Eye of Sauron: Long-Range Hidden Spy Camera Detection and Positioning with Inbuilt Memory EM Radiation

Qibo Zhang and Daibo Liu, *Hunan University;* Xinyu Zhang, *University of California San Diego;* Zhichao Cao, *Michigan State University;* Fanzi Zeng, Hongbo Jiang, and Wenqiang Jin, *Hunan University*

https://www.usenix.org/conference/usenixsecurity24/presentation/zhang-qibo

This paper is included in the Proceedings of the 33rd USENIX Security Symposium.

August 14–16, 2024 • Philadelphia, PA, USA

978-1-939133-44-1

Eye of Sauron: Long-Range Hidden Spy Camera Detection and Positioning with Inbuilt Memory EM Radiation

Qibo Zhang
Hunan University

Daibo Liu*
Hunan University

Xinyu Zhang
University of California San Diego

Zhichao Cao
Michigan State University

Fanzi Zeng*
Hunan University

Hongbo Jiang
Hunan University

Wenqiang Jin
Hunan University

Abstract

In this paper, we present ESauron — the first proof-of-concept system that can detect diverse forms of spy cameras (i.e., wireless, wired and offline devices) and quickly pinpoint their locations. The key observation is that, for all spy cameras, the captured raw images must be first digested (e.g., encoding and compression) in the video-capture devices before transferring to target receiver or storage medium. This digestion process takes place in an inbuilt read-write memory whose operations cause electromagnetic radiation (EMR). Specifically, the memory clock drives a variable number of switching voltage regulator activities depending on the workloads, causing fluctuating currents injected into memory units, thus emitting EMR signals at the clock frequency. Whenever the visual scene changes, bursts of video data processing (e.g., video encoding) suddenly aggravate the memory workload, bringing responsive EMR patterns. ESauron can detect spy cameras by intentionally stimulating scene changes and then sensing the surge of EMRs even from a considerable distance. We implemented a proof-of-concept prototype of the ESauron by carefully designing techniques to sense and differentiate memory EMRs, assert the existence of spy cameras, and pinpoint their locations. Experiments with 50 camera products show that ESauron can detect all spy cameras with an accuracy of 100% after only 4 stimuli, the detection range can exceed 20 meters even in the presence of blockages, and all spy cameras can be accurately located.

1 Introduction

Hidden spy cameras placed in sensitive locations such as hotels and dressing rooms are increasingly a threat to individual privacy [1, 2, 3]. A recent survey of travelers in the US revealed major concerns about hidden spy cameras, and an alarming percentage (11%) of Airbnb users have found a hidden camera in their short-term rental [4]. Modern spy cameras are highly miniaturized and can be easily installed anywhere in a private space or implanted into daily objects, such as smoke detectors, bathroom light fixtures, USB chargers, or power outlets (examples shown in Figure 1). Worse still, recent research [5] reported that a significant number of private devices, e.g., home security and surveillance cameras, are at risk of being manipulated by hackers, thus becoming spy cameras [6, 7, 8].

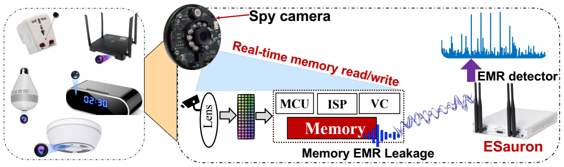


Figure 1: Examples of commodity hidden spy cameras and the principle of employing ESauron to detect them.

Unfortunately, it is difficult to detect such spy cameras due to their small form factors and/or stealthy working mode. Depending on the transferring methods of real-time video captures, a spy camera can be wired/wireless to a remote receiver or onboard storage. Recent research proposed to detect hidden cameras by searching for the tell-tale signs of tiny reflections from camera lens [9], or analyzing the traffic patterns of RF signals emitted from their radio interfaces [10, 11, 12, 13, 14]. However, these techniques still bear a few major limitations. First, the light-reflecting-based methods assume users are aware of the approximate location of the cameras, which is not always the case. A user usually has to carry a detecting device and scrutinize the entire private space, which requires a slow and meticulous sweep while exhibiting high false positives due to ambient reflections. Second, RF scanning-based methods are only applicable to hidden cameras that wirelessly stream their recorded videos, whereas a substantial portion of spy cameras rely on wired transmission or local storage, meaning that they do not initiate any wireless connections. All of the above evidence

[*] Corresponding Author.

suggests that there is still a lack of effective method to detect diverse forms of spy cameras.

