You Can Hear But You Cannot Record: Privacy Protection by Jamming Audio Recording

Xiaosong Ma[†*§], Yubo Song[†*§], Shang Gao[‡], Bin Xiao[‡], Aiqun Hu[†*§]

[†]*School of Cyber Science and Engineering, Southeast University*

[*]*Key Laboratory of Computer Network Technology of Jiangsu Province*

[§]*Purple Mountain Laboratories*

[‡]*Department of Computing, The Hong Kong Polytechnic University*

220194602@seu.edu.cn, songyubo@seu.edu.cn, goldensaintgao@qq.com, csbxiao@polyu.edu.hk, aqhu@seu.edu.cn

***Abstract*—Unauthorized voice recording via smartphones can leak the talking content stealthily. This would be a serious security threat to those individuals, enterprises and the government who need to keep the conversation confidential. Furthermore, due to the size miniaturization of smartphones, it is hard to find the covert recording from malicious attendees. Existing solutions usually jam the recording with audible noise or electromagnetic emitting. However, the audible noise will seriously interfere with conversation and the effect of electromagnetic emitting will be limited by the distance. In this paper, we propose UltraArray, a pioneering silent ultrasonic anti-recording jammer, which can covertly block recording for a long distance. The principle of covert blocking is inspired by acoustic parametric array theory, which suggests that the audible frequency wave can be spread through the air silently while it is modulated to an inaudible ultrasonic frequency. The modulation used in this paper is double sideband (DSB) modulation. The microphone on the phone will record the audible frequency and filtering out the ultrasonic frequency. The jammer we developed uses an acoustics array to form a beam to spread the signal further. The evaluation shows that the device has a good jamming effect on more than 5 meters for most Android smartphones. It will also work well with more than 2.5 meters effective distance on iPhone XR, which has the active noise control (ANC) function. Those results achieve ten times the interference ability of existing solutions.**

Keywords—Acoustic parametric array, anti-recording, ultrasonic

I. Introduction

Now we are in an era of the rapid development of network information, the importance of information is self-evident. There is so many information which is attractive to criminals. That information involves the private information of hundreds of millions of ordinary citizens, including basic information, device information, account information, privacy information, social relationship information, network behavior information and the secret information of groups, enterprises or even countries. The leakage, theft and trafficking of private information are growing barbarously, even creating a striking black industry chain. Unintentional actions in our lives may have leaked a lot of our information. There are various ways to steal information. Common methods include eavesdropping on recordings, spyware on the PC, keyloggers, and advertisers' records of browsing information. In recent years, more and more researchers have paid attention to information leakage caused by voice-based smart devices.

Some researchers have realized that some smart devices can monitor and save voice information, such as smartphones, smart watches and smart speakers. These smart devices can execute commands by recognizing vocal keywords. In order to respond to voice commands at any time, they will record and analyze in real time. In this process, the user's voice data may be leaked, then the attacker can analyze it and get the user's privacy information [1-3]. Some researchers found that ultrasound can affect the microphone of smart devices. Because ultrasonic waves are inaudible, this phenomenon can be used covertly on smart devices, and has high research value and application value. An attack method is proposed [4,5]. By sending a special ultrasonic wave to the microphone of the smart voice device, the voice command to the smart voice device can be realized. Some experiment indicates that this method can activate and control the smart device at a distance of 2-3 meters. This attack method has been analyzed from the feasibility to prove its effectiveness [6]. Other researchers try to realize precisely positioned the attack by cleverly arranging ultrasonic speakers [7]. Inspired by this phenomenon, some researchers want to use a more powerful ultrasound to cover the human voice in the microphone recording [8]. They designed a bracelet covered with ultrasonic speakers and wanted to use it to block the recording of the microphone. However, the effective interference distance of this scheme is too close, and the equipment is too large. There is still a long distance for practical use.

Among the information leakage caused by all smart devices, the easiest and most direct way is voice recording by smartphones, which does not require complicated equipment and method, just need a smartphone carried by the attendee and turn on the recording function. With the rapid development of smart devices, it is more convenient and hidden to use some recording devices, such as smartphones and recording pens, to steal information, which is a hidden danger to individuals, enterprises and governments. Imagine the following scenario: In a company's meeting, the company's executives are exchanging confidential information. However, one of the participants was bought by the competitors. He only needs to open the recording function of his smartphone or secretly place hidden recording equipment in the venue. Then he records all the content delivered by voice in the meeting Without anyone knowing. So it is urgent to study an effective and practical technology against covert voice recording. At present, common anti-eavesdropping methods for recording devices include [9]: audio jamming, playing some background noise to override the voice signal, which is obviously restricted by environmental conditions and will bring great disturbance and inconvenience to people. Electromagnetic wave jamming, by emitting high-power electromagnetic wave and lead to the digital equipment to generate interference signal, in order to destroy its recording function, but now the digital recording device is mostly designed with the anti-interference circuit, it is difficult to use a corresponding electromagnetic signal to interfere internal electrical circuits, the jamming distance is very short, so the effect is limited.

We propose a scheme based on the acoustic modulation broadband parametric array model. This theory points out that if the broadband signal is modulated to high-frequency sound waves and enters the nonlinear system, it will be demodulated and generate low-frequency noise, which can jam the normal voice input into the phone. The ultrasonic frequency we use is more than 30 kHz, while the audible frequency range of human ear is less than 20 kHz. Because it is the ultrasound which is transmitted in the air, the process of interference is inaudible; while the left the audible noise in the recording.

Based on the parametric array theory, this paper presents UltraArray, a silent ultrasonic jammer against covert voice recording, then tests and evaluates our equipment. The main contributions of this paper include: 1) Design a scheme of ultrasonic jamming, which solves the problem that the existing sound recording jamming scheme cannot have good jamming effect and concealment at the same time. Firstly, a set of random number sequence is generated, which is used as the time function of the modulation signal in Double sideband (DSB) amplitude modulation. Then, we use DSB modulation to generate the ultrasonic signal for transmission. Finally, the power of the generated ultrasonic signal is amplified and transmitted to the target device by a parametric loudspeaker array. 2) Based on the proposed recording jamming scheme, the ultrasonic silent jammer was designed and had an experiment. Some smartphones from Apple, Huawei and other brands were used for testing. Our ultrasonic silent jammer was evaluated in the experiment.

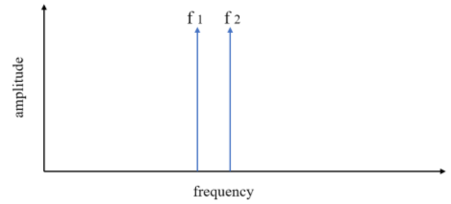
II. Background

A. Westervelt parametric array theory

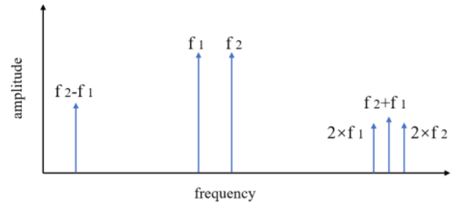
The sound wave does not change to nonlinearly during linear transmission. However, when the sound wave travels nonlinearly or enters the nonlinear medium, the sound wave will produce harmonics, sum-frequency waves and difference-frequency waves. The superposition principle of waves is no longer valid. At this time, the principle of nonlinear acoustics is needed to be studied.

The theory of acoustic parametric array belongs to the category of nonlinear acoustics. A basic acoustic parametric array is a device that uses two high-frequency initial waves propagating in the same direction to obtain the directional propagation ability in the far field. It is based on the nonlinear characteristics of sound waves propagating in the medium. Acoustic parametric array has following advantages: Firstly, parametric array works with short wavelength, which can get high directivity sound beam; Secondly, at low frequencies, a relatively wider acoustic bandwidth can be obtained. Thirdly, there is almost no side lobe, which can improve the precision of the equipment. Fourthly, with good penetrability, it can be widely used in underwater detection and audio directional technology.

In late December 1962, Westervelt published a famous paper named "Parametric Acoustic Array" in the journal of the acoustical society of America [10], which first proposed the concept of " acoustic parametric array theory". In 1965, Berktay published a paper about acoustic parametric array, he explained the theory of acoustic parametric array more clearly [11], and proposed the method of modulating input



(a) The frequency of the original ultrasonic signal emitted



(b) The signal frequency after nonlinear transformation

Figure 1. Effect of acoustic parametric array theory

audio, which provided a foundation for the practical application of parametric array. In 1975, Blackstock and Mary Beth Bennett used an oil-filled hydrophone experiment to prove that acoustic parametric array could work in air for the first time [12]. Their work laying a foundation for the subsequent development of acoustic parametric loudspeaker array system.

The simplest Westervelt acoustic parametric array transmits two ultrasonic waves of different frequencies through a small-sized loudspeaker or a loudspeaker array. These two strong sound waves of different frequencies have nonlinear effects in the interactive region of the sound field, producing low-frequency difference frequency waves and some high-frequency signals, as shown in Figure 1. Note that Figure 1(b) shows the result of two high frequency ultrasonic sine signal after nonlinear transformation. There will be a signal with a low-frequency single frequency of f_2-f_1. It can be used to intuitively understand why the ultrasonic signal can produce low-frequency vocalizations component to interfere with the sound recording.

However, a single frequency signal cannot completely override the voice information in the recording. The actual scheme adopted in this paper is an improvement on the Westervelt parametric array. We propose an enhanced modulation broadband parametric array model based on KZK equation. It is more complex and suitable to the reality. We innovatively modulate the noise signal to the ultrasonic frequency. Those original signal can eventually form a low-frequency broadband signal covering the entire vocal band. The interference effect of our scheme far exceeds the single frequency signal.

B. KZK equation and its important corollary

"Modulation broadband parametric array model" describes the propagation of modulated ultrasonic signals under the parametric array theory. It can be deduced from an important corollary of KZK equation. The principle is as follows:

$$\frac{\partial^2 p}{\partial z \partial \tau} = \frac{c_0}{2} \nabla_\perp^2 p + \frac{\delta}{2c_0^3} \frac{\partial^3 p}{\partial \tau^3} + \frac{\beta}{2\rho_0 c_0^3} \frac{\partial^2 p}{\partial \tau^2} \tag{1}$$

Eq. 1 known as KZK equation, which is proposed by Khokhlov, Zabolotskaya and Kuznetsov. It can describe the phenomena such as diffraction, nonlinearity and absorption during sound propagation. Where p is the sound pressure, z is the coordinate of the positive axis of the beam, τ is the time delay, c_0 is the sound velocity of sound waves with small amplitude, $\nabla_\perp^2$ is the Laplace operator, δ is the acoustic diffusivity, β is the nonlinear coefficient, and ρ_0 is the static medium density.

The most direct solution to KZK equation is to use the method of quasilinear equation. The KZK equation can be reduced to the sum of the following two terms: $p = p_l + p_{nl}$, where p_l is the result of a linear field and p_{nl} is the result of nonlinear field. p_{nl} signal will produce difference frequency signal wave, which is the focus of this paper. For p_{nl} on the positive axis, we can finally get:

$$p_{nl}(x,R,\tau) = \frac{\beta}{2\rho_0 c_0^4}\frac{\partial^2}{\partial\tau^2}\int_0^x\int_0^\infty p_1^2\left(x',r',\tau - \frac{r'^2}{2c_0(x-x')}\right)\frac{r'dr'dx'}{x-x'} \quad (2)$$

Eq. 2 is an important corollary of the KZK equation and can be used to derive many other important conclusions, such as the modulation broadband parametric array model and the modulation far-field solution used in this paper. Modulation far-field solution is the most important conclusion obtained by the modulation broadband model. It describes the relationship between the original signal and the low-frequency signal finally generated by the modulated ultrasonic signal. The specific ultrasonic signal generation will be introduced in the next section.

III. IMPLEMENTATION

Our equipment realizes the covert transmission of ultrasonic signal to the mobile phone microphone. With the nonlinear transformation, the low-frequency noise jamming signal will be generated according to the parametric array theory and the modulation far-field solution theory. Finally, low-frequency noise overrides the human voice message in the recording to protect the speaker's speech from being recorded by the recording device.

A. *The system structure*

The system structure of our equipment is shown in Figure 2. Our system mainly consists of three modules: random amplitude-signal generation module; ultrasonic signal generation module and parametric loudspeaker array module. Their specifications are as follows:

- Random amplitude-signal generation module, its main work is the generation and encryption of random numbers. First, a random number generation algorithm is used to generate a set of random numbers. Then, It is encrypted with the AES algorithm to enhance its randomness. Finally, random low-frequency amplitude signal is generated, which is used for DSB amplitude modulation.

- Ultrasonic signal generation module, its main work includes: (a) Generation and storage of ultrasonic carrier: high-frequency ultrasonic wave is generated by the external sinusoidal wave generation module, they will be used as the carrier. The on-off of the carrier is controlled by the switching sub-module. (b) Modulation of low-frequency amplitude signal: with ultrasonic wave of different frequencies as the carrier, DSB modulation method is adopted. ADC signal sampling, DAC signal generation modules and single-chip microcomputer are used to realize DSB amplitude modulation of low-frequency amplitude signal, as well as multiplex generation and output of ultrasonic jamming signal.

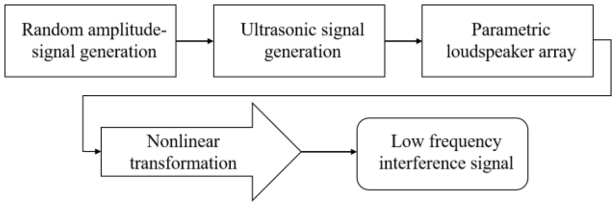


Figure 2. Ultrasonic anti-recording system structure

- Parametric loudspeaker array module, its main work includes: power amplifying the ultrasonic jamming signal, then transmitting it to the target recording equipment through the parametric loudspeaker array.

After the processing of the above three modules, the ultrasonic signal emitted by the parametric loudspeaker array will enter the air for transmission. According to the acoustic parametric array theory, the ultrasonic signal will occur nonlinear transformation and self-demodulate, which will generate a low-frequency noise signal. This low-frequency noise signal will be recorded by the microphone of the phone, and override the human voice signals, so as to interfere with the recording function.

B. *Random amplitude-signal generation*

The sequence of random numbers generated by this module will be used for DSB modulation as part of the amplitude function, which will be modulated with the ultrasonic signal. The process is divided into two steps: 1) With ANSI X9.82 standard, generating an initial 128bit random number. 2) The AES encryption algorithm is used to encrypt it to further improve the randomness of the number sequence. As the initial random number is only generated by single function iteration, the randomness is still limited. When the random number is encrypted in a way similar to text encryption, the original random number will be changed into a set of numbers with more complex calculations and very strong randomness.

C. *Ultrasonic signal generation*

The ultrasonic recording jamming scheme adopted in this paper uses DSB modulation to modulate the low-frequency audible noise signal to the ultrasonic frequency band and transmit it through the parametric loudspeaker array. The purpose of our scheme is to produce low-frequency sound signals to interfere with the recording function of recording devices such as smartphones. It is an application of the modulation broadband parametric array model in acoustic parametric array theory. The modulation far field solution can be derived from this model and explain the relationship between the resulting low-frequency signal and the original ultrasonic signal.

Assuming that the ultrasonic signal emitted by the loudspeaker has been amplitude modulated and the carrier wave is a sine wave, the equation of the signal emitted by the initial sound source is:

$$p_{out}(0,R,t) = P_0 E(t)\sin(\omega_0 t) \quad (3)$$

$\sin(\omega_0 t)$ is the carrier wave with ultrasonic frequency, $E(t)$ is a broadband signal in the low frequency band. p_{out} is the result of Amplitude modulation. By substituting Eq. 3

equation into Eq. 2, the equation of self-demodulation sound wave can be obtained:

$$p_B = \frac{\beta p_0^2 a^2}{16\rho_0 \alpha c_0^4 x} \frac{\partial^2}{\partial \tau^2} E^2(\tau) \quad (4)$$

Eq. 4 is the modulation far-field solution, and p_B is a low-frequency broadband signal, Eq. 4 illustrates an important fact: after nonlinear transformation, the ultrasonic signal emitted by the parametric loudspeaker array can self-demodulate to the low-frequency acoustic signal, which can be recorded by recording equipment such as the microphone of smartphones. According to Eq. 4, the sound pressure relationship between the modulation envelope function and the low-frequency sound wave generated by self-demodulation can be obtained, as shown in the Eq. 5.

$$p_B \propto \frac{\partial^2}{\partial \tau^2} E^2(\tau) \quad (5)$$

Next, we will introduce how the ultrasonic signal in our scheme is generated and analyze how it interferes with the recording function of the mobile phone. The whole ultrasonic signal generation process is shown as follows:

In the previous step, a sequence of random numbers is generated, which is denoted as $g(t)$. As the time function of the modulated signal in DSB amplitude modulation, $g(t)$is DSB modulated with the ultrasonic signal generated by the ultrasonic wave generation module, and the expression of the amplitude modulated wave is Eq. 6.

$$s(t) = [A_0 + kg(t)] \cos(\omega_c t + \varphi_0) \quad (6)$$

A_0 is the amplitude of the incremental carrier; k is the coefficient representing the proportional relationship between the signal strength and amplitude increment; ω_c is the ultrasonic carrier frequency; φ_0 is the initial phase of the carrier; both ω_c and φ_0 are constants.

Since the modulated signal $g(t)$ is a complex signal, it can be expressed by Fourier series. $g(t)$ can be expressed as Eq. 7:

$$g(t) = \Sigma G_n \cos(\omega_n t + \varphi_n) \quad (7)$$

G_n, ω_n and φ_n are respectively the amplitude, frequency and initial phase of the nth-degree harmonic in the modulated signal, if kG_n is denoted as ΔA and $m_n = \frac{\Delta A_n}{A_0}$, $s(t)$ can be expressed as Eq. 8:

$$s(t) = A_0[1 + \Sigma m_n \cos(\omega_n t + \varphi_n)] \cos(\omega_c t + \varphi_0) \quad (8)$$

m_n is the partial amplitude modulation coefficient for the nth-degree harmonic of the modulating signal. In order to facilitate analysis and discussion, the modulation of a simple tone is first discussed, and then extended to complex waves. For a simple tone, Its modulation envelope function $E(t)$ is given:

$$E(t) = [1 + m\cos(\omega t + \phi_0)] \quad (9)$$

ω is the frequency of the simple tone and ϕ_0 is the initial phase. Applying $E^2(t)$ to the modulation far-field solution Eq. 5. Eq. 10 will be derived:

$$p_B \propto 2m^2\omega^2 \cos[2(\omega t + \phi_0)] + 2m\omega^2 \cos(\omega t + \phi_0) \quad (10)$$

Eq. 10 shows that the first term on the right is the harmonic with a frequency of 2ω, while the second term on the right is the low-frequency acoustic signal function, which is exactly the noise we want to be recorded by the microphone of the recording equipment. For the complex waves used in our actual scheme, each component of the modulated signal with a frequency of ω_n will generate a low-frequency acoustic signal with a frequency of ω_n. Therefore, the noise signal record by the microphone will be the superposition of many low-frequency noise components. The waveform is complex and the interference of the recording is significant and effective.

D. *Parametric loudspeaker array*

In this module, the modulated ultrasonic signal is amplified by a power amplifier and then transmitted to the target device through a parametric loudspeaker array. The jammer designed in this paper use 12 ultrasonic loudspeakers to form a loudspeaker array. Each loudspeaker transmits ultrasonic signals with different carrier frequencies. Ultrasonic signals' frequencies range from 30kHz to 40kHz, which has the best effect [13]. Ultrasonic waves from different speakers are superimposed in the air. Finally, the loudspeakers array can form a beam to spread the signal further. In addition, the distance between the speakers will affect the beam formed. Thus, the distance is also designed to form a better beam, which has a higher sound pressure level at a long distance. The ultrasonic beam can generate complex and high-power noise signals in low frequency band after self-demodulation, and can cover human voice well in recording equipment.

IV. EXPERIMENT AND EVALUATION

Based on the acoustic parametric array theory, we design an ultrasonic anti-recording scheme and manufactures UltraArray, an ultrasonic silent jammer against covert voice recording. Inside it, we integrated all three modules mentioned in chapter 2. The appearance of our equipment is shown in Figure 3. The size of UltraArray is 145×150×55mm, which is easy to carry. The device is battery-powered, with a maximum power consumption of 15W, comply with the requirements of the acoustic safety code. We have carried out experiments on our equipment. The experiment scheme, results and evaluation are shown below.

A. *experiment scheme*

The experiment mainly needs to study three aspects: 1) testing the influence of distance, sound source volume, equipment's direction and other factors on acoustic jamming effect; 2) detecting the jamming effect of parameter loudspeaker array on different types of mobile phones; 3) estimating the maximum distance that different mobile phones' recording function can be jammed. To solve these three problems, the following experiment scheme is designed: Place the device and the dialogue audio source in the same place, facing the same direction, the relative position of them is unchanged, the sound play direction points to the recording device. In the experiment, the distance between the recording

Figure 3. UltraArray: silent ultrasonic jammer prototype

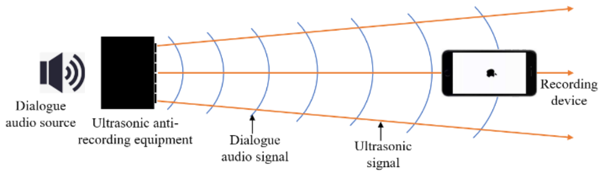


Figure 4. Schematic diagram of the experimental scheme

device and our equipment was changed. The dialogue audio source is played by a PC. The recording devices tested in this experiment are smartphones of common brands such as Apple and Huawei. The experiment was conducted in an empty room with a relatively wide environment and low background noise. The schematic diagram of the experimental scheme is shown in Figure 4.

In the experiment, the PC playing dialogue audio was fixed at a distance of 30cm from our ultrasonic anti-recording equipment. As a variable, the distance between the recording smartphone and our equipment increases from 0.5m to the weak jamming effect, that is, the voice information in the original dialogue can be distinguished from the recorded audio. In the following section, experimental results on several typical distances will be selected for comparative analysis. The specific experimental process is described below.

For a distance, first only play the original dialogue audio, using a decibel meter to measure the average sound intensity, at the same time using a smartphone recording as a reference; Then turn on the ultrasonic anti-recording equipment, play the ultrasonic sound for recording interference, also record the decibel meter number, at the same time using a smartphone recording. The experiment obtained the decibel of the recording before and after the interference, two audio segments. In addition, the audio analysis software on the phone is used to analyze the real-time audio recording of the phone, and two spectrum analysis charts are obtained before and after the interference. The experiment recruited 20 volunteers to listen to the resulting recording audio. When more than five volunteers thought they could distinguish the original voice from the recording audio, the interference was judged a failure.

B. *results and evaluation*

The Figure 5 shows the comparison of noise signal intensity and sound source signal intensity of Apple iPhone 6s Plus at different distances. It can be seen from Figure 5 that after the use of ultrasonic anti-recording equipment for recording interference, the signal intensity of the generated noise signal is much stronger than the signal intensity of the dialogue sound source's signal in the effective range, so the human voice signal can be overwritten in the recording of the smartphone.

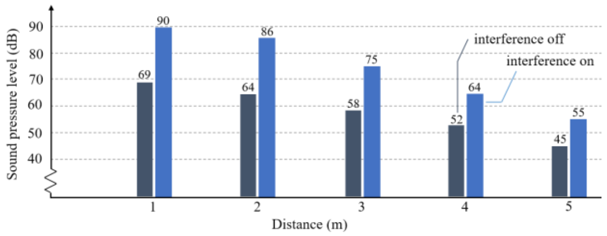
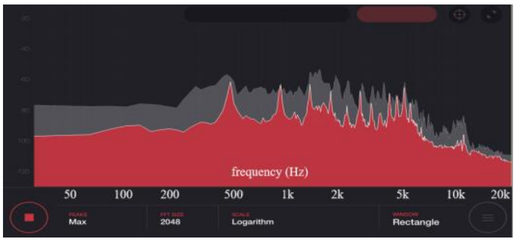
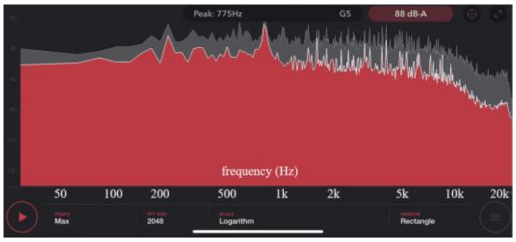


Figure 5. The comparison of noise signal intensity and sound source signal intensity in iPhone 6s Plus

Then, analyzing the spectrum difference between the recordings of smartphones when there is interference or not. The real-time audio spectrum analysis diagram of audio analysis software in iPhone XR is shown in the Figure 6.



(a) The spectrum of the original dialogue audio

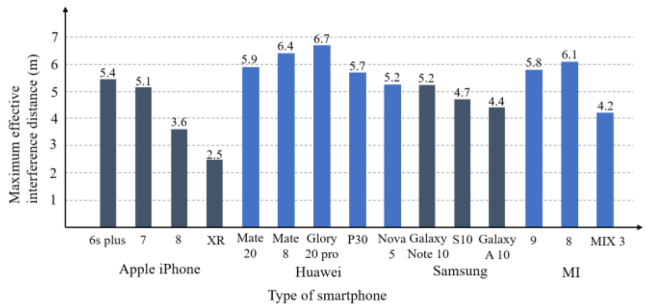


(b) The spectrum of acoustical signal with interference

Figure 6. The spectrum of recorded signal in iPhone XR

The frequency distribution of most human voice signals is within the range of 3.4KHz. As can be seen from Figure 6, the spectrum distribution of noise signals covers human voice signals well in every effective frequency band. It also confirmed that within the interference range of our equipment, the human voice signal could not be distinguished from the recording file.

There were 15 types of smartphones tested in the experiment, which's brands include Apple, Huawei, Samsung and MI. The specific type can be seen in Figure 7. According to Figure 5, as the distance between the recording device and the sound source increases, the sound intensity will decrease. However, the effect of the ultrasound decays more than normally audible sound, so the interference will fail when it exceeds a certain distance. Different types of smartphones also influence experimental results. The experiment also carried out a more detailed test, measured the maximum effective interference distance of different types of smartphones. It should be noted that there are many factors that affect the limit distance, including the sound clarity of the dialogue audio itself, the sound intensity of the dialogue audio source, the volunteers who test and judge the recording data file, and the ambient noise, etc. Therefore, the maximum effective interference distance will be different in different experimental environments. In addition, when experiments



are carried out in close distance, the audio results produced by

Figure 7. The maximum distance of different types of smartphones

the recording are very similar, so the test of the limit distance moves the minimum of 10cm each time. The results of our experiment can only be used as one of the evaluation factors to compare the interference effect of different smartphones. The maximum distance of different types of smartphones is shown in Figure 7.

As can be seen in Figure 7, the maximum effective interference distance of our equipment to most smartphones is more than 5 meters. Huawei Glory 20 Pro has the best interference distance which is 6.7 meters; and iPhone XR has the worst interference effect with a distance of 2.5 meters. it should be paid attention that those results of our equipment far exceeds that of other existing solutions. The experimental results shown in Figure 7 prove the effectiveness of our equipment, which can interfere with the recording function of the current mainstream brands of smartphones, and this effectiveness will not disappear with the change of smartphone types. The maximum effective interference distance is different between different smartphones, which may be related to many factors, such as microphone type, audio processing software and hardware.

In addition to the most obvious and critical distance factor, there are many other factors that can affect the jamming effect. We also tested three of them: the direction of the phone's microphone, the range of the angle of the ultrasonic jamming signal, and the sound intensity of the dialogue audio source.

a) Direction of the phone's microphone placement: after testing, it is found that when the microphone is completely turned away from our equipment, the interference effect will become worse. This is because the directivity of the ultrasonic wave is very strong, and the microphone turned away from our equipment will naturally receive less ultrasonic signal. on the other hand, ultrasound decays quickly in the air, and the ultrasonic signal bounces off the walls to the microphone less than normal dialogue audio. These reasons together influenced the results of the experiment.

b) Ultrasonic jamming signal’s angle range: the ultrasonic signal has a strong directivity. If the recording device was placed outside the angle range, the interference effect will become worse. This limitation, however, can be addressed by the placement of ultrasonic anti-recording equipment. Multiple ultrasonic anti-recording equipments can be placed toward where those recording devices are typically placed, such as on and below the table in a conference room.

c) Sound intensity of the dialogue audio source: when the signal intensity of the dialogue audio source is higher, the maximum interference distance for the smartphone recording will be shorter, which is obvious, because the larger signal intensity of human voice will be more easily heard from the recording file under the same noise intensity. It should be noted that the sound intensity of the main dialogue audio source tested in this experiment reaches 67dB at 1 meter, which is much higher than the sound signal intensity of general dialogue. Therefore, the experimental results in this paper are convincing.

V. CONCLUSION

This paper proposed, engineered and validated an ultrasonic anti-eavesdropping device, which is composed of random amplitude-signal generation module, ultrasonic signal generation module and parametric loudspeaker array module. Firstly, the random number generation algorithm and AES algorithm are used to generate a set of random number sequences, which are used as the modulation signal. Then, it will be DSB modulated with ultrasonic carrier. Finally, the power of the modulated signal is amplified and transmitted to the target recording device through the parametric loudspeaker array.

Our equipment is based on acoustic parametric array theory, which takes advantage of the fact that ultrasonic signal emitted by parametric loudspeaker arrays can generate low-frequency noise in the recording devices such as smartphones. The human voice in the recording is overwritten so that it cannot be distinguished. Compared to the current anti-recording scheme, our equipment has a huge advantage on the concealment and performance: for the common smartphone brands, our equipment's maximum effective interference distance is more than 5 meters on most types. Our equipment is small in size and battery-powered, so it is portable and convenient to set. Because ultrasound is inaudible, our device can quietly interfere the recording without blocking the process of conversation. We verified these advantages by experiments.

ACKNOWLEDGMENT

This work is supported in part by National Key R&D Program of China under Grant Nos. 2018YFB2202200, 2018YFB2100403. This research is partially supported by HK PolyU ZVRG P0033725.

REFERENCES

[1] Chung, Hyunji and Sangjin Lee, “Intelligent Virtual Assistant knows Your Life.” *ArXiv* abs/1803.00466, 2018.

[2] L. Malina, J. Hajny, and Z. Martinasek, "Privacy-preserving authentication systems using smart devices," in *2016 39th International Conference on Telecommunications and Signal Processing (TSP)*, 2016, pp. 11-14.

[3] X. Dai, W. Lou, P. Liu *et al.*, "Speaker tracking based on microphone cross arry in the smart conference system," in *2014 IEEE International Conference on Consumer Electronics - China*, 2014, pp. 1-4.

[4] G. Zhang, C. Yan, X. Ji *et al.*, “DolphinAttack: Inaudible Voice Commands,” in *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, New York, NY, USA, 2017, pp. 103–117.

[5] Song, Liwei , and P. Mittal., "POSTER: Inaudible Voice Commands." *the 2017 ACM SIGSAC Conference* ACM, 2017.

[6] C. Yan, G. Zhang, X. Ji *et al.*, "The Feasibility of Injecting Inaudible Voice Commands to Voice Assistants," *IEEE Transactions on Dependable and Secure Computing*, pp. 1-1, 2019.

[7] R. Iijima, "Audio Hotspot Attack: An Attack on Voice Assistance Systems Using Directional Sound Beams and its Feasibility," *IEEE Transactions on Emerging Topics in Computing*, pp. 1-1, 2019.

[8] Y. Chen, H. Li, S. Teng *et al*., Wearable Microphone Jamming, 2020.

[9] B. Thakker, A. Mathew, and D. Doriwala, "Design and Implementation of Voice Recording System Prototype," in *2018 IEEE Punecon*, 2018, pp. 1-5.

[10] Westervelt P J, "Parametric Acoustic Array, " *Journal of the Acoustical Society of America*, 1963, 35(4):535-537.

[11] Berktay H O., "Possible exploitation of non-linear acoustics in underwater transmitting applications, " *Journal of Sound & Vibration*, 1965, 2(4):435-461.

[12] Bennett M B, Blackstock D T., "Parametric array in air, " *Journal of the Acoustical Society of America*, 1998, 57(3):562-568.

[13] D. Xu-dan, "Simulation Study on Directional Dispersion of Strong Noise," in *2019 International Conference on Modeling, Analysis, Simulation Technologies and Applications (MASTA 2019)*, 2019.