

Inducing Wireless Chargers to Voice Out for Inaudible Command Attacks

Donghui Dai^{*†}, Zhenlin An^{*†} and Lei Yang[†]

^{*}Co-first Authors

[†]Department of Computing, The Hong Kong Polytechnic University
{dai,an,young}@tagsys.org

Abstract—Recent works demonstrated that speech recognition systems or voice assistants can be manipulated by malicious voice commands, which are injected through various inaudible media, such as ultrasound, laser, and electromagnetic interference (EMI). In this work, we explore a new kind of inaudible voice attack through the magnetic interference induced by a wireless charger. Essentially, we show that the microphone components of smart devices suffer from severe magnetic interference when they are enjoying wireless charging, due to the absence of effective protection against the EMI at low frequencies (100 kHz or below). By taking advantage of this vulnerability, we design two inaudible voice attacks, HeartwormAttack and ParasiteAttack, both of which aim to inject malicious voice commands into smart devices being wirelessly charged. They make use of a compromised wireless charger or accessory equipment (called parasite) to inject the voice, respectively. We conduct extensive experiments with 17 victim devices (iPhone, Huawei, Samsung, etc.) and 6 types of voice assistants (Siri, Google STT, Bixby, etc.). Evaluation results demonstrate the feasibility of two proposed attacks with commercial charging settings.

I. INTRODUCTION

Voice assistants have become an increasingly popular human-computer interaction approach in smart devices (e.g., smartphones or wearables) with the recent incredible advances achieved in the field of speech recognition. For example, Apple Siri [1] and Google Now [2] allow users to initiate phone calls and launch apps through their voices; Alexa [3] even allows users to instruct an Amazon Echo to control their entire smart home. With the spread of voice assistants, a built-in microphone (as a compulsory component for a smart device) has become a new vulnerability under sneaky and malicious *inaudible voice attacks*. In these attacks, inaudible voice commands, which are unintelligible and unnoticeable to human listeners, can take control of the victim devices [4].

The known voice command attacks can be initiated via different types of inaudible media, such as the ultrasound [4]–[9], laser [10], and EMI [11]–[13]. Particularly, a large number of works attacked computer systems through the EMI [14]–[16] in the last decades. Recently, these potential EMI were reused to initiate inaudible voice attacks on smartphones through the external wireless circuitry [11], headphone cables [12] and power lines [13]. As an important countermeasure against the potential EMI [17], the industry equips today’s microphones with Faraday cages (a kind of EM shield) and EMI filters, especially against the 3G/4G signals operating at 800-900 MHz [10]–[12]. Unfortunately, a recent study shows that



Fig. 1: Illustration of wireless chargers in public. With the fast spread of wireless charging technology, wireless chargers are becoming public facilities everywhere.

magnetic fields can still penetrate Faraday cases due to the immunity of magnetic fields to electromagnetic shields [18], which is also confirmed by our preliminary tests.

By taking advantage of the above vulnerability, we explore a new type of inaudible voice attack through the wireless chargers, which produce the well-modulated magnetic interference to inject the voice commands into the microphones as if they were recorded from a physical sound. Wireless charging delivers power from an energy supply to smart devices without contact. Wireless charging also encourages the production of completely sealed or even waterproof device casing, which substantially improves convenience, usability, and reliability. Thus, wireless charging is becoming a de facto power supply solution for a vast number of smart devices, especially for wearables (such as Apple Watch or AirPods). Fig. 1 shows some typical public wireless charging stations, where numerous free wireless chargers are deployed in public everywhere and hundreds of millions of people benefit from them every day. Nevertheless, these public wireless chargers are becoming potential security breaches.

Achieving magnetic-inductive sound (MIS) at microphones is very challenging because of a fundamental communication issue, that is, there exists an about 80 kHz frequency gap between microphones and chargers. Specifically, a microphone can only record the voice below 22 kHz, where higher frequencies will be completely filtered out, whereas a wireless charger produces magnetic fields at 100 kHz to 200 kHz. To address this issue, we propose two attacking approaches, HeartwormAttack and the ParasiteAttack, as shown in Fig. 2.

- **HeartwormAttack:** We envision that an adversary can install the malware called *heartworm* into a wireless charger during the manufacturing phase. The compromised charger can opportunistically trigger a victim device to execute an expected command through the MIS, as shown in Fig. 2(a). Such an attack uses the nonlinearity of the amplifier in a microphone to downconvert the MIS from charging frequency into an audible spectrum.

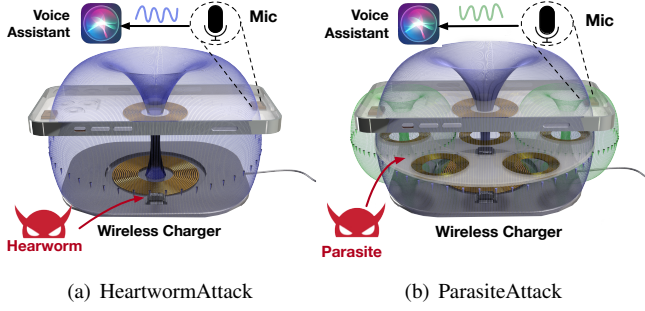


Fig. 2: Attack scenarios. The blue curves represent the magnetic field generated by the TX coil of the wireless charger, while the green curves represent the fields generated by the TX coils of the parasite.

- **ParasiteAttack:** We envision that an adversary attaches small and thin accessory equipment called *parasite* onto a public wireless charger, as shown in Fig. 2(b). As a Near-field Communication (NFC) card, the parasite uses a receiving (RX) coil to “steal” power from the host charger and drives one of the transmitting (TX) coils to directly generate the magnetic field at the voice frequency, which further produces MIS at microphones.

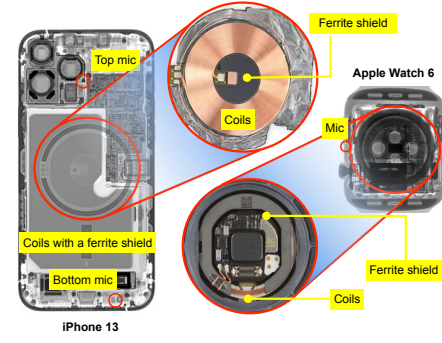
Both attacking approaches have pros and cons. First, the HeartwormAttack must intrusively hack the wireless chargers in advance, whereas the ParasiteAttack can be launched anytime after parasites are deployed. Second, commercial wireless chargers are usually not equipped with wireless communication modules (e.g., Wi-Fi and Bluetooth). Thus, the HeartwormAttack can only work offline using pre-stored voice commands. In contrast, the ParasiteAttack allows the adversary to inject an on-demand voice in real-time through the 4G/5G functionality of the parasite equipment.

We have tested the two attacks on 17 device models including smartphones, smartwatches, tablets, and add-on microphones, which involve 6 voice controllable systems or speech recognition systems. Each attack is successful on at least one SR system. The attacking demos can be found at [19]. We believe this list is by far not comprehensive. Nevertheless, this study serves as a wake-up call to consider the security breach caused by the magnetic interference and reconsider what functionality shall be introduced in voice assistant systems.

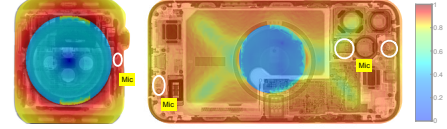
Totally, we made the contributions as follows. First, we discover the potential security threat of magnetic interference to most audio-capable devices and demonstrate such a security threat by using wireless chargers. Second, we show that adversaries can inject a sequence of inaudible voice commands into microphones through two approaches that are, HeartwormAttack and ParasiteAttack. Both two attacks are validated on 17 popular smart devices and 6 common speech recognition systems. Third, we suggest both hardware-based and software-based countermeasures to alleviate the attacks. We also raise a practical concern on the negative consequence resulting from the excessive demand for faster wireless charging.

II. FEASIBILITY ON MAGNETIC-INDUCTIVE SOUND

Our core idea is to skip the microphone diaphragm, but *directly induce* an acoustic signal at the onboard circuits of



(a) X-ray imaging of an iPhone 13 and Apple Watch 6



(b) Distributions of magnetic field over two devices
Fig. 3: Magnetic interference over two smart devices

a microphone by using a manipulated magnetic field. In this section, we conduct the feasibility analysis and verification experiments to answer three key questions: (1) Can the microphone receive magnetic interference when the device is being wirelessly charged? (2) Can the leaked magnetic field violate the current EMI protections? (3) Which part of a microphone is interfered?

A. Magnetic Interference in Smart Devices

In the field of electronic engineering, EMI is a phenomenon that may occur when an electronic device is exposed to an electromagnetic field. The magnetic field generated by a wireless charger will induce an eddy current in the circuits of a nearby device and cause magnetic interference (i.e., a type of EMI). To address such potential hazards, existing smart devices usually use a ferrite shield to protect the motherboard from the EMI. Fig. 3(a) shows the X-ray imaging of an iPhone 13 and Apple Watch 6 from the backside. The RX coils are arranged in the back center of device bodies. A circular ferrite shield is inserted between the RX coil and the motherboard to protect the internal circuits. However, the ferrite shield only covers the area right behind the RX coil, where the magnetic field is at the strongest. The remaining areas beyond the coverage of the shield still receive magnetic interference from the TX coil of the wireless charger.

To quantify the magnetic interference, we compute the distributions of the magnetic field over the whole motherboards of the two devices by using Ansys Maxwell [20]. In the simulation, we use the shell models from [21], [22] and manually model the internal circuitry and main components including the RX coil with a ferrite sheet. The exciting source is the Ansys built-in standard A2 coil model [23]. On the basis of the Biot-Savart Law, the magnetic field strength $\mathbf{B}(r)$ at the distance r from the center of the transmitting coils, is computed as follows:

$$\mathbf{B}(r) = \frac{\mu_0}{4\pi} I \int_C \frac{d\mathbf{l} \times \mathbf{r}}{|\mathbf{r}|^3} \quad (1)$$

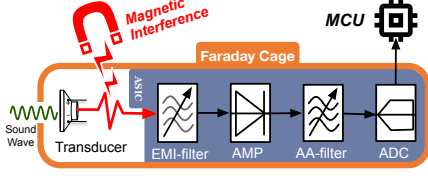


Fig. 4: Architecture of a MEMS microphone

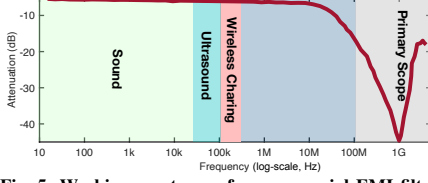


Fig. 5: Working spectrum of a commercial EMI filter

where μ_0 is the vacuum permeability, I is the charging current, C is the current's flow path in the coil, and $d\ell$ is a vector along the path. We adopt the recommended charging settings specified in the Qi standard [23], which is the most widely adopted wireless power transfer (WPT) protocol on the market. Specifically, the charging frequency and the power are set to 100 kHz and 15 W, respectively. The A2 TX coil [23] with 20 mm inner diameter and 40 mm outer diameter is utilized for the charger. The Qi Example 4 RX coil with 28 mm diameter and 47 mm outer diameter is used for the receiver. The magnetic shield made of Mn-Zn ferrite is 1 mm thick. The TX and RX coils are spaced by 1 cm.

Fig. 3(b) shows the distribution results. The magnetic strength behind the ferrite shield is mostly reduced to zero. However, the remaining areas, especially the marginal areas where microphones are located, are fully exposed to the strong magnetic field. This phenomenon is caused by the functional principle of ferrite shield, i.e., ferrite materials cannot weaken magnetic fields but distract the field lines from them to the nearby areas [24] as observed in the figure. As a result, the magnetic strength is actually enhanced in the areas beyond the shield. Therefore, the ferrite shield cannot protect the microphone from magnetic interference. Worse, the industry is quite “aggressive” in raising the charging power to pursue faster charging. To date, 50 W and even 80 W chargers are found on the market [25], which further intensify the magnetic interference to other components like microphones.

B. Magnetic Interference to Microphones

A microphone is a component that can convert sound waves into electrical signals. There are two types of microphones, electret condenser microphone (ECMs) and Micro electro mechanical system (MEMS), available on the market. Due to the miniature package size and lower power consumption, MEMS microphones dominate smart devices. Thus, this paper focuses mainly on MEMS microphones. Nevertheless, MEMS and ECMs work similarly. Fig. 4 shows the internal structure of a digital MEMS microphone, which consists of two main components, an acoustic transducer, and an ASIC chip. When a sound wave presents, the air pressure triggers the mechanical vibrations of the diaphragm and further causes a capacitive change of a connected capacitor. In this way, air pressure is

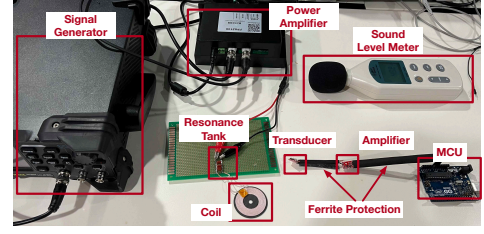


Fig. 6: Experimental setup for feasibility verification

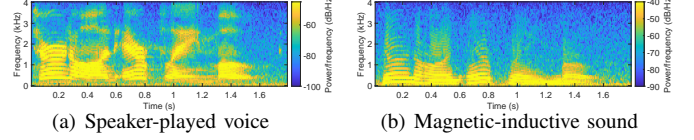


Fig. 7: Spectrograms of the voice signals. (a) shows the original voice signal; (b) and (c) show the voice recorded by iPhone 8 but injected by a speaker and a TX coil, respectively.

converted into an analog acoustic signal for further processing. The acoustic signal is then amplified, filtered, and digitalized by the following ASIC chip.¹ Finally, the acoustic signals are transmitted to other components like MCU. The industry has made efforts to protect the microphone from the previously reported EMI [10]–[12] as follows:

(1) **Faraday Cage:** The whole microphone component is protected by a Faraday cage from disturbance of EM signals, except a small hole reserved to capture the sound wave from the air. A Faraday cage is formed by a continuous covering of conductive material, such as copper [26]. EM signals outside are prevented from going into the cage due to the skin effect [27], and their energy is mostly dissipated in the form of heat. A Faraday cage can only attenuate EM signals with wavelengths shorter than the skin depth. Mathematically, a Faraday cage acts as a low-pass filter to move out EM signal above frequency f :

$$f \geq \frac{\rho}{\pi\mu\delta^2} \quad (2)$$

where δ denotes the skin depth, ρ denotes the resistivity of the conductor, and μ denotes the permeability of the conductor. In accordance with the datasheet [17], the Faraday cages of MEMS microphones are made of copper (i.e., $\mu = 1.256 \times 10^{-7}$ H/m and $\rho = 1.68 \times 10^{-8} \Omega \cdot m$), and their depths are approximately $2.06 \mu m$. Substituting these settings into Eqn 2, we find out $f \geq 1$ GHz. Therefore, the current Faraday cages can only shield 1 GHz or above EMI caused by common wireless communications, such as FM Radio, Bluetooth, WiFi, Cellular, and GPS. They fail to defend against the magnetic interference at 100 kHz caused by wireless chargers unless $100\times$ thicker Faraday cages than the current are adopted.

(2) **EMI Filter:** As shown in Fig. 4, the analog sound signal is pre-processed by an EMI filter before being amplified. The filter aims to eliminate the potential EMI. We show the working spectrum of an EMIF02-MIC03F2 filter in Fig. 5. This filter is fabricated by STMicroelectronics [28] and dominates the market of MEMS microphones. It can be seen that these EMI filters focus on shielding 100 M or above EMI. As stated

¹MEMS microphones can be further categorized into two types, analog MEMS and digital MEMS, regarding if the output of the microphone is an analog or digital signal. More details refer to Appx. B.

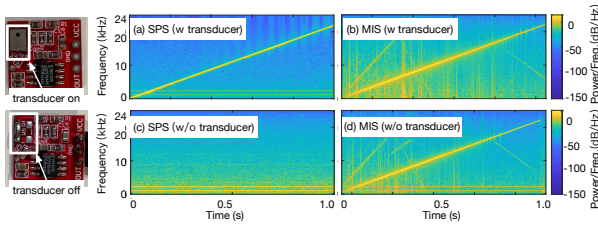


Fig. 8: Spectrograms of chirp signals recorded by a MEMS microphone with and without the transducer component.

in [17] (Sec. 4, pp.15), the filters mainly aim to suppress the EMI from GSM communications (e.g., TDMA noise) at 800 to 900 MHz and 1800 to 1900 MHz, which is far higher than the operating frequency of wireless charging.

In summary, the two existing industrial countermeasures against the EMI reported previously well protect microphones from interference at 100 MHz or above. They fail to eliminate the magnetic interference at a low frequency.

C. Real-life Verification

Given the above theoretical analysis, we verify the feasibility of generating a clear magnetic-inductive sound (MIS) on real microphones by using charging coils. The experimental setup is shown in Fig. 6. We use a vector signal generator to produce a voice signal below 20 kHz. The voice signal is then amplified by a power amplifier to 15 W. The amplified voice signal is then passed through a resonant tank circuit and broadcasted by a standard A2 TX coil into the air. The distance between the microphone and the coil is about 3 cm.

First, we use the signal generator plus the TX coil to transmit a voice clip of “turn on airplane mode”. Meanwhile, we used an iPhone 8 to record the clip. We also play the signal by a loudspeaker directly and use the recorded version as the baseline. As shown in Fig. 7, the spectrograms of two recordings exhibit similar patterns as that of the original voice. We also notice that the components at higher frequencies (e.g., > 5 kHz) are more attenuated than those at lower frequencies in the MIS. Actually, the energy at frequencies above 5 kHz is ignored in speech perception systems because human speech mainly concentrates on the lower frequencies [29]. This experiment fully demonstrates the feasibility of MIS at the microphone. Second, we use an external MEMS microphone (i.e., TDA1308 from Knowles [30]) to record chirp signals with or without the transducer, as shown in Fig. 8. In each case, we play the chirp signal sweeping from 100 Hz to 22 kHz by the loudspeaker and TX coil, respectively. Fig. 8(a) and (b) show the results of speaker-played sound (SPS) and MIS with the transducer. Both methods can generate the chirp signals in the spectrograms. Then, we forcibly remove the transducer from the microphone. In the figure, (c) and (d) show the results without the transducer. Consequently, none SPS is observed in the spectrogram but the MIS still presents. This experiment fully demonstrates that MIS is completely not produced by the acoustic vibrations, which should be captured by the transducer only. Finally, we assemble an external microphone using several separated components to mimic a MEMS microphone. As shown in Fig. 6, the microphone is

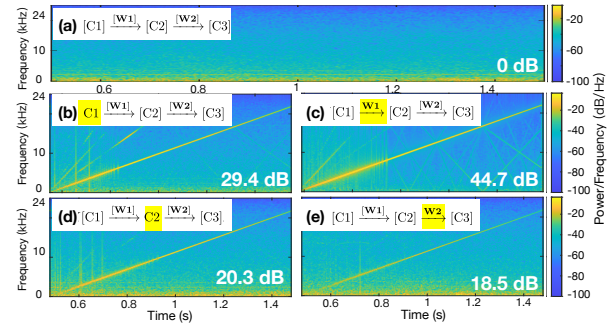


Fig. 9: Spectrograms of chirp signals recorded by an assembled microphone. The microphone consists of three separated components (C1, C2, and C3) and two wires (W1 and W2). (a) shows the absence of the signal when all components are wrapped by ferrite sheets. (b)-(e) show the presence of chirp signals when C1, W1, C2, and W2 are exposed to the interference in turn.

composed of a Knowles SPV1840LR5H-B [31] transducer (C1), a MAX9812 [32] amplifier integrating an internal low-pass filter (C2), and an onboard ADC (C3). These three components are connected through two wires (W1 and W2) as follows:

$$C1 \xrightarrow{W1} C2 \xrightarrow{W2} C3 \quad (3)$$

Now, we repeat to play the chirp signals using the TX coil. At each time, one of the C1, C2, W1, and W2 is exposed to the magnetic interference, and others remain wrapped with ferrite sheets. Fig. 9(a) shows the baseline when everything is wrapped with ferrite sheets, and no signal is recorded. Fig. 9(b)-(e) shows that MISs are always detected when C1, W1, C2, or W2 is exposed to the interference in turn. This demonstrates that magnetic interference affects all circuitry instead of a specific component or a wire. Even a small fragment of wire can still capture the MIS. However, we ensure that the interference only works before the ADC because the injected voice is an analog signal that cannot be recognized by the follow-up digital components. Thus, we should consider the microphone as a whole to defend against magnetic interference.

III. OVERVIEW

The preliminary experiments fully verify the presence of MIS caused by a charging TX coil. These positive results encourage us to conduct further studies on leveraging MIS to inject inaudible voice attacks. In this section, we introduce a general attack model and then present two attack approaches from a high level.

A. Threat and Attack Model

We adopt the similar threat model used in previous inaudible voice attacks, like Dolphinattack [4] and so on [6]–[9], [33]. The goal of the adversary is to inject malicious voice commands into voice assistants equipped on mobile devices, such as Apple Siri [1], Google Assistant [34] and Samsung Bixby [35]. The smart soundboxes like Alex are not considered since they are usually powered by cables. Through these commands, the adversary can execute unauthenticated actions, such as visiting a malicious website to launch a subsequent drive-by-download attack, making fraud calls to the victim’s

friends and family, injecting fake information such as fake instant messages, emails, online posts, and even calendar events, or turning on airplane mode to deny all incoming connections [4].

- **Adversarial Abilities.** We assume the adversary has no direct access permission to or cannot inject malware into the victim’s devices. The settings of these devices cannot be altered. However, the adversary is fully aware of the characteristics of all targeted smart devices. He/she can identify the victim device model and manufacturer model, which can be overheard during the handshake phase when the two sides exchange WPT data. The activation commands (like “Hey Siri”) might be voice fingerprinted. We assume that the adversary can eavesdrop on the victim through a hidden microphone deployed nearby the wireless charger or other similar side channels. The eavesdropped speech might contain the activation command, which can be applied for the activation directly; otherwise, the adversary can use the AI-powered voice synthesis technique to forge the voice-fingerprinted activation commands [36], [37] with the eavesdropped speech.²

- **Attack Conditions.** We assume that adversaries can modify the firmware of a wireless charger or attach accessory equipment nearby to a wireless charger. Our attack is initiated when the victim’s devices are being wirelessly charged using a public or private wireless charger. The chargers might be deployed in a cafe, street, park, or mounted in a car [39]. One goal of the adversary is to attack victim devices without being noticed. The voice commands generated through the magnetic field are apparently inaudible to humans. Correspondingly, the first command is initiated to turn down the volume to the extent that users cannot hear feedback clearly from the voice assistant. Without loss of generality, we assume that the victim devices are placed a few centimeters away from a wireless charger. The majority of voice assistants are allowed to be waked up and conduct many security-sensitive tasks (e.g., making phone calls, reading messages, or turning on Bluetooth [40]) even when their screens are locked. This makes sense because the ultimate goal of the voice assistant systems is to free users’ hands and accomplish major tasks through the voice when smartphones are placed far away and locked by default. We list all security-sensitive and privacy-sensitive tasks that can be accomplished by voice assistants in Appx. F. The voice injection introduces some Gaussian noises unavoidably. The voice assistants are assumed to equip with de-noising algorithms, which can well deal with the Gaussian noise introduced by the background, internal circuitry or our voice injection.

B. Attacking Approaches

The feasibility of generating MIS on microphones has been fully verified in benchmark experiments. However, launching inaudible voice attacks via wireless chargers still remains challenging due to the working frequency gap between the charger and the microphone. The sensitivity spectrum of microphones

²We demonstrate how to synthesize the fingerprinted voice with Vo-cloner [38] and use it to activate the assistant in the demo video.

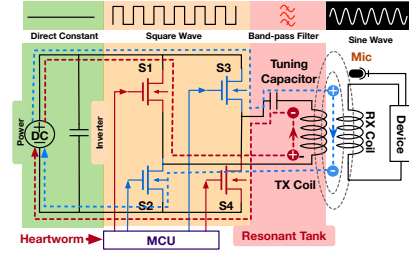


Fig. 10: Schematic diagram of a wireless charger

targets between 20 Hz to 22 kHz and ideal signals beyond this spectrum should be filtered. Thus, an anti-aliasing filter (AA-filter) is adopted after the amplifier, as shown in Fig. 4. The MIS between 100~200 kHz will be removed. To resolve this problem, we propose two attacking approaches on account of the intrusiveness, that is, HeartwormAttack and ParasiteAttack. Specifically, HeartwormAttack utilizes the nonlinearity of the amplifier inside a microphone to downconvert the MIS. On the contrary, ParasiteAttack utilizes the parasite device to harvest energy at a high frequency but generates MIS at the voice band. In the following, we will elaborate on the two attacks.

IV. DESIGN OF HEARTWORMATTACK

In this section, we introduce the HeartwormAttack where malware called *heartworm* is implanted into a public wireless charger in advance. The heartworm takes control of the MCU to inject the voice commands by using the existing hardware components in accordance with the Qi standard, i.e., the de facto standard for wireless charging on the market.

A. Architecture of a Wireless Charger

Fig. 10 shows the schematic diagram of a wireless charger, which contains five main components: a power supply, an inverter, a resonance tank, a TX coil, and an MCU. The charger accepts a 12 voltage direct current (DC) as input. The inverter can convert the DC into an alternating current (AC) signal at some frequencies. Internally, the inverter contains four power field-effect transistor (FET) switches, denoted by S1~S4, which can be turned on or off by the MCU. The inverter is controlled by the MCU to toggle between two states:

- *Positive.* When the switches of S1 and S4 are on, but S2 and S3 are off, the current flow (highlighted in red) passes through the coil from the bottom to the top, leading to an upward magnetic field around the TX coil.
- *Negative.* When the switches of S2 and S3 are on, but S1 and S4 are off, the current flow (highlighted in blue) passes through the coils from top to bottom, leading to a downward magnetic field around the TX coil.

A square wave is created when the MCU controls the inverter to toggle periodically between the two states at a fundamental frequency f , as shown in the figure. The resonant tank acts as a frequency-selective network that only allows the wave at or around the fundamental frequency to pass. It is well known that a square wave can be decomposed into an infinite set of harmonic sinusoidal waves, that is, $\sin(2\pi \cdot ft)$, $\sin(2\pi \cdot 2ft)$, $\sin(2\pi \cdot 3ft)$, \dots . As a result, only the sinusoidal wave at

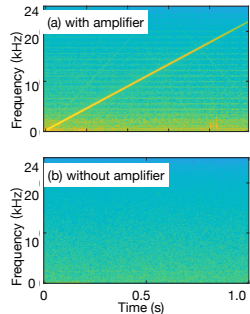


Fig. 11: Nonlinearity effect

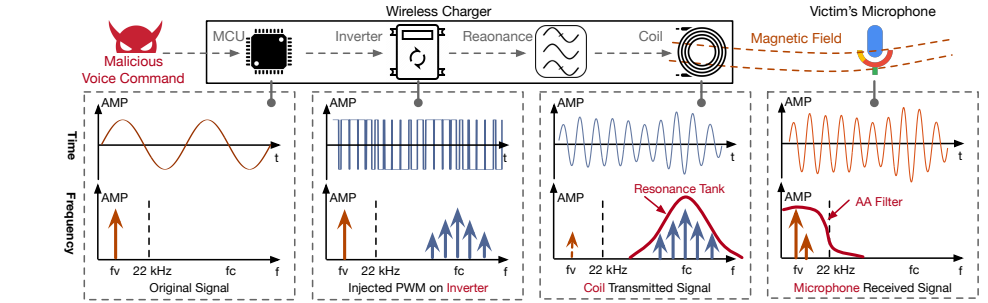


Fig. 12: Workflow of HeartwormAttack. The heartworm manipulates the MCU to generate the predefined PWM-emulated voice commands at the charging frequency (~ 100 KHz) using the inverter, which is further propagated into the air. Finally, the MIS can be received and downconverted at the microphone (< 24 kHz). Finally, the voice commands are executed by the voice assistants. The figure plots the time and frequency domain at different states.

frequency f , i.e., $\sin(2\pi ft)$, can successfully pass through the resonant tank and arrive at the TX coil and trigger an alternative magnetic field at frequency f .

B. Generating Analog Signals using a Digital MCU

A voice is an analog signal full of fine-grained amplitudes, but the MCU can only turn on or off the four switches to produce a digital signal with two amplitude levels (i.e., a positive level or a negative level). To address this challenge, we adopt pulse-width modulation (PWM), which can emulate any analog signal with digital means. Acting as an amplitude-based modulation scheme, PWM generates variable-width pulses to represent the amplitude of an analog signal. PWM is a powerful technique for controlling analog circuits. It has been used in many applications, ranging from communications to power control and conversion. For example, the wireless charger MCU of SWBTC [41], [42] which holds the highest market shares, allows adjusting the duty cycle by 10% to 50%. The analog voice signal is emulated by using a series of pulses with different widths. The pluses are created by toggling the inverter at a fast rate. Let $S_v(t)$, and $S'_v(t)$ denote the real voice signal and the emulated voice signal. In addition, PWM also requires a carrier signal denoted by $S_c(t)$. Their relationship can be formulated as follows:

$$S'_v(t) = \text{PWM}(S_v(t), S_c(t)) \approx S_v(t)S_c(t) \quad (4)$$

where PWM represents the modulation scheme. Mode detailed underlying principle of PMW scheme is introduced in Appx. C. For simplification, we can consider the charger as a wireless transmitter, which can modulate the signal generated via the MCU onto a carrier. Actually, the chargers can use this way to exchange the data with the devices being charged.

C. Using Nonlinearly Effect as a Downconverter

Suppose the microphone receives a combination of two sinusoidal signals at the frequencies of f_1 and f_2 , which is formalized as follows:

$$S_{\text{in}}(t) = \cos(2\pi f_1 t) + \cos(2\pi f_2 t) \quad (5)$$

Amplifiers inside microphones are expected to be linearly proportional to the input S_{in} , but known to exhibit the nonlin-

earity [4], [5]:

$$\begin{aligned} S_{\text{out}}(t) &= \underbrace{As_{\text{in}}(t)}_{\text{Linear}} + \underbrace{Bs_{\text{in}}^2(t)}_{\text{Nonlinear}} \\ &= A(\cos(2\pi f_1 t) + \cos(2\pi f_2 t)) + B(\cos(2\pi f_1 t) + \cos(2\pi f_2 t))^2 \\ &= A\cos(2\pi f_1 t) + A\cos(2\pi f_2 t) \\ &\quad + B + 0.5B\cos(2\pi 2f_1 t) + 0.5B\cos(2\pi 2f_2 t) \\ &\quad + B\cos(2\pi(f_1 + f_2)t) + B\cos(2\pi(f_1 - f_2)t) \end{aligned} \quad (6)$$

where $S_{\text{out}}(t)$ is the output signal from the amplifier, A is the gain for the input signal, and B is the gain for the quadratic term. It can be seen that the above signal has frequency components at f_1 , f_2 , $2f_1$, $2f_2$, $f_1 + f_2$ and $f_1 - f_2$. Before digitizing and recording, the microphone applies a low pass filter (LPF) to remove frequency components above its cutoff frequency 24 kHz. To put this into perspective, when $f_1 = 110$ kHz and $f_2 = 100$ kHz, f_1 , f_2 , $2f_1$, $2f_2$, $f_1 + f_2$ are all greater than 24 kHz and thereby removed totally except $f_1 - f_2 = 10$ kHz. Consequently, what remains becomes:

$$S_{\text{out}}(t) = B + B\cos(2\pi(f_1 - f_2)t) \quad (7)$$

Therefore, we can view the amplifier plus the LPF as a frequency downconverter, which downconverts a combination of two signals at higher frequencies of f_1 and f_2 into a lower frequency at $f_1 - f_2$. Next, we will use this downconverter to pull down the high charging frequencies to the voice band.

On the charger side, we use the PWM to transmit the following signal:

$$\begin{aligned} \text{PWM}(S_v(t) + 1, S_c(t)) &= (S_v(t) + 1)S_c(t) = S_c(t) + S_v(t)S_c(t) \\ &= \cos(2\pi f_c t) + S_v(t)\cos(2\pi f_c t) \end{aligned} \quad (8)$$

where 1 represents the DC component. On the microphone side, after the above signal passes through the “downconverter” (i.e., substituting Eqn. 8 into Eqn. 7), we obtain the finally recorded signal as follows:

$$S_{\text{out}}(t) = B + BS_v(t)\cos(2\pi(f_c - f_c)t) = B + BS_v(t) \quad (9)$$

where B is a constant and only affects the voice volume. In this way, we successfully close the frequency gap between the charger and the microphone.

To verify the nonlinearity effect, we transmit the chirp signal sweeping from 100 \sim 122 kHz using the PWM and the TX coil. Meanwhile, we use the previously assembled microphone to capture the MIS. Fig. 11(a) shows that the

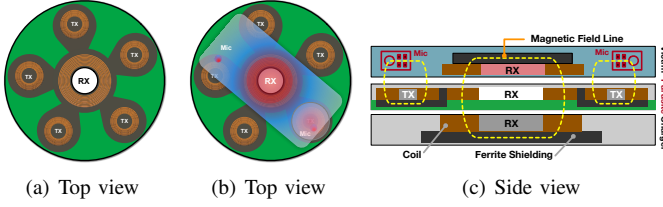


Fig. 13: Architecture of a parasite label. (a) and (b) show the top view of a parasite label without and with a phone respectively. (c) shows the side view of the attack scenario where the parasite presents between a phone and a charger. They are stacked in a pile.

microphone can successfully capture the downconverted chirp signal even if it is modulated onto a carrier at 100 kHz. Then, we use a flying wire to short-circuit the amplifier and repeat the experiment. As a result, no signal is detected anymore as shown in Fig. 11(b). This fully demonstrates the presence of the nonlinearity effect of the amplifier.

D. Summary

We summarize the workflow of heartworm attack in Fig. 12. To inject a voice command, the heartworm takes control of the MCU to generate a PWM-emulated voice command and then manipulates the GPIO pins by using high-level programming instructions to generate the digital signals, which drive the TX coil to produce an amplitude-varying magnetic field. Consequently, an MIS that piggybacks the voice commands is captured by a nearby microphone. Then, the MIS is automatically downconverted to audible voice commands by the downconverter (i.e., amplifier plus LPF) due to the nonlinearity effect. Finally, the command is executed by the voice controllable systems. The voice commands modulated onto the magnetic field are injected into the microphone without triggering any mechanical vibration, so no in-air voice can be detected or heard.

V. DESIGN OF PARASITEATTACK

In this section, we introduce a non-intrusive attacking approach called ParasiteAttack, which launches the attack through accessory equipment called parasite. The battery-free parasite is as thin and small as an NFC tag. The adversary adheres to the parasite on the top of a charger and disguises it as a sticker by printing some signs, such as “Free Charging” or “Quick Charging,” which mislead users into viewing a parasite label as a part of the real wireless charger. More deployment examples are shown in Appx. G.

A. Parasite in a Nutshell

The parasite is deployed between the host wireless charger and the smart device. We design the parasite as a battery-free device to be small, compact, and not eye-catching. Fig. 13 shows the architecture of a parasite. Specifically, a parasite label is composed of an inner RX coil and several outer TX coils. After the power transfer contract is established, the parasite uses the inner RX coil to steal power from the underneath charger and boosts the attack using outer TX coils. The center of the RX coil is empty without a ferrite shield such that the magnetic field created by the host charger can reach the RX coil of the victim device with minimal attenuation.

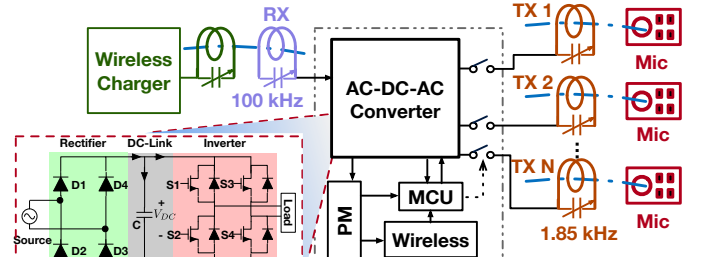


Fig. 14: Schematic diagram of a parasite

Multiple TX coils are deployed on a ring to ensure at least one TX coil is located nearby the victim’s microphone even if the device’s posture is uncertain. The TX coils are shielded from the RX coil and the host charger’s TX coil by using ferrite sheets to avoid potential mutual interference.

B. Parasite Architecture

Fig. 14 shows the schematic diagram of a parasite’s circuits. The RX coil and its corresponding resonant tank are designed to work at 100 kHz in accordance with the physical-layer guideline of the Qi standard. The harvested power is stored in the module of power management (PM), which boosts an MCU and a wireless communication module such as a Wi-Fi or Bluetooth transceiver, allowing the adversary to initiate the controllable voice attack in real-time fashion. The key module is the power converter, that is, AC-DC-AC converter. It has two purposes: first, it can rectify AC to DC for boosting the MCU and the communication; second, it also converts the high-frequency current at 100 kHz harvested from the RX coil down to a low-frequency current at 1.85 kHz for TX coils. Therefore, the additional downconversion on microphones is needless because the parasite exactly transmits the voice in the operating range of a microphone directly. In the following, we will elaborate on the design of a parasite.

C. Stealing Power from a Host Wireless Charger

We adopt a Qi Example 4 coil as the RX coil, which consists of 66 strands of 0.88 mm diameter Litz wires. The inner and outer diameters of the coil are 47 and 28 mm, respectively. Thus, a 615 mm² hole is found in the center of the RX coil, which allows the above victim device to absorb energy as usual. In the view of the victim, the parasite is totally transparent. Inspired by the electric power transmission system [43] and motor control [44], we use an AC-DC-AC power converter to convert the AC induced by the RX coil to DC. The power converter has three main components: a rectifier, a DC-link, and an inverter. Specifically, the rectifier consists of four diodes to provide full-wave rectification, that is the whole of the input waveform to one of constant polarity at its output. The DC-link consists of a capacitor, which can remove ripples caused by the rectifier and absorb the power surges between the RX coil and the TX coils. The power surges are inevitable because the voice signal modulated on the TX coils might cause varying strength of the output power. The inverter inside the converter is used to generate a new AC

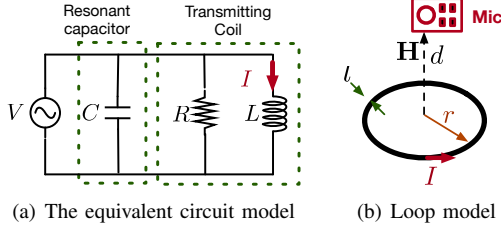


Fig. 15: The equivalent circuit model of a TX coil

at 1.85 kHz as a carrier to modulate the voice signal. A reverse diode should be added in parallel to each FET switch to protect it from reverse surges in case no-load disperses the redundant energy at TX coils.

D. Voicing Out through TX Coils

Similar to the HeartwormAttack, the MCU manipulates the inverter inside the power converter to create a PWM-emulated voice signal, which is further propagated into the air through the TX coils. Here, we skip the modulation procedure (which is as same as the heartworm attack) and focus on the key question: *how could we design the TX coils to maximize the magnetic strength at a victim's microphone?*

The input current is evenly distributed on the coil because the size of the coil is relatively small compared to the carrier wavelength. Hence, the coil itself can be regarded as a combination of an inductor and a resistor. An additional capacitor is added across the coil to form a resonator. In this way, the transmitter is usually modeled as the equivalent RLC parallel circuitry, as shown in Fig. 15(a). Correspondingly, the natural oscillation frequency of the circuitry is determined by

$$f = \frac{1}{2\pi\sqrt{LC}} \quad (10)$$

We can tune the resonant frequency to any wanted frequency by changing the value of the capacitor. The efficiency of a magnetic transmitter is measured by using a physical parameter called *quality factor* denoted by Q , which is defined as follows:

$$Q = \frac{2\pi fL}{R} \quad (11)$$

When a current flows into a resonator, the resonant current passing through the coil will be amplified by $Q \times$ of the input current. Correspondingly, the magnetic field strength can be enhanced by $Q \times$ compared with a resistive load although the input power stays the same [45]. However, a trade-off is found between the fundamental frequency f , half-power bandwidth B , and quality Q in an inductive system. Their relationship is represented as follows:

$$B = \frac{f}{Q} \quad (12)$$

The above equation suggests that the quality factor of a resonator is inversely proportional to the communication bandwidth. Considering that the narrowest band of the recognizable human voice is between 300 Hz and 3.4 kHz [29], then $f = 1.85$ kHz (i.e., the center frequency) and $B = 2.46$ kHz. Thus, the maximum Q that we can achieve is 0.75; otherwise, the voice might be incompletely propagated out.

Given an input power P , the magnetic field H generated by the TX coil at a victim's microphone to the coil can be modeled as follows [45]:

$$H \propto \sqrt{\frac{PQ}{r \left[\ln \left(\frac{8r}{l} \right) - 2 \right]}} \frac{r^2}{(d^2 + r^2)^{\frac{3}{2}}} \quad (13)$$

where r is the radius of the coil, l is the wire radius, and d is the distance between the victim microphone and the coil. Clearly, we may increase P , r , l , Q , or decrease d to achieve a stronger magnetic field at the microphone. All the parameters are illustrated in Fig. 15. Our purpose is to enhance H by choosing appropriate parameters. Q cannot be increased because of the bandwidth constraint. l logarithmically contributes to H . Thus, even a significant raise in l only leads to minimal improvement at H . Moreover, l determines the thickness of the label, which should be as small as possible for better concealment. r is also constrained by the area of the top surface of the host charger. Thus, the only method of enhancing H is to shorten distance d . In other words, we should put the TX coil of a parasite tightly close to the victim's microphone as much as possible. However, the horizontal orientation of the victim device is unclear although knowing its RX coil must be overlapped with the RX of the parasite label for best-efficient charging. To address this issue, we deploy a number of TX coils on a ring around the RX coil, as shown in Fig. 13(a). As a result, we always find a close TX coil that is placed right below the microphone regardless of which horizontal angle the device is toward. The practical implementation might adopt a denser arrangement.

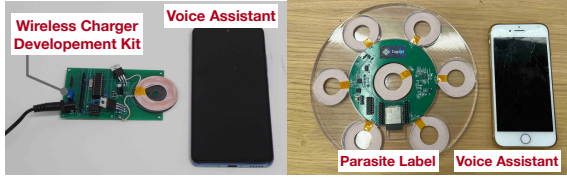
Transmitting voice commands at multiple TX coils concurrently might disperse the magnetic power. Thus, only a single TX coil is chosen to connect to the power converter each time using switches, as shown in Fig. 14. However, how can we know which TX coil should be chosen? When any metallic object is close to a TX coil, it will absorb some energy due to the vortex effect, resulting in the current or voltage change in the TX coil. In this way, we can determine which TX coil is under the victim device. The parasite polls all TX coils to choose the one that can cause vortex effect as the best TX coil for the voice attack.

E. Avoiding Foreign Object Detection

In the Qi standard, the wireless charger is required to detect foreign objects for safety charging [23], which prevents deformation or damage from occurring due to excessive heat generation in the event a metallic object is placed between the TX and RX coils. Such detection might consider our parasite label as a foreign object and stop the charging. Two methods are used for foreign object detection (FoD) in the Qi standard [46]. (1) *Checking quality factor*. Similar to our previous approach for detecting the microphone location, the wireless charger can examine the frequency or quality factor of its resonant tank to determine the presence of the foreign object. The RX coil of the parasite label is well-tuned to align with the TX coil of the wireless charger already. Thus, the parasite will not trigger the report of a foreign object. (2) *Checking power loss*. The smart device approximates the

TABLE I: Experimental devices, speech recognition and results. The Qi column indicates if the smart device is compatible with wireless charging.

Type	Manuf.	Model	Rel.Date	OS	SR	Qi	ParasiteAttack			HeartwormAttack		
							Recog.	Power	Dist.	Recog.	Power	Dist.
Smartphone	Xiaomi	V9 Pro	2019/09	MIUI 12.0	Xiaoai	Yes	✓	≥ 15W	< 1cm	✓	≥ 15W	3cm
Smartphone	Xiaomi	V11	2021/01	MIUI 12.5	Xiaoai	Yes	✓	≥ 15W	< 1cm	✓	≥ 30W	4cm
Smartphone	Huawei	Mate 20 Pro	2018/10	Harmony OS 2.0	Xiaoyi	Yes	✓	≥ 15W	< 1cm	✓	≥ 15W	3cm
Smartphone	Huawei	Honor V30 Pro	2019/11	Magic UI 3.0	YOYO	Yes	✓	≥ 15W	< 1cm	✓	≥ 15W	5cm
Smartphone	Apple	iPhone 8	2017/09	iOS 15.1	Siri	Yes	✓	≥ 15W	< 1cm	✓	≥ 15W	2cm
Smartphone	Apple	iPhone 11	2019/09	iOS 15.0	Siri	Yes	✓	≥ 50W	< 1cm	×	N.A	3cm
Smartphone	Apple	iPhone 12	2020/10	iOS 15.1	Siri	Yes	✓	≥ 50W	< 1cm	×	N.A	3cm
Smartphone	Samsung	Galaxy 21	2021/01	One UI 3.1	Bixby	Yes	✓	≥ 50W	< 1cm	×	N.A	3cm
Smartwatch	Samsung	Galaxy Watch 4	2021/08	One UI 3.0	Bixby	Yes	✓	≥ 15W	< 1cm	✓	≥ 15W	< 1cm
Smartwatch	Huawei	GT2 Pro	2020/12	Harmony OS 2.0	Xiaoyi	Yes	✓	≥ 15W	< 1cm	✓	≥ 15W	< 1cm
Smartwatch	Apple	Watch 7	2021/10	WatchOS 8.0.1	Siri	Yes	✓	≥ 15W	< 1cm	✓	≥ 30W	< 1cm
Tablet	Apple	iPad(6th)	2018/03	iPadOS 15.0.1	Siri	No	✓	≥ 15W	< 1cm	✓	≥ 15W	3cm
Tablet	Apple	iPad Air 4	2020/10	iPadOS 15.0.1	Siri	No	✓	≥ 50W	< 1cm	✓	≥ 50W	3cm
Tablet	Apple	iPad mini 6	2021/09	iPadOS 15.1.0	Siri	No	✓	≥ 15W	< 1cm	✓	≥ 50W	3cm
Add-on Mic	SparkFun	ADMP401	2010/04	Windows 10	Google STT	No	✓	≥ 15W	< 1cm	✓	≥ 15W	3cm
Add-on Mic	Knowles	SPH0690LM4H-1	2019/06	Windows 10	Google STT	No	✓	≥ 15W	< 1cm	✓	≥ 15W	3cm
Add-on Mic	Joy-IT	KY-037	2017/06	Windows 10	Google STT	No	✓	≥ 15W	< 1cm	✓	≥ 15W	3cm



(a) HeartwormAttack

(b) ParasiteAttack

Fig. 16: Prototypes. (a) a wireless charger development kit is employed for the HeartwormAttack; (b) The prototype of parasite for ParasiteAttack.

received power and sends it to the wireless charger, which computes the power loss between the transmitted and the received power. If the loss exceeds a threshold (i.e., 500 mW), a foreign object is reported. The parasite label can forge a false power report or interfere with the report from the smart device to avoid detection. As demonstrated in the evaluation, we find that it is easy to avoid the FoD since the commercial chargers usually adopt a relatively higher threshold.

VI. IMPLEMENTATION AND EVALUATION

A. Implementation

We developed the heartworm malware using a development kit for wireless charging and prototypes the parasite label:

HeartwormAttack. Fig. 16(a) shows the development kit for wireless charging. This kit comprises of a STC12C2052AD MCU [47] with 72 MHz clock frequency (i.e., f_{CLK}), an inverter with JRF540N FET switches [48], and a standard A11 coil from TKD [49]. The maximum output power is 30 W. We also developed a heartworm malware using the PWM timer API in MCU. The carrier frequency of PWM (i.e., f_{PWM}) is 1 MHz, while the duty cycle can be varied from 0% to 100%. Therefore, we use $\log_2(f_{CLK}/f_{PWM}) = 6$ bits to control the inverter. The analog voice commands are firstly sampled with a 16 kHz rate and then PWM-emulated. The commands are finally generated by the MCU.

ParasiteAttack. As shown in Fig. 16(b), we prototype the parasite label using an annular PCB, which holds the main circuits (e.g., MCU, rectifier, and inverter). The inner hole accommodates a Qi Example 4 RX coil from TDK [50] to harvest power from the real charger, while six TX coils are arranged along the outer margin of the PCB to launch voice attacks at microphones. All coils comprise 2 mm thin Litz wires.

The full-wave rectifier comprises 1N400x diodes [51] and TS61005 power FET switches [52]. The power converter (i.e., AC-DC-AC converter) is a reconfigurable platform controlled by an external low-power ESP32-C3 microcontroller [53], which is integrated with a Wi-Fi module internally. Therefore, we can control the parasite label remotely through Wi-Fi. An LM7805 voltage regulator [54] is adopted for power management, which can stabilize the rectified current and power up the ESP32. Similarly, the analog voice commands are sampled with a 16 kHz rate. The carrier frequency of PWM is 100 kHz, and 10 bits are used to control the inverter. The thickness of the parasite label is less than 2 cm.

B. Experiment Setup

We test our attacks across eight types of smartphones and three types of smartwatches, which were all released after 2018 when the functionality of wireless charging was ready to be generally adopted among smart devices. The three types of tablets are from the series of Apple iPad series. Unfortunately, iPads have not offered the wireless charging function to date. Nevertheless, we still test the attacks on these tablets considering a generalized scenario, where they are accidentally placed near a wireless charger. We finally test three types of add-on microphones, which are the most popular components for developing smart wearables. We connect these microphones to an Arduino for voice recording. We test the following three power levels: 15, 30, and 50 W. The Qi standard specifies the 15 W default power and the 30 W maximum power, while the enhanced quick chargers on the market adopt 50 W power [25]. Unless specified, the distance between the charger and the victim devices is less than 5 cm, the transmitting power is set to 15 W, and iPhone 8 is employed by default. We use an AR824 sound level meter [55] to measure the environmental noise strength in the unit of dBA, which represents the relative loudness of sounds in air as perceived by the human ear on average. All experiments are conducted in a quiet lab room with a background noise level of 45 dBA.

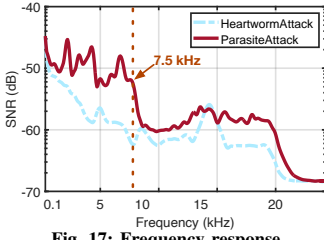


Fig. 17: Frequency response

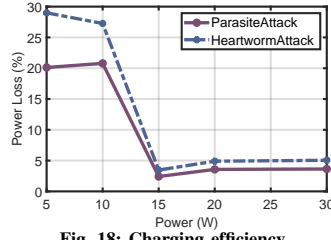


Fig. 18: Charging efficiency

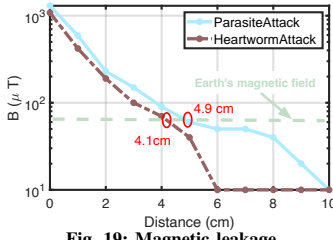


Fig. 19: Magnetic leakage

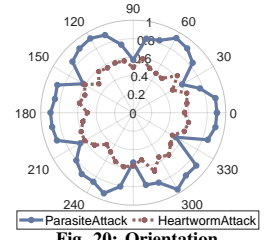


Fig. 20: Orientation

C. Feasibility Results

Table I summarizes the experiment results. The attack is viewed as a “success” (ticked with \checkmark) once the speech recognition (SP) system can successfully recognize the short wake-up voice commands (e.g., “Hey Siri”, “Hey Google” and “Hi Xiaoi”) recorded by the microphones under attacks. The table reveals the following findings:

- First, the attacks are successfully conducted in 91% cases (31 out of 34), where each device can be successfully attacked in at least one case. We believe this list is by far not comprehensive. Nevertheless, the results serve as a warning to consider the security breach caused by wireless charging.

- Second, the ParasiteAttack performs better than the HeartwormAttack. The average minimum power that the HeartwormAttack and ParasiteAttack require are about 28.2 W and 23.2 W, respectively. Namely, the HeartwormAttack needs additional 5 W power for the voice injection. This is because magnetic interference generated by the HeartwormAttack experiences two extra attenuations, which are caused by the nonlinearity of a microphone’s amplifier and the $2 \times \sim 3 \times$ further distance compared with ParasiteAttack.

- Third, high-end smartphones adopt substantially effective packaging techniques and EMI shield materials, which is reflected in HeartwormAttack failures on iPhone 11, 12, and Galaxy 21. They are well protected against the HeartwormAttack even when using 50 W transmitting power. Unfortunately, these devices still suffer from ParasiteAttack when the TX coil of the parasite is placed close (< 1 cm) to the microphone.

Overall, regardless of the types of models, manufacturers, and speech recognition systems, the commercial off-the-shelf devices all fail to defend against the proposed attacks when given sufficient power. Particularly, the transmitting power plays a key role in attacking. In the industry, the current aggressive pursuit of faster charging via raising the transmitting power will intensify the potential magnetic interference and thus considerably increase the rate of attacking success. This reminds us of the potential negative consequence of faster or quick charging. In addition, whether victim devices support wireless charging is not a prerequisite for our attacks. On the contrary, those devices without WPT functionality become more vulnerable when getting close to a wireless charger because they have very little protection against magnetic interference.

D. Frequency Response Analysis

Next, we quantify the quality of the injected voice by analyzing the frequency response, that is, the quantitative measure of the output spectrum of a system or device in

response to a stimulus. In the experiment, we transmit a chirp sound sweeping from 100 Hz to 22 kHz with the two attacking approaches. Fig.17 compares the two frequency responses. We find that: (1) ParasiteAttack achieves a higher SNR in the range of 100 Hz to 7.5 kHz, which exactly covers the human speech spectrum. This demonstrates that the parasite label is rationally designed and meets the attacking requirements. (2) By contrast, the response of a HeartwormAttack is relatively unsatisfactory. Notably, a wireless charger’s TX coil is not designed for our attack but to provide more power. Thus, the quality factor of its coil is over 77, and the operating frequency is 100 kHz. Based on Eqn. 12, its half-power bandwidth is as small as $100/77 \approx 1.2$ kHz [56]. Theoretically, the power of frequency components above 2.4 kHz is nearly zero.

Fig. 30 shows the spectrograms of two recorded voices, which are injected by HeartwormAttack and ParasiteAttack, respectively. As analyzed above, we clearly observe a cutoff at 2.4 kHz in Fig. 30(a). This suggests that the optimal voice that the HeartwormAttack injects into microphones should be less than 2.4 kHz. Nevertheless, the current voice recognition systems usually ignore high-frequency components and exhibit strong fault tolerance. Thus, HeartwormAttack still works well in practice for keywords based commands.

E. Stealthiness Analysis

We would like to evaluate the stealthiness of our attacks, that is, if being cautious of the attacks is easy. First of all, we invited 20 subjects including 10 males and 10 females, who are aged between 15 and 40. These subjects were requested to stay nearby the wireless charger (< 1 m) and report if any voice or noise was heard during the attack. As a result, none of them heard any sound and realized the attacks had occurred. Thus, the attacks or the voice commands are completely inaudible.

However, the attacks might still be discovered through the two implicit side channels. (1) **Power Efficiency.** Both attacks consume a certain amount of power from the wireless charger and thereby affect the charging efficiency, making the user feel that it takes a longer time to charge than usual. In the experiment, we charge a victim smartphone *with* or *without* attacks. In each trial, the initial percentage of the remaining battery power is fixed at 50%. We then observe how much power percentage can be recharged after half an hour. Let $P_{w/o}$ and P_w respectively denote the increased power percentage without and with a sustained attack, i.e., playing wake-up commands continuously. Apparently, $P_{w/o} > P_w$ because the attack consumes additional power. Choosing $P_{w/o}$ as the baseline, we compute power loss as the ratio of $1 - P_w/P_{w/o}$. Fig. 18 shows the power loss of the two attacking

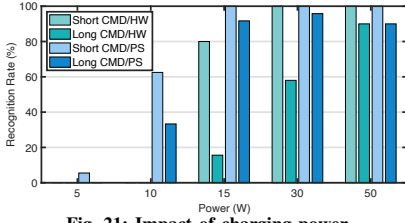


Fig. 21: Impact of charging power

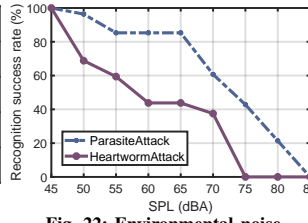


Fig. 22: Environmental noise

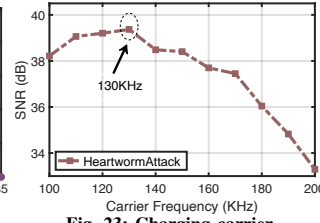


Fig. 23: Charging carrier

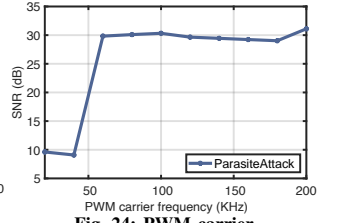


Fig. 24: PWM carrier

approaches. As a result, the power loss is below 5%, which is substantially small to be noticed by users when the charging power is larger than 15 W. Interestingly, the HeartwormAttack consumes more power than the ParasiteAttack because the PMW-emulated signals do not fit the follow-up resonate tank effectively and additional energy is dissipated into the air in the form of heat. **(2) Magnetic Leakage.** The RX and TX coils of the parasite label might introduce a stronger magnetic field around the victim device, which might alert users of abnormal magnetic flux. To this end, we use an HT201 Gaussmeter [57] to measure the magnetic flux density (denoted by B) as a function of the distance to the charger. The measurements are repeated 20 times, and the median is reported for each distance. Fig. 19 shows the median density with a varying distance from 0 to 10 cm. Consequently, the magnetic density caused by ParasiteAttack is almost similar to that of the HeartwormAttack when the distance is less than 4.9 cm because the TX coil of the charger is far larger than that of TX coil of the parasite label, thereby dominating the magnetic field. In short, the abnormal magnetic fields caused by the two attacks are drowned in the charger’s field when the distance is less than 4.9 cm. Beyond this distance, the density caused by a parasite label attenuates to an extremely small value (i.e., 65 μ T) close to the geomagnetic field [58]. Overall, being cautious of the attacks is difficult for users regardless of the power loss or magnetic leakage.

F. Avoiding Foreign Object Detection

The current wireless charging techniques require the chargers to detect foreign objects for safety charging, which considers the parasite label as the foreign object by “mistake”. Therefore, we test if the detection across five commercial wireless chargers: Fast Charge 2.0 (C1) [59] from Samsung, Charging Stand (C2) [60] from Baseus, Charging Pad (C3) [61] from UGREEN, MagSafe (C4) [62] and MagSafe Duo (C5) [62] from Apple. In the experiment, we place the parasite label on the top of a charger and use a wireless charging power test module [63] to check whether FoD is triggered and measure the charging power loss. The FoD results are shown in Table II where \times denotes the detection failure. Consequently, the average power loss caused by ParasiteAttack is only around 1.08% of the overall power consumption and all these charges failed to detect the presence of our parasite label. This result shows that the current FoD mechanism is remarkably easy to dodge. More testing details refers to Appx. E.

VII. IMPACT ANALYSIS

The performance is further measured using the recognition success rate, which is defined as the percent of words success-

fully recognized by the Google Speech-to-Text toolkit [64].

A. Impact of Horizontal Orientation

In wireless charging, the horizontal orientation of smart devices is uncertain in practice, which may affect the attack performance. In this experiment, we launched the two attacks in this experiment by placing the phone at different angles under a 30 W charging power. Fig. 20 shows the recognition rate in different angles. The rate under the HeartwormAttack remains highly similar at all angles because the distance between the TX coil and the microphone is irrelevant to the orientation in this attack. By contrast, the rate under the ParasiteAttack fluctuates remarkably in angles, resulting from the petal-shaped TX design. (see Fig. 16(b)). Consequently, even the minimum rate measured in the ParasiteAttack when the microphone is located in the space between two adjacent TX coils is slightly greater than that of HeartwormAttack.

B. Impact of Charging Power

The previous validation experiment summarizes the minimum charging power required to recognize the wake-up commands successfully. Here, we further evaluate the impact of charging power by testing short and long voice commands, “Open the Door” (3 words, Short CMD) and “Call 1234567890” (11 words, Long CMD). The commands are repeated ten times, and the mean rate is reported. Fig. 21 shows the recognition rate in the five power settings where HW and PS are short for HeartwormAttack and ParasiteAttack, respectively. The results confirm that the charger power is the most important parameter for the two attacks. Specifically, the recognition rate increases as the charging power is raised; second, the minimum power to recognize the short and the long commands are 5 W and 10 W in the ParasiteAttack, while that is 15 W and 15 W in the HeartwormAttack. Namely, the HeartwormAttack requires more power due to the attenuations caused by the downconversion and the longer distance.

C. Impact of Environmental Noise

Speech recognition is known to be sensitive to background noises and is recommended to be used in a quiet environment. Thus, we examine the inaudible voice command injection with the two attacks regarding the controllable and the uncontrollable Environment noises. **(1) Controllable Environment:** In this experiment, we used a speaker to play recorded traffic noise and controlled the noise strength by adjusting the speaker volume. Fig. 22 shows the recognition results of the two attacks under different environmental sound pressure levels (SPL). As desired, the rate linearly decreases as the noise strength increases. Certainly, the results also depend on how

strong the speech recognition system is resilient to noise. Considering the injection, the recognition rates under the HeartwormAttack and ParasiteAttack are over 60% when the noise is lower than 55 and 70 dBA, respectively. **(2) Uncontrollable Environment.** We then conducted the experiment in five real-life environments: library, park, cafe, subway station, and bus. The recognition results are listed in Table. III. Specifically, ParasiteAttack performs well in the various scenarios except for the bus, whereas HeartwormAttack is more adapted to a relatively quiet environment like the library or park. The noise reaches a maximum level of 70-85 SPA on a bus where both two attacks underperform. Overall, noise is a long-standing challenge that voice recognition systems are facing. Both attacks more or less introduce extra noise into the injected voice, leading to a lower recognition rate.

D. Impact of Carrier Frequency

(1) Charging Carrier. Qi standard allows a charger to choose a frequency between 100 - 200 kHz to transfer power [23], which might affect the performance of the HeartwormAttack because the voice command is piggybacked by the charging carrier. In this experiment, the heartworm transmits a 1 kHz signal tone but at different-frequency charging carriers. Fig. 23 shows the signal-to-noise (SNR) of the recorded MIS as a function of the carrier frequency. We observe that the 130 ± 30 kHz carriers best fit the attack. This result confirms our previous assumption that a charger improves the charging efficiency at the cost of bandwidth. Therefore, the HeartwormAttack should be launched during 100 ~ 130 kHz around. **(2) PWM Carrier.** The frequency of PWM carrier is a key parameter to emulate the voice command in the ParasiteAttack. There is a trade-off in choosing the carrier frequency, namely, a higher frequency better emulates the voice signal but consumes much power. Fig. 24 shows the SNR of the recorded MIS as a function of PWM frequency. It can be seen that the SNR is maintained at the maximum (i.e., ~ 30 dB) when the carrier frequency is above 60 kHz, which is triple higher than the maximum frequency of the voice (i.e., 24 kHz). However, the clock frequency of the MCU is 1 MHz, and the carrier frequency must be an integral division of the clock frequency. Thus, the optimal frequency is 100 kHz.

VIII. LIMITATIONS & COUNTERMEASURE

A. Limitations of Proposed Attacks

We successfully inject magnetic-inductive inaudible commands into smart devices via their microphones. However, the two proposed attack approaches are still limited in many aspects as follows. These limitations will help us to find the corresponding countermeasures. **(1) Short attack range.** The intensity of the magnetic field degrades rapidly at $\mathcal{O}(1/d^6)$ with the distance d [45]. Thus, the attacks are launched successfully only when the smartphone is placed exactly on the wireless charger where the distance is less than 5 cm as reported. **(2) Voice fingerprinted activation.** Nowadays, iPhone or other models have started to use the fingerprinted voice to activate the assistant, which is an effective way to

defend against all inaudible voice attacks including ours. Thus, the adversary must obtain the voice fingerprint and imitate the wake-up commands in practice. Voice cloning [38] can mitigate this constraint but require short voice clips from the victim through a side channel. **(3) Higher charging power.** As mentioned early, the charging power plays a key role in the attack. To achieve a higher success rate, both attacks require a minimum charging power of 10 W, 15 W, or even 50 W. **(4) Aimless attack.** All inaudible voice attacks are unidirectional, that is, the attack device (e.g., charger) cannot obtain any feedback from the victim devices. As a result, the attacks are initiated aimlessly. For example, “turn speaker off”, “download a software” and “transfer money”. The three consecutive voice commands are less logically connected. **(5) Need for extra equipment.** The ParasiteAttack must be initiated by the parasite label, which is stuck onto the wireless charger. Even if installed snugly, the parasite label is still likely to be found by careful users because the depth of charging is increased by $2 \sim 3$ cm. **(6) Additional protection for high-risk tasks.** Some high-risk tasks, especially for financial issues (such as payment, bank transfer, photo access, etc) require the double-check with the password or the face-based authentication although they can be initiated by the VA. **(7) Background noise.** The noises introduced by our attack may impact the quality of the injected voice. While, in most cases, the adversarial commands are still clean enough to pass the voice biometrics.

B. Countermeasure Recommendations

Inspired by the limitations, we design the following countermeasures to defend against the potential magnetic interference. **(1) Upgrading hardware design.** The magnetic interference has not attracted enough attention at present because wireless charging has been quickly spread in recent years. As mentioned early, the outdated EMI countermeasures (including the Faraday cage and the EMI filter) can only protect the microphones from interference at 1 MHz above. Thus, the most effective countermeasure is to upgrade the hardware design for newly developed smart devices. For example, mounting the MEMS chip and the ASIC chip inside a ceramic substrate; sealing a microphone with a polymer foil [65]; adopting the material (e.g., mu-metal for oscilloscope protection) with higher magnetic permeability for the device’s casing. **(2) Throughout fingerprinted.** To date, the verification of voice fingerprints is limited to the wake-up commands such as “Hi, Siri!”. Once the voice assistant is activated successfully, the follow-up voice interactions will skip the verification for a quick response. We should take fingerprint verification all the time. **(3) Abnormal voice detection.** Both HeartwormAttack and ParasiteAttack introduce some particular acoustic charac-

TABLE II: FoD Results

Model	Loss(mW)	Loss(%)	FoD
C1	337.0 mW	1.12%	×
C2	331.8 mW	1.10%	×
C3	320.5 mW	1.06%	×
C4	155.6 mW	1.03%	×
C5	163.3 mW	1.08%	×

TABLE III: Noise Results

Scene	SPL(dBA)	HW.	PR.
Library	35-45	100%	100%
Park	45-55	68.8%	96.4%
Cafe	55-65	43.8%	85.3%
Subway	60-75	37.5%	60.7%
Bus	70-85	0	21.4%

teristics, which can be utilized to detect the presence of the attacks. For example, the voice injected by HeartwormAttack and ParasiteAttack is cut off at 2.4 kHz and 4 kHz, respectively, where the absence of higher frequencies can be used as a typical feature of the abnormal voice command. Actually, a similar countermeasure has been studied in some previous work [66]–[68] to defend against other attacks. **(4) Abnormal commands.** As mentioned, the attacks cannot obtain feedback from smart devices. They have to attempt the pre-defined commands successively. These commands are not logically connected. We can detect malicious commands via analyzing the purposes of these commands and their connections. **(5) Sophisticated FoD.** The current FoD algorithms fail to detect the presence of a parasite label. We could develop a more sophisticated algorithm through various parameters, including power consumption, magnetic density, and so on. We could also encrypt the data exchanged between the charger and the device. **(6) Disabling microphone.** The microphones or the voice assistants shall be automatically disabled when the devices are being charged. **(7) Limitation on charging power.** Finally, the industry is advised to properly limit the charging power to avoid the negative consequence of quick charging.

IX. RELATED WORK

A. Inaudible Voice Attacks

A number of recent studies have demonstrated the feasibility and the negative consequences of inaudible voice commands. **(1) Vibration-based Attacks.** The first group attempts to trigger the mechanical vibrations of a microphone’s membrane and inject the voice commands [4]–[9], [33]. Backdoor [5] shows that a microphone that is originally designed to record human voice only can receive ultrasonic signals because the membranes can be vibrated by ultrasound as well. DolphinAttack [4] uses this physical characteristic to demonstrate inaudible voice attacks toward many popular voice controllable systems by injecting ultrasound signals over the air. The follow-up work named LipRead [6] further extends the attack range from 5 to 25 ft by using an array of ultrasonic transducers. Unlike these existing studies, our attacks never trigger any mechanical vibrations but directly induce the voice signals on the circuits of a microphone, thereby successfully bypassing the defense method of detecting abnormal vibrations [68]. **(2) Coupling-based Attacks.** The second group leverages the EMI to launch the inaudible voice attack [10]–[13]. EMI is an intervention generated by external excitations that affect peripheral sensitive electrical circuits and sensors [69], [70] by electromagnetic induction, energy coupling, or conduction. GhostTalk [11] firstly finds that microphones suffer from the bogus audio signals caused by the EMI from the wireless communication at the ultra-high frequency (i.e., 800–900MHz). The follow-up work [12] utilizes the front-door coupling on headphone cables to capture the forged AM-modulated signals at high frequency (i.e., 80–108 MHz) from the air. These previous EMI-based attack methods at high or ultra-high frequency are now completely disabled by Faraday cages equipped for microphones [17]. Another line of work

uses light to induce the coupling current at microphones and initiate the inaudible voice attack [10]. However, keeping in line with the sight of the victim device is required despite the long range. By contrast, our attacks utilize the ubiquitous wireless chargers to attack smart devices when not in use, which is more unnoticeable to users. Recently, the authors in [13] even proposed that the audio signals can be injected directly into the power line through a modified charging cable wire. We move forward and show that wireless chargers can also be manipulated to perform inaudible attacks. To the best of our knowledge, we are the first to use the wireless chargers’ magnetic leakage as an EMI to induce the coupling current in microphones for injecting inaudible commands.

B. Electromagnetic Compatibility (EMC)

EMC is the ability of electrical systems to function acceptably in their electromagnetic environment. Here, we mainly review the studies on EMC in microphones and chargers. **(1) EMC for Microphones.** A white book [17], [71] from Infineon Technologies (a worldwide semiconductor manufacturer) shows that the EMC for current MEMS microphones is achieved through three main measures (Sec. 3, pp.16), namely, using capacitors to filter low and high-frequency interference, adding a series resistor or ferrite, and connecting microphone grounds to the circuit board ground plane with vias. After a complete analysis of the EMI of current microphone MEMS, Ko et al. [26] points out that the EMI suppression on MEMS microphone can be further improved by 14 dB by increasing the number of micro bumps, adding the ground via in the substrate, and applying metal coating around the acoustic port. Reitsema et al. [72] emphasized suppressing the down-conversion of high-frequency disturbances to audio frequency and compensating for the remaining disturbance signals. **(2) EMC for Wireless Chargers.** Wireless charging utilizes electromagnetic induction to transfer power and inevitably causes EMI in peripheral circuits of devices to be charged, such as smartphones, watches, and electric vehicles. However, the state-of-the-art studies [73]–[78] are only limited to the EMI suppression in the scenario where electronic vehicles are being charged because vehicles require extremely high power (KW) to transfer sufficient energy. The EMC for wireless charging is absent at present.

X. CONCLUSION

In this work, we first demonstrated that magnetic interference is a practical threat to the microphone system and can be manipulated to inject malicious voice commands into smart devices. This work will raise awareness of such great potential safety hazards on smart devices.

ACKNOWLEDGMENT

We thank all anonymous reviewers and shepherd for the insightful feedback. This study is supported by NSFC Key Program (No. 61932017), NSFC Excellent Young Scientists Fund (Hong Kong and Macau) (No. 62022003), NSFC General Program (No. 61972331), UGC/GRF (No. 15204820, 15215421), and National Key R&D Program of China 2019YFB2103000.

REFERENCES

- [1] "Find out what siri can do on iphone," <https://support.apple.com/en-hk/guide/iphone/ipha48873ed6/ios>, 2022.
- [2] "Google Now," <https://www.androidcentral.com/google-now>, 2016.
- [3] "Amazon Alexa," <https://developer.amazon.com/en-US/alexa>, 2017.
- [4] G. Zhang, C. Yan, X. Ji, T. Zhang, T. Zhang, and W. Xu, "Dolphinattack: Inaudible voice commands," in *Proc. of ACM CCS*, 2017, pp. 103–117.
- [5] N. Roy, H. Hassanieh, and R. Roy Choudhury, "Backdoor: Making microphones hear inaudible sounds," in *Proc. of ACM MobiSys*, 2017, pp. 2–14.
- [6] N. Roy, S. Shen, H. Hassanieh, and R. R. Choudhury, "Inaudible voice commands: The long-range attack and defense," in *Proc. of NSDI*, 2018, pp. 547–560.
- [7] C. Yan, G. Zhang, X. Ji, T. Zhang, T. Zhang, and W. Xu, "The feasibility of injecting inaudible voice commands to voice assistants," *IEEE Transactions on Dependable and Secure Computing*, 2019.
- [8] M. Zhou, Z. Qin, X. Lin, S. Hu, Q. Wang, and K. Ren, "Hidden voice commands: Attacks and defenses on the vcs of autonomous driving cars," *IEEE Wireless Communications*, vol. 26, no. 5, pp. 128–133, 2019.
- [9] Q. Yan, K. Liu, Q. Zhou, H. Guo, and N. Zhang, "Surfingattack: Interactive hidden attack on voice assistants using ultrasonic guided waves," in *Network and Distributed Systems Security (NDSS) Symposium*, 2020.
- [10] T. Sugawara, B. Cyr, S. Rampazzi, D. Genkin, and K. Fu, "Light commands: laser-based audio injection attacks on voice-controllable systems," in *29th USENIX Security Symposium (USENIX Security 20)*, 2020, pp. 2631–2648.
- [11] D. F. Kune, J. Backes, S. S. Clark, D. Kramer, M. Reynolds, K. Fu, Y. Kim, and W. Xu, "Ghost talk: Mitigating emi signal injection attacks against analog sensors," in *2013 IEEE Symposium on Security and Privacy*. IEEE, 2013, pp. 145–159.
- [12] C. Kasmi and J. L. Esteves, "Iemi threats for information security: Remote command injection on modern smartphones," *IEEE Transactions on Electromagnetic Compatibility*, vol. 57, no. 6, pp. 1752–1755, 2015.
- [13] Y. Wang, H. Guo, and Q. Yan, "Ghosttalk: Interactive attack on smartphone voice system through power line," *arXiv preprint arXiv:2202.02585*, 2022.
- [14] L. Palisek and L. Suchy, "High power microwave effects on computer networks," in *10th International Symposium on Electromagnetic Compatibility*. IEEE, 2011, pp. 18–21.
- [15] F. Sabath, "Classification of electromagnetic effects at system level," in *Ultra-Wideband, Short Pulse Electromagnetics 9*. Springer, 2010, pp. 325–333.
- [16] C. Kasmi, J. Lopes-Esteves, N. Picard, M. Renard, B. Beillard, E. Martinod, J. Andrieu, and M. Lalande, "Event logs generated by an operating system running on a cots computer during iemi exposure," *IEEE Transactions on Electromagnetic Compatibility*, vol. 56, no. 6, pp. 1723–1726, 2014.
- [17] Infineon Technologies, "AN558: MEMS microphone electrical implementation.pdf," <https://www.infineon.com/>, 2018, Last accessed October 22, 2021.
- [18] M. Guri, B. Zadov, and Y. Elovici, "Odini: Escaping sensitive data from faraday-caged, air-gapped computers via magnetic fields," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 1190–1203, 2019.
- [19] "Demo," <https://anplus.github.io/magsound/>, 2022.
- [20] "Ansys maxwell," <https://www.ansys.com/products/electronics/ansys-maxwell>, 2021.
- [21] "Apple smart watch model," <https://grabcad.com/library/apple-smart-watch-1>, 2021.
- [22] "Apple iphone 13 model," <https://grabcad.com/library/tag/iphone13>, 2021.
- [23] "Qi Standard v1.3," <https://www.wirelesspowerconsortium.com/knowledge-base/specifications/download-the-qi-specifications.html>, 2020.
- [24] M. Maaß, A. Griessner, V. Steixner, and C. Zierhofer, "Reduction of eddy current losses in inductive transmission systems with ferrite sheets," *Biomedical engineering online*, vol. 16, no. 1, pp. 1–18, 2017.
- [25] "Mi 80W Wireless Charging Stand," <https://www.mi.com/global/product/mi-80w-wireless-charging-stand/>, 2020.
- [26] C.-H. Ko, H.-L. Lee, and C.-H. Wang, "The emi suppression of ultra thin mems microphone package," in *2010 5th International Microsystems Packaging Assembly and Circuits Technology Conference*. IEEE, 2010, pp. 1–3.
- [27] A. Vander Vorst, A. Rosen, and Y. Kotsuka, *RF/microwave interaction with biological tissues*. John Wiley & Sons, 2006, vol. 181.
- [28] "ST EMIF02-MIC03F2 datasheet," <https://www.st.com/en/emi-filtering-and-signal-conditioning/emif02-mic03f2.html>, 2021.
- [29] B. B. Monson, E. J. Hunter, A. J. Lotto, and B. H. Story, "The perceptual significance of high-frequency energy in the human voice," *Frontiers in psychology*, vol. 5, p. 587, 2014.
- [30] "Knowles tda1308 silicon microphone," <https://www.aliexpress.com/item/1005001415977347.html>, 2021.
- [31] "Zero-height sisonic mems microphone," <https://www.knowles.com/docs/default-source/model-downloads/spv1840lr5h-b-rev-b-datasheet.pdf>, 2014.
- [32] "Max9812: fixed-gain microphone amplifiers," <https://www.maximintegrated.com/en/products/analog/audio/MAX9812.html>, 2022.
- [33] X. Ji, J. Zhang, S. Jiang, J. Li, and W. Xu, "Capspeaker: Injecting sounds to microphones via capacitors," in *Proc. of ACM CCS*, 2021.
- [34] "Google assistant," https://assistant.google.com/explore?hl=en_us, 2022.
- [35] "Samsung bixby," https://www.samsung.com/hk_en/apps/bixby/, 2022.
- [36] "Ultra-realistic voice cloning and text-to-speech," <https://lyrebird.ai/>, 2021.
- [37] S. Arik, J. Chen, K. Peng, W. Ping, and Y. Zhou, "Neural voice cloning with a few samples," *Advances in neural information processing systems*, vol. 31, 2018.
- [38] "Vocloner: standard voice cloning tool," <https://vocloner.com/>, 2022.
- [39] "Guidelines for automotive aftermarket qi chargers," <https://www.wirelesspowerconsortium.com/data/downloadables/9/5/3/20121001-guidelines-for-automotive-aftermarket-chargers-v-10.pdf>, 2012.
- [40] "Get google assistant on your android lock screen," <https://support.google.com/assistant/answer/9134021?hl=en>, 2022.
- [41] "Ti Qi Standard Wireless Charger BQ500210," <https://www.ti.com.cn/product/cn/BQ500210?qgpn=bq500210>, 2021.
- [42] "STWBC: Digital controller for wireless battery charger transmitters supporting Qi A11 topology," <https://www.st.com/en/power-management/stwbc.html>, 2021.
- [43] L. Malesani, L. Rossetto, P. Tenti, and P. Tomasin, "Ac/dc/ac pwm converter with reduced energy storage in the dc link," *IEEE Transactions on Industry Applications*, vol. 31, no. 2, pp. 287–292, 1995.
- [44] X. Chen and M. Kazerani, "Space vector modulation control of an ac-dc-ac converter with a front-end diode rectifier and reduced dc-link capacitor," *IEEE Transactions on power electronics*, vol. 21, no. 5, pp. 1470–1478, 2006.
- [45] R. Zhao, P. Wang, Y. Ma, P. Zhang, H. H. Liu, X. Lin, X. Zhang, C. Xu, and M. Zhang, "Nfc+ breaking nfc networking limits through resonance engineering," in *Proc. of ACM SIGCOMM*, 2020, pp. 694–707.
- [46] V. MURATOV, "Methods for foreign object detection in inductive wireless charging. qi developer forum, 2017."
- [47] "STC12C2052AD," <http://www.stmicro.com/STC/STC12C2052AD.html>, 2021.
- [48] "IRF540NPBF," https://www.mouser.com/datasheet/2/196/Infineon_IRF540N_DataSheet_v01_01_EN-1732489.pdf, 2021.
- [49] "TDK A11 Coil," https://product.tdk.com/en/search/wireless-charge/wireless-charge/tx-coil-module/info?part_no=WT505090-10K2-A11-G, 2021.
- [50] "TDK Rx Coil," https://product.tdk.com/en/search/wireless-charge/wireless-charge/rx-coil-module/info?part_no=WRM483265-10F5-12V-G, 2021.
- [51] "1N4002G," <https://www.mouser.hk/ProductDetail/Taiwan-Semiconductor/1N4002G-R0G?qs=%2FQ2q2Z%2FWFyOCLsW3PjwYQ%3D%3D>, 2018.
- [52] "Wireless Charging PWM Controller," <https://www.semtech.com/products/wireless-charging/linkcharge-ics/ts61005>, 2021.
- [53] "ESP32-C3-DevKitC-02," <https://docs.espressif.com/projects/esp-idf/en/latest/esp32c3/hw-reference/esp32c3/user-guide-devkitc-02.html>, 2021.
- [54] "LM7805," <https://www.ti.com/lit/ds/symlink/lm340.pdf>, 2021.
- [55] "AR824 Sound Level Meter," <https://www.amazon.com/SENSOR-Digital-Handheld-Decibel-Monitor/dp/B07RJDWK93>, 2019.
- [56] "Designing a Qi-compliant receiver coil for wireless power systems," <https://www.mouser.com/pdfDocs/TI-Designing-a-Qi-compliant-receiver-coil.pdf>, 2018.
- [57] "HT201 Tesla Meter Surface Magnetic Field Gauss Meter Tester," <https://www.amazon.com/Surface-Magnetic-Tester-Digital-Gaussmeter/dp/B018MPFF6I>, 2018.

[58] C. C. Finlay, S. Maus, C. Beggan, T. Bondar, A. Chambodut, T. Chernova, A. Chulliat, V. Golovkov, B. Hamilton, M. Hamoudi *et al.*, "International geomagnetic reference field: the eleventh generation," *Geophysical Journal International*, vol. 183, no. 3, pp. 1216–1230, 2010.

[59] "Samsung 15W Fast Charge," <https://www.amazon.com/Samsung-Charge-Wireless-Charger-Stand/dp/B07VG9JMG1>, 2019.

[60] "Baseus Wireless Charger Station 15W," <https://amz.run/56lZ>, 2018.

[61] "UGREEN Qi Wireless Charger," <https://www.amazon.com/UGREEN-Qi-Wireless-Charger-Compatible/dp/B096SC2ZQK>, 2021.

[62] "Apple MagSafe Charger," <https://www.amazon.com/Apple-MHXX3-AM-A-MagSafe-Charger/dp/B08L5NP6NG>, 2020.

[63] "Wireless charging test module tester," <https://www.aliexpress.com/item/1005002014025696.html>, 2021.

[64] "Google Cloud: Speech-to-Text," <https://cloud.google.com/speech-to-text>, 2021.

[65] "Product Overview MEMS Microphone T4064/T4081," <https://product.tdk.com/en/techlibrary/productoverview/mems-microphone.html>, 2021.

[66] L. Blue, L. Vargas, and P. Traynor, "Hello, is it me you're looking for? differentiating between human and electronic speakers for voice interface security," in *Proceedings of the 11th ACM Conference on Security & Privacy in Wireless and Mobile Networks*, 2018, pp. 123–133.

[67] K. S. Tharayil, B. Farshtindiker, S. Eyal, N. Hasidim, R. Hershkovitz, S. Houri, I. Yoffe, M. Oren, and Y. Oren, "Sensor defense in-software (sdi): Practical software based detection of spoofing attacks on position sensors," *Engineering Applications of Artificial Intelligence*, vol. 95, p. 103904, 2020.

[68] G. Zhang, X. Ji, X. Li, G. Qu, and W. Xu, "Eararray: Defending against dolphinattack via acoustic attenuation," in *Network and Distributed Systems Security (NDSS) Symposium*, 2021.

[69] M. Li and Y. Liu, "Underground coal mine monitoring with wireless sensor networks," *ACM Transactions on Sensor Networks (TOSN)*, vol. 5, no. 2, pp. 1–29, 2009.

[70] L. M. Ni, Y. Liu, Y. C. Lau, and A. P. Patil, "Landmarc: Indoor location sensing using active rfid," in *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications, 2003. (PerCom 2003)*. IEEE, 2003, pp. 407–415.

[71] Infineon Technologies, "IM69D130: High performance digital XENSIVTM MEMS microphone.pdf," <https://www.infineon.com/cms/en/product/sensor/mems-microphones/mems-microphones-for-consumer/im69d130/#!documents>, 2017, Last accessed October 22, 2021.

[72] G. Reitsma, M. Kouwenhoven, and A. Mosterd, "A low-power microphone preamplifier with emi canceling," in *Proc. of the 26th European Solid-State Circuits Conference*. IEEE, 2000, pp. 296–299.

[73] T. Shijo, K. Ogawa, M. Suzuki, Y. Kanekiyo, M. Ishida, and S. Obayashi, "Emi reduction technology in 85 khz band 44 kw wireless power transfer system for rapid contactless charging of electric bus," in *2016 IEEE Energy Conversion Congress and Exposition (ECCE)*. IEEE, 2016, pp. 1–6.

[74] M. Suzuki, K. Ogawa, F. Moritsuka, T. Shijo, H. Ishihara, Y. Kanekiyo, K. Ogura, S. Obayashi, and M. Ishida, "Design method for low radiated emission of 85 khz band 44 kw rapid charger for electric bus," in *2017 IEEE Applied Power Electronics Conference and Exposition (APEC)*. IEEE, 2017, pp. 3695–3701.

[75] K. Inoue, K. Kusaka, and J.-I. Itoh, "Reduction in radiation noise level for inductive power transfer systems using spread spectrum techniques," *IEEE Transactions on Power Electronics*, vol. 33, no. 4, pp. 3076–3085, 2017.

[76] B. Sim, S. Jeong, Y. Kim, S. Park, S. Lee, S. Hong, J. Song, H. Kim, H. Kang, H. Park *et al.*, "A near field analytical model for emi reduction and efficiency enhancement using an nth harmonic frequency shielding coil in a loosely coupled automotive wpt system," *IEEE Transactions on Electromagnetic Compatibility*, vol. 63, no. 3, pp. 935–946, 2020.

[77] L. Tan, Z. Tang, S. Wang, Z. Li, W. Zhao, R. Zhong, and X. Huang, "Design and optimization of parameters of electric vehicle's wireless power transmission system to decrease the influence of coexistence interference," in *2019 IEEE 2nd International Conference on Electronics Technology (ICET)*. IEEE, 2019, pp. 315–319.

[78] C. Song, H. Kim, Y. Kim, D. Kim, S. Jeong, Y. Cho, S. Lee, S. Ahn, and J. Kim, "Emi reduction methods in wireless power transfer system for drone electrical charger using tightly coupled three-phase resonant magnetic field," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 9, pp. 6839–6849, 2018.

[79] "Tdk analog mems microphone t4086," <https://invensense.tdk.com/download-pdf/t4086-datasheet/>, 2021.

[80] "Infineon digital mems microphone im69d130," 2021.

[81] "Tdk corporation overview mems microphone t4064/t4081," <https://product.tdk.com/en/techlibrary/productoverview/mems-microphone.html>, 2021.

[82] J. Sun, "Pulse-width modulation," in *Dynamics and control of switched electronic systems*. Springer, 2012, pp. 25–61.

APPENDIX A

BACKGROUND OF WIRELESS POWER TRANSFER

A. Principles of Wireless Power Transfer

Wireless power transfer (WPT, aka wireless charging) works on the principle of electromagnetic induction. Both the transmitter (i.e., a wireless charger) and the receiver (i.e., a smart device) are equipped with coils. Coils of wire in the transmitter create a magnetic field as the current passes through. Alternating magnetic fields can then induce an electrical current in any close-loop conductor nearby. If the conductor is the coil of a device's charging circuit, then the transmitter and the receiver are inductively coupled with each other via the magnetic fields. Thus, they effectively form a transformer with a specific coupling coefficient. The transmitter delivers power to the receiver through the transformer.

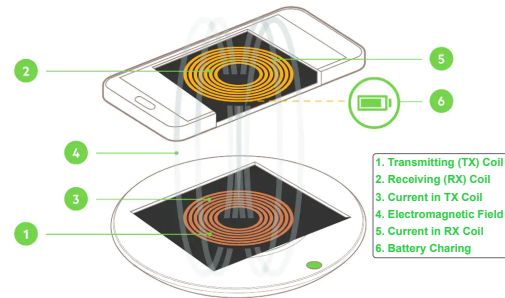


Fig. 25: Illustration of wireless power transfer.

As shown in Fig. 25, the transmitter (a charging station) and the receiver (a smart device) are inductively coupled with each other by the two coils to form a transformer with a specific coupling coefficient. The transmitter is composed of a TX coil and an inverter that is used to convert a DC low voltage (5 to 20V) power source to an AC high voltage of 50 to 100V. This AC voltage is used to energize the TX resonance tank circuit to create a tuned magnetic field frequency in the range of 100 kHz to 200 kHz. The receiver (RX) is also composed of an RX coil and a rectifier that converts the power harvested from the magnetic field back to DC power that can then be used to charge a battery. The receiver creates a resonant LC tank circuit to improve the power transfer efficiency by matching the TX response frequency. Accordingly, the maximum efficiency is given by

$$\eta_{\max} = \frac{k^2 Q_1 Q_2}{1 + \sqrt{1 + k^2 Q_1 Q_2}} \quad (14)$$

where Q_1 and Q_2 represent the quality factor of TX and RX coils, respectively, k is the coupling coefficient between the TX and RX coils. Clearly, the efficiency can be improved by

raising the quality factor. However, a higher Q will reduce the bandwidth.

B. Charging Workflow in Qi Standard

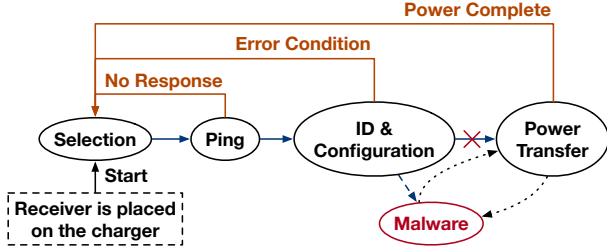


Fig. 26: The charging procedure specified in Qi standard

We must integrate the voice injection into a victim device in accordance with the Qi standard seamlessly so that the charging process is uninterrupted by the voice injection, which is key to keeping a dying victim device alive and preventing a smart device from detecting abnormalities. The Qi standard specifies a complicated charging workflow. For clarity, we show a simplified workflow in Fig. 26. Initially, the wireless charger starts with extremely low power for safe charging. By default, it is in a selection state, during which it checks the placement of a receiver. Once a smart device is placed, the charger turns into the ping state, where it sends out a digital ping pulse and listens for a response from a receiver. This step aims to detect the presence of a receiver rather than non-rechargeable matter. The charger then transits to the identification and configuration state, during which the two sides exchange the configuration information, such as the manufacturing model, battery capacity, acceptable maximum power, etc. Finally, the charger creates a power transfer contract (PTC), including the parameters of the power transfer. Afterward, the charger raises the power based on the PTC and starts to transfer power to the receiver. During the transfer state, the receiver can also adjust the PTC parameters to meet the battery requirement. From the figure, we could find that the best attack timing should be after the PTC has been established.

APPENDIX B BACKGROUND OF MICROPHONES

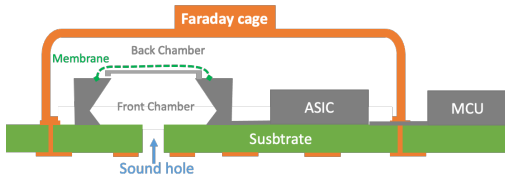


Fig. 27: On-board structure of a digital MEMS microphone

The MEMS microphones can be further divided into two types, analog MEMS and digital MEMS microphones. If the output of the microphone is an analog signal, it is called an analog MEMS microphone (like ADMP 401 and TDK

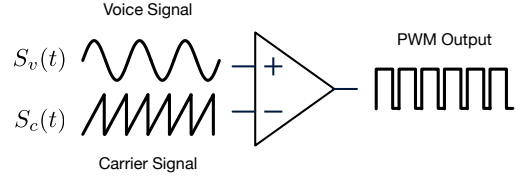


Fig. 28: The implementation of PWM modulation.

4086 [79]); otherwise, it is called a digital MEMS microphone (like Infineon IM69D130 [80]). The main difference is in that the analog MEMS microphone contains all components (i.e., transducer, filter, and amplifier) except ADC [81], while the digital MEMS microphone further integrates the ADC. For example, all Apple iPhones use analog MEMS microphones, and most Samsung phones use digital MEMS microphones. Either way, the workflows of the two types of microphones are similar.

A Fig. 27 shows a typical digital MEMS microphone contains a membrane and a complementary perforated back-plate. When a sound wave presents, the air pressure traveling through the holes triggers the mechanical vibrations at the diaphragm, which is a thin solid membrane that flexes in response to the change in air pressure. This mechanical vibration triggers a capacitive change of a capacitor, resulting in an AC signal. In this way, air pressure is converted into an analog acoustic signal for further processing. The acoustic signal is then amplified, filtered, and digitalized by the following ASIC chip. Finally, the digitalized acoustic signal goes out to an external acoustic microchip.

APPENDIX C PRINCIPLE BEHIND PWM SCHEME

The PWM scheme can be modeled as a function, which accepts two inputs. One input is the analog signal to be emulated, denoted by $S_v(t)$, and the other one is the carrier denoted by $S_c(t)$. Fundamentally, the PWM function is implemented using a comparator as shown in Fig. 28. The output of PWM is a digital signal with high and low amplitudes. Formally, the function is defined as follows:

$$\text{PWM}(S_v(t), S_c(t)) = \begin{cases} 1 & S_v(t) \geq S_c(t) \\ 0 & S_v(t) < S_c(t) \end{cases} \quad (15)$$

The most common PWM carrier is a sawtooth carrier, which ramps upward and then sharply drops periodically at a frequency f_c . The sawtooth carrier can be formally defined as follows:

$$S_c(t) = 2\pi f_c t - \lfloor 2\pi f_c t \rfloor$$

For clarity, we assume to emulate a single-tone sinusoid signal, i.e., $S_v(t) = \cos(2\pi f_v t)$. According to the derivation in [82], the output signal is:

$$\begin{aligned} \text{PWM}(S_v(t), S_c(t)) &= A_0 + A_1 \cos(2\pi f_v t) \\ &+ \sum_{m=1}^{+\infty} \frac{1}{m\pi} \{ \sin[m(2\pi f_c t)] - B_m \sin[m(2\pi f_c t) - m\phi_d] \} \\ &+ \sum_{m=1}^{+\infty} \sum_{n=\pm 1}^{\pm \infty} \frac{C_m}{m\pi} \sin \left[\frac{n\pi}{2} - m(2\pi f_c t) - n(2\pi f_v t) + m\phi_d \right] \end{aligned}$$

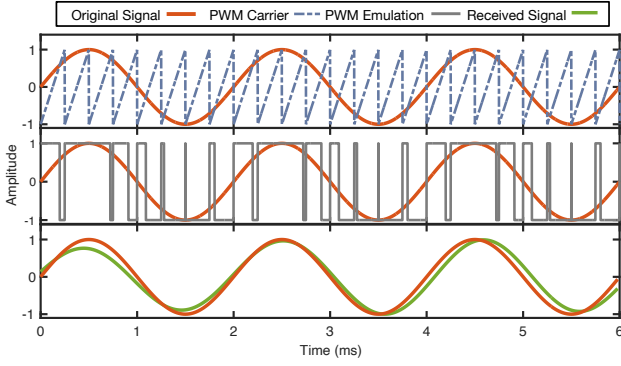


Fig. 29: Pulse-width Modulation (PWM)

where A_0, A_1, B_m, C_m are the constant amplitude gain and ϕ_d is the phase shift. To verify the effectiveness of the PWM-emulated voice, we use the signal generator to transmit a single-tone signal through a TX coil. Fig. 29 compares the original signal, the PMW-emulated signal, and the signal recorded by the MEMS microphone. The amplitude of all signals is normalized because we focus on if the voice signal can be successfully recovered here. It can be seen that the signal is almost recovered without any loss.

APPENDIX D ATTACK VOICE SPECTRUM

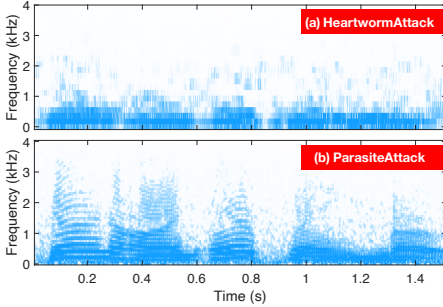


Fig. 30: Spectrograms of the voice signals. (a) and (b) show the voice injected HeartwormAttack and ParasiteAttack

APPENDIX E MEASUREMENT OF POWER CONSUMPTION

We employ a wireless charging power meter [63] to measure how much power is transferred from the wireless charger to the smartphone with or without parasite attacks. We show a measurement log in the below table. We can clearly observe that the overall power loss caused by the parasite attack is smaller than 500 mW. Thus, the presence of the parasite label is hard to be detected as a foreign object by using the power loss only.

APPENDIX F TASKS AVAILABLE TO VAS ON THE LOCKED SCREEN

Voice assistants can help launch various applications or execute operations. We investigate the security- and privacy-sensitive tasks available to mainstream VAs under the locked screen status. The detailed comparison result can be found in Table. V. Specifically, we explore the performance of four

TABLE IV: Wireless charging power test log.

Without parasite attack					
Time	Mode	U (V)	I (A)	Power (W)	Freq (kHz)
14:48:05	EPP	8.54	1.66	14.1	139.6
14:48:10	EPP	8.54	1.66	14.1	139.4
14:48:15	EPP	8.54	1.6	13.6	140
14:48:20	EPP	8.44	1.66	14.0	139.8
With parasite attack					
Time	Mode	U (V)	I (A)	Power (W)	Freq (kHz)
14:50:05	EPP	8.54	1.59	13.6	137.5
14:50:10	EPP	8.19	1.63	13.4	138.0
14:50:15	EPP	8.19	1.72	14.1	137.9
14:50:20	EPP	8.54	1.59	13.6	138.0

TABLE V: Available tasks for different VAs when the screen is locked.

Commands \ VA models	Apple Siri	Google Assistant	Samsung Bixby	Xiaomi Xiaoi
Make phone calls	✓	✓	✓	✓
Read messages or emails	✓	✓	✓	×
Send messages	✓	×	✓	✓
Send emails	✓	×	✓	×
Search the websites	✓	✓	✓	✓
Turn on/off WiFi	✓	✓	✓	✓
Turn on/off Bluetooth	✓	✓	✓	✓
Turn on airplane mode	✓	✓	✓	✓
Mute/Unmute the phone	✓	✓	✓	✓
Set/Delete alarms	✓	✓	✓	✓
Set/cancel appointments	✓	✓	✓	✓
Get location	✓	×	×	×
Open payment apps	×	×	×	×
Open social apps ¹	×	×	×	×
Access photos	×	×	×	×

¹ Social apps including Facebook, Whatsapp, and Wechat.

different mainstream voice assistants, i.e., Apple Siri, Google Assistant, Samsung Bixby, and Xiaomi Xiaoi respectively. We assume all voice assistants have been pre-set to allow interaction when locked by default. A total of 15 security-sensitive commands are selected as the test commands. We have the following findings: (1) Apple Siri has the largest range of capabilities, followed by Samsung Bixby. Siri and Bixby can execute 12 and 11 out of 15 commands, respectively even when the phone is locked. (2) VAs on all platforms can fully control the phone call, WiFi/Bluetooth connection, and schedule management when locked. These commands can raise potential security issues by either violating personal privacy or stealing sensitive information. (3) All platforms prohibit VAs access sensitive tasks, including opening payments, social apps, and accessing photos when the phone is locked. These tasks are highly sensitive and require passwords to proceed. (4) Besides Siri, other systems additionally restrict VAs access to the message and email. Overall, voice assistants can conduct many security-sensitive tasks even when the screen is locked. This fully validates our insight that voice assistants manipulated by malicious commands can be a severe danger to mobile security.

APPENDIX G

REAL-LIFE ATTACK SCENARIOS



Fig. 31: ParasiteAttack in our library. (a) the public wireless charger; (b) the parasite label stuck onto the wireless charger; (c) the parasite label with a cover-up is disguised as a signboard; (d) a victim is using our parasite label to charge his smartphone.



Fig. 32: ParasiteAttack in a subway station. (a) the public wireless charger stuck with a parasite label; (b) the wireless and the parasite label with a signboard cover-up; (c) a victim is using the parasite label to charge his smartphone.