Due to the mismatched bits in S_A and S_B, they need to reconciliate these mismatched bits.

Then, we consider using extended Binary Gray Code G_{24} to reconciliate S_A and S_B. G_{24} can encode 12 bits to 24 bits, correcting any 3 error bits and monitoring 7 error bits [5]. Due to the high correlations between S_A and S_B, S_A and S_B can be treated as two codewords, which are both distorted from a common bit sequence, according to the encoding theory.

Next, we illustrate the process of reconciliation. Atfi rst, both Alice and Bob encode their bit stream to Gray Code sequence. For example, Alice gets Gray Code sequence $W_A = E(S_A) = [S_A, F_A]$, where $E(\cdot)$ is an encoding function, and F_A is parity check sequence. Then, Alice sends the difference Z between S_A and F_A to Bob. Bob calculates the codeword $\bar{W}_B = [S_B, S_B - Z]$ and decodes it to $\bar{S}_A = D(\bar{W})$. Next, Bob calculates the number of mismatched bits between $\bar{S}_A$ and $\bar{S}_B$. If it is greater than 3, Bob will discard the sequence S_B and notify Alice to discard sequence S_A. Otherwise, Bob will replace S_B with $\bar{S}_A$, and generate a common sequence $K = S_A = \bar{S}_A$ with Alice.

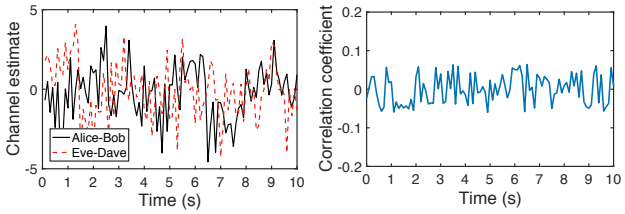
4.2.4 Privacy amplification. In the above reconciliation stage, Alice sends some information such as Z to Bob through public wireless channel. Thus, Eve can deduce some privacy information about secret sequence. Privacy amplification can mitigate this problem by reducing the length of secret sequence K. We can adopt some methods based on $left over hash lemma$ [10, 21], which is a universal hash function. Obviously, privacy amplification generates shorter bits with higher entropy. After going through the privacy amplification, Alice and Bob acquire thefi nal secret key.

5 SECURITY ANALYSIS

In this section, we analyze the security performance of FREE.

5.1 Against eavesdropping attack

In eavesdropping attack, Eve can overhear all communication transmitted through public acoustic channel and wireless channel.



(a) Channel estimates (b) The correlation coefficients of channel estimates

Figure 6: The channel estimates and their correlation coefficients of Alice-Bob and Eve-Dave at the same location.

Then we analyze what information can be listened by Eve. In channel estimation stage, Eve can listen and receive all of the acoustic signal transmitted by Alice and Bob. Since Eve receives the acoustic signal through different acoustic channel, Eve cannot estimate the same channel parameters as Alice or Bob. In quantization stage, he can know the exchanged index tables, which records the positions of removed bits. Since Eve does not know the bit stream S_A and S_B, he cannot know the content of removed parts either. In information reconciliation stage, Eve can listen the difference Z between S_A and S_B, and the discarding notice. But privacy amplification eliminates this leakage risk by removing the parts. Thus, Eve cannot deduce the secret key.

5.2 Against approaching attack

In approaching attack, Eve approaches to Alice or Bob to receive the acoustic signals transmitted from another side. Eve wants to use his proximity to Alice or Bob to estimate similar acoustic channel parameters. But the spatial decorrelation of acoustic channel presented in Section 3, illustrates that the correlation is lower than 0.2 when the distance is greater than one wavelength λ. For example, the center frequency is 20k Hz and the sound speed is 340 m/s, and the wavelength $\lambda = \frac{340m}{20k} = 1.7$ cm. In real life scenarios, we cannot allow others to put their mobile devices so close to our own mobiles. As a result, Eve cannot guess the similar secret key by approaching legitimate users.