This paper aims to reliably detect diverse forms of spy cameras (i.e., wireless, wired and offline devices) and quickly pinpoint their locations despite of these challenges. Our key insight is that despite diversified models and forms, all cameras for video recording will inevitably digest the captured raw images in real-time before transferring to target receiver or storage medium. The digestion process takes place in an inbuilt read-write memory (as illustrated in Figure 1) whose operations cause electromagnetic radiation (EMR). Therefore, we may be able to detect a functioning hidden spy camera if we can receive and identify its unique EMR traces over the air. Despite recent studies, e.g., EarFisher [15], Memscope [16], DeHiREC [17] and CamRadar [18], on EMR-based hidden device detection, however, due to ignorance of the critical dynamic traits (frequency drift) of memory EMRs in camera devices, they are ineffective in distinguishing and continuously tracking the time-varying EMRs from even a single device, not to mention common scenarios deployed with multiple cameras probably of the same model. Moreover, due to the weak of ADC EMRs produced by simple switched-capacitor circuits, DeHiREC and CamRadar's detection range is limited within dozens of centimeters, bringing insurmountable hurdle for device identification and localization when spy cameras are deployed high up in a room. Due to lack of thorough characterizing on memory EMR from camera devices, the inability to continuously differentiate and track camera devices of the same model, and lack of EMR-based positioning capability, existing works are still far from effective hidden camera detection in real scenarios. To address these problems and achieve the goal of ESauron, we need to answer the following research questions.

- **RQ 1:** What are the characteristics of the EMR traces leaked from a spy camera?
- **RQ 2:** How to effectively identify the extremely weak EMR leakage from a spy camera especially when it is at a distance and interfered by other devices also producing EMRs in similar spectrum?
- **RQ 3:** Given the EMR traces of an identified spy camera, can the weak and memory workload-relevant EMR signals be used for pinpointing its location under the multipath and shadowing effects?

We begin by conducting an analytical model and extensive measurements to address **RQ 1**. This involves analyzing memory EMRs resulting from image data processing, which emits EMR signals at the memory clock frequency. Our spectrum analysis identifies harmonic components around the memory clock. To address **RQ 2** and overcome the challenge of detecting weak EMR signals, we utilize the unique EMR spectrum characteristics. Our enhanced signal processing algorithm aggregates energy from multiple harmonic components, significantly improving the signal-to-noise ratio. This enhancement enables the accurate detection of weak EMR emissions even from considerable distances, even in the presence of obstacles. In the presence of interference from non-camera devices emitting EMRs in a similar spectrum, we establish a causal relationship between camera scene changes and responsive EMR patterns. This analysis eliminates the impact of interference, ensuring accurate spy camera detection. To distinguish mixed EMRs from multiple spy cameras, we conduct extensive measurements on 50 camera products. These measurements reveal unique memory clock fingerprints in the EMR spectrum, both static and dynamic. These fingerprints are then used to separate and track each spy camera's memory EMRs. Addressing **RQ 3**, we employ a received signal strength (RSS)-based iterative-approximation search algorithm to guide the receiver toward the spy camera's location. EMRs, operating at relatively low frequencies, are less susceptible to multipath and shadowing effects. This approach enables the accurate pinpointing of spy camera locations by following the RSS gradient.

We built a prototype platform of ESauron, with a laptop PC equipped with a plug-in miniature strobe light and a USRP B210 with a log-periodic antenna. Note that we are working on hosting our system on Raspberry Pi to miniaturize the platform equipment of ESauron.

Summary of contribution. Our contributions are summarized below:

- We propose ESauron, a first generalized system to detect and pinpoint all kinds of hidden spy cameras, including wireless/wire-connected and storage-based offline devices.
- By digging out the underlying cause of memory EMRs leaked by spy cameras and characterizing a series of unique properties, we design techniques to sense and differentiate memory EMRs, and assert the existence of spy cameras.
- We present a searching algorithm to heuristically search the appropriate direction to move forward to approach the spy camera. In this way, ESauron can quickly guide the user to the spy camera's location.
- We design a proof-of-concept system of ESauron with USRP B210 with a log-periodic antenna. We demonstrated the effectiveness of the proposed system design through a full-fledged testbed implementation and comprehensive experiments in real environments.

2 Background and Motivation

In this section, we first introduce the hardware architecture and processing logic of consumer-grade cameras. Then, we investigate the memory models used in spy cameras and explain the source of memory EMR.

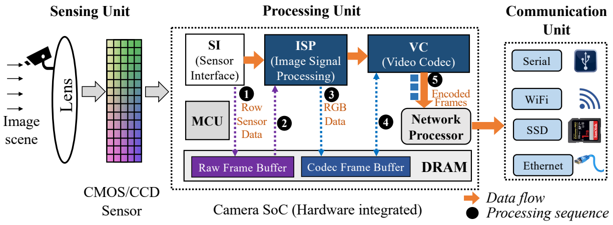


Figure 2: Hardware modules inside a customer-grade camera. A scene image through a lens is first digitized by a CMOS/CCD sensor. A dedicated video SoC chip is used for multimedia preprocessing, i.e., image signal processing (ISP), codec and transmission (networking processor)

2.1 Basic Operation of Video Cameras

Hardware architecture. Figure 2 depicts the hardware architecture of general consumer-grade cameras, comprising a sensing unit, processing unit, and communication unit. The sensing unit, typically a CMOS/CCD image sensor, captures and digitizes the target scenes to generate raw image data, and then passes them to the processing unit that is located on the camera system on chip (SoC). The processing unit contains: (1) an image signal processor (ISP) module that performs operations such as demosaicing, noise reduction, auto exposure, autofocus, auto white balance, and image sharpening designed to convert raw images into high-quality video frames; (2) a video codec (VC) module that compresses the video frames following standard protocols such as H.264/H.265. After that, the communication unit transfers the encoded frames to a target receiver or storage medium through wireless or wired paths, e.g., Ethernet, WiFi, or serial bus.

Table 1: Memory models used by different camera products.

Brand	Model	SOC chip	Memory frequency	Transfer Mode
Google Nest	Cam Indoor	Ambarella S2LM	DDR3-1600	Wireless\Storage
Logitech	Circle 2	Ambarella S2LM	DDR3-1600	Wireless
Arlo	Pro 3	OV 00798	DDR2-667	Wireless
Amcrest	IP2M-841W	Ambarella S2LM	DDR3-1600	Wireless\Storage
Hikvision	DS-IA	HK-2019	DDR3-1700	Wired
Hikvision	EzvizC2C	Hisilicon 3518	DDR3-1700	Wireless
Xiaoyi	Y1	QG2101A	DDR3-1033	Wired\Wireless
Dahua	LC-TC7	Hisilicon 3518	DDR3-1700	Wireless\Storage
Xiaomi	SXJ01ZM	Grain 8136	DDR2-1600	Wireless
Skyworth	c10	Hisilicon 3518	DDR3-1700	Wireless\Storage

Essential memory operation. From the perspective of the data flow, as illustrated in Figure 2, the raw images need to be first transferred and cached to the SoC memory. Then, the ISP module reads out the SoC memory and, together with a microprocessor (MCU), executes the image processing operations. The resulting RGB data (i.e., pixels) are further stored in the frame buffer within the SoC memory. On this basis, the VC module loads the pixel data from the memory and performs encoding, which involves scene analysis, motion estimation, macroblock classification, intra/inter-frame encoding, deblocking filter, etc. Almost every encoding step will exchange a mass of intermediate data with the SoC memory, among which motion estimation is most overwhelming. In other words, a dynamic change of the scene can drive a new round of motion estimation, which brings a notable increase in data exchange between the VC and the SoC memory, namely memory workload.

Note that existing consumer-grade cameras, no matter spy cameras, surveillance or ordinary camera, are with similar hardware architecture and processing logic.

2.2 Memory Models on Spy Cameras

Does any camera-device have an inbuilt read-write memory? For a camera device, despite diversified models and forms, the captured raw images must be first digested (e.g., encoding and compression) before transferring to target receiver or storage medium. Such encoding/compression processes have to exchange (read and write back) a mass of intermediate data with a specified memory. Thus a read-write memory is indispensable to any form of camera-devices, even the miniaturized spy cameras (as shown in Figure 16 below).

DRAM used in modern camera devices. DRAM has been widely adopted by consumer electronics due to its low cost, low power consumption, and high efficiency. Generally, DRAM can be divided into SDRAM (Synchronous Dynamic Random-Access Memory) and DDR SDRAM (Double Data Rate SDRAM, denote as DDR memory). Owing to its much higher data transfer efficiency, DDR memory has dominated the memory market [19]. Specific to the camera memory models, our survey in Table 1 reveals the same trend. We have purchased and examined 50 camera products from a major e-commerce portal. Our survey covers a variety of camera types, including miniature spy cameras, webcams, surveillance cameras, and home security cameras; and a variety of brands, such as Hikvision, Logitech, Arlo, Dahua, and some unknown models. We found all the cameras, without exception, are equipped with DDR memory (e.g., DDR2-667/1600, DDR3-1033/1600/ 1700) in their SoC.

2.3 Memory EMR Leakage

We now demystify the DDR memory EