

SoK: Acoustic Side Channels

Ping Wang
Xidian University
pingwangyy@foxmail.com

Shishir Nagaraja
Newcastle University
shishir.nagaraja@ncl.ac.uk

Aurélien Bourquard
Massachusetts Institute of Technology
aurelien@mit.edu

Haichang Gao
Xidian University
hchgao@xidian.edu.cn

Jeff Yan
University of Strathclyde
jeff.yan@strath.ac.uk

Abstract—We provide a state-of-the-art analysis of acoustic side channels, cover all the significant academic research in the area, discuss their security implications and countermeasures, and identify areas for future research. We also make an attempt to bridge side channels and inverse problems, two fields that appear to be completely isolated from each other but have deep connections.

Index Terms—acoustic side channel, covert channel, inverse problem, acoustic eavesdropping, attack, countermeasure

1. Introduction

Acoustic side channels (ASCs) have had a long history of interest. In the 1950’s, the British intelligence spied on acoustic emanation of an Egyptian embassy’s cipher machine [75]. This was a case of using sounds emitted by a Hagelin rotor machine for a side-channel attack, thereby recovering its secret key.

The National Security Agency in the USA also had a curious and keen interest in acoustic emanation for long. It was a part of their TEMPEST program, although unsurprisingly much of the program was on leaking electromagnetic emanations. According to the NSA’s NACSIM 5000 document [57], produced in 1982 and now partially unclassified, the TEMPEST documents NACSEM 5103, 5104 and 5105 are about acoustic emanations. But they remain classified. It was stated in [57] that ‘*Keyboards, printers, relays – these produce sound, and consequently can be sources of compromise*’, but no further details are provided.

In the unclassified world (academia and beyond), Briol [12] showed in 1991 that acoustic emanations of matrix printers carry, and thus leak substantial information about the printed text. Ten years later, UC Berkeley researchers Song et al. [70] observed that time intervals between consecutive keystrokes leak information about the keys typed. This would make an ASC, if and only if the inter-keystroke timing is captured via acoustics. Instead, the Berkeley team exploited their neat insight for a timing side-channel attack on SSH, which relied on (network) packet timing and would give them about a factor of 50 advantage in guessing a password. The study of keyboard emanation by Asonov and

Agrawal [4] in 2004 was a landmark paper on ASCs. In the same year, Adi Shamir et al. [65] announced in the rump session at Eurocrypt’04 that RSA decryption/signature operations running on a PC would sound differently for different secret keys. This suggested acoustic cryptanalysis become possible. It was unclear by then how to extract individual key bits from such acoustic emanations, until Shamir’s team (Genkin et al. [27]) figured out the technical details in 2014. Since 2004, the field of ASCs started to grow rapidly, with many academic papers being published in the years to come.

Our paper represents the first (comprehensive) effort in systematising knowledge of ASCs discovered to date. We aim to make the following contributions.

First, we will clarify some conceptual ambiguity within side-channel literature. Basic and key concepts are not defined adequately, or not at all. Consequently, the literature as a whole presents a confusing and sometimes chaotic picture. Some attacks are in fact not ASCs, but were treated as such; others are indeed ASCs but were not perceived as such. For example, does the Dolphinattack [79] exploit an ASC? Is Lamphone [55] an ASC attack? How do ASCs and acoustic covert channels (ACCs) differ? A number of authors have presented different and even contradicting views. To tidy up things, we will introduce intuitive definitions that are simple, clear-cut and easy to operationalise. We will also introduce rigorous formal definitions, when necessary. Moreover, we will put ASCs in perspective, clarifying ASC vs ACC vs signal injection attacks, and elaborate the boundary between similar-looking but fundamentally different attacks.

Second, we will establish a taxonomy to map out, structure and evaluate the ASCs discovered to date. We will also apply a structured framework to analyse countermeasures proposed to address these ASCs.

Third, we will perform a meta analysis of the state of the art, identifying its strengths and weaknesses. We will also offer new insights, and identify future research directions.

Moreover, we make an attempt to bridge side channels and inverse problems, two fields that appear to be completely isolated from each other but have deep connections.

2. Tidy Up the Mess

2.1. Ambiguity, Confusion and Possible Root Causes

It is not always straightforward to determine whether an attack is a side channel or not. Sometimes it can be tricky. Misconceptions have scattered around in the literature. For example, a well-cited paper on voice assistant security [22] mistakenly treated the Dolphinattack [79] as a side channel, although it is a signal injection attack which involved with no side channel. Similarly, the long-range dolphin attack [63] and the attack of ‘light commands’ [71] were classified as side-channel attacks in [22]. In fact, they are both not. On the other hand, some attacks (e.g. [21], [80], [81], [82]) were indeed ACSs, but their authors did not make it explicit at all. More examples can go on and on. One cannot help wondering: what have caused such ambiguity, confusion or even mistakes? Our contemplation leads to three possible root causes as follows.

Root cause 1: lack of a definition of side channels that is both widely applicable and easy to operationalize.

Many papers in the literature simply used the term of ‘side channels’ without any definition. This practice would work at early stages of the field, when the attacks were either a straightforward side channel or not, and many lookalike or related attacks were not invented yet. However, without a generally accepted and widely applicable definition, it will for sure invite for ambiguity and confusion.

On the other hand, many definitions of side channels are available in the literature, but they are different from each other, and are not very useful. Some are too narrow; perhaps more importantly, some are not *operational*—you cannot readily apply it to determine whether an attack is a side channel or not. We quote several definitions from the literature as follows.

‘An attack enabled by leakage of information from a physical cryptosystem. Characteristics that could be exploited in a side-channel attack include timing, power consumption, and electromagnetic and acoustic emissions.’ [56]. This NIST definition was driven by side-channel cryptanalysis, and it did not cover non-cryptanalytic side channels. It is also difficult to operationalize this definition.

‘Physical side-channel attacks extract information from computing systems by measuring unintended effects of a system on its physical environment.’ Used in a recent Oakland paper [26], this definition is hard to operationalize.

‘This can often be accomplished by means of a side-channel attack, whereby an unintended information source is leveraged.’ This definition was introduced in a recent Oakland SoK paper [51]. It is neat, but too brief, too abstracted, and operationally not very helpful.

‘... a side-channel attack is any attack based on information gained from the implementation of a computer system, rather than weaknesses in the implemented algorithm itself (e.g. cryptanalysis and software bugs).’ From Wikipedia, this definition is clearly driven by cryptanalysis and of a limited scope.

Root cause 2: side channels and covert channels have subtle differences, and some new attack class can further complicate this subtlety.

First, side channels and covert channels are two concepts that are related and easy-to-confuse. For example, Covert-Band [53] examined the privacy implication of tracking human movements with acoustics. It makes a clever covert channel leaking people’s privacy information, e.g. whether someone was in a room or not, or whether she was moving or standing still. But this is not a side-channel attack, as the leakage was not unintentional but on purpose.

Second, the definitions of side channels quoted earlier ALL fail to give an angle to differentiate between side channels and covert channels.

Third, as we will clarify later, some new class of attacks (e.g. active acoustic side channels) make it harder than before even for experts to tell whether they are a side channel or a cover channel.

Root cause 3: The surge of similar looking but different acoustic attacks has further complicated the conceptual ambiguity and confusion in the field.

Acoustic security has expanded rapidly and substantially in the recent years. Acoustic attacks such as the Dolphinattack [79], the long-range dolphin attack [63] and the ‘light commands’ attack [71], discussed earlier, represent only a single class of sources for confusions. There are more.

Another set of acoustic attacks eavesdrop and recover human speech by picking up vibrations via motion sensors, cameras, laser or lidar, e.g. [3], [32], [50], [51], [55], [62]. They represent another class of confusion sources. These attacks involved with side channels, but not necessary acoustic ones. For example, a gyroscope’s reading is sensitive to sound vibrations, and Stanford researchers Michalevsky et al. [50] used it to recover human speech. This is a side-channel attack, but not an acoustic one. Only when the vibration frequency is in a certain range (20~20KHz), the signal is acoustic. The Lamphone attack [55] recovers human speech by measuring vibrations of a light bulb caused by acoustic waves. However, it exploits an optical side channel, rather than an acoustic one, to recover the sound.

2.2. Our Definitions

A key aspect of side channels is unintended communication. Acoustic energy is present as wave energy within an air medium, or as vibrations within solid media. Formally, a *side channel is a communication channel which allows one-way information transfer from the targeted system to the attacker. A side channel is defined as a functional mapping $S : I \times M \mapsto O$, where I is the valid input system inputs, M is the attacker’s influence on S , and O are the observations made by an attacker monitoring the channel. The attacker’s goal is to infer I from observations O .* In an ASC, the system leaks information acoustically, i.e. observations O are made on an acoustic medium regardless of any adversarial influence M , the influence mechanism or the influence medium. Not all side channels involve adversarial influence, in which case S is termed as a *passive*

side channel (when $|M| = 0$). However, in the presence of adversarial influence $|M| > 0$, S is termed as an *active* side channel. Note that key distinguishing characteristic of an ASC is that the attacker can only observe the victim over the acoustic channel. The scenario where adversarial influence is over an acoustic channel whilst observations are made on a non-acoustic channel is not an ASC. When no information is leaked, i.e. O is *NULL*, then there is no side channel in existence, even if the adversary is able to influence the system. This is the dual of the active side channel and is termed as a *signal-injection attack*.

To address the ambiguity and confusion in the field (see Section 2.1), we have developed definitions by first organising the research landscape on the basis of attacker and defender capabilities or *threat-models* (See Figure 1). Threat models can be classified based on two factors namely the physical channel the attacker can access (eg. acoustic) and transmissions (receive-only (Rx), transmit-only (Tx), or send-and-receive (Rx,Tx)). An attacker is denoted as M_F^C where $C \in U$ is the set of channels the attacker can access out of the universal set of possibilities U , and F is a subset of $\{tx, rx\}$. The target $T_{F'}^{C'}$ is similarly defined in terms of channels accessed $C' \in U$ and channel functions $F' \in \{tx, rx\}$. The combinations of possible attacker and target profiles define the threat landscape.

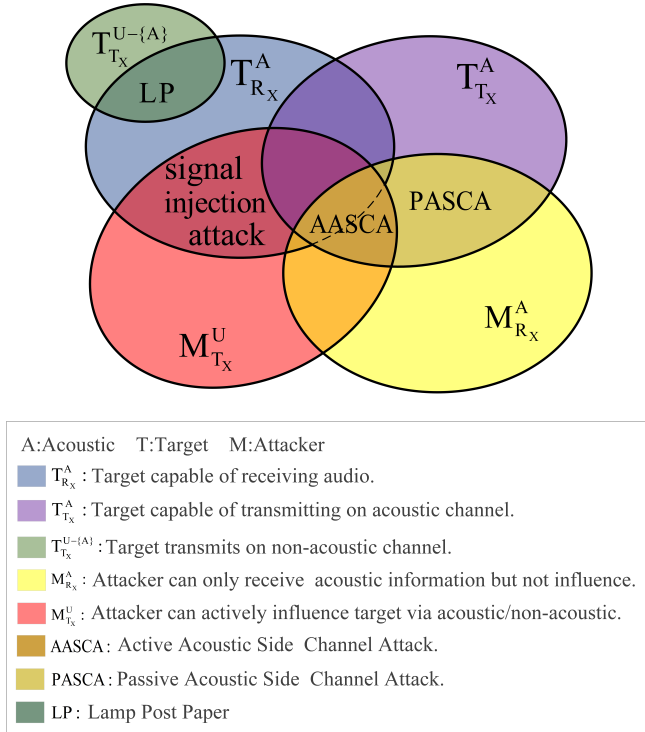


Figure 1. Venn diagram mapping threat models to side channel taxonomy

Side channels can be exploited either for *offensive* [4], [16], [21], [52], [81] or *defensive* purposes [7], [8], [58]. When used for attack, one of the channel endpoints, typically the source of information leakage is the defender,

while the sink is the attacker. Side channels as defenses are possible where the attacker is replaced by a defender.

A *Passive Acoustic Side-Channel Attack (PASCA)* is characterised by an attacker M_{Rx}^A who can only monitor the acoustic channel while the victim T_{Tx}^A only transmits. A PASCA is a receiver-only channel for the attacker and a transmission-only channel for the victim. Therefore, the victim cannot be influenced by the attacker. Figure 1 visualises the landscape and suggests boundaries between the various notions of abusing unintended communication channels on the basis of threat models. An *Active Acoustic Side-Channel Attack (AASCA)* is characterised by a victim $T_{Rx,Tx}^A$, who is unintentionally transmitting information over an acoustic channel and an attacker who can make observations. Additionally, the attacker can also influence the victim via another channel which can either be acoustic or non-acoustic to induce a change in leakage behaviour i.e. change the rate of leakage or what is leaked through the acoustic channel. The crucial difference from PASCA is that the attacker influences the victim. Note that influence can be either via non-acoustic or acoustic means, as long as the target leaks information acoustically, we have an ASC attack.

AASCA is relatively more powerful than its PASCA counterpart. The ability to influence a victim means that an attacker can induce leakage to optimise inference. On the other hand, a PASCA is stealthier since the attacker is not transmitting any information that can be used by the defender to detect and isolate the attacker. For example, the transmission of ultrasound or mechanical vibration by the attacker may be observed by the defender, thus making active attacks relatively detectable. A successful side-channel attacker must therefore draw a balance between the active and stealth components of their attack.

Different from side channels but related, an *Acoustic Covert Channel Attack (ACCA)* involves two or more attackers who are communicating over a channel that is unintentionally present i.e. the endpoints are intentional but the channel is unintentionally present. Thus a covert channel differs from side channels primarily in the functional mapping $S : I \times M \mapsto O$ in the following important way: in a side channel the function S is defined by the victim and the attacker has no control over it. In a covert channel, both ends are under attacker control, therefore the attacker can optimally define and implement S such that hidden information I can be readily inferred from observations O . In a covert channel, the leak is deliberate as the attacker controls both channel endpoints, whereas in a side channel the attacker does not control the source endpoint. Covert channels were first described by Lampson [43].

Due to their similarity, side channels and covert channels are often confused for one another. As one example, SonarSnoop [16] is a side channel rather than a covert channel attack. In SonarSnoop, speakers are used to emit human inaudible acoustic signals and the echo is recorded via microphones, turning the acoustic system of a smartphone into a sonar system. The echo signal from a user's finger movements can be inferred to steal Android phone unlock patterns. In this attack, indeed acoustic signals were

intentionally induced, but the researchers measured only echoes from finger movements, which did not deliberately leak information i.e. source endpoint is not under attacker control. As the transmission was accidental, SonarSnoop is a side-channel attack rather than a covert-channel attack.

Another comparative point is that in the case of side-channel-as-defense, the defender has no control over source behaviour. For instance, they cannot make changes to the keyboard in order to enable the generation of optimal acoustic signatures. However, that changes when we consider an *Acoustic Covert-Channel-as-Defense* (ACCaD). An example ACCaD would be the use of an unintentional communication channel between systems at the same security level perhaps to fulfill a monitoring function. To the best of our knowledge, no ACCaD has been proposed thus far.

Often, a direct measurement of the output from a side channel does not immediately give the information leaked via the channel. And the channel output is more like meta data, from which attackers deduce the leaked information in a sensible way to complete their attacks. An exception is transient execution attacks such as Meltdown [44] and Spectre [39], which are side channels that leak actual data, instead of meta data. In contrast, traditional micro-architectural side-channel attacks leak only meta data, such as memory access patterns.

3. Acoustic Side Channels: A Taxonomy

To classify ASCs, we consider leaking devices, the leaked signals, the media via which the leakage occurs, as well as the information leaked. We also consider various features of each ASC, such as whether it is an active or passive attack, whether it is used for offensive or defensive purposes, the attacker’s distance, and the signal properties. We propose a taxonomy as in Table 1, whereas its high-level structure is illustrated in Figure 2. Our taxonomy categories highlight the most interesting ASC characteristics.

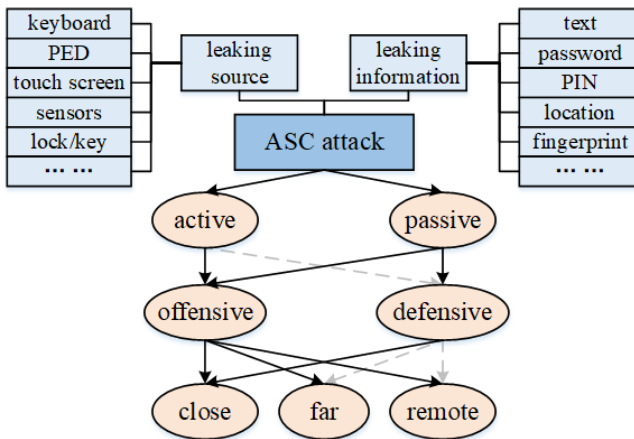


Figure 2. The structure of our ASC taxonomy: a high-level view (dash lines represent possible combinations but no such papers published yet)

3.1. Keyboard Emanation

Asonov and Agrawal [4] was the first to observe that each physical key has a unique acoustic (sound) signature as a fundamental property of keyboard design. Their main insight was that the physical plate beneath the keys causes each key to produce a different sound (frequency) depending on its location on the plate (similar to hitting a drum at different locations) thus these keystroke sounds can be used to steal what is being entered. Zhuang et al. [84] combined per-key acoustic fingerprints with a language model in an unsupervised learning setting (KMeans+HMM) improving inference efficiency from 52% to 67%. Berger et al. [10] introduced a comprehensive language model via a password dictionary.

An alternate to acoustic frequency spectrum is to leverage signal timing. Zhu et al. [83] observed that the relative time-of-arrival of an acoustic signal is dependent on the distance between the sensor and the originating keypress measured as the time-difference-of-arrival (TDoA) at attacker microphones placed 1m apart. Reported inference accuracy is 72%.

Combining both signal timing and acoustic features, Liu et al. [45], report a recovery rate of 94% of keystrokes. Their main insight was that combining signal warfare (TDoA) techniques with the frequency spectrum (MFCC) effectively replaced the benefits accorded by a language model, and simply running K-Means over the fingerprint vector was enough to cluster them by the key. This is significant since security practices around password construction may not permit content that is compatible with a language model.

Halevi et al. [31] evaluated the impact of typing styles in key recovery rates. They observed that while keys have unique sound signatures, touch typing significantly reduces the signal-to-noise ratio reducing recovery rates to 56% in the supervised case. They also found a significant decrease in key recovery rates when training and testing writing styles differ. The extent to which this applies to the unsupervised learning approaches above is unknown. Martinasek et al. [47] and Slater et al. [68] utilized neural networks to complete classification and Slater et al. found that deep learning approaches are well suited to the task of key recovery in noisy environments.

Specialist keyboards such as Pin Entry Devices (PEDs) and ATM/PoS keypads are equally vulnerable to key transcription attacks via sound side-channels and the attacks leverage the sound produced by a keypress on ATM keypads [61] and Enigma keyboards [72]. Cardaioli et al. [15] found that using inter-key delays extracted from signal arrival information works well too. This is an important improvement over Asonov’s sound-of-the-key approach, since it only uses signal timing information via a single sensor (as opposed to the multi-sensor TDoA approach of Zhu et al. [83]). Panda et al. [58] also recovered PIN keys from the keypress acoustic emanation, but they used the interval between two keystrokes as the main feature. In addition to exploiting this ASC for offensive purposes, the researchers in [58] also explored it for defensive purposes. Namely,

TABLE 1. ACOUSTIC SIDE CHANNELS: A TAXONOMY

Categories	Ref.	Source	Leaking Information	Audible	Purpose	Active	Intrusive	Proximity ¹	Sampling frequency
Keyboard emanation	Asonov'04 [4]	Physical keyboard	Typed text	✓	offensive	✗	✗	close, far	44.1KHz
	Zhuang'05 [84]	Physical keyboard	Typed text	✓	offensive	✗	✗	close, far	0.4~12KHz
	Berger'06 [10] Zhu'14 [83] Helavi'15 [31] Slater'19 [68]	Physical keyboard	Typed text	✓	offensive	✗	✗	close	44.1KHz
	Liu'15 [45]	Physical Keyboard	Typed text	✓	offensive	✗	✗	close	48KHz 192KHz
	Martinasek'15 [47]	Physical keyboard	Typed text	✓	offensive	✗	✗	close	48KHz
	Ranade'09 [61]	PED	Key taps	✓	offensive	✗	✗	close	44.1KHz
	Cardaioli'20 [15]	PED	Key taps	✓	offensive	✗	✗	close	48KHz
	Panda'20 [58]	PED	Key taps & User identity	✓	offensive & defensive	✗	✗	close	0.04~20KHz
	Enigma'15 [72]	Enigma keyboard	Key taps	✓	offensive	✗	✗	close	44.1KHz
Acoustic finger-tapping emissions	Narain'14 [54]	Touch screen	Typed text	✗	offensive	✗	✓	close	48KHz
	PIN Skimmer'13 [67]	Touch screen	Typed text	✗	offensive	✗	✓	close	16KHz
	Shumailov'19 [66]	Touch screen	Typed text	✗	offensive	✗	✓	close	44.1KHz
	Zarandy'20 [78]	Touch screen	Typed text	✗	offensive	✗	✓	close	48kKHz
Acoustic motion detection	SonarSnoop'18 [16]	Human-Computer Interaction	Gesture password	✗	offensive	✓	✓	close	48KHz
	KeyListener'19 [46]	Human-Computer Interaction	Typed text	✗	offensive	✓	✓	close	20kKHz
	PatternListener'18 [81] PatternListener+'19 [80]	Human-Computer Interaction	Gesture password	✗	offensive	✓	✓	remote	48KHz
VoIP hitchhiking ASC	Skype & Type'17 [20]	Keyboard	Key taps	✓	offensive	✗	✗	remote	44.1KHz
	Anand'18 [2]	Keyboard	Key taps	✓	offensive	✗	✓	close, remote	44.1KHz
	LendMeYourEar'22 [25]	EM fields (via acoustics)	Computation dependent leakage	✓	offensive	✗	✗	remote	48KHz
Physical location fingerprinting	Jeon'18 [35]	Electricity network	Physical location	✓	offensive	✗	✗	remote	1KHz
	VoIPLoc'21 [52]	Rooms	Physical location	✓	offensive	✗	✗	remote	44.1KHz
Acoustic device fingerprinting	Das'14 [21]	Internal sensors	Device ID	✓	offensive	✗	✓	close, far	8KHz 22.05KHz 44.1KHz
	Zhou'14 [82]	Internal sensors	Device ID	✗	offensive	✗	✓	close, far	44.1KHz
	Kotopoulos'14 [41]	Internal sensors	Phone module	✓	offensive	✗	✗	close	16KHz
ASC based on Device Hum	Briol'91 [12] Backes'10 [6]	Dot-matrix printer	Printed text	✓	offensive	✗	✗	close	96KHz
	Hojjati'16 [33]	3D printer & CNC mill	Proprietary IPR info	✓	offensive	✗	✗	close	44.1KHz
	Song'16 [69]	3D printer	Proprietary IPR info	✓	offensive	✗	✗	close	44.1KHz
	Faruque'16 [24] Chhetri'18 [19]	3D printer	Proprietary IPR info	✓	offensive	✗	✗	close	96KHz
	KCAD'16 [18]	3D printer	Control signals	✓	defensive	✗	✗	close	>40KHz
	Bayens'17 [7]	3D printer	Fill pattern	✓	defensive	✗	✗	close	44.1KHz
	Belikovetsky'19 [8]	3D printer	Audio fingerprint	✓	defensive	✗	✗	close	44.1KHz
	Synesthesia'19 [26]	LCD monitor (power bank)	Display contents	✗	offensive	✗	✗	close, far, remote	40KHz 192KHz
	Islam'18 [34]	Cooling fan	Electrical load	✓	offensive	✗	✗	close	8KHz
Physical-key leakage	SpiKey'20 [59]	Mechanical lock and key	Physical key	✓	offensive	✗	✗	close	44.1KHz
	Keynergy'21 [60]	Mechanical lock and key	Physical key	✓	offensive	✗	✗	close	44.1KHz 192KHz
Acoustic cryptanalysis	Genkin'14 [27] Genkin'17 [28]	Motherboard	Crypto keys	✗	offensive	✗	✓	close	21, 40, 48, 200, 350KHz
DNA synthesis	Oligo-Snoop'19 [23]	DNA synthesizers	DNA sequence	✓	offensive	✗	✗	close	48KHz

¹ The proximity between the attacker and the target. Close: the attacker is physically near the target (up to 3 meters). Far: typically 10 to 100 meters. Remote: the attacker can only access the target remotely, usually through a network connection.

the keystroke dynamics emitted via acoustics could work as behavioural biometrics for each user, offering additional protection for their PINs in theory.

3.2. Acoustic Finger-tapping Emissions

This category of attacks targets touchscreen keyboards on smartphones and tablets, instead of physical keyboards. When a user taps the screen, a fixed glass plate, with a finger, the tap generates a sound wave that propagates on the screen surface and in the air. Although signal strength is weaker than keystrokes from physical keyboards, it is well above the noise floor.

Early efforts were multi-modal—they combined acoustic information with other sources to isolate keypresses. Narain et al. [54] proposed a passive attack method to infer the text content created by taps on a touchscreen keyboard by using a Trojan application to capture sensed data from stereoscopic microphones and gyroscope. Simon et al. [67] developed PIN Skimmer which combines device microphones to detect touch events and device orientation information from the video camera inputs, to estimate the position of the tapped number.

The first to propose a fully acoustic passive ASC attack was Shumailov et al. [66] on touchscreen keyboards. They observed that acoustic waves passing through the glass bounce off the screen sides creating unique acoustic patterns observable from the internal microphones. Authors record the audio through the built-in microphones and demonstrate that simple TDoA allows the attacker to decipher PIN rows, while more complex machine learning models can use acoustic information to recover the actual PIN code, as well as, the text typed in.

Building on findings of [66], Zarandy et al. [78] observed that voice assistants such as Amazon Alexa and Google Home can be abused by an attacker to echolocate the sounds of a key tap on a different device. The authors demonstrate that it is possible to perform the attack up to half a meter away from the voice assistant.

3.3. Acoustic Motion Detection

An *active* attacker can exploit system behaviour by introducing a *new side-channel*. SonarSnoop [16] is the first such ASC attack for detecting finger motion. The attacker deploys malware on a victim’s smartphone to generate ultrasound chirps. By analysing echoes (chirp reflection), the dynamic motion of the fingers can be reconstructed in a fine-grained resolution to support recovery of pattern passwords. In this attack, the active component is the introduction of a stealthy sound-field outside human-audible range. The attacker exploits the property that the victim unintentionally modulates the attacker signal with confidential information. The unintentional transmission is a key characteristic of a side-channel. Zhou et al [80], [81] explored a similar approach to recover gesture passwords. Acoustic motion detection can also be used to localise virtual keyboard inputs. In 2019, KeyListener [46] developed an active ASC

attack that leveraged the change in Doppler effects due to finger movement within an induced sound field, to isolate touchscreen taps. All three works are active ASC as they require an active agent (malware or external device) to induce the sound field.

3.4. VoIP Hitchhiking ASC

It is natural to explore whether side channels can span (hitch-hike over) Voice over Internet Protocol (VoIP) sessions. Theoretically, this should be possible as human-voice frequency (20-20KHz) overlaps with keyboard sound frequency range (2-4KHz). Compagno et al. [20] confirm this via real-world experiments over the Skype network (Opus Codec) as long as the bandwidth is more than 20bps. The technical mechanism is largely based on the same attack components as prior art (MFCC-based acoustic signature features mated with a supervised learning inference mechanism). Anand et al. [2] confirm that keypads and ATM PEDs are equally vulnerable to key transcription side-channel attacks over VoIP sessions as they are close-proximity attacks. This means that scammers who get victims to hand over account information and then persuade them to walk over to an ATM to ‘check balance’ whilst on a call to the scammer, may steal their victim’s PIN as well as their account information.

More recently, Genkin et al.’s [25] observed that the built-in microphones of PCs can inadvertently capture computation-dependent leakage with electromagnetic (EM) fields within the computer even at a remote distance. It is possible because CPU computation leaks through audio signals. They demonstrated the efficacy by exploiting the leakage to perform attacks in three different scenarios—website identification, cryptographic key recovery, and multiplayer games cheating, via remote VoIP communication.

3.5. Physical-location Fingerprinting

When using VoIP to communicate, the created audios and data streams always include electrical network frequency (ENF) signals and other acoustic-reflection signals except for audible sounds. These signals always have specific characteristics and some important information, such as time and location. Therefore, it is possible to use those signals as signatures for location inference. Jeon et al. [35] proposed an attack to identify the physical location of where a target video or sound was recorded or streamed from. To achieve the attack, they first created a reference map of ENF signals extracted from the multimedia streaming data from a victim’s device via the microphone and then extracted the location information from the map by a two-step estimation. This work belongs to a passive way and is considered an ASC attack because all the targeted information is essentially leaked from the acoustic signals of multimedia streaming data. Different from those that require installing a specific malicious application on a victim’s device, this attack can be performed with existing VoIP applications or

online streaming services, which means the only data needed is a target multimedia file and it is non-intrusive.

Nagaraja et al. [52] proposed a passive attack (location fingerprinting technique) for a location inference on VoIP calls via ASCs, called VoIPLoc. Specifically, it exploited the acoustic-reflection characteristics of the physical space of a VoIP user. Using the speaker voice as the impulse signal, it extracted signals and then utilized a multi-layer classifier to map the fingerprint to a location.

3.6. Acoustic Device Fingerprinting

Microphones and speakers can be fingerprinted by variations in sensing and actuation respectively, introduced by variations in their physical properties. Das et al. [21] note that variations in the chemical compositions of diaphragm material, aging-related changes in the mount point, the glue used, wear-and-tear in manufacturing machines, humidity, and temperature levels during manufacturing all play a role in ensuring that no two microphones or speakers come off the assembly line working identically. Given an audio sample, they were able to trace 98% of the samples to the sensing device by using MFCC features of recorded audio. Both Zhou et al. [82] and Kotropoulos et al. [41] independently discovered the same phenomena and devised a speaker fingerprinting method based on high-frequency power spectrum. In summary, manufacturing imperfections have been successfully exploited to attribute audio recordings to specific devices.

3.7. ASC based on Device Hum

Printer hum: Often, electro-mechanical devices with moving physical parts are vulnerable to ASCs. Moving mechanical parts create vibrations that leak into the surroundings either as sound or as acoustic vibrations through the body of the device. In many cases, the movement of the mechanical components such as motors, fans, base plates, pins, and drums, is a function of user input leading to information leakage through acoustic channels. Briol [12] was the first to report an ASC in dot-matrix printers. Dot-matrix printers use multiple rows of needles. When printing a character, a subset of needles strike the paper surface mounted on a backing plate, a mechanical action that generates a sound wave. It turns out that printed characters generate a unique sound for each character printed (just as keyboards). It is therefore natural to expect that the approach and techniques developed for key transcription attacks are applicable to printer inference attacks. Backes et al. [6] confirm this—recording the sound from a microphone close enough to the printer, and passing it through a standard pipeline of basic signal processing to extract the short-term power-spectrum features (MFCC) in the relevant frequency band ($> 20\text{KHz}$). The main difference with keyboards, is the characters are printed at a higher rate than human keypresses. Due to this, acoustics of keys get mixed up due to time-overlapping signals. Interestingly, the sound of printers is above 20KHz band whereas keyboards emit sound at $2\sim 4\text{KHz}$ band. This

means key transcription and printer inference do not interfere with each other, and can be executed simultaneously, if required. In comparison with key-transcription attacks, printer information leakage is relatively less developed. We know of no works that apply time-difference-of-arrival of printer sound, learning-based inference, and signal-timing information (inter-character delay period). The application of these ideas may improve the state-of-the-art in printer transcription attacks, especially the issue of separating overlapped signals.

3D printer hum: Different from toner-based printers, 3D printers use a motorised filament extruder which deposits layers of material via an extrusion arm, whose location is controlled by multiple stepper motors to precisely control where filament is delivered on a base plate. The amount of current supplied to the various motors depends on the (confidential) printer input. Fundamentally, motors emit sound waves as a direct result of the current applied [14], arising first from *magnetostriction*: change in material dimensions in proportion to passing current in fixed electromagnets in the motor; *electrostriction*: change in dimensions of the conducting coil within the motor in proportion to current passing in rotor coil; and, third, in certain brushless and stepper motors, the air gap between rotor (rotating part) and stator (fixed part), varies drastically with rotor rotation while the radial forces causing rotation vary with current. In all three causes, the current applied (confidential printer input) causes a proportional change in the size of an air column, resulting the production of sound waves with frequency components originating from motor hum, stator hum, and coil hum. Faruque et al. [24] exploited this sound to propose the first attack against 3D printers. Using similar tools as keyboard side-channel attacks, namely the use of signal frequency features and supervised learning, they could extract the 3D printer style files corresponding to various objects with a recovery rate of 78% in FDM printers. This approach of exploiting motor acoustics to infer inputs applies to all 3D printers based on motors including FDM, laser sintering, and laser sintering. Note that unlike the sound of a key (on a keyboard), the sound of a 3d printed object does not have a fixed frequency fingerprint—motor, stator, and coil hum frequencies change based on current applied. For this reason, using MFCC to extract the frequency component is not the best approach. In follow up work, Chhetri et al. [19] replaced MFCC with MODWT (Maximal Overlap Discrete Wavelet Transform) to capture a better fingerprint, increasing recovery rate from 78% to 86%. Song et al. [69] use a smartphone stereo microphone and magnetometer together to better capture signal characteristics (Hojjati et al. [33] proposed the same for CNC milling machines). This approach has only incremental benefits since all motor inputs are already converted into acoustic sound due to magnetostriction, electrostriction, and radial forces on the rotor. Therefore combining acoustic with magnetic side-channels results in no fundamental improvement over audio side-channels. A number of works leverage acoustic side-channels to defend 3D printers. KCad [18] were the first to observe that integrity compromising attacks—false inputs in

STereoLithography (STL) files that encode the CAD model), the GCodes, or firmware compromise—necessarily lead to acoustic emissions. They successfully isolated 3D integrity compromising attacks through supervised models. Bayens et al. [7] leveraged acoustic and other spatial layers emanations to verify the unseen internal fill structure present in 3D printed objects. They used microphones to record the sounds leaking from printer mounts and housing and trained an audio classifier to recover GCodes using peak frequency and its temporal location within recorded acoustic data. Their defense can verify 40–60% of fill-pattern modification attacks. Belikovetsky et al. [8] build on both the above approaches, to extend the defense coverage to 100% of fill-modification attacks using a PCA over the spectrogram of recorded sound.

Display hum: The instantaneous power consumption of a display unit is a function of the screen content (processed in the raster sequence). This creates variations in the power supply causing power-circuit components to vibrate due to electrostriction resulting in a power-acoustic transducer. This property generalises well beyond display circuits to all digital circuits where current varies as a function of the workload. Synesthesia [26] developed a passive ASC attack that leverages power-acoustic transduction to extract images from the audio traces of the display power supply, captured by a microphone and accessed by a remote attacker over a VoIP channel. However, they use specialist equipment (a large parabolic signal collector).

Fan hum: A simple power-acoustic transduction occurs when heat triggers system cooling. Islam et al. [34] analyse fan noise to determine power consumption thus developing a timing power attack rooted in acoustic signal analysis.

3.8. Physical-key Leakage

Pin tumbler locks are widely used to secure homes and office spaces around the world. Recent work has developed methods to clone physical keys from the sounds emitted when a key is inserted. Ramesh et al. [59] proposed SpiKey, which exploits the fact that each pin in the tumbler makes a unique sound when depressed (just like a keyboard key). In follow up work, Ramesh et al. [60] combined the acoustic signal with visual information to achieve a key recovery rate of up to 75%.

3.9. Acoustic Cryptanalysis

Genkin et al. [27] introduced a passive acoustic cryptanalysis attack to extract full 4096-bit RSA keys with using the sound generated by the computer during the decryption of some ciphertxts. Using a phone or a sensitive microphone to record the sounds, the processed signals were then computed through a designed modular exponentiation which was based on the mathematical analysis of GnuPG (GNU Privacy Guard). Although this work has shown that different RSA keys induce different sound patterns that can be used to attack the keys, it was still not clear how to extract individual key bits. To address this issue, Genkin et

al. [28] further expanded [27]. The main improvement of the key extraction is the time decision computation when performing the additional multiplication for every key bit. Compared to the previous version, this work built more detailed experiments to analyze the relevant code of GnuPG and experimentally showed that this acoustic key distinguishability is also possible on other ciphers, such as AES and DES, and other versions of GnuPG.

3.10. DNA Synthesis

Faezi et al. [23] proposed the first ASC attack on DNA synthesizer, where compromising confidentiality will leak valuable information on nucleotide sequences. Two sound sources were leveraged: 1) the unstable noise radiation caused by vibration when the DNA synthesizer transports materials through the pipeline, 2) the audible click produced by the DNA synthesizer when it opens and closes the flow of material. In the threat model, the DNA synthesizer can be connected to computers, external drives, and Ethernet cables, and it is impossible to tamper with the machine or access the output DNA sequence. The attacker must place at least one microphone to the DNA synthesizer within close physical proximity, which is a passive but non-invasive ASC.

4. Countermeasures

To analyse countermeasures against ASC in a structured way, we use a three-dimensional framework, namely *Impediment*, *Interference*, and *Obfuscation*. They represent three different defense principles respectively: preventing access to the ASCs, interfering with the observed signals, and obfuscating the original sound pattern with noise. We summarise these countermeasures in Table 2, and note whether each of them was evaluated empirically.

4.1. Impediment

Considering that getting access to target devices/systems or collecting useful acoustic signals is a necessary precondition for ASC attacks, to stop attackers from acquiring such acoustics, i.e. Impediment, is naturally an intuitive defense. Approaches include noise-dampening material or blocking the malicious application before access.

Asonov et al. [4] explore impediment defenses based on keyboard structure. They observed that keys located at different positions on a single mechanical plate will produce unique acoustic fingerprints, like tapping a drum in different places. They suggested developing *silent* keyboards with multiple sound-dampening plates and locating keys in acoustically equivalent locations to mitigate the attack. Zhuang et al. [84] and Zarandy et al. [78] also discussed these ideas and claimed that for mechanical keyboard emanations, the use of a silent keyboard is not an effective countermeasure, as the signal is still above the noise floor, unless each key is mounted on a separate plate. Zarandy et al. [78] also mentioned that using phone cases or screen

TABLE 2. ACOUSTIC SIDE CHANNELS: COUNTERMEASURES

Acoustic side channels	Countermeasures									Evaluation
	Principles			Techniques						
	Im	In	Ob	Acoustic shielding	Stricter access control	Alert	Add noise	Randomization	Other techniques	
Asonov'04 [4]	✓			✓					Place the keys not in one plate	✓
Zarandy'20 [78]	✓	✓		✓			✓		Use phone cases or screen protectors	✗
Backes'10 [6]	✓			✓	✓				Longer distance	✓
Faruque'16 [24]	✓		✓	✓				✓	Make the motor loads similar	✗
Hojjati'16 [33]	✓	✓		✓			✓		Enlarge machines's enclosures	✓
Keynergy'21 [60]	✓	✓		✓			✓			✗
PIN Skimmer'13 [67]	✓				✓	✓				✗
Narain'14 [54]	✓				✓				Reduce sampling rate of the sensors	✗
SonarSnoop'18 [16]	✓	✓				✓	✓		Disable the sound system; modify sensor design	✗
PatternListener'18 [81] PatternListener+'19 [80]	✓		✓		✓	✓		✓	Limit the frequency range of the speaker and mic	✗
Shumailov'19 [66]	✓	✓				✓			Inject fake taps; introduce timing jitter	✗
Synesthesia'19 [26]	✓		✓	✓					Make variations on software mitigations	✗
Genkin'17 [28]	✓	✓	✓	✓			✓	✓		✗
KeyListener'19 [77]	✓		✓		✓			✓		✗
Oligo-Snoop'19 [23]	✓	✓	✓		✓		✓	✓		✗
Zhuang'05 [84]	✓	✓		✓			✓			✗
Anand'16 [1]		✓					✓			✗
Skype & Type'17 [20]		✓					✓		Perform a short random transformation	✓
Anand'18 [2]		✓					✓			✓
VoIPLoc'21 [52]		✓	✓						Use acoustic jitter and network jitter	✗
Song'16 [69]	✓	✓	✓	✓			✓	✓	Inject additional dummy tasks	✗

Im:Impediment, In:Interference, Ob:Obfuscation, ✓:partially evaluated.

protectors may provide some measure of protection against acoustic side-channel snooping.

In the case of 3D printers and physical locks (both low-frequency ASC), noise reduction is a direct and effective measure. Regarding countermeasures against ASC attacks on printers, Backes et al. [6] tested the effectiveness of using acoustic shielding foam, placing the microphone at a larger distance, and placing the printer in another room. They found that ensuring the absence of sound collections in the printer's room is sufficient to resist most eavesdropping. A similar countermeasure was also considered in DNA synthesizer defense by Faezi et al. [23]—prevent unauthorized person from entering the room. Faruque et al. [24] and Song et al. [69] also suggested that shielding the 3D printer with a sound-proofing material can be considered as a countermeasure. Hojjati et al. [33] recommended improving shield motors, such as using composites to cover the stepper motors in manufacturing equipment, can help protect it from broadcasting sensitive information to an adversary. They also stated that enlarging the machines' enclosures could help since magnetometer readings drop off with the cube of the distance from the source. In the case of physical keys, Ramesh et al. [60] suggested modifying the lock design, such as making the key with noise-reducing material and removing the vulnerable key.

Early approaches to implementing the impediment have been crude—both these works suggest notifying users of the existence of side channels—in effect, asking the user to solve the sensor deadlock problem. To impede PIN inference

attacks, Simon et al. [67] suggested using activity detection components at the OS level. When an activity is used to collect sensitive information from users, the component informs the OS and the OS will deny access to shared resources from other applications. Narain et al. [54] suggested blocking sensors in a mutually exclusive manner when a sensitive app runs. Cheng et al. [16] also proposed similar countermeasures to disable the sound system or notify users of a present sound signal in the high frequency range during sensitive operations to deal with gesture unlocking attacks which actively emit sound signals and use echoes to attack. Zhou et al. [80], [81] discussed preventing the microphone from being used in the background and limiting the frequency range of the speaker and microphone. However, all these works fail to discuss how to deal with deadlocks that will naturally arise such as when app A has locked the accelerometer and waiting for the camera and app B does the same in reverse order. Another defense proposed by [16] is to modify sensor design to limit the supported frequency range, but this is challenging because deciding the threshold for cutoff is hard. A third approach as Zhou et al. [80], [81], Yu et al. [77] and Shumailov et al. [66] proposed is to notify the user and let them deal with it by disabling sound and/or sensors except touch screen during sensitive operations, this also seems inappropriate indicating that there is much further work to be done in impediment-based access control research. For attacks of cryptographic key leaking and desktop display leaking, Genkin et al. [26], [28] propose acoustic shielding, however, this does not sit

well with the need for air circulation to cool the heat.

4.2. Interference

The working principle of interference defences is to drive the signal features the attack relies upon to well under the noise floor.

The ASC attack for keyboard input has reached a certain degree of accuracy—attackers are exploring different advanced signal processing and classification algorithms to continuously improve the effectiveness of the attack, therefore disrupting the feature construction and classification process is a basic way for defenders. Zhuang et al. [84] pointed out that quieter keyboards (Impediment) are useless. They believe that the ASC attack can be resisted by reducing the quality of the sound signal that the attacker may obtain, that is, increasing the noise. However, noise may also be separated, especially when faced with a microphone array attack, which records and distinguishes multiple channels of sound based on the location of the sound source. When an attacker is able to collect more data, this defense may also be ineffective. A smarter way proposed to add noise is to add a short noise window at each predicted peak, which may be more acceptable to users than continuous noise shielding. Anand et al. [1] proposed a defense mechanism against keyboard attacks which had good performance in the face of geometric measurement, feature classification, and other attack methods. The specific measure is to use background sounds to cover up the audio leakage.

The same is true for defense against remote attacks via VoIP. Compagno et al. [20] proposed to perform a short random transformation of the sound when a keystroke is detected. The intuitive method is to apply a random multi-band equalizer on multiple small frequency bands of the frequency spectrum or mix the victim’s microphone with a masking signal to prevent remote attacks. Anand et al. [2] also believed that a noisy defense mechanism is feasible by generating a masking signal with speakers at the victim’s end, and those strategies were experimentally proved to be effective in protecting victims’ important information.

Nagaraja et al. [52] also discussed a countermeasure for ASC attack on VoIP calls, while their target is to prevent location fingerprint leakage. Defenders may use acoustic jitter to damage the fingerprint information, such as using a constant amplitude signal at a room’s characteristic frequencies (50-2KHz) can cause a decrease in VoIPLoc’s performance. But it is hard to deploy because even small amounts of audible noise will negatively impact the voice quality, which is the first issue to be considered in VoIP.

In fact, this defense strategy of interfering with the original audio is effective for other different attack scenarios. Shumailov et al. [66] introduced timing jitter, or decoy tap sounds, into the microphone data stream to prevent attackers from reliably identifying tap locations when using virtual keyboards. As the taps themselves are pretty unnoticeable for humans, this should not disturb applications that run in the background. Another feasible countermeasure is to randomly play some distracting noises that are close to

pressing when the virtual keyboard is used [78]. Cheng et al. [16] suggested a possible countermeasure against active ASC attacks is to block the propagation of inaudible sounds, such as generating inaudible noise to interfere, and when possible, refuse to receive low-frequency or high-frequency sound signals.

The interference can still be applied to ASC attacks on 3D printers and physical key leaking. To protect 3D printing, Hojjati et al. [33] obfuscated the ASC emissions from manufacturing equipment by playing audio recordings of similar but flawed processes during production. Their experiments showed that such interference can make it harder for the attacker to separate the target audio stream from the others and reconstruct the object’s exact dimensions or process parameters. Song et al. [69] also suggested introducing more interference during printing. Ramesh et al. [60] thought that injecting noise to corrupt key insertion sounds is also a hopeful direction to improve security. Placing the machine in a noise environment has been discussed in Genkin et al.’s work [28], but the noise is easily filtered by a high-pass filter due to the low frequency (below 10kHz) of the generated noise. In the DNA synthesizer ASC scenario, Faezi et al. [23] also suggested introducing additional noise by adding redundant physical components.

4.3. Obfuscation

One significant factor that causes keyboard acoustic attacks is that the keyboard always has a unified key layout, which makes an attacker easily infer the keys since the fixed location results in a distance pattern. Creating some similar noise with the target acoustics or randomizing the keys’ location (soft keyboard) can obfuscate the signals, thus hampering an adversary to infer the information correctly.

This countermeasure is very useful and convenient to implement for the virtual keyboard on the touch screen, and it will not seriously affect the user experience. Compared with the physical keyboard, the layout of the touch screen virtual keyboard is easier to be customized, especially when inputting the PINs, the user’s input habits can be temporarily ignored. For KeyListener, it needs prior knowledge of QWERTY keyboard layout to map localized keystroke positions to accurate characters. Therefore, Yu et al. [77] proposed that generating a random layout of the QWERTY keyboard is an effective way to resist touchscreen keystroke eavesdropping attacks. For the on-screen gesture unlocking leakage, a similar defense is to randomize the layout of the pattern grid [80].

In addition to changing the position of the keys, randomization also plays a role in the defense against other attacks, such as cryptographic key leaking. Genkin et al. pointed out that their attack aimed at cryptanalysis can be prevented by some algorithmic countermeasures, such as ciphertext normalization and randomization [28].

As for computer screen leaking, attacks can be defended against by changing the screen content. Genkin et al. [26] proposed that a more promising approach is software mitigation. Specifically, these programs cover leaks by

changing the content on the screen, such as font filtering. By changing the font, all letters on the screen project the same horizontal intensity, avoiding the loss of information within a single pixel line. They also proposed two ways of shielding (impediment) and masking (interference), but these countermeasures are more difficult to achieve.

In fact, the defense strategy of obfuscation is also to prevent an attacker from extracting reliable information with distinct distinguishing characteristics. Nagaraja et al. [52] proposed a similar strategy, which is to use network jitter to induce packet latencies encouraging standard codec implementations to drop packets containing reverberant components, thus preventing the sender from extracting a credible room fingerprint. Moreover, Obfuscation can also be used for 3D printer and DNA synthesizer attacks. Faruque et al. [24] suggested that creating similar loads on each motor and incorporating random motor movements can obfuscate the acoustic emissions. Song et al. [69] considered adopting dynamic printing configurations in the process of G-code generation and injecting additional dummy tasks (e.g. use random trajectories). Faezi et al. [23] suggested that operators can randomly select redundant steps of varying time length prior to delivery or randomly select and execute steps unrelated to base delivery to obfuscate signals.

5. Discussions

We draw a number of interesting observations, which either reflect the strengths and weaknesses of the state of the art, or shed light on promising future research directions.

Ever expanding attack surfaces. Early work largely concentrated on physical keyboard emanation, and therefore targeted devices were PCs, laptops, payment devices and the like. The range of attack surfaces has been significantly expanded to date, covering smartphones, LCD displays, motherboards, mechanical locks, specialised equipment such as 3D printers and DNA synthesizers, and even computer-human interactions. Particularly, smartphones and 3D printers have attracted considerable attention in recent years.

Overall, keyboard emanations have been the most studied among the ASCs. The second most studied is touch-screen leaking; followed by 3D printer leaking. Those less-studied categories are likely to offer more opportunities for future research. Where else to look for new ASCs? New devices and equipment where noise and sound are emitted will deserve a look.

Data analysis and machine learning. The power of data analysis is critical for ASCs, as it hinges on the capability of extracting signals from often noisy data. There is a clear trend that ASC research evolved from simpler machine learning methods (e.g. probabilistic neural network, k-nearest neighbors, support vector machines) to more sophisticated deep learning (like convolutional neural network and recurrent neural network). As machine learning advances, it helps advance side-channel research.

However, it is unnecessary that the more sophisticated the machine learning methods, the better. The nature of signals and the features of datasets collected all play an

important role in choosing appropriate analysis methods. For example, Gohr [29] reported at CRYPTO'19 some impressive cryptanalysis results achieved by deep learning. However, Benamira et al [9] showed at Eurocrypt'21 that, after stripping down Gohr's deep neural network to a bare minimum, they achieved a similar accuracy using simple standard machine learning tools.

In cases where deep learning does outperform simple machine learning methods, the black-box nature of the former can cause interpretability issues. For example, it may be unclear why the deep learning method has worked. What is its weakness? And, how to improve it? For example, Benamira et al. [9] achieved a complete interpretability of their method and the decision process, whereas Gohr [29] fared poorly in explainability.

More nuanced nature of ASCs. Early ASCs were passive ones, but recently active ASCs emerged [16], [46], [81]. Active ASCs are intriguing, as they involve with both intentional and accidental elements. Although acoustic signals were intentionally introduced by an attacker in active attacks, the signal-responses from the victim unintentionally leak information.

Overall, most ASCs identified to date are passive ones, and only a few are active ones. Research into active ASCs is an interesting direction for future research.

We would not be surprised if many real-world attacks in the future will exploit a combination of active and passive ASCs, or exploit a combination of acoustic and other side channels, or simply amplify an ASC with non-side-channel attacks or vice versa. Certainly, researchers with imagination and creativity will be able to discover exciting new attacks along these directions, and only the sky is the limit.

Constructive applications of ASCs. Most research in this area employed ASCs for offensive purposes only, and several exceptions such as [7], [8], [18], [58] looked into constructive or defensive applications of ASCs. Panda et al [58] investigated both offensive and defensive aspects of ASCs, where they attempted PIN guessing via keyboard emanation, as well as user verification via keystroke dynamics, which is a known behavioural biometric. The basic idea of using ASCs to build security defenses is that acoustic signals emitted by devices can also be considered a fingerprint of the system or the program and used to protect the identification systems. It can be used alone or in combination with other protection mechanisms. This can be an exciting and promising direction for future research.

Imbalance in attack and defence research. The literature has put significant effort into discovering new ASCs and their exploitation, rather than investigating countermeasures to them. In fact, we could only name a small portion that covered and discussed countermeasures. For this very reason, Table 2 is significantly shorter than Table 1.

Inadequate evaluations of countermeasures. What is worse, among those investigating countermeasures, only a small portion attempted empirical evaluations. Most countermeasures proposed remain theoretical. Practical implementations and empirical evaluations are often limited, if any.

Clearly, countermeasure investigations, in particular their empirical evaluations, have been under-appreciated and inadequate. Countermeasures lag behind attacks, and this may well suggest that the former may be much harder to deliver than the latter. However, all these no doubt warrant fertile grounds for future research.

Research methodology. Experimentation is an intrinsic element of ASC research. However, experimental details are often under-reported in the literature. Thus, reproducibility can be a significant challenge.

Moreover, many studies were mostly controlled experiments, conducted in strict laboratory settings or similar environments. There was inadequate effort in considering or pursuing whether the results could be generalized to other settings, in particular to the naturalistic real-world setting. Still much effort is required to demonstrate the ecological validity of these ASC studies.

In terms of rigor and validity, ASC experiments in general are far behind the area of keystroke dynamics. Via a series of solid works including [37], [48], [49], [73], Maxion’s team at Carnegie Mellon meticulously examined and explored keystroke dynamics, and they achieved a high standard for repeatable, reproducible, well-grounded and generalizable experiments in security research. There is much for ASC researchers to learn from them.

Common metrics, reusable high-quality datasets, and standardized experimental setups and procedures (e.g. as shared operational protocols for experiments) all help to improve reproducibility. They will enable direct comparisons of attack or countermeasure research conducted by different teams. These will improve the rigor, validity and scientific foundation of ASC research, and advance the state of the art in an efficient and cost-effective way.

Lack of human, social and economic perspectives. Only a few papers (e.g. [1], [66]) considered usability and human factors, although some ASC countermeasures may potentially impact many users. On the other hand, monetary and computational costs incurred by potential countermeasures are rarely considered.

Side channels could be hugely serious, with a far-reaching social and economic impact at a large scale, e.g. multi-billion dollar consequences. For example, following the discovery of differential power analysis [40], smart cards had to be redesigned for banking and other stakeholders all over the world. The microarchitectural side-channels like Meltdown [44] and Spectre [39] suggested a major revisit of CPU designs, too. ASCs do not appear to be as serious.

However, how serious can and will ASCs be in the future? Some security economic analysis can be relevant and interesting. To have an answer, it is critical to understand the severity, practicality, and impact of the various acoustic side channels in the real world. Which acoustic side channels pose a real threat? Or, most of them will remain of academic interest only? There are many interesting open problems.

6. Bridging Side Channels and Inverse Problems

In unclassified worlds, side channels are a young field, with a history of less than forty years. Inverse problems have been studied for more than a century. However, side channels and inverse problems appear to be two fields that are completely isolated from each other¹.

A problem is *inverse* because it starts with the observable effects to calculate or infer the causes, such as determining causal factors and unknown parameters from a set of measurements of a system of interest. It is the inverse of a forward—or direct—(physical) problem, which starts with the causes and then deduces or calculates the effects, such as modelling a system from known parameters.

The field of inverse problems has deep and historical roots in mathematics, pioneered by giants like Hermann Weyl and Jacques Hadamard [30], [38], [74]. The main source of inverse problems is science and engineering. These problems have pushed not only the development of mathematical theories and tools, but also scientific and technological innovations in a wide range of disciplines, including astronomy, geophysics, biology, medical imaging, optics, and computer vision, among others. Classical achievements of inverse problems include computed tomography (CT) and magnetic resonance imaging (MRI), where the inverse Radon transform is foundational.

6.1. Side Channels versus Inverse Problems

In a side channel, information leaks accidentally via some medium or mechanism that was not designed or intended for communication. Often, a direct measurement of the output from a side channel does not immediately give away the information leaked. Instead, the direct output measurement is akin to metadata, from which attackers deduce the leaked information.

Therefore, **every side channel implies or involves an inverse problem, but not vice versa.**

In some instances, a side channel may involve a relatively straightforward inverse problem. For example, Kuhn demonstrated a classical optical side-channel, where the information displayed on a computer monitor could be reconstructed remotely by decoding the light scattered from the face or shirt of a user sitting in front of the computer [42]. A sophisticated attack was required to successfully exploit this side channel. However, its key insight was the fact that the whole screen information was available as a time-resolved signal, rather than solving a complex inverse problem. On the other hand, not all inverse problems involved in side channels are straightforward to solve. For example, active acoustic side channels such as SonarSnoop [16], KeyListener [46], and PatternListener [81] all involved a rather complex inverse problem.

1. Some analysis in this section were initially developed for [11].

6.2. Potential Impact on Side Channels

How do the fields of inverse problems and side channels inform each other? We believe that the problem-formalisation strategies, theoretical models, mathematical techniques, algorithms, and concepts developed in inverse problems have significant potential to benefit and inspire future research of side channels (including acoustic ones).

First, it helps to properly navigate between the languages used in both fields. This will, for instance, help to identify similarities and differences, to clarify misconceptions, and to unify terminologies. For example, *information*, which is the set of relevant parameters approximated by the solution to the inverse problem, conceptually differs from *measurements*, which are the physically leaked raw-data input of the inverse problem and which can contain various amounts of useful information.

In a unified language that is understandable to both communities, blocking a side-channel attack essentially amounts to making the corresponding inverse problem unsolvable, intractable, harder to model, or at least harder to compute efficiently. Accordingly, there are the following three scenarios where one could: (a) prove that the inverse problem becomes impossible to solve by getting rid of the information that is present in the measurements, in such a way that the analysed measurements contain nothing relevant; (b) make the inverse problem much harder to model mathematically or solve computationally; (c) get rid of the leakage (e.g. physically) so that there are no measurements to exploit whatsoever, regardless of whether the said measurements would have contained meaningful information or not. Adding random perturbations such as noise is an example of a classical mechanism that makes an inverse problem unsolvable or harder to model.

Second, the perspective of inverse problems offers a new lens for examining side channels. As first elaborated by Jacques Hadamard, a fundamental challenge in inverse problems is they are typically ill posed in terms of the solution's *existence*, *uniqueness*, and *stability*, whereas their corresponding forward problems may be well posed in all these regards [38]. The stability property means that a solution depends continuously on the available measurements (i.e. the observed data). Accordingly, a problem lacks stability if adding or removing data implies a radically different solution. If a computed solution lacks stability, it will simply depart from the true solution.

Some studies of side channels (e.g. [16], [17]) may amount to only proving the existence of a solution for the corresponding inverse problem, rather than investigating the two related properties, namely, uniqueness and stability. Therefore, looking into these other properties, as studied from the perspective of inverse problems, will likely give security researchers a new lens for examining side channels, as well as their countermeasures.

For example, examining the stability property alone warrants interesting research to answer the following questions. How will the side channel be impacted if less, or more, measurement data are collected for experiments? How much

measurement data is necessary for the side channel to be stable, in such a way that the retrieved information depends continuously on the data, as opposed to varying abruptly across nearly similar datasets? Could specific countermeasures, such as adding some type of physical disturbance or interference, influence the observed output from the side channel in such a way that stability decreases? Answers to these questions could allow better optimising side-channel countermeasures, accurately simulating their expected effect before implementing them (e.g. in the case of optical side channels as demonstrated in [11]), quantifying their efficiency, and providing a robust framework to compare them in a systematic and rigorous manner.

Third, some theoretical results on inverse problems are relevant to side channels. One such result is reconstruction guarantees for several types of problem structures, such as lower bounds on reconstruction errors (Cramér-Rao bounds [76]). These reconstruction guarantees are often only tied to the forward model mapping the relationship between the information of interest and measurements, in the sense that they do not depend on any specific algorithm or solution used. Another useful result is the extent to which the recovery is affected by noise or other non-idealities [5], [13]—which amount to mitigating side-channel attacks in security and cryptanalysis. Such results could inform one on how to best characterise various side channels—including acoustic, EM, and optical ones—and how to best design and evaluate their countermeasures. In particular, the interference and obfuscation countermeasures elaborated in Section 4 can substantially benefit from the perspective of inverse-problem research due to their operational nature, even though impediment and some elements of obfuscation countermeasures may be out of scope for inverse problems.

To solve challenging inverse problems, mathematics has been applied to accurately describe the forward model as well as assumptions on the solution, if any. For instance, sound statistical modelling allows reducing the dimensionality of the parameter spaces and producing accurate solutions [36], [64], and specific algorithms also allow maximizing computational efficiency. These may prove inspiring for side channel research, too.

Finally, it will be intriguing to explore possible connections between the optimality² of a side channel in a given scenario and the uniqueness and stability of the solution to the corresponding inverse problem. In some cases, it appears that the latter indeed implies an optimal side channel. However, in many other scenarios, whether such a connection holds or not has no straightforward answers. Instead, these will be interesting areas for future research.

7. Conclusions

We have seen steady progress in ASC research in the past twenty years. Some creative or even surprising results

2. By optimality, we mean that the maximum amount of information that can in theory be leaked from a side channel is fully extracted.

have emerged, such as acoustic cryptanalysis [27], keyboard emanation [4] and Synesthesia [26], to name a few.

We have laid down some foundations to clear conceptual chaos, and put together a framework to structure our collective understanding of existing ASCs and their countermeasures. We have also identified gaps in the research, which point to promising future directions.

We hope this paper sounds the marching bugle, attracting ambitious and creative researchers to further grow the field of ASCs, where imagination can make a difference.

Finally, we have made an attempt to bridge side channels and inverse problems. In general, every side channel implies (or involves) an inverse problem, but not vice versa. Although it may be a small step forward at this stage, it is perhaps the start of an aspiration that will grow in the future. We believe that this bridge has the potential to foster cross-field collaboration and inspire several new research directions, for example, building a more rigorous and effective scientific foundation for side channel research, and encouraging the possibility for ideas and techniques originated in one field to enjoy a wider applicability than was previously anticipated.

Acknowledgments

We thank Ilia Shumailov for his contribution, and Roy Maxion (Carnegie Mellon) for discussing experimental methods. PW and HCG were supported in part by the Natural Science Foundation of China under Grant 61972306 and by SongShan Laboratory under Grant YYJC012022005. This work was conceived and led by JY.

References

- [1] S. A. Anand and N. Saxena, “A sound for a sound: Mitigating acoustic side channel attacks on password keystrokes with active sounds,” in *International Conference on Financial Cryptography and Data Security*. Springer, 2016, pp. 346–364.
- [2] —, “Keyboard emanations in remote voice calls: Password leakage and noise (less) masking defenses,” in *Proceedings of the Eighth ACM Conference on Data and Application Security and Privacy*, 2018, pp. 103–110.
- [3] —, “Speechless: Analyzing the threat to speech privacy from smartphone motion sensors,” in *2018 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2018, pp. 1000–1017.
- [4] D. Asonov and R. Agrawal, “Keyboard acoustic emanations,” in *Security and Privacy, 2004. Proceedings. 2004 IEEE Symposium on*. IEEE, 2004, pp. 3–11.
- [5] R. C. Aster, B. Borchers, and C. H. Thurber, *Parameter estimation and inverse problems*. Elsevier, 2018.
- [6] M. Backes, M. Dürmuth, S. Gerling, M. Pinkal, and C. Sporleder, “Acoustic Side-Channel Attacks on Printers,” in *Proc. USENIX Security’10*, 2010.
- [7] C. Bayens, T. Le, L. Garcia, R. Beyah, M. Javanmard, and S. Zonouz, “See no evil, hear no evil, feel no evil, print no evil? malicious fill patterns detection in additive manufacturing,” in *26th {USENIX} Security Symposium ({USENIX} Security 17)*, 2017, pp. 1181–1198.
- [8] S. Belikovetsky, Y. A. Solewicz, M. Yampolskiy, J. Toh, and Y. Elovici, “Digital audio signature for 3d printing integrity,” *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 5, pp. 1127–1141, 2018.
- [9] A. Benamira, D. Gerault, T. Peyrin, and Q. Q. Tan, “A deeper look at machine learning-based cryptanalysis,” in *Advances in Cryptology—EUROCRYPT 2021: 40th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zagreb, Croatia, October 17–21, 2021, Proceedings, Part I 40*. Springer, 2021, pp. 805–835.
- [10] Y. Berger, A. Wool, and A. Yeredor, “Dictionary attacks using keyboard acoustic emanations,” in *Proceedings of the 13th ACM conference on Computer and communications security*, 2006, pp. 245–254.
- [11] A. Bourquard and J. Yan, “Differential imaging forensics: a feasibility study,” *arXiv preprint arXiv:2207.04548*, 2022.
- [12] R. Briol, “Emanation: How to keep your data confidential,” in *Proceedings of Symposium on Electromagnetic Security For Information Protection*, 1991, pp. 225–234.
- [13] L. Bungert, M. Burger, Y. Korolev, and C.-B. Schönlieb, “Variational regularisation for inverse problems with imperfect forward operators and general noise models,” *Inverse Problems*, vol. 36, no. 12, p. 125014, 2020.
- [14] D. Cameron, J. Lang, and S. Umans, “The origin and reduction of acoustic noise in doubly salient variable-reluctance motors,” *IEEE Transactions on Industry Applications*, vol. 28, no. 6, pp. 1250–1255, 1992.
- [15] M. Cardaioli, M. Conti, K. Balagani, and P. Gasti, “Your pin sounds good! augmentation of pin guessing strategies via audio leakage,” in *European Symposium on Research in Computer Security*. Springer, 2020, pp. 720–735.
- [16] P. Cheng, I. E. Bagci, U. Roedig, and J. Yan, “SonarSnoop: Active acoustic side-channel attacks,” *CoRR*, vol. abs/1808.10250, 2018. [Online]. Available: <http://arxiv.org/abs/1808.10250>
- [17] —, “Sonarsnoop: Active acoustic side-channel attacks,” *International Journal of Information Security*, vol. 19, no. 2, pp. 213–228, 2020.
- [18] S. R. Chhetri, A. Canedo, and M. A. Al Faruque, “Kcad: kinetic cyber-attack detection method for cyber-physical additive manufacturing systems,” in *2016 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*. IEEE, 2016, pp. 1–8.
- [19] S. R. Chhetri, A. Canedo, and M. A. A. Faruque, “Confidentiality Breach Through Acoustic Side-Channel in Cyber-Physical Additive Manufacturing Systems,” *ACM Transactions on Cyber-Physical Systems*, vol. 2, no. 1, p. 3, 2018.
- [20] A. Compagno, M. Conti, D. Lain, and G. Tsudik, “Don’t skype & type! acoustic eavesdropping in voice-over-ip,” in *Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security*, 2017, pp. 703–715.
- [21] A. Das, N. Borisov, and M. Caesar, “Do you hear what I hear? Fingerprinting smart devices through embedded acoustic components,” in *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, 2014, pp. 441–452.
- [22] J. S. Edu, J. M. Such, and G. Suarez-Tangil, “Smart home personal assistants: a security and privacy review,” *ACM Computing Surveys (CSUR)*, vol. 53, no. 6, pp. 1–36, 2020.
- [23] S. Faezi, S. R. Chhetri, A. V. Malawade, J. C. Chaput, W. Grover, P. Brisk, and M. A. Al Faruque, “Oligo-snoop: A non-invasive side channel attack against dna synthesis machines,” in *Network and Distributed Systems Security (NDSS) Symposium 2019*, 2019.
- [24] A. Faruque, M. Abdullah, S. R. Chhetri, A. Canedo, and J. Wan, “Acoustic Side-Channel Attacks on Additive Manufacturing Systems,” in *Proc. ICCPS’16*, 2016.
- [25] D. Genkin, N. Nissan, R. Schuster, and E. Tromer, “Lend me your ear: Passive remote physical side channels on {PCs},” in *31st USENIX Security Symposium (USENIX Security 22)*, 2022, pp. 4437–4454.

- [26] D. Genkin, M. Pattani, R. Schuster, and E. Tromer, "Synesthesia: Detecting screen content via remote acoustic side channels," in *2019 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2019, pp. 853–869.
- [27] D. Genkin, A. Shamir, and E. Tromer, "RSA Key Extraction via Low-Bandwidth Acoustic Cryptanalysis," in *CRYPTO'14*. Springer, 2014, pp. 444–461.
- [28] —, "Acoustic cryptanalysis," *J. Cryptology*, vol. 30, pp. 392–443, 2017.
- [29] A. Gohr, "Improving attacks on round-reduced speck32/64 using deep learning," in *Advances in Cryptology—CRYPTO 2019: 39th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 18–22, 2019, Proceedings, Part II 39*. Springer, 2019, pp. 150–179.
- [30] J. Hadamard, *Lectures on Cauchy's problem in linear partial differential equations*. Yale university press, 1923, vol. 15.
- [31] T. Halevi and N. Saxena, "Keyboard acoustic side channel attacks: exploring realistic and security-sensitive scenarios," *International Journal of Information Security*, vol. 14, no. 5, pp. 443–456, 2015.
- [32] J. Han, A. J. Chung, and P. Tague, "Pitchln: eavesdropping via intelligible speech reconstruction using non-acoustic sensor fusion," in *Proceedings of the 16th ACM/IEEE International Conference on Information Processing in Sensor Networks*, 2017, pp. 181–192.
- [33] A. Hojjati, A. Adhikari, K. Struckmann, E. Chou, T. N. Tho Nguyen, K. Madan, M. S. Winslett, C. A. Gunter, and W. P. King, "Leave Your Phone at the Door: Side Channels That Reveal Factory Floor Secrets," in *Proc. CCS'16*, 2016.
- [34] M. A. Islam, L. Yang, K. Ranganath, and S. Ren, "Why some like it loud: Timing power attacks in multi-tenant data centers using an acoustic side channel," *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, vol. 2, no. 1, pp. 1–33, 2018.
- [35] Y. Jeon, M. Kim, H. Kim, H. Kim, J. H. Huh, and J. W. Yoon, "I'm listening to your location! inferring user location with acoustic side channels," in *Proceedings of the 2018 World Wide Web Conference*, 2018, pp. 339–348.
- [36] J. Kaipio and E. Somersalo, *Statistical and computational inverse problems*. Springer Science & Business Media, 2006, vol. 160.
- [37] K. S. Killourhy and R. A. Maxion, "Comparing anomaly-detection algorithms for keystroke dynamics," in *2009 IEEE/IFIP International Conference on Dependable Systems & Networks*. IEEE, 2009, pp. 125–134.
- [38] A. Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems*, 3rd ed. Springer, 2021.
- [39] P. Kocher, J. Horn, A. Fogh, D. Genkin, D. Gruss, W. Haas, M. Hamburg, M. Lipp, S. Mangard, T. Prescher *et al.*, "Spectre attacks: Exploiting speculative execution," *Communications of the ACM*, vol. 63, no. 7, pp. 93–101, 2020.
- [40] P. Kocher, J. Jaffe, and B. Jun, "Differential power analysis," in *Advances in Cryptology—CRYPTO'99: 19th Annual International Cryptology Conference Santa Barbara, California, USA, August 15–19, 1999 Proceedings 19*. Springer, 1999, pp. 388–397.
- [41] C. Kotropoulos and S. Samaras, "Mobile phone identification using recorded speech signals," in *2014 19th International Conference on Digital Signal Processing*. IEEE, 2014, pp. 586–591.
- [42] M. G. Kuhn, "Optical time-domain eavesdropping risks of crt displays," in *Proceedings 2002 IEEE Symposium on Security and Privacy*. IEEE, 2002, pp. 3–18.
- [43] B. W. Lampson, "A note on the confinement problem," *Communications of the ACM*, vol. 16, no. 10, pp. 613–615, 1973.
- [44] M. Lipp, M. Schwarz, D. Gruss, T. Prescher, W. Haas, J. Horn, S. Mangard, P. Kocher, D. Genkin, Y. Yarom *et al.*, "Meltdown: Reading kernel memory from user space," *Communications of the ACM*, vol. 63, no. 6, pp. 46–56, 2020.
- [45] J. Liu, Y. Wang, G. Kar, Y. Chen, J. Yang, and M. Gruteser, "Snooping keystrokes with mm-level audio ranging on a single phone," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, 2015, pp. 142–154.
- [46] L. Lu, J. Yu, Y. Chen, Y. Zhu, X. Xu, G. Xue, and M. Li, "Keylistener: Inferring keystrokes on qwerty keyboard of touch screen through acoustic signals," in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE, 2019, pp. 775–783.
- [47] Z. Martinasek, V. Clupek, and K. Trasy, "Acoustic attack on keyboard using spectrogram and neural network," in *2015 38th International Conference on Telecommunications and Signal Processing (TSP)*. IEEE, 2015, pp. 637–641.
- [48] R. Maxion, *Making experiments dependable*. NSA, 2012, vol. 19, no. 2. [Online]. Available: <https://www.nsa.gov/portals/75/documents/resources/everyone/digital-media-center/publications/the-next-wave/TNW-19-2.pdf>
- [49] —, "Reproducibility: Buy low, sell high," *IEEE Security & Privacy*, vol. 18, no. 6, pp. 33–41, 2020.
- [50] Y. Michalevsky, D. Boneh, and G. Nakibly, "Gyrophone: Recognizing speech from gyroscope signals," in *23rd {USENIX} Security Symposium ({USENIX} Security 14)*, 2014, pp. 1053–1067.
- [51] J. V. Monaco, "Sok: Keylogging side channels," in *2018 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2018, pp. 211–228.
- [52] S. Nagaraja and R. Shah, "Voiploc : passive voip call provenance using acoustic side-channels," in *14th ACM Conference on Security and Privacy in Wireless and Mobile Networks 2021*, ser. WiSec '21, May 2021.
- [53] R. Nandakumar, A. Takakuwa, T. Kohno, and S. Gollakota, "Covert-Band: Activity Information Leakage using Music," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 3, p. 87, 2017.
- [54] S. Narain, A. Sanatinia, and G. Noubir, "Single-stroke Language-agnostic Keylogging Using Stereo-microphones and Domain Specific Machine Learning," in *Proc. WiSec'14*, 2014.
- [55] B. Nassi, Y. Pirutin, A. Shamir, Y. Elovici, and B. Zadov, "Lam-phone: Real-time passive sound recovery from light bulb vibrations," *Cryptology ePrint Archive*, 2020.
- [56] NIST, "Side-channel attack," [EB/OL], https://csrc.nist.gov/glossary/term/side_channel_attack Accessed Oct 17, 2021.
- [57] NSA, "NACSIM 5000: TEMPEST fundamentals," *National Security Agency, Fort George G. Meade, Maryland*, 1982. [Online]. Available: <http://cryptome.org/nacsim-5000.htm>
- [58] S. Panda, Y. Liu, G. P. Hancke, and U. M. Qureshi, "Behavioral acoustic emanations: Attack and verification of pin entry using keypress sounds," *Sensors*, vol. 20, no. 11, p. 3015, 2020.
- [59] S. Ramesh, H. Ramprasad, and J. Han, "Listen to your key: Towards acoustics-based physical key inference," in *Proceedings of the 21st International Workshop on Mobile Computing Systems and Applications*, 2020, pp. 3–8.
- [60] S. Ramesh, R. Xiao, A. Maiti, J. T. Lee, H. Ramprasad, A. Kumar, M. Jadhwal, and J. Han, "Acoustics to the rescue: Physical key inference attack revisited," in *30th USENIX Security Symposium*, Aug. 2021, pp. 3255–3272.
- [61] V. Ranade, J. Smith, and B. Switala, "Acoustic side channel attack on atm keypads," 2009.
- [62] N. Roy and R. Roy Choudhury, "Listening through a vibration motor," in *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services*, 2016, pp. 57–69.
- [63] N. Roy, S. Shen, H. Hassanieh, and R. R. Choudhury, "Inaudible voice commands: The long-range attack and defense," in *15th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 18)*, 2018, pp. 547–560.

- [64] J. A. Scales and L. Tenorio, "Prior information and uncertainty in inverse problems," *Geophysics*, vol. 66, no. 2, pp. 389–397, 2001.
- [65] A. Shamir and E. Tromer, "Acoustic cryptanalysis: On noisy people and noisy machines," *Eurocrypt2004 Rump Session, May*, 2004.
- [66] I. Shumailov, L. Simon, J. Yan, and R. Anderson, "Hearing your touch: A new acoustic side channel on smartphones," *arXiv preprint arXiv:1903.11137*, 2019.
- [67] L. Simon and R. Anderson, "PIN skimmer: Inferring pins through the camera and microphone," in *Proceedings of the Third ACM Workshop on Security and Privacy in Smartphones & Mobile Devices (SPSM'13)*. New York, NY, USA: ACM, 2013, pp. 67–78.
- [68] D. Slater, S. Novotney, J. Moore, S. Morgan, and S. Tenaglia, "Robust keystroke transcription from the acoustic side-channel," in *Proceedings of the 35th Annual Computer Security Applications Conference*, 2019, pp. 776–787.
- [69] C. Song, F. Lin, Z. Ba, K. Ren, C. Zhou, and W. Xu, "My smartphone knows what you print: Exploring smartphone-based side-channel attacks against 3d printers," in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, 2016, pp. 895–907.
- [70] D. Song, D. A. Wagner, and X. Tian, "Timing analysis of keystrokes and timing attacks on SSH," in *Proc. of 10th USENIX Security Symposium*, 2001.
- [71] 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 Association, Aug. 2020, pp. 2631–2648.
- [72] E. Toreini, B. Randell, and F. Hao, *An Acoustic Side Channel Attack on Enigma*. Computing Science Technical Report, Newcastle University, 2015.
- [73] M. A. Wetherell, S.-H. Lau, and R. A. Maxion, "The effect of socially evaluated multitasking stress on typing rhythms," *Psychophysiology*, vol. 60, no. 8, p. e14293, 2023.
- [74] H. Weyl, "Über die asymptotische verteilung der eigenwerte," *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, vol. 1911, pp. 110–117, 1911.
- [75] P. Wright, *Spy Catcher: The Candid Autobiography of a Senior Intelligence Officer*. Viking Adult, 1987.
- [76] J. C. Ye, Y. Bresler, and P. Moulin, "Cramer-rao bounds for parametric shape estimation in inverse problems," *IEEE transactions on image processing*, vol. 12, no. 1, pp. 71–84, 2003.
- [77] J. Yu, L. Lu, Y. Chen, Y. Zhu, and L. Kong, "An indirect eavesdropping attack of keystrokes on touch screen through acoustic sensing," *IEEE Trans. on Mobile Computing*, 2019.
- [78] A. Zarandy, I. Shumailov, and R. Anderson, "Hey alexa what did I just type? decoding smartphone sounds with a voice assistant," *arXiv preprint arXiv:2012.00687*, 2020.
- [79] G. Zhang, C. Yan, X. Ji, T. Zhang, T. Zhang, and W. Xu, "Dolphinattack: Inaudible voice commands," in *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 2017, pp. 103–117.
- [80] M. Zhou, Q. Wang, J. Yang, Q. Li, P. Jiang, Y. Chen, and Z. Wang, "Stealing your android patterns via acoustic signals," *IEEE Transactions on Mobile Computing*, 2019.
- [81] M. Zhou, Q. Wang, J. Yang, Q. Li, F. Xiao, Z. Wang, and X. Chen, "Patternlistener: Cracking android pattern lock using acoustic signals," in *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, 2018, pp. 1775–1787.
- [82] Z. Zhou, W. Diao, X. Liu, and K. Zhang, "Acoustic fingerprinting revisited: Generate stable device id stealthily with inaudible sound," in *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, 2014, pp. 429–440.
- [83] T. Zhu, Q. Ma, S. Zhang, and Y. Liu, "Context-free attacks using keyboard acoustic emanations," in *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security, CCS'14*. ACM, 2014, pp. 453–464.
- [84] L. Zhuang, F. Zhou, and J. Tygar, "Keyboard acoustic emanations revisited," in *Proceedings of the 12th ACM conference on Computer and communications security*, 2005, pp. 373–382.