5.3 Against repeating imitation attack

Under repeating imitation attack, after legitimate users left the site, Evefi nds a partner Dave to imitate the communication between Alice and Bob. Their goal is to estimate similar acoustic channel to generate same secret key. But the acoustic channel varies all the time due to the moving of mobile devices and the dynamic environment, as presented in Section 3. Therefore, Eve and Dave cannot capture the same acoustic channel randomness even in the same place.

This has been validated by our experiments. We studied the channel estimates of Eve and his partner, and compared them to the channel estimates between Alice and Bob, then computed the channel estimate correlation coefficients, please refer to Figure 6. We recorded the channel estimates of Alice-Bob in 10 seconds indoor. After both Alice and Bob left from the original positions, we recorded the channel estimates of Eve-Dave in the next 10 seconds. We aligned them in Figure 6 (a) for a better comparison. We can observe that channel estimates are different in different periods.

The outdoor experiments show the same results. We also computed the channel estimate correlation coefficients of Alice-Bob and Eve-Dave, as shown in Figure 6 (b). We can see that the correlation coefficients range from -0.1 to 0.1, showing the irrelevance between the channel estimates of Alice-Bob and Eve-Dave. Thus, Eve cannot get similar channel estimates as legitimate users.

6 EVALUATION

In this section, we evaluate the performance of FREE.

6.1 Methodology

We have conducted extensive experiments with four participants, named Alice, Bob, Eve, and Dave (Eve's partner). Each of them holds a mobile device (e.g., Nexus 7, MEIZU MX 6, Xiaomi 3), equipped with microphone and speaker. The illegitimate users, Eve and Dave, both are located more than 5 cm away from Alice and Bob, respectively.

6.1.1 Implementation and settings. We implement Android-based prototype of FREE on the mobile devices. We use Bluetooth to offer the public wireless channel between mobile devices. We call the Android API, AudioRecord(*) and AudioTrack(*), to transmit and receive acoustic signal, and the sampling rate is 44.1k Hz. We set up Alice's device stands facing Bob's device. At the beginning, Alice, as an initiator, sends Bob a synchronization signal. After receiving Bob's acknowledgement (ACK), Alice starts to transmit acoustic signal with band from 18k to 22k Hz, in every 16 ms. After receiving Alice's acoustic signal, Bob starts to transmit his own acoustic signal with the same band, in every 16 ms. At the same time, they shake the mobile devices together to generate more randomness. Then, Alice and Bob start to compute the channel estimates and quantize them into secret bit stream. They exchange their index table tofi nish the quantization by Bluetooth. Next, they exchange the reconciliation information to generate the same secret bit stream and perform the privacy amplification to extract a common secret key.

We conduct experiments in different scenarios, i.e., indoor, outdoor, mobile, and static, please refer to Table 1. Mobile means that users move around, the movement speed ranges from 0.5m/s to 1.5m/s; static means that the users keep still when extracting secret key.

6.1.2 Metrics. To evaluate the performance, we consider the following three metrics.

Bit Generation Rate (BGR): BGR is defined as the number of generated secret bits in a second. The more secret bits generated, the higher the efficiency of key extraction.

Bit Mismatch Ratio (BMR): BMR is defined as a ratio (the number of mismatch bits over the number of all generated bits in a second). The lower bit mismatch ratio, the higher the robustness of key extraction.

Randomness and Entropy (RE): RE is used to evaluate the quality of generated key. We use an extensively used randomness tool NIST test to measure the randomness of generated key. Besides, we compute the entropy of generated secret key. The higher entropy, the better quality of generated secret key.

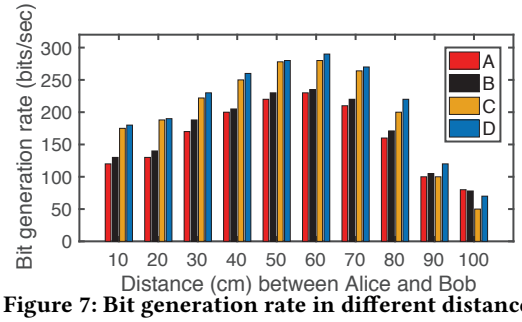


Figure 7: Bit generation rate in different distance.

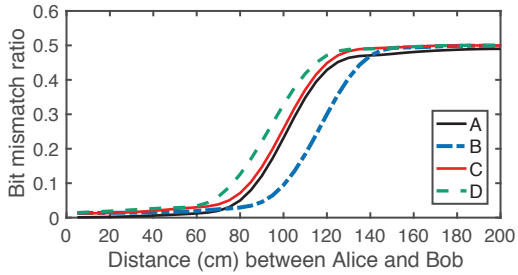


Figure 8: Bit mismatch ratio in different distance.

Table 1: Experiments scenarios.

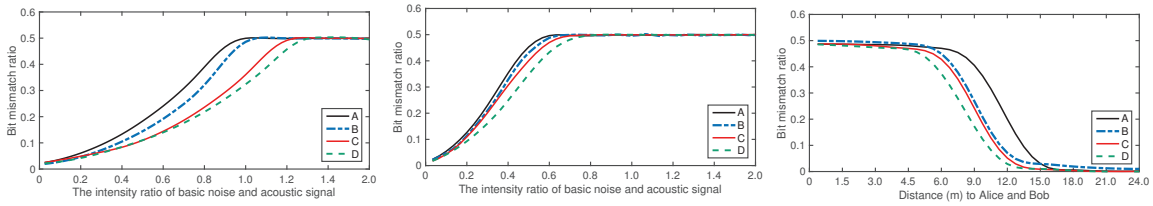
Index	State	Environment
A	Static	Indoor
B	Static	Outdoor
C	Mobile	Indoor
D	Mobile	Outdoor

Table 2: NIST statistical test results.

Test	A	B	C	D
Monobit Frequency	0.662	0.745	0.911	0.773
Longest Run of 1s	0.714	0.654	0.843	0.892
FFT	0.509	0.782	0.838	0.737
Approximate Entropy	0.801	0.783	0.903	0.887
Cumulative Sums (Fwd)	0.570	0.642	0.915	0.793
Cumulative Sums (Rev)	0.773	0.752	0.902	0.917
Block Frequency	0.717	0.736	0.825	0.914
Runs	0.753	0.796	0.821	0.833
Serial	0.505	0.674	0.818	0.839
	0.602	0.718	0.772	0.790

6.2 Randomness of extracted key

Wefi rst conduct experiments to validate the randomness of the key extracted by FREE. For experiment setting, the distance between Alice and Bob is 50 ~ 100 cm, and we use a single bit quantization method and gray coding. We utilize NIST test to measure the generated 300 sequences, and compute their p-values for 8 types of tests, listed in Table 2. The sequence is marked as random if all p-values are greater than 0.05. We can see that the generated secret keys pass all types of tests. Thus, the extracted keys have good quality in randomness.



(a) The bit mismatch ratio in different noise intensities.

(b) The bit mismatch ratio when Alice and Bob receive different basic noises.

(c) The bit mismatch ratio when the basic noise source in different distances.

Figure 9: The bit mismatch ratio in different basic noise conditions.

6.3 The influence of distance

To study how the distance influences the performance of FREE, we conduct the experiments with different distances. Figure 7 illustrates the secret bit generation rates of legitimate users under different distances and four kinds of scenarios. Wefi nd that the key generation rate is lower than 100 bits/sec when the distance is greater than 90 cm. The reason is that the correlation of Alice's and Bob's channel estimates decrease with distance growing. In real scenarios, the signals transmitted by Alice and Bob are not transmitted exactly along the same paths, due to the distance between the microphone and speaker embedded on the mobiles. To reduce the system error, uses let their microphones and speakers face to face, respectively. But the system error still increases with distance growing. We alsofi nd that the key generation rate is lower than 200 bits/sec when the distance is lower than 40 cm. The reason is that there is not enough randomness for key extraction when Alice and Bob are located too close. Wefi nd that the bit generation rate is greater than 260 bits/sec, when the distance is between 50 cm and 70 cm. The reason is that Alice and Bob can capture more channel randomness and the system error is small in this distance range. To generate a 512-bit cryptographic key, FREE only needs a couple of seconds. We alsofi nd that the key generation rate in outdoor environment is higher than that in indoor environment. The bit generation rate at mobile state is higher than that at static state. The reason is that the outdoor environment and mobile state offer more channel diversity and sufficient randomness.

Figure 8 illustrates the bit mismatch ratios of legitimate users in different distance and the same four kinds of scenarios. Wefind that the bit mismatch ratio increases obviously when the distance is greater than 60 cm. The main reason is also that the correlation between Alice's and Bob's channel estimate decreases with distance growing, due to the system error growing. The mismatch ratio is around 0.5 when the distance is greater than 120 cm. Thus, FREE cannot work well in this distance range. 120 cm can be set as the authenticate distance, and a longer distance (e.g., 2 m) can be set as the safe distance, referring to [32]. Thus, a device has maximum bit mismatch ratio 0.5 when it is out of the safe distance. The bit mismatch ratio under the scenario of outdoor and mobile increases earlier than that of other scenarios due to more complicated and changing acoustic channel.

6.4 The influence of basic noises

We also evaluate the performance of FREE in the environment with basic noises. To study the influence of basic noises, another mobile device is located 5 m away from both Alice and Bob, sending noises with band from 18k to 22k Hz to interfere the key extraction of Alice and Bob. Figure 9 (a) illustrates the bit mismatch ratio of Alice and Bob in different noise intensities and scenarios. The X axis denotes the intensity ratio of the basic noises and the acoustic signal transmitted between Alice and Bob. We can see that the bit mismatch ratio increases with the intensity ratio growing. The bit mismatch ratio is under 0.2 when the intensity ratio is smaller than 0.5. But the mismatch ratio approaches the maximum value of 0.5 when the intensity ratio is larger than 1. Therefore, we can turn the intensity of transmitted acoustic signal up as high as possible to reduce the bit mismatch ratio of FREE in the environment with obvious basic noises.

Figure 9 (b) illustrates the bit mismatch ratio of key extraction when Alice and Bob receive different basic noises due to their location difference. To realize different basic noises, another mobile device (the basic noise source) is located 100 cm away from Alice, and 50 cm away from Bob. As shown in Figure 9 (b), the X axis denotes the intensity ratio of the basic noises transmitted by the basic noise source and the acoustic signal transmitted between Alice and Bob. We can see the bit mismatch ratio approximates to 0.5 when the intensity ratio is larger than 0.6. The bit mismatch ratio grows rapidly when the two devices in different basic noises. But we can also turn up the intensity of transmitted acoustic signal to reduce the bit mismatch ratio.

Figure 9 (c) illustrates the bit mismatch ratio when the basic noise source in different distances to Alice and Bob. The basic noise source transmits the acoustic signal with the same intensities of Alice and Bob. As shown in Figure 9 (c), the X axis denotes the basic noise source's distance to Alice and Bob. We can see the bit mismatch ratio decreases with the distance growing. When the distance is larger than 15 m, the bit mismatch is smaller than 0.05.

6.5 The influence of quantization and encoding methods

We also evaluate the performance of FREE in different quantization and encoding methods. In this experiment, the quantization mainly includes 1-bit quantization method and 2-bit quantization method; the encoding methods mainly include binary encoding method and gray encoding method. The distance range is 50 ~ 70 cm.

Figure 10 shows the comparisons of key generation rate of FREE in different quantization, encoding methods, and scenarios. We can see that the key generation rate of 2-bit quantization is higher than that of 1-bit quantization. This is because 2-bit quantization can generate more bits to extract more secret bits. The key generation rate of gray encoding is higher than that of binary encoding. This is because gray encoding can generate more similar bit to reduce the mismatch ratio.

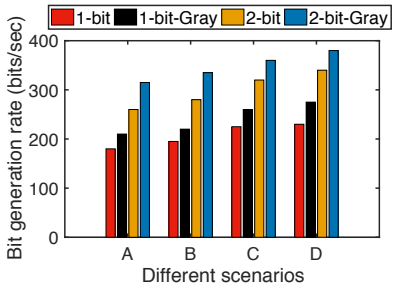


Figure 10: Bit generation rate in different quantization and encoding methods.

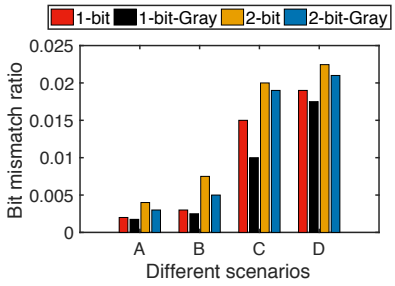


Figure 11: Bit mismatch ratio in different quantization and encoding methods.

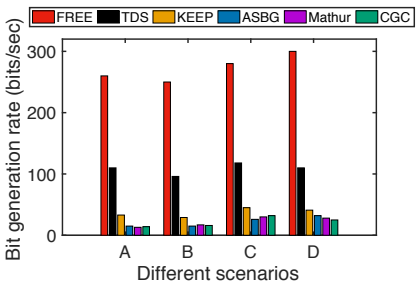


Figure 12: Comparison of bit generation rate.

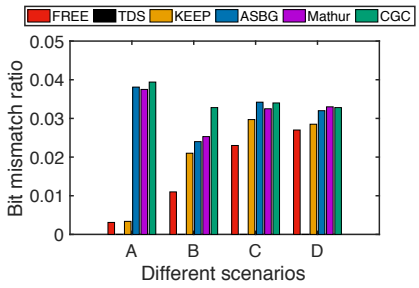


Figure 13: Comparison of bit mismatch ratio.

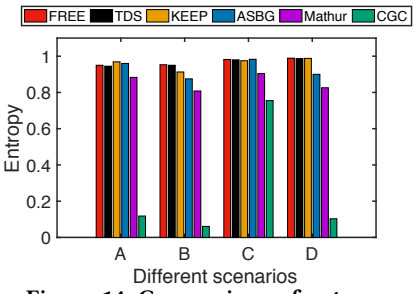


Figure 14: Comparison of entropy.

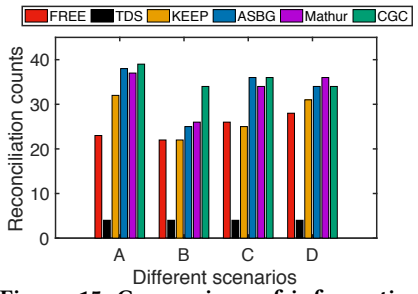


Figure 15: Comparison of information reconciliation counts.

Figure 11 shows the comparisons of bit mismatch ratio of FREE in different quantization, encoding methods, and scenarios. We can see that bit mismatch ratio of 2-bit quantization is higher than that of 1-bit quantization. It is because that 2-bit quantization causes more quantization errors. The bit mismatch ratio of gray encoding is lower than that of binary encoding. It is also because that gray encoding generates more similar bits and reduces the number of inconsistent bits.

6.6 Comparisons of existing key extraction approaches

We compare FREE with existing key generation approaches, i.e., Mathur et al. [17], ASBG [10], CGC [14], KEEP [31], TDS [32]. First, we need to align the baseline of comparisons. In approach proposed by Mathur et al., there are two parameters, α and m. We set α and m as 0.35 and 2, respectively, to guarantee more secret bits. For ASBG, CGC, and KEEP, We set α and fragment size as 0.35 and 50, respectively, to ensure a lower bit mismatch ratio. For TDS, we set block size β as 6 in static sate and as 4 in mobile state. The distance between Alice and Bob is within 4 cm. For FREE, we set *block_size* as 30. The distance between Alice and Bob is within 80 cm.

We report the bit generation rates of different approaches in Figure 12. We can see FREE has obvious higher bit generation rate than other approaches.

We report the bit mismatch ratio of different approaches in Figure 13. FREE has around 0.5% to 3.0% bit mismatch ratio, which is lower than other approaches except for TDS.

We report the entropy in Figure 14. The entropy can represent the randomness of extracted key from the perspective of uncertainty. FREE, TDS, and KEEP have higher entropy, CGC has the lowest entropy.

We report the information reconciliation counts in Figure 15. FREE needs exchange reconciliation information for 22~28 times in a second. This is because FREE needs one time reconciliation for every 12 bits in bit sequence.

In our experiments, we count the mean values of FREE's key generation rate and bit mismatch ratio in different scenarios. The statistical results show that, compared with the state of art methods, FREE improves the key generation rate by 38.46% and reduces the bit mismatch ratio by 42.34%.

In summary, FREE has significant bit generation rate and great performance on the entropy, bit mismatch ratio and information reconciliation counts.

7 RELATED WORK

Secret key extraction has been studied for many years. In wireless network, the security of data transmission is guaranteed by the security protocols of upper layers. Physical layer security also needs to be guaranteed by encryption schemes. To achieve information theoretic-security at physical layer, many existing works exploit the unpredictable and random characteristics to establish cryptographic key [4, 16, 35]. Ahlswede et al. and Maurer, et al. discussed the key generation theoretically in [1] and [18]. Hershey et al.fi rstly proposed the idea of using channel measurements to extract secret key [9]. Then, plenty of works exploit wireless channel measurements to extract secret key. The channel measurements includes arrival of angle [2], phase[25], and received signal strength (RSS) [10, 17]. Then, channel state information (CSI) has been extensively exploited for key extraction with higher key generation

rate. But CSI-based key extraction schemes need the assist from special devices [12, 14, 15, 31, 32].

Some works use the environment sensing to extract secret key. Bichler et al. and Mayrhofer et al. exploited acceleration data of shaking process for key generation and secure device pairing [3, 19]. MAGIK uses the dynamic geomagneticfi eld sensing to extract secret key [23]. But they cannot resist against the imitation attack.

In addition, some works use audio signals to authenticate legitimate users [11, 13, 26–28, 33]. Schürmann, et al used the similar ambient audio pattern to authenticate legitimate users and secure communication [26], but their method underperform the key generation efficiency and cannot resist against imitation attack. Spartacus uses audio to establish spontaneous interactions between mobile devices, but cannot secure their transmission privacy [28]. SoundProof uses ambient audio to validate the proximity to authenticate users, but cannot generate secret key [11]. GeneWave uses acoustic signal for authentication and key agreement [33]. However, it relies on public key system to exchange the secret key. Some works use ambient audio to secure pairing [27], but they must use Diffie-Hellman protocol to generate secret key.

8 CONCLUSION

In this paper, we have studied how to achieve high key generation efficiency with commodity mobile devices. We have proposed FREE, which is a fast and robust key extraction mechanism that uses the randomness of inaudible acoustic channel to establish a secure wireless channel between two mobile devices. We have carefully studied and validated the feasibility of utilizing acoustic channel randomness for key extraction through theoretical analysis and extensive experiments. We also have implemented FREE on mobile devices, e.g., Nexus 7, MEIZU MX 6, Xiaomi 3. The results of experiments show the high efficiency and satisfactory robustness of FREE. Compared with existing solutions for key establishment, FREE has several advantages: First, FREE has significantly higher key generation rate, for a 512-bit cryptographic key, FREE only needs two seconds to generate it; Second, FREE only requires off-the-shelf mobile devices like a smartphone; Third, FREE can resist against certain attacks; Last, FREE works without disturbing nearby people.

REFERENCES

[1] Rudolf Ahlswede and Imre Csiszár. 1993. Common randomness in information theory and cryptography. I. Secret sharing. *IEEE Transactions on Information Theory* 39, 4 (1993), 1121–1132.

[2] Tomoyuki Aono, Keisuke Higuchi, Makoto Taromaru, Takashi Ohira, and Hideichi Sasaoka. 2005. Wireless secret key generation exploiting the reactance-domain scalar response of multipath fading channels: RSSI interleaving scheme. In *Proceedings of IEEE EuMA*.

[3] Daniel Bichler, Guido Stromberg, Mario Huemer, and Manuel Löw. 2007. Key generation based on acceleration data of shaking processes. In *Proceedings of Ubicomp*.

[4] Matthieu Bloch and Joao Barros. 2011. *Physical-layer security: from information theory to security engineering*. Cambridge University Press.

[5] A Robert Calderbank, G David Forney, and Alexander Vardy. 1999. Minimal tail-biting trellises: The Golay code and more. *IEEE Transactions on Information Theory* 45, 5 (1999), 1435–1455.

[6] Whitfield Diffie and Martin Hellman. 1976. New directions in cryptography. *IEEE transactions on Information Theory* 22, 6 (1976), 644–654.

[7] Andrea Goldsmith. 2005. *Wireless communications*. Cambridge university press.

[8] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. 2011. Tool release: Gathering 802.11 n traces with channel state information. *ACM SIGCOMM Computer Communication Review* 41, 1 (2011), 53–53.

[9] John E Hershey, Amer A Hassan, and Rao Yarlagadda. 1995. Unconventional cryptographic keying variable management. *IEEE Transactions on Communications* 43, 1 (1995), 3–6.

[10] Suman Jana, Sriram Nandha Premnath, Mike Clark, Sneha K Kasera, Neal Patwari, and Srikanth V Krishnamurthy. 2009. On the effectiveness of secret key extraction from wireless signal strength in real environments. In *Proceedings of ACM MobiCom*.

[11] Nikolaos Karapanos, Claudio Marforio, Claudio Soriente, and Srdjan Capkun. 2015. Sound-Proof: Usable Two-Factor Authentication Based on Ambient Sound.. In *Proceedings of USENIX Security*.

[12] Havish Koorapaty, Amer A Hassan, and Sandeep Chennakeshu. 2000. Secure information transmission for mobile radio. *IEEE Communications Letters* 4, 2 (2000), 52–55.

[13] Xiaohui Liang, Tianlong Yun, Ronald Peterson, and David Kotz. 2017. LightTouch: Securely connecting wearables to ambient displays with user intent. In *Proceedings of IEEE INFOCOM*.

[14] Hongbo Liu, Yang Wang, Jie Yang, and Yingying Chen. 2013. Fast and practical secret key extraction by exploiting channel response. In *Proceedings of IEEE INFOCOM*.

[15] Yanpei Liu, Stark C Draper, and Akbar M Sayeed. 2012. Exploiting channel diversity in secret key generation from multipath fading randomness. *IEEE Transactions on Information Forensics and Security* 7, 5 (2012), 1484–1497.

[16] Suhas Mathur, Alex Reznik, Chunxuan Ye, Rajat Mukherjee, Akbar Rahman, Yogendra Shah, Wade Trappe, and Narayan Mandayam. 2010. Exploiting the physical layer for enhanced security [security and privacy in emerging wireless networks]. *IEEE Wireless Communications* 17, 5 (2010).

