PriWhisper+: An Enhanced Acoustic Short-Range Communication System for Smartphones

Xiao Zhang, *Graduate Student Member, IEEE*, Jiqiang Liu, *Member, IEEE*, Si Chen, *Member, IEEE*, Yongjun Kong, *Student Member, IEEE*, and Kui Ren, *Fellow, IEEE*

***Abstract*—The recent proliferation of Internet of Things (IoT) device coupled with the demand for an inexpensive way of transmitting data has been the primary factors behind the popularity of the short-range acoustic communication. It enables a seamless transfer of digital information via soundwaves, using device's loudspeaker and microphone only and the whole interaction takes place without the need for network connection. Due to the limited computational power of the IoT device, the short-range acoustic communication system has adopted the friendly jamming technology to save energy. However, the security of the friendly jamming technology in mobile applications has not been thoroughly studied. When propagating sound in public, the soundwaves are subject to eavesdropping by nature. In particular, the friendly jamming technology is vulnerable to the separation attacks which can separate data signals from mixed signals. In this paper, we propose PriWhisper+—a secure acoustic short-range communication system between IoT devices. We analyze the security of PriWhisper+ via information theory and propose physical security enhancement mechanisms for acoustic communication by combining device mobility with secret sharing scheme. We then design a secure data communication scheme that transmits data in acoustic signals. This scheme can be applied in many security-sensitive situations, such as device pairing, contactless payments, and privacy data sharing. At last, we evaluate the performance of our proposed PriWhisper+ by extensive experiments on Android smartphones. The results of the experiment show that the PriWhisper+ can protect the data confidentiality within thirty centimeters of communication range.**

***Index Terms*—Acoustic short-range communication, blind signal segmentation (BSS), friendly jamming, independent component analysis (ICA), Internet of Things (IoT), security and privacy, smartphone wireless communication.**

I. Introduction

RECENTLY, the growing demand for short-range communications for Internet of Things (IoT) devices. The short-range communication not only makes the revolutionary progress in devices interaction approach but also affects many aspects of our daily life. For instance, both near field communication (NFC) and QR code are popular short-range communication technologies. NFC can build an ad hoc communication between two embedded NFC chip devices by putting them together. Apple Inc provides the NFC-based Apple Pay system that renders you purchases using an iOS device [1]. People can just scan QR code on new friend's smartphone instead of typing his/her ID to add a new friend on Wechat. Moreover, acoustic communication becomes an emerging short-range communication method. An audio data transmission scheme called the Chirp can modulate the data to the audio signal, such as enabling Hijinx toys to interact with content on any IoT device [2]. Unlike QR code using the directional channel called line-of-sight (LOS) channel, the sound wave is practically omnidirectional propagating in the air medium, which means it can ignore the obstacle, illumination, and device stability. Meanwhile, only the acoustic communication system can achieve full duplex short-range communication on smartphone platform because the speakerphone and microphone can work simultaneously. Comparing with NFC system which needs additional hardware chips. All smartphones and massive IoT devices configure the microphone and speakerphone by default.

Due to the short communication distance, NFC was initially considered secure. However, some studies prove that attacker can eavesdrop the NFC communication from a few meters away [3], [6]. On the other hand, the visual nature of barcode-based short-range communication makes them extremely vulnerable to shoulder sniffing [15]. Based on the above security issues of the popular short-range communication method, some research proposes acoustic-based secure short-range communication scheme by adopting the friendly jamming technology from radio communication. The theory of the friendly jamming technology is that the receiver propagates the artificial noise while the sender sends the data. The attacker cannot extract the data from the recorded noisy signal, but the receiver has the artificial noise and thus can remove the noise from the mixed signal. To the best of our knowledge, NFC and barcode system can either receive or transmit a signal at a certain moment. Hence, the acoustic system can exploit friendly jamming technology to secure the data. But Dhwani's [14] design cannot defend some passive attack, such as blind source separation (BSS) attack. Priwhisper can defend such attack by limiting the communication distance less than 0.5 cm, which challenges utilization in practice.

Manuscript received March 24, 2018; revised May 19, 2018; accepted June 13, 2018. Date of publication June 25, 2018; date of current version February 25, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant 61672092 and in part by the Fundamental Research Funds for the Central Universities of China under Grant 2018JBZ103. *(Corresponding author: Jiqiang Liu.)*

X. Zhang and J. Liu are with the Beijing Key Laboratory of Security and Privacy in Intelligent Transportation, Beijing Jiaotong University, Beijing 100044, China (e-mail: zhangx@bjtu.edu.cn; jqliu@bjtu.edu.cn).

S. Chen is with the Computer Science Department, West Chester University of Pennsylvania, West Chester, PA 19382 USA.

Y. Kong is with the College of Password Engineering, Engineering University of CAPF, Xi'an 716000, China.

K. Ren is with the Computer Science and Engineering Department, University at Buffalo, State University of New York, Buffalo, NY 14260 USA.

Digital Object Identifier 10.1109/JIOT.2018.2850524

To address these above challenges, we design an enhanced secure short-range communication system, named PriWhisper+, to extend the secure communication distance by using portable device's mobility for smartphones. Different from the existing work, PriWhisper+ has more flexibility on the transmission range which is from a few centimeters to dozens of centimeters while the data confidentiality can be guaranteed. There are several inherent advantages of using acoustic short-range communication. First of all, the transmission of the acoustic signal can achieve omnidirectional communication while the barcode signal needs the LOS channel, which gives PriWhisper+ more flexibility on device physical location. This feature ensures the data transmission when devices are moving. Second, the acoustic signal decays quickly in the air medium. This terrible feature can prevent the transmission from eavesdropping attacks and thus provides the data confidentiality in the short-range communications scenario. Finally, the human ear can hear the sound when the acoustic signal carrier frequency lies within the audible bandwidth. The user can easily detect the location of the adversary who executes an active attack in practice, such as man-in-the-middle attack, DoS attack, and replaces attack.

We implemented a prototype called PriWhisper+ on the Android platform for the smartphone that has speakers and microphones to secure the transmitted data. We use the mobility of the device to randomize the channel condition to prevent the separation attack. We address several challenges to enable the security purposes of PriWhisper+. For example, are the impacts of device movement on the channel sufficiently resistant to separation attacks? In this paper, we analyze this problem and found the threshold between moving speed and secure distance. The experiment result shows the randomize channel can prevent the separation attack in most cases. For higher security, we utilize the secret sharing scheme in our system to eliminate possible data leaks. Finally, we propose the PriWhisper+ scheme with friendly jamming technology, devices mobility and secret sharing scheme. Theoretical analysis and experimental results show a positive conclusion that PriWhisper+ is a secure, easy-to-deploy system without additional hardware.

Our contribution can be summarized as the following three aspects.

1) We design and implement PriWhisper+, an enhanced acoustic short range communication system for IoT portable device that exploits devices mobility, friendly jamming technology, and secret sharing scheme for data confidentiality.
2) Unlike most prior works for achieving the secure communication distance less than 1 cm, our scheme supports longer secure communication distance from few centimeters to dozens of centimeters (30 cm). To best of our knowledge, it is the first work on extending secure communication distance via the acoustic short range communication system for smartphones.
3) We prove the security of our system through theoretical analysis and experimental verification. In particular, we show that the attacker cannot extract the data signal from the movement mixed signal even if the adversary uses the state-of-art blind signal separation technology.

The rest of this paper is organized as follows. Section II provides some related work. In Section III, we introduce our system model, threat model, and design goal. Section IV presents the PriWhisper+ system details. We give the thorough security analysis about our system in Section V. Section VI estimates the system performance. Finally, Section VII summarizes the whole paper and gives the future work.

II. Related Work

NFC is a short range data communication technology mainly used in mobile devices, and it is derived from the radio frequency identification technology. The key advantages of the NFC are low energy consumption, wide application range, and high security. But NFC requires additional chip support which is not cost-effective. It is considered safe because the communication distance is only a few centimeters. Some works [3]–[5] eavesdropped the communication beyond the NFC working range. Moreover, Diakos *et al.* [6] used antenna to eavesdrop the NFC at the distance of 1 m. Although NFC [7] introduced Diffie–Hellman key exchange protocol to secure their communication in new standard protocols, most short-range applications just require only a few rounds of message exchange. Therefore, the process of the key exchange may spend most of the time in the entire communication session.

Barcode uses a particular geometric pattern, such as black and white square, to record information. For many years, the barcode has been developed from the 1-D code to the 2-D code. The common 2-D code has QR code, matrix code. Wang *et al.* [8] designed a novel and improved color barcode-based visual communication system which features a carefully designed high-capacity barcode layout design to allow flexible frame synchronization and accurate code extraction. Zhang *et al.* [9] proposed a secure system for barcode-based visible light communication between smartphones. They utilize the 2-D/3-D screen geometric model and the mobility of the smartphones to prevent the eavesdropping attacks. On the other hand, several recent studies seek to achieve unobtrusive screen-to-camera communication, such as InFrame++ [10]. InFrame++ enables concurrent, dual-mode, full-frame communication for both users and devices. In contrast, all above barcode-based work may suffer the eavesdropping attacks because of LOS channel features. Moreover, A 2-D barcode only contains an insufficient amount of information and hence cannot adopt advanced encryption primitives.

Acoustic short-range communication is becoming a hot topic in recent years. Lopes and Aguiar [32] provided audible sound as a means for aerial acoustic wireless communications. They used musical sounds replace the modem sounds in 2001. In their work, an audible sound is used as a means for wireless device communications. Mostafa [33] implements a software called minimodem that convenience researcher is handling traditional modem protocols. Mao *et al.* [11] presented a high-precision acoustic tracker, which aims to replace a traditional mouse and let a user play games by moving a smartphone in the air. Wang *et al.* [12] develop a novel form of real-time

acoustics-based dual-channel communication, which uses a speaker and the microphones on off-the-shelf smartphones to achieve concurrent audible and hidden communication. Roy *et al.* demonstrated that such signals remain inaudible to humans but are record-able by unmodified off-the-shelf microphones.

To guarantee the confidentiality of the transmitted data, Nandakumar *et al.* [14] proposed an acoustic-based NFC system called *Dhwani*. The key idea of Dhwani is using the acoustic jamming technology, which uses self-jamming at the receiver to secure the communication channel between the devices. Dhwani designs an acoustic jamming technology that the receiver instead of the sender propagates the jamming signal called JamSecure. The JamSecure is a self-jamming technology used by the receiver to cloak the message being transmitted by the sender, hence preventing eavesdropping attacks. Friendly jamming technology can prevent eavesdropping attacks but cannot defend the separation attack which can retrieve the data signals from the mixed signals, which include data signals and jamming signals, by using BSS technology. Zhang *et al.* [15] presented a keyless secure acoustic communication system called Priwhisper, they also use the friendly jamming technology to prevent the eavesdropping attack. Furthermore, they considered the separation attack then using the channel similarity to defend it. They studied the separability of the mixed signal against the BSS attacks. Their result shows the secure communication range is only 0.5 cm which is hard to apply in practice.

Unlike Bluetooth or WiFi, acoustic short-range communication is designed for small ad-hoc data transfers. The key advantage of acoustic short-range communication is that it gets rid of the requirement for complex network configuration efforts necessary to build a secure communication channel. The user needs to set up the connection in few seconds, which requires the communication delay as small as possible. Although Bluetooth has a higher data rate standard, it needs pairing before use, which means it costs more time to set up a secure channel. In addition, The battery consumption of acoustic short-range communication generally is negligible because of short communication time. However, Bluetooth will cost more energy because it needs to maintain a period of connection.

III. Problem Formulation

A. System Model

PriWhisper+ is designed to enhance secure acoustic short-range communication that ensures data confidentially under the premise of extending the communication range from few centimeters to dozens of centimeters. Due to the mobility of smartphone and handful devices, PriWhisper+ tries to use this feature to randomize the channel condition to defend the separation attack. Fig. 1 illustrates the PriWhisper+'s working scenario. When an ad-hoc short communication begins, the sender starts with propagating a preamble signal, and the receiver starts to continuously send the jamming signals and move with random path simultaneously when he receives the preamble signal. The sender transmits the data signal and moves at the same time when the jamming signals arrive at the sender. The sender and receiver remain the random movement during communication.

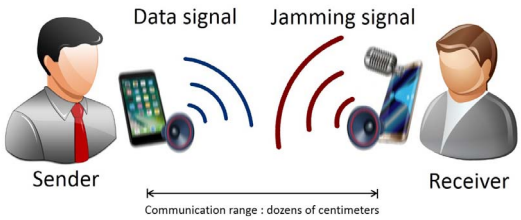


Fig. 1. System application scenario.

B. Threat Model

The security goal of PriWhisper+ is to guarantee the confidentiality and integrity of information exchanged between two mobile devices located within flexible short range distance (a few centimeters to dozens of centimeters). In this section, we make the following assumptions about the security model.

1) *Trusted Devices:* The transmitter and the receiver are trusted entities. We assume that these devices run our scheme honestly and run steadily. The attacker cannot hack the devices directly. Any failure is considered accidental, such as system crash.
2) *Adversary Capability:* The attacker has single or multiple devices to execute an active attack (e.g., the man-in-the-middle attack) or a passive attack (e.g., eavesdropping).
3) *Attack Distance:* The attacker can deploy sensors on arbitrary location. The distance of the communication devices is flexible from a few centimeters to dozens of centimeters.

The above assumptions are consistent with the general short-range communication model. Below we list the acoustic attacks that we consider.

1) Placement Attack: The placement attack is a location-related eavesdropping passive attack, which seeks the location, where jamming signals cannot completely cover the data signals. Hence, attackers can eavesdrop the secret messages. Intuitively, the intensity of the data signal has maximum strength at sender's location. And from above assumptions, the adversary could be anywhere, so if the adversary cannot obtain messages at sender's location, we can believe that the placement attack is invalid to the system. We will discuss the detail in Section V.

2) Separation Attack: The separation attack is an advanced eavesdropping attack. Assuming that Alice tries to have a short-range communication with Bob, and Eve is curious the communication content. Alice and Bob use the acoustic short-range communication scheme with friendly jamming technology to secure the channel. Eve has two microphones that can record the mixed signals which merge jamming signals and data signals, and then she can retrieve the data signals by using some BSS algorithm. The BSS problem has become a popular topic in the last few years [17]. Imagine the following scenario; many guests are chatting with each other at a noisy cocktail party, you accidentally clearly hear that Alice and Bob

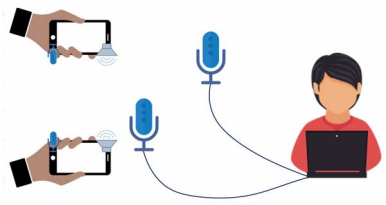


Fig. 2. BSS attack scenario.

mention your name in their conversation. The problem of finding a single sound of content from all of the sound sources is the cocktail problem. Generally, the BSS problem is to separate one or several original audio signals from the recorded mixed audio signals by one or several microphones [18], [19]. From Fig. 2, the adversary executes the BSS attacks by using two microphones x_1 and x_2. Let s_1 be the data signal from transmitter and s_2 be the jamming signal from receiver. Let x_1 and x_2 be the received mixed signal which is linear mixture from s_1 and s_2, respectively. Assume that the channel between x_1 and s_1 are h_{11}, similarly, we have h_{12}, h_{21}, and h_{22}. Let the vector $\mathbf{x} = [x_1, x_2]^T$ denote the attacker's recorded mixed signals, vector $\mathbf{s} = [s_1, s_2]^T$ denotes the data signal from the sender and the jamming signal from receiver, which can be expressed as

$$\mathbf{x} = \mathbf{H} \cdot \mathbf{s} + \mathbf{e} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \cdot \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + e \tag{1}$$

where $\mathbf{H}$ denotes the channel mixing matrix and e is the ambient noise vector. The adversary attempts to recover date signal from the mixed signal $\mathbf{x}$ by using BSS technology, such as independent component analysis (ICA). The ICA algorithm can find the original sources by assuming that the linear mixed signals are statistically independent.

3) Man-in-the-Middle Attack: The man-in-the-middle attack is one of the active attacks, where adversary needs to propagate the jamming signal or data signal to achieve its purpose. Assuming that Eve pretends to be Bob that propagates the jamming signals to Alice, and then Alice transmits the data signals to Eve. He can retrieve the data signals because he knows the jamming signals. After that, Eve send the preamble signals to Bob make him believe Eve is Alice, then Bob will receive the tampered data.

4) DoS Attack: We also give an example to explain this attack. Eve propagates another jamming signals while Alice and Bob communicate with each other. If Eve's jamming signals are strong enough, it will cover Alice's data signals and Bob's jamming signals. Hence, Bob cannot retrieve the data signals from the mixed signals anymore.

C. Design Goal

The goal of PriWhisper+ is to enable secure easy-to-deploy and configuration-free acoustic short-range communication within the dozen of centimeters of working distance. Our system can deploy on any existing mobile devices with speaker and microphone. The designed working distance is from a few centimeters to dozens of centimeters, which is more flexible than the PriWhisper. As depicted in Fig. 1, the transmitter and receiver play acoustic signals to each other within secure communication distance.The transmitter sends data signals while the receiver sends jamming signals.

IV. Proposed System

In this section, we present our enhanced acoustic short-range communication system. We first introduce the architecture of our system and then describe our system by the module. At last, we integrate all modules and implement our prototype.

A. System Overview

One key goal of PriWhisper+ is to extend the secure communication range while being easy-to-deploy in public, such as cafes and malls, where the noise level is large enough to interfere with the acoustic signals transmitting. The architecture of PriWhisper+ is depicted in Fig. 3. For easy-deployment concern, our scheme adopts software-based modulator and demodulator module. The raw data is first divided into n pieces and then channel-encoded. The Shamir's secret sharing scheme is utilized as our data segmentation algorithm. Before the data is being transmitted, the modulator modulates the data signals to the acoustic signals with M-FSK. Then sender's speaker transmits the data signals while detecting the jamming signals by receiver's speaker. Then receiver's microphone collects the mixed signals through the air medium. After retrieving the acoustic data signals from the receive signals, the demodulator demodulates the acoustic data signals and then channel-decodes it. Finally, the receiver unpacks the packages and integrates the segmented data to get the raw data.

B. Secret Sharing Module

Secret sharing is one of the key security measures of our scheme. The advantage of the secret sharing scheme is that it does not require a secure channel for key exchange. Furthermore, the position of error data transmitted in the air medium by the sound signal is relatively concentrated rather than randomly distributed, which is more suitable for the secret sharing scheme. Assume the data that the transmitter tries to send is D. If D is long, we should break it into shorter blocks to avoid multiprecision arithmetic operations [31]. The blocks cannot be arbitrarily short since the smallest usable value of p is $n + 1$, where p is a big prime number which the secret scheme randomly selected. The sender divides the D into m blocks according to the length of the D. For the experiment convenience concern, we choose 8 bytes as the block size in our prototype. Therefore, $m = (D + 7)/8$. If the length of D is indivisible by 8, we append zeros to the end of D. After segmenting data, let each $\sum_{j=1}^{m} D_j$ run secret sharing scheme (k, n) then we get $m \times n$ pairs of $(x_{ji}, f(x_{ji}))$ and denote as D_{ji}.

As shown in Fig. 4, we divide the data package into two parts: 1) the header and 2) the data. The header contains a preamble bit and a count number of the segment data. The preamble bit marks a start of a data package, and the count number is telling the receiver this package belongs to the jth data segment. The data contains a pair of numbers which is $(x_i, f(x_i))$ in jth data segment.

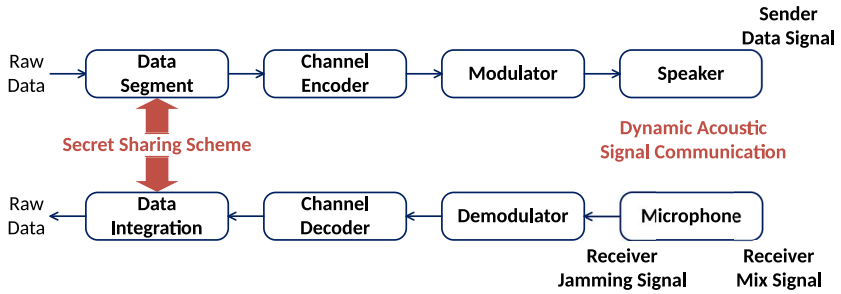


Fig. 3. System architecture of PriWhisper+.

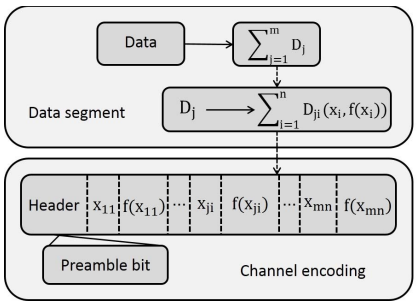


Fig. 4. Package structure of the secret sharing module.

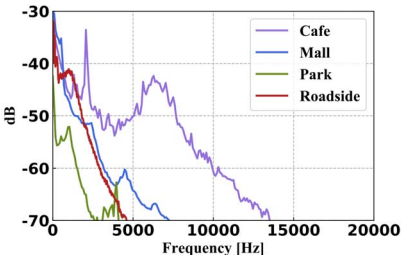


Fig. 5. Ambient noise level in different locations.

The goal of this module is to divide D_j into n pieces $D_{j1}, \ldots, D_{jn}$ with having any k or more D_{ji} pieces can easily recover D_j, whereas any $k-1$ pieces or fewer cannot rebuild the D_j with all its possible values are equally likely [31]. We choose the Shamir's threshold secret sharing scheme to secure the data in our prototype. The transmitter has the data D_j and determines k arbitrary figure $a_0, \ldots, a_{k-1}$. Let $a_0 = D_{ji}$ then construct polynomial as follows:

$$a(\mathbf{x}) = a_0 + a_1\mathbf{x} + a_2\mathbf{x}^2 + \cdots + a_{k-1}\mathbf{x}^{k-1}$$

note that all calculations are done in the finite field $\mathbb{F}$. Assuming $f(\mathbf{x}) = a(\mathbf{x}) \bmod p$ where p is a random big prime number. The sender substitutes arbitrary k numbers $x_1, \ldots, x_k$ into polynomial then we can get $f(\mathbf{x}_1), \ldots, f(x_k)$. The transmitter then has all subdata $(x_1, f(\mathbf{x}_1)), \ldots, (x_k, f(x_k))$ which are ready to be sent.

The receiver utilizes the Lagrange's interpolation formula to reconstruct the data as follows. First of all, the receiver picks k pairs of subdata $((x_{i1}, y_{i1}), \ldots, (x_{ik}, y_{ik}))$ which can and only can determine a polynomial which the orders are $k-1$. According to the Lagrange's interpolation formula

$$a(x) = \sum_{l=1}^{k} y_{il} \prod_{1 \le m \le tm \ne l}^{k} \frac{(x - x_{im})}{x_{il} - x_{im}} (\mathrm{mod} q).$$

It is easy to prove that: $y_{il} = a_{il}$, because the data $D_j = a_0$

$$D_j = \sum_{l=1}^{k} y_{il} \prod_{1 \le m \le tm \ne l}^{k} \frac{x_{im}}{x_{il} - x_{im}} (\mathrm{mod} q).$$

From the above equation, we can reconstruct the data D_j completely.

C. Carrier Frequency Selection

The PriWhisper+ is designed to deploy in practice, where there may be a noisy environment. For the communication effectiveness concern, we need to consider various types of noise in different application scenario. We analyze the frequency of noise in four main real-life situations, which are traffic noise near a road, musical noise in a pub, vocal noise in a market, and nature noise in a park. From Fig. 5, we can find that most of the noise frequency is below 8 kHz. In order to avoid the background noise band, it is reasonable to set our carrier frequency beyond 8 kHz. On the other hand, the working spectrum of the speaker and microphone of the commercial smartphone is between 20 and 40 000 Hz [34]. However, the commercial smartphone is designed for vocal which means the frequency response curve is much stable between 20 Hz and 20 kHz. Therefore, limited by the current conditions of smartphone hardware, we need to choose our carrier frequency in the audible band. Moreover, the frequency response is another parameter we need to think about. The frequency response specification attempts to describe the range of frequencies a speaker can reproduce, measured in Hertz. Fig. 6 shows the frequency response curve of different models of smartphones. Taking into account the above restrictions, the appropriate carrier frequency spectrum is from 8 to 20 kHz, and we choose 9 kHz as our carrier frequency for higher decoding success rate (DSR).

D. Adaptive Signal Strength Selection

PriWhisper+ is designed for the smartphone in the atmospheric environment, which limits the propagation range because the acoustic signal decays fast in the air medium and restricts jamming signal strength because of the smartphone hardware. In our system, the receiver's speaker determines the maximum intensity of the jamming signal we can get.

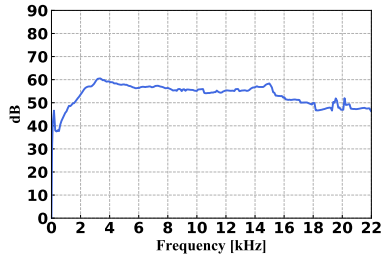


Fig. 6. Frequency response curve of a smartphone.

The maximum volume of the speaker is about 110 dB for the most commercial smartphone. It is user-unfriendly and waste-to-energy that simply choose the maximum strength to propagate the jamming signal. Hence, it is necessary that system selects the appropriate jamming signal level according to the actual situation. To guarantee the data confidentiality, the jamming signal strength $\mathbf{P}_{\text{jam}}$ must be strong enough to cover the data signal. Moreover, the environmental noise strength $\mathbf{P}_{\text{ambient}}$ needs to be considered for data integrity. In Fig. 5, we investigate the noise level in four places which can represent most of the scenario in our daily life. The figure shows that noise level is below 80 dB, except for a few extreme situations. Then we test our system bit error rate performance based on different signal-to-noise ratio (SNR) in Fig. 5. The signal strength of the sender's speaker needs no less than the sum of ambient noise level and the minimum SNR $\mathbf{P}_{\text{minSNR}}$ which is sufficient to decode the data. Also, BSS technology can improve the SNR by dividing part of the jamming signal, which denotes as $\mathbf{P}_{\text{BSS}}$. The impact of the BSS attack will discuss more specifically in Section VI. The attenuation of the jamming signal from the receiver to sender is another factor that can influence the data confidentiality. In summary, The jamming signal strength of our system $\mathbf{P}_{\text{jam}}$ can be calculated from the following equation:

$$\mathbf{P}_{\text{jam}} = \mathbf{P}_{\text{ambient}} + \mathbf{P}_{\text{minSNR}} + \mathbf{P}_{\text{BSS}} + \mathbf{P}_{\text{range}}. \tag{2}$$

E. Jamming Signal Generation and Remove

Friendly jamming technology is designed to protect the sender's data without public key cryptography or symmetric cryptography. Different from the traditional friendly jamming-based radio communication systems that generate the jamming signal by the sender, jamming signal of our system are generated and propagated by the receiver. The jamming signal is prepared before a communication session beginning due to the limitation of computing capacity for the IoT device platform. The receiver uses the fast Fourier transform (FFT) to map the random white Gaussian noise signal to the frequency domain. Then he minimizes all the strength except some special frequency ranges which is the carrier frequencies range. At last, the receiver gets the jamming signal by taking the inverse FFT algorithm. For example, if the carrier frequencies range is 9–10 kHz, the jamming signal should cover the range is 8.5 kHz-10.5 kHz. Hence, the jamming signal can cover the data signal perfectly. In the radio communication systems, the jamming signal is canceled by an antidote signal transmitted by a special transmit chain connected with the receive chain through a so-called self-looping channel, where the antidote signal is carefully chosen to cancel the jamming signal at the receive antenna's front end [15], [30]. It is unacceptable to deployed an additional antidote antenna on the off-the-shelf smartphone platform. Hence, we choose the receiver to propagate the jamming signal to avoid the specific hardware requirements.

The receiver must estimate the jamming signal which its own generated to remove the jamming signal in the received mixed signal. The received jamming signal from the receiver's microphone is distortion because of the imperfection of the manufacturing process of the smartphone's hardware. For example, the microphone's frequency response will cause the magnitude and phase change. To solve this problem, we utilize the frequency selective fading estimation to estimate the jamming signal. In the stage of the system initialization, the receiver's microphone records the acoustic signal while its speaker propagates the signal. Then we can get the acoustic signal's selective fading factor $p(f_i)$ at the frequency f_i. Due to the factor $p(f_i)$ depends on the smartphone's hardware, we obtain the value of $p(f_i)$ from the training data at system initialization stage. We take the short-time Fourier transformation (STFT) and the Inverse STFT to estimate the received jamming signal. For each frequency track, we introduce an independent frequency-selective fading function from STFT of the original jamming signal. The signal transfers to frequency domain after STFT which we can estimate the signal distortion. Then the adjusted signals are combined to the estimate of the received jamming signal by the inverse STFT [15]. And we estimate the frequency-selective fading when the smartphone first run our system. After initialization, the system can estimate the random jamming signal in the following communication. In addition, we also introduce the preamble sound to the jamming signal to assist the synchronization process. The mixture signal can easily divide between the data signal and estimated jamming signal.

F. System Integration

Now, we are ready to integrate all modules together, and we develop the PriWhisper+ system on Android 4.4 OS. The sender needs to synchronize with the receiver before transmitting the data signal. The preamble signal is repeatedly played until the receiver's response signal is received. Then the sender starts to estimate the strength of the jamming signal. In the meantime, both sender and receiver approximately estimate their speed by using acceleration sensor. If the strength of the jamming signal and the moving speed both beyond the threshold, the sender starts to transmit the data while the receiver continues to propagate the jamming signal. The communication process of our prototype as follows. First of all, the sender calculates the length of the data and divide the data into a suitable part of the length. If the length of the last part is insufficient, fill the zero. According to throughput and security requirements, sender determines the value of (k, n, p) and get subdata from the secret share modules. Then sender encodes and modulates the data with parameters $(f, \Delta f, M)$. Then sender recorded 0.01-s audio to evaluate the ambient

Algorithm 1 Transmitter

$\sum_{j=1}^{m} D_j \leftarrow Segment(Data)$
for $j = 1 \rightarrow n_b$ **do**
    $\sum_{i=1}^{n} D_{ji} \leftarrow Secret_$Share$(D_j, k, n)$
end for
$x \leftarrow$ Channel_Encoding$(\sum_{j=1}^{m} \sum_{i=1}^{n} D_{ji})$
$y \leftarrow$ Modulate$(x, , f, \Delta f, M)$
$s \leftarrow$ Audio_Record$(0.01second)$
$e \leftarrow$ Evaluate_Ambient_Noise_level(s)
$z \leftarrow$ Adjust(y, e)
Playback$(Identifier_Sound)$
while true do
    $s \leftarrow$ Audio_Record$(0.01second)$
    $v \leftarrow$ Motion_Sensor_Estimation
    if Jamming_Signal_Detect$(s) =$ **true then**
        if $v > v_{min}$ **then**
            break
        end if
    end if
end while
Audio_Playback(z)

Algorithm 2 Receiver

Playback_$Frequency_Sound(f_a)$
$r \leftarrow$ Audio_Record
$f_r \leftarrow$ Evaluate_Frequency_Response(f_a, r)
$j_s \leftarrow$ Generate_$Jamming_Signal$
$s \leftarrow$ Audio_Record$(0.01second)$
$e \leftarrow$ Evaluate_Ambient_Noise_Level(s)
$j \leftarrow$ adjust(j_s, e)
while true **do**
    $i \leftarrow$ Audio_Record$(0.01second)$
    $v \leftarrow$ Motion_Sensor_Estimation
    if Identifier_Signal_Detect$(i) =$ **true then**
        if $v > v_{min}$ **then**
            break
        end if
    end if
end while
Playback(j)
$r_m \leftarrow$ Audio_Record
$y \leftarrow$ Remove_Jamming_Signal(r_m, j, f_r)
$x \leftarrow$ Demodulate$(y, f, \triangle f, M)$
$Data \leftarrow$ Secret_Recover(x, m, n)

noise level. Adjust the intensity of the data signal according to the intensity of the background noise. After preparing the data signal, the sender plays the preamble sound to wake up the receiver. The sender starts to loop recording 0.01-s audio until the strong enough jamming signal detected. In the meantime, through calculating the data from the accelerator sensor and the orientation sensor, the sender estimates the speed of the mobile devices. If the speed is fast enough and the strength of the jamming signal is sufficient, the sender starts to play the data signal. For the receiver, he needs to play and record a frequency sound simultaneously. The receiver evaluates the frequency response from the recording. The next step is to generate a random jamming signal. Then the receiver adjusts the strength of the jamming signal base on the evaluate ambient noise level. After the above preparations, the receiver starts to loop recording audio until the preamble signal detected. Then receiver plays the jamming signal and records the mix signal at the same time. Based on estimated frequency response previously, the receiver removes the jamming signal from the mix signal and demodulate the signal. At last, receiver recovers the data based on decoding subdata. Algorithms 1 and 2 show the pseudo code of our PriWhisper+ system.

V. Security Analysis

In this section, we discuss major security attacks on the PriWhisper+, assuming that device S tries to safely send a message M to R, while there is a malicious attacker E right next door.

A. Placement Attack

Our approach in PriWhisper+ combines jamming technology with Shamir's secret sharing scheme to guarantee data confidentiality. We first analyze the jamming technology which can protect against eavesdropping attack very well. Different from the common methods for achieving secure communication through cryptographic methods, the jamming technology uses information theory to secure communication. Shannon's information theory is the foundation of the information-theoretic approach. The one-time pad (OTP) encryption is a classic algorithm which has perfect secrecy proved by Shannon. The core idea of the algorithm is to encrypt the message by using a completely random key that never reuses. We assume that the sender S tries to send a confidential message M to the receiver R and there is an eavesdropper E that can eavesdrop on all the messages. In the first place, the sender S and receiver R need to share a random secret key ω which is the same length as the message, S and R share ω via the independent channel to ensure that the eavesdropper E cannot get the ω. Then S uses ω to encrypt the M to get $M' = M \oplus \omega$, then the sender transmits the encrypted message M' to the receiver with the presence of E. The receiver R can then extract the M from the M' because he holds the ω while the eavesdropper E cannot retrieve the M without ω.

There are obvious practical difficulties in S and R finding a secure channel to share ω although the OPT approach is secure theoretically. The friendly jamming technology takes a different method, S transmits M to R via the channel CH_{SR} that is less noisy than the channel CH_{SE} which E eavesdrops. The key point in this model is that the receiver propagates the noise N_r strong enough to cover the message M, both CH_{SR} and CH_{SE} have too much noise to decode the message. However, the CH_{SR} can cancel noise by using noise canceling mechanism, such as self-interference-cancelation (SIC) mechanism. Receiver R can apply the SIC to cancel the jamming signal because he exactly knows what the jamming signals look like, and he can extract the M by estimating the ω from M'. The receiver propagates jamming signal instead of sending random key from the sender in OTP. This

approach may cause some security vulnerability which we will continue the discussion in the security attack analysis. From the above the paragraph, we have demonstrated that the jamming technology can perfectly prevent the eavesdropping attack. The attacker E can get the mixed signal M' easily, but he cannot retrieve the M without the original jamming signal. However, that may change if E is some special location. Assuming that the eavesdropper E's position is between M and R, E may recover the message M because of the insufficient jamming signal strength. We denote the distance between S and R by d and between S and E by L. The acoustic signal power decaying with distance x is $x^{-\alpha}$ [14]. We can get the SNR at the eavesdropper's location by

$$\text{SNR} = \frac{P_s}{P_j}\left[\frac{L^2 + d^2 - 2Ld\cos\theta}{L^2}\right] \tag{3}$$

where the P_S and P_J, respectively, denote the sender's signal power and the receiver's jamming signal power. The θ denotes the angle between L and d. From (3), we can find the maximum value of SNR when $d = 0$. Thus, the closer E is getting to S, the higher recovering success rate E can get. However, as discussed in Section IV, S does not transmit the message unless the jamming signal is strong enough. In our system, the sender detects the strength of the jamming signal before sending the data signal. The power of the jamming signal should ensure that the jamming signal can cover the data signal at the location of S. If E fails to extract the data signal at the location of S, E cannot succeed anywhere else. Because E cannot get better SNR than S's location, E cannot extract the message from the mixed signal. E cannot decode the data signal as long as R propagates the jamming signal follows (2). In (2), R calculates the required strength of the jamming signal according to the ambient noise level and the attenuation of the audio wave in the air medium. So E cannot perform the placement attack if E does not have the data. We get the secure communication distance by combining the experiment in the practical environment with (2). We find that the 30 cm is a safe communication distance, and the attacker probability to implement placement attacks because the strength of the jamming signal cannot cover the data signal.

B. Blind Source Separation Attack

In the previous part, we show that the system can prevent the placement attack. In this part, we assume a more powerful adversary who has two or more microphones and can place any location to record the communication so that he can use the BSS technology to separate the data signal from the mixed signal. The ICA is one of the most effective technology in the BSS field.

ICA has become one of the most effective technologies on BSS problem. The ICA attempts to find the original non-Gaussian sources by assuming that the linear mixed signals are statistically independent. Imagine that you are in a room, where two smartphones are playing different music simultaneously, which we could denote by s_1 and s_2. You have two microphones that give you two recorded signals, which we could denote by x_1 and x_2. Each recorded signals consists of all these music signals, which have different weights according to their distance to the microphone. Ideally, we omit the channel attenuation and ambient noise so that we could express this as a mixing model [20]. There are plenty of research on ICA algorithm [21]–[29]. We choose a state-of-art ICA algorithm for the adversary in our experiment

$$\begin{aligned} \mathbf{x}_1 &= a_{11}\mathbf{s}_1 + a_{21}\mathbf{s}_2 \\ \mathbf{x}_2 &= a_{12}\mathbf{s}_1 + a_{22}\mathbf{s}_2. \end{aligned} \tag{4}$$

The basic model of the ICA technology is described in (4), a_{i1} can be replaced by the frequency-selective fading function $p(f_c, L_{i1})$ in the following situation. Fig. 7(a) describes two typical movement routes in the short-range communication scenario. We first perform the security analysis of the case that the sender and receiver move along the red line in this figure, then extend to the general scenario. Let d denote the distance between two legitimate communication devices S_i, l denote the distance between two eavesdropping devices E_j. The distance between S_i and E_j is denoted by L_{ij}. Assuming that the fading function is homogeneous and uniform for simplicity. Let v be the movement speed of S_i, we have

$$\Delta L_{ij} = \frac{\text{dLP}}{\sqrt{8L_{ij}^2 - 4d(L + \text{vt}) + d^2}}.$$

From the above equation, we get the relationship between the distance L_{ij} and the speed v. The L_{ij}'s change will lead to the change of Δa_j because the strength of speaker is constant. We denote Δa_j as the change of a

$$\Delta a_j = \left|\frac{a'_{1j}}{a'_{2j}} - \frac{a_{1j}}{a'_{2j}}\right|.$$

Then we bring the ΔL_{ij} into the above equation, and we can get a relationship between Δa_j and v. Fig. 7(b) shows a successful separation result while the sender and receiver are stationary. Fig. 7(c) shows the BSS algorithm is failed when the v is up to 60 cm/s. Furthermore, the adversary will need to address more challenges when the direction of v is random, which will cause the channel condition to become more randomized.

Because the BSS model is a second- or fourth-order statistic-based method to separate the signal by estimating the value of a, it needs adequate samples for iteration to get a. The adversary cannot divide the signal into the arbitrarily small segment [27]. If a's change is not negligible in the smallest segment that the attacker can divide. The BSS algorithm cannot estimate the value of a while the v is fast enough. We can extend the above situation to the general scenario and the only difference is the direction of the smartphone's movement. Furthermore, the moving acoustic signal will also bring a doppler effect and more complex multipath impact, which makes the channel conditions even more randomized. In our design, the sender and receiver move with a random path, which combines the randomized channel and the changeable mixed signal. And our experiments in the next section verify the security of our proposed system.

Consider that the user may move their smartphone below the secure speed occasionally during communication. Intuitively, the attacker may get part of data when the v is not big enough.

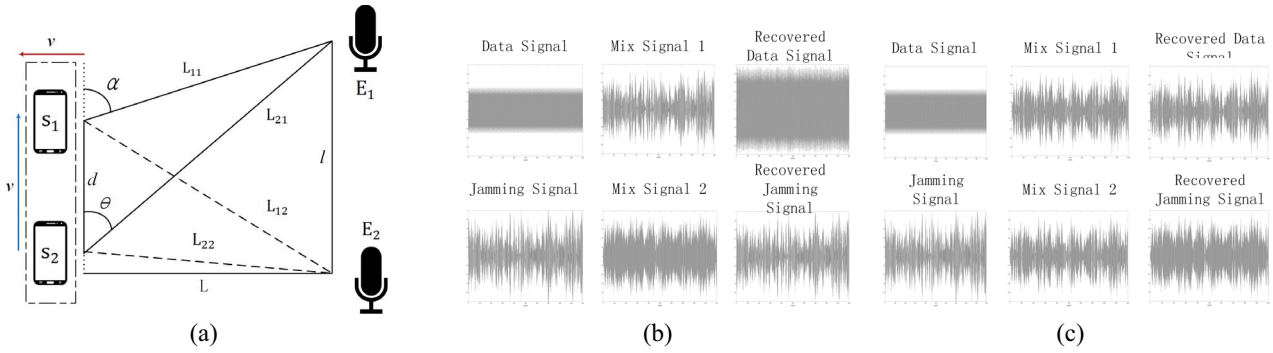


Fig. 7. (a) BSS attack model. (b) Successful static separation attack experiment. (c) Failed static separation attack experiment.

Fig. 8. Error data distribution in acoustic communication.

Furthermore, consider the complexity of the acoustic channels in practice. Normal acoustic communication between the sender and the receiver may get some error bit. For instance, Fig. 8 shows the error bit position when there is a sudden loud noise. We can find that the error data focus in few parts. Hence, we leverage the secret sharing scheme to improve the confidentially of the data and the success rate of data decoding.

Shamir and Blakley independently proposed the concept of secret sharing and gave a $(k,\ n)$ threshold secret sharing scheme in 1979. The threshold scheme divides a secret S into n members in the control of n segments. A size k (or more) of the participants can cooperate to reconstruct the secret, and the subsets of size less than k participants cannot cooperate to rebuild the secret. Shamir's secret sharing scheme is based on the polynomial interpolation. The scheme intends to divide a secret data D into n parts that any k parts can retrieve D, but any $k-1$ parts understand nothing about D. The scheme will share secret S as a point in the k-dimensional space. The sender S performs the following steps to share a secret data D.

1) Let $D \epsilon GF(p)$, where p is a prime and $p > n$. S secretly chooses $k-1$ random points of $GF, a_1, a_2, \ldots, a_{k-1}$.
2) S gets a polynomial, where

$$f(x) = D + \sum_{k-1}^{j=1} a_j x^j.$$

3) For $1 \le i \le n$, S computes $y_i = f(x_i)$.
4) For $1 \le i \le n$, S sends (x_i, y_i) to R.

The receiver R can get D while he receive no less than k points by calculating

$$a(\mathbf{x}) = a_0 + a_1\mathbf{x} + a_2\mathbf{x}^2 + \cdots + a_{k-1}\mathbf{x}^{k-1}.$$

Adversary cannot retrieve D unless he has more than $k-1$ points. Suppose the attacker has $k-1$ points and bring it into the above polynomial. For every hypothesized value D of the secret, there is a unique polynomial $y(x)$. Hence, no value of the secret can be ruled out. The [31] describes more specific details about the secure analysis.

C. DoS Attack

A malicious device E can generate and propagate its jamming signals in order to disturb the normal communication between S and R. E needs to generate enough power of the jamming signal if it tries to disallow the regular communication successfully. The E can be easily detected because the working spectrum of communication is between 8 and 10 kHz in practice. Some prior work use ultrasound to create a shadow in the audible frequency range [13], [34], but it has a lot of limit in practice, such as working distance is very short because of the fast attenuation on ultrasound band, impact frequency band is below 5 kHz, and the sound pressure is less than 50 dB [34]. Our system can adaptively adjust data and jamming signal strength to handle ambient noise below 60 dB. Even if it paralyzed the communication between S and R without exposure, E still cannot extract the data from the mixed signal.

D. Man-in-the-Middle Attack

Due to the feature of the acoustic communication, the attacker needs to propagate audible jamming signal which means exposing its position. In fact, almost all active attacks, including man-in-the-middle attack, are easy to find because of the protocol rule that receiver needs to send the jamming signal to the sender. If the attacker sends the jamming signal, his position will expose. And the masquerade that attacker pretends to be the sender will fail because the different positions between the attacker and the real sender will be found when attacker start sending data signal.

In the above argument, we can find that the active attacks, such as DoS attack and man-in-the-middle attack, are hard to attack because of the audible feature in acoustic short-range communication. In contrast, the passive attacks, such as placement attack and separation attack, are hard to detect. Hence, the security analysis for the passive attack is our primary objective.

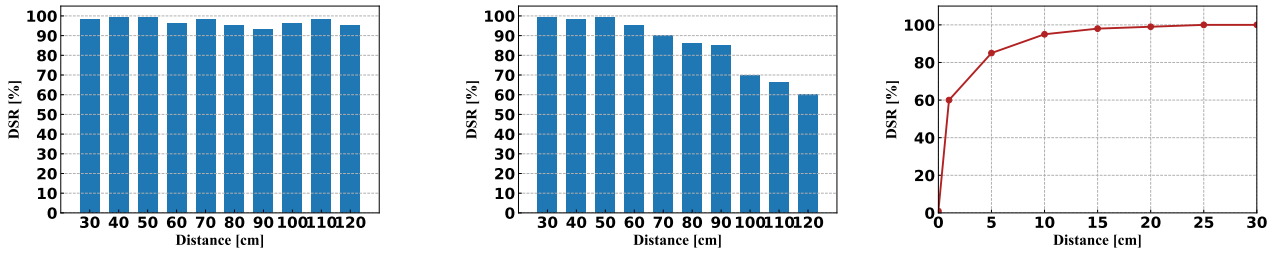


Fig. 9. Static experiment for decode success rate versus change distance (a) between two eavesdropping microphones, (b) from the microphones of eavesdropper to victim smartphones, and (c) from sender to receiver.

VI. Experimental Evaluation

In this section, We quantify the aspects of PriWhisper+, such as its secure communication range, the performance in different locations, and data DSR under the impact of the movement.

A. Experiment Settings

1) Hardware: We run our prototype on many models of Android phones, such as Samsung Note 5, Huawei P9, and LG G4. The result shows that the type of the smartphone does not affect the experiment. We choose an LG Nexus 5X and a Samsung Galaxy S4 as our sender and receiver in the following experiment. For the adversary, we assume that he can have an advanced capability. Hence, we use commercial microphones which type is SF—666 instead of the smartphones as the eavesdropper's recording devices. We link the microphone to the laptop and run state-of-the-art ICA algorithm on it for the separation attack.

2) Environment: In order to test the robustness of our system prototype in a more noisy practice environment, we test it in four practical environments: 1) roadside; 2) mall; 3) cafe; and 4) park. Fig. 10 shows the performance of PriWhisper+ in various environments. This result demonstrates that our system can work perfectly, wherever in a traffic noisy roadside or a mall with vocal noise. Most types of ambient noises are hard to influence the DSR of PriWhisper+.

B. Performance Evaluation

We conduct massive experiments to evaluate the performance of our system prototype and verify its security level. We will gradually test our system in this section. First of all, we conduct the attenuation of the acoustic signal that is propagated by commercial smartphone in practice. We select four representative locations: a community park includes many nature noises, such as singing birds, shaking leaves, and barking dogs; a busy roadside includes many traffic noise, such as brakes, horn sound, and roar engine; a large mall includes footsteps noise, trolley noise, and talking noise; a busy cafe includes background music and chatting noise. Due to our system works in the audible band, users may feel bothered about the sound of the communication. Fortunately, the process of the short-range communication is only a few seconds under normal conditions. Moreover, our system can adjusts the sound based on the volume of the ambient noise, which ensures that the communication sound does not bother the users. Then we evaluate the communication distance of the acoustic signal propagating by smartphone, which we do not introduce the jamming signal in Fig. 11. We use the Galaxy Note 5 to test the communication distance without jamming signal to verify the attenuation of the acoustic signal in the air medium. The result shows the DSR is below 10% while the communication distance exceeds two meters. Thanks to the fast decay of the acoustic waves in the air medium, the strength of the data signal is lower than the ambient noise. The acoustic signal is suitable for applying in the NFC.

C. Capability of the Separation Attack

Before evaluating the PriWhisper+'s performance against the separation attack, we analyze the capability of the separation attack against acoustic friendly jamming technology first.

1) Distance Between Eavesdropping Sensors: From Fig. 9(a), we can see that the distance between eavesdropping sensors which is denoted by l have a negligible impact on DSR. The experiment results show that all of the DSR is above 90% from 5 to 50 cm. Hence, we conclude that the influence of different distance between eavesdropping sensors can be omitted in our experiment. We select a fixed value $l = 10$ cm in our later experiment.

2) Distance Between Eavesdropping Sensors and Transaction Devices: Assuming the distance between eavesdropping sensors and transaction devices is L. Because of the complexity of the aerial channel, the L is one of the important factors that affect the DSR. In this experiment, we use $d = 30$ cm, and the sender speaker's volume is max. From Fig. 9(b), we can see that when the distance is larger than 90 cm, the DSR is under 80%, so we will control our displacement distance shorter than 90 cm to guarantee the accuracy of our experimental results.

3) Distance Between Transaction Devices: We choose the state-of-art ICA algorithm as our attack method. We conclude a BSS attack experiment, where transaction devices are static. In this experiment, we set the $L = 50$ cm. From Fig. 9(b), the eavesdropper utilizes the ICA technology to separate the mixed signal from x_1 and x_2. We can see that when distance d

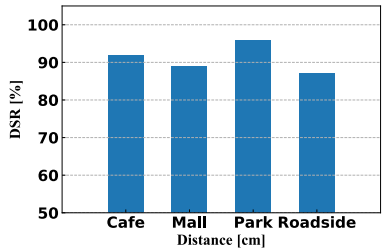


Fig. 10. Performance of PriWhisper+ in various environment.

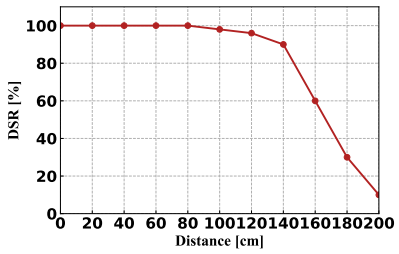


Fig. 11. Outdoor acoustic signal attenuation experiment.

is greater than 5 cm, the DSR is above 90%, and the separation results compared with the original signal is shown in Fig. 9(c).

D. Performance of Preventing the Separation Attack

We start with the fixed distance d to explore the relationship between smartphone speed and the DSR. Fig. 13 shows two lines we choose to move. From Section V, we can see that the speed changing rate and the direction changing rate are the most important factors. In order to accurately evaluate the performance of signal movement for preventing the separation attack, we test our system without using secret sharing module at first.

1) Relative Movement: The biggest speed changing rate is moving along the relative direction. The experiment result is shown in Fig. 12(a). The decode success rate drops quickly with increasing speed. We can see that the attacker can hardly decode the data while the moving speed is over 60 cm/s.

2) Parallel Movement: On the other hand, the fastest direction changing rate is moving along the parallel direction. Fig. 12(b) shows that when speed is up to 60 cm/s, the attacker will fail to retrieve the secret data. Moreover, if we compare the result between Fig. 12(a) and (b), we can find that parallel movement has more impact on DSR than relative movement. The reason is that the parallel movement of the mixing ratio changes more.

3) Random Movement: Then we evaluate the scenario that users move their smartphone along an arbitrary path and variable speed. In random movement, sender and receiver move in an arbitrary direction with their mobile device, they can parallel move in a specific direction or walk to the respective path. But the distance between two communication devices must be less than the secure distance which is 30 cm. Sender and receiver can move freely as long as the secure communication distance and speed are guaranteed. Fig. 12(c) shows the result that both sender and receiver along the random path communicate with each other. From the experiment result, we can see that the user's actions are unpredictable which introduces increasing difficulty of eavesdropping.

TABLE I
THROUGHPUT EVALUATION OF PRIWHISPER+

N	1	3	5	10
Data rate (bps)	1024	351	212	134

4) Discussion: We can see that the direction of the user's movement relative to the attacker's position has an impact on the attack success rate from above experiment result. The parallel movement has a more significant effect on DSR. However, we are not clear the attacker's position in practice. That is the reason why we assume adversary can eavesdrop the communication at any location in Section III-B. Hence, we also evaluate the DSR in random path scenario which closer to the real life. The result shows the attacker cannot extract the data no matter how the user moves.

E. Choice of the Parameter (k, n)

Fig. 14 shows the result that six people move along a random line with different speed. It reflects the relationship between the choice of (k, n) and the DSR of the attacker. We can find that attacker's DSR is below 40% in most situations when the speed is around 30 cm/s, and we can set the speed threshold at 60 cm/s, and set the threshold of the $(k/n) = 60\%$ just in case. For example, we choose $k = 3, n = 5$ in our prototype and test it in many times, and the attacker cannot get enough subdata to recover the communication data. If the user's speed cannot achieve the 60 cm/s, we can set $(k = 7, n = 10)$ or $(k = 4, n = 5)$ to guarantee the data confidentiality. If the user has a higher security requirement, we can choose the speed of 60 cm/s for communication. From Fig. 14, we can find the attacker cannot extract any data when the speed is around 100 cm/s. So the user can choose different communication speed according to different levels of security requirement. According to the species of the IoT portable devices, we can choose different threshold of the speed and the choice of the parameter (k, n) to ensure the security of the system communication.

F. Throughput

The throughput of our prototype depends on the value of n. In Table I, the maximum throughput is almost 1 kb/s when the $n = 0$, which means that the secret sharing module is not used. We choose a bigger n for higher data confidentially, which leads to a tradeoff between security and throughput. The throughput decreases as the number n increases, for example, if user chooses $n = 5$ to protect the data, then the data will be partitioned into five pieces by using secret sharing scheme, which means that the system will use five packets to deliver one packet data and hence the throughput of our system will drop to one-fifth of the maximum transmission speed. Throughput is not our primary objective because our system places more values on the data confidentiality. Fortunately, it is easy to improve the throughput by using some advanced technology, such as orthogonal frequency-division multiplexing (OFDM). The QR code is a common method for wireless

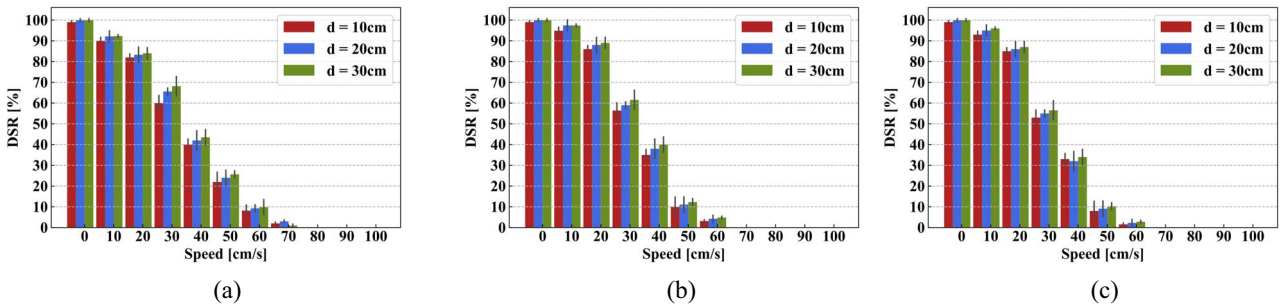


Fig. 12. Dynamic experiment for (a) relative movement versus decode success rate, (b) parallel movement versus decode success rate, and (c) move along random path by users in practice.

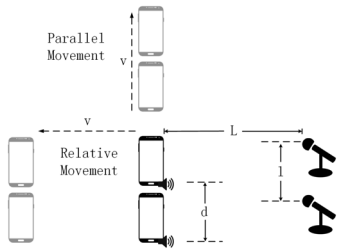


Fig. 13. Experimental parameter settings.

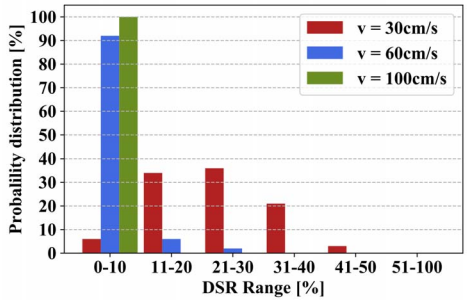


Fig. 14. Attacker's DSR range while the users move along a random path with different speed.

payment field [35]. A normal QR code pattern (version 4) can contain 807 bytes data [36]. Our prototype throughput can reach 2 kb/s by using OFDM technology. So we believe that our system throughput can be fully applied in the wireless payment field.

G. Energy Consumption

Battery consumption also needs to be considered in the mobile applications. We study the remaining battery percentage after 1 h acoustic communication by PriWhisper+ between two Samsung Note 5 smartphones. Fig. 15 shows the result that the sender has 88% power left while the receiver has approximately 86% power left. We find that the receiver consumes more power than the sender. Due to the receiver needs computing to remove the jamming signal. It also needs more computing that recovers the data from the subsecrets of the secret sharing scheme. It explains the reason why receiver consumes more battery. Comparing with the PriWhisper system, we can find the PriWhisper+ consumes more power. That is

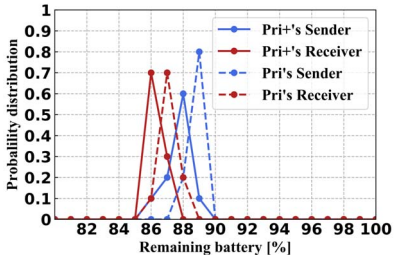


Fig. 15. Battery drain experiment.

because we introduce the secret sharing scheme and the speed evaluation module.

VII. Conclusion

In this paper, we design the PriWhisper+, an enhanced secure acoustic short-range communication system for IoT portable devices. We analyze the performance of the system against a variety of attacks in short-range communication scenario, especially against BSS attack. The results of the experiment show that the PriWhisper+ achieves a high-security level in the dozens of centimeters range. The system's throughput can achieve as high as 1 kb/s, and we will improve the throughput in future work. The system currently works in the band that is audible to humans and may bother the user. Another further improvement will be the feasibility and security of communication in the ultrasound band.

Acknowledgment

The authors would like to thank the editors and anonymous reviewers for their insightful comments and constructive feedback.

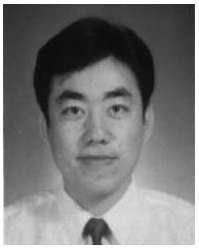
References

[1] J. Parker and N. Ralph. *Everything You Want to Know About Apple Pay*. Accessed: Jul. 14, 2015. [Online]. Available: https://www.cnet.com/news/everything-you-want-to-know-about-apple-pay/
[2] Chirp. *Chirp: Why Sound*. Accessed: 2018. [Online]. Available: https://www.chirp.io/why-sound/
[3] J. Guerrieri and D. Novotny, "HF RFID eavesdropping and jamming tests," Electromagnetics Div., Electron. Elect. Eng. Lab., NIST, Gaithersburg, MD, USA, Rep. 818-7-71, 2006.
[4] H. Kortvedt and S. Mjølsnes, "Eavesdropping near field communication," in *Proc. Norwegian Inf. Security Conf. (NISK)*, 2009, p. 27.

[5] M. M. A. Allah, "Strengths and weaknesses of near field communication (NFC) technology," *Glob. J. Comput. Sci. Technol.*, vol. 11, no. 3, pp. 50–56, 2011.
[6] T. P. Diakos, J. A. Briffa, T. W. C. Brown, and S. Wesemeyer, "Eavesdropping near-field contactless payments: A quantitative analysis," *J. Eng.*, vol. 2013, no. 10, pp. 48–54, 2013.
[7] *NFC-SEC: NFCIP-1 Security Services and Protocol*, Standard ECMA-385, 2010. [Online]. Available: http://www.ecma-international.org/publications/files/ECMA-ST/ECMA-385.pdf
[8] Q. Wang *et al.*, "Rain bar: Robust application-driven visual communication using color barcodes," in *Proc. IEEE 35th Int. Conf. Distrib. Comput. Syst. (ICDCS)*, 2015, pp. 537–546.
[9] B. Zhang, K. Ren, G. Xing, X. Fu, and C. Wang, "SBVLC: Secure barcode-based visible light communication for smartphones," *IEEE Trans. Mobile Comput.*, vol. 15, no. 2, pp. 432–446, Feb. 2016.
[10] A. Wang *et al.*, "InFrame++: Achieve simultaneous screen-human viewing and hidden screen-camera communication," in *Proc. 13th Annu. Int. Conf. Mobile Syst. Appl. Services*, 2015, pp. 181–195.
[11] W. Mao, J. He, and L. Qiu, "CAT: High-precision acoustic motion tracking," in *Proc. 22nd Annu. Int. Conf. Mobile Comput. Netw.*, 2016, pp. 69–81.
[12] Q. Wang *et al.*, "Messages behind the sound: Real-time hidden acoustic signal capture with smartphones," in *Proc. 22nd Annu. Int. Conf. Mobile Comput. Netw.*, 2016, pp. 29–41.
[13] N. Roy, H. Hassanieh, and R. R. Choudhury, "BackDoor: Making microphones hear inaudible sounds," in *Proc. 15th Annu. Int. Conf. Mobile Syst. Appl. Services*, 2017, pp. 2–14.
[14] R. Nandakumar, K. K. Chintalapudi, V. Padmanabhan, and R. Venkatesan, "Dhwani: Secure peer-to-peer acoustic NFC," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 43, no. 4, pp. 63–74, 2013.
[15] B. Zhang *et al.*, "PriWhisper: Enabling keyless secure acoustic communication for smartphones," *IEEE Internet Things J.*, vol. 1, no. 1, pp. 33–45, Feb. 2014.
[16] V. Gerasimov and W. Bender, "Things that talk: Using sound for device-to-device and device-to-human communication," *IBM Syst. J.*, vol. 39, nos. 3–4, pp. 530–546, 2000.
[17] P. Smaragdis, C. Fevotte, G. J. Mysore, N. Mohammadiha, and M. Hoffman, "Static and dynamic source separation using nonnegative factorizations: A unified view," *IEEE Signal Process. Mag.*, vol. 31, no. 3, pp. 66–75, May 2014.
[18] E. Vincent, R. Gribonval, and F. C. Votte, "Performance measurement in blind audio source separation," *IEEE Trans. Audio, Speech, Language Process.*, vol. 14, no. 4, pp. 1462–1469, Jul. 2006.
[19] D.-T. Pham, C. Servière, and H. Boumaraf, "Blind separation of convolutive audio mixtures using nonstationarity," in *Proc. ICA*, 2003, pp. 981–986.
[20] J. Shlens. (Apr. 2014). *A Tutorial on Principal Component Analysis*. [Online]. Available: http://arxiv.org/abs/1404.1100
[21] S. Ding, A. Cichocki, J. Huang, and D. Wei, "Blind source separation of acoustic signals in realistic environments based on ICA in the time-frequency domain," *Int. J. Pervasive Comput. Commun.*, vol. 1, no. 2, pp. 89–100, 2005.
[22] A. Koutvas, E. Dermatas, and G. Kokkinakis, "Blind speech separation of moving speakers in real reverberant environments," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, vol. 2, 2000, pp. II1133–II1136.
[23] H. Saruwatari, T. Kawamura, T. Nishikawa, A. Lee, and K. Shikano, "Blind source separation based on a fast-convergence algorithm combining ICA and beamforming," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 14, no. 2, pp. 666–678, Mar. 2006.
[24] L. Faivishevsky and J. Goldberger, "ICA based on a smooth estimation of the differential entropy," in *Proc. Adv. Neural Inf. Process. Syst.*, 2009, pp. 433–440.
[25] G. E. Naik and D. K. Kumar, "An overview of independent component analysis and its applications," *Informatica*, vol. 35, no. 1, pp. 63–81, 2011.
[26] K. E. Hild, D. Erdogmus, and J. C. Principe, "On-line minimum mutual information method for time-varying blind source separation," in *Proc. Int. Workshop Independ. Component Anal. Signal Separation (ICA)*, 2001, pp. 126–131.
[27] R. Mukai, H. Sawada, S. Araki, and S. Makino, "Blind source separation for moving speech signals using blockwise ICA and residual crosstalk subtraction," *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, vol. 87, no. 8, pp. 1941–1948, 2004.
[28] T.-W. Lee, A. J. Bell, and R. Orglmeister, "Blind source separation of real world signals," in *Proc. IEEE Int. Conf. Neural Netw.*, vol. 4, 1997, pp. 2129–2134.
[29] S. Araki, R. Mukai, S. Makino, T. Nishikawa, and H. Saruwatari, "The fundamental limitation of frequency domain blind source separation for convolutive mixtures of speech," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 2, pp. 109–116, Mar. 2003.
[30] S. Gollakota, H. Hassanieh, B. Ransford, D. Katabi, and K. Fu, "They can hear your heartbeats: Non-invasive security for implantable medical devices," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 41, no. 4, pp. 2–13, 2011.
[31] A. Shamir, "How to share a secret," *Commun. ACM*, vol. 22, no. 11, pp. 612–613, 1979.
[32] C. V. Lopes and P. M. Q. Aguiar, "Aerial acoustic communications," in *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust.*, 2001, pp. 219–222.
[33] K. Mostafa. *Minimodem*. Accessed: Jan. 1, 2013. [Online]. Available: http://www.whence.com/minimodem/
[34] G. Zhang *et al.*, "DolphinAttack: Inaudible voice commands," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security*, 2017, pp. 103–117.
[35] Wikipedia. (Mar. 2018). *QR Code*. [Online]. Available: https://en.wikipedia.org/wiki/QR code
[36] British Standards Institution(BSI), *Information Technology. Automatic Identification and Data Capture Techniques. QR Code 2005 Bar Code Symbology Specification*, BS Standard ISO/IEC 18004:2006. [Online]. Available: https://shop.bsigroup.com/ProductDetail/?pid=000000000030201420
[37] K. E. Hild, D. Erdogmus, and J. C. Principe, "Blind source separation of time-varying instantaneous mixtures using an on-line algorithm," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, vol. 1, 2002, pp. I-993–I-996.

Xiao Zhang (GS'18) received the B.S. degree from Beijing Jiaotong University, Beijing, China, in 2013, where he is currently pursuing the Ph.D. degree with the School of Computer and Information Technology.

His current research interests include short-range communication security, privacy preserving, and Internet of Things security.



Jiqiang Liu (M'14) received the B.S. and Ph.D. degrees from Beijing Normal University, Beijing, China, in 1994 and 1999, respectively.

He is currently a Professor with the School of Computer and Information Technology, Beijing Jiaotong University. He has authored or co-authored over 100 scientific papers in various journals and international conferences. His current research interests include trusted computing, cryptographic protocols, privacy preserving, and network security.

Si Chen (GS'12–M'17) received the Ph.D. degree in computer science and engineering from the University at Buffalo–SUNY, Buffalo, NY, USA, in 2016.

He is an Assistant Professor with the Computer Science Department, West Chester University, West Chester, PA, USA. His research is supported by Microsoft. His current research interests include security and privacy in mobile sensing and cyber-physical systems.

Dr. Chen was a recipient of the Best Student Paper Award of IEEE ICDCS 2017. He is a member of the ACM.

Yongjun Kong (S'18) received the B.S. degree in information research and security form the Engineering University of the Chinese Armed Police Force, Xi'an, China, in 2017, where he is currently pursuing the M.S. degree.

His current research interests include information security, reversible data hiding, and image application based on generator neural networks.

Kui Ren (F'16) received the Ph.D. degree from the Worcester Polytechnic Institute, Worcester, MA, USA.

He is currently a SUNY Empire Innovation Professor of Computer Science and Engineering with the University at Buffalo, State University of New York, Buffalo, NY, USA, where he is also the Director of the UbiSeC Laboratory. His research is supported by the NSF, DoE, AFRL, MSR, and Amazon. He has authored extensively in peer-reviewed journals and conferences. His current research interests span cloud and outsourcing security, wireless and wearable systems security, and mobile sensing and crowdsourcing.

Dr. Ren was a recipient of several Best Paper Awards of IEEE ICDCS 2017, ACM/IEEE IWQoS 2017, and IEEE ICNP 2011. He currently serves as an Associate Editor for the IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING, the IEEE TRANSACTIONS ON SERVICES COMPUTING, the ACM/IEEE TRANSACTIONS ON NETWORKING, *IEEE Wireless Communications*, and the IEEE INTERNET OF THINGS JOURNAL and as an Editor for *Springer Briefs on Cyber Security Systems and Networks*. He is a Distinguished Member of the ACM and a Past Board member of the Internet Privacy Task Force, State of Illinois.