[17] Suhas Mathur, Wade Trappe, Narayan Mandayam, Chunxuan Ye, and Alex Reznik. 2008. Radio-telepathy: extracting a secret key from an unauthenticated wireless channel. In *Proceedings of ACM MobiCom*.

[18] Ueli M Maurer. 1993. Secret key agreement by public discussion from common information. *IEEE transactions on information theory* 39, 3 (1993), 733–742.

[19] Rene Mayrhofer and Hans Gellersen. 2009. Shake well before use: Intuitive and secure pairing of mobile devices. *IEEE Transactions on Mobile Computing* 8, 6 (2009), 792–806.

[20] Alan V Oppenheim. 1999. *Discrete-time signal processing*. Pearson Education India.

[21] Sriram Nandha Premnath, Suman Jana, Jessica Croft, Prarthana Lakshmane Gowda, Mike Clark, Sneha Kumar Kasera, Neal Patwari, and Srikanth V Krishnamurthy. 2013. Secret key extraction from wireless signal strength in real environments. *IEEE Transactions on Mobile Computing* 12, 5 (2013), 917–930.

[22] Markku Pukkila. 2000. Channel estimation modeling. *Nokia Research Center* (2000).

[23] Fudong Qiu, Zhengxian He, Linghe Kong, and Fan Wu. 2017. MAGIK: An efficient key extraction mechanism based on dynamic geomagneticfi eld. In *Proceedings of IEEE INFOCOM*.

[24] Theodore S Rappaport et al. 1996. *Wireless communications: principles and practice*. Vol. 2. prentice hall PTR New Jersey.

[25] Akbar Sayeed and Adrian Perrig. 2008. Secure wireless communications: Secret keys through multipath. In *Proceedings of IEEE ICASSP*.

[26] Dominik Schürmann and Stephan Sigg. 2013. Secure communication based on ambient audio. *IEEE Transactions on Mobile Computing* 12, 2 (2013), 358–370.

[27] Stephan Sigg, Yusheng Ji, Ngu Nguyen, and An Huynh. 2012. AdhocPairing: Spontaneous audio based secure device pairing for Android mobile devices. In *Proceedings of IWSSI/SPMU*.

[28] Zheng Sun, Aveek Purohit, Raja Bose, and Pei Zhang. 2013. Spartacus: spatially-aware interaction for mobile devices through energy-efficient audio sensing. In *Proceedings of ACM MobiSys*.

[29] David Tse and Pramod Viswanath. 2004. Fundamentals of wireless communication. In *Cambridge University Press*.

[30] Qian Wang, Hai Su, Kui Ren, and Kwangjo Kim. 2011. Fast and scalable secret key generation exploiting channel phase randomness in wireless networks. In *Proceedings of IEEE INFOCOM*.

[31] Wei Xi, Xiang-Yang Li, Chen Qian, Jinsong Han, Shaojie Tang, Jizhong Zhao, and Kun Zhao. 2014. KEEP: Fast secret key extraction protocol for D2D communication. In *Proceedings of IEEE IWQoS*.

[32] Wei Xi, Chen Qian, Jinsong Han, Kun Zhao, Sheng Zhong, Xiang-Yang Li, and Jizhong Zhao. 2016. Instant and robust authentication and key agreement among mobile devices. In *Proceedings of ACM CCS*.

[33] Pengjin Xie, Jingchao Feng, Zhichao Cao, and Jiliang Wang. 2017. GeneWave: Fast authentication and key agreement on commodity mobile devices. In *Proceedings of IEEE ICNP*.

[34] Sangki Yun, Yi-Chao Chen, Huihuang Zheng, Lili Qiu, and Wenguang Mao. 2017. Strata: Fine-Grained Acoustic-based Device-Free Tracking. In *Proceedings of ACM MobiSys*.

[35] Junqing Zhang, Trung Q Duong, Alan Marshall, and Roger Woods. 2016. Key generation from wireless channels: A review. *IEEE Access* 4 (2016), 614–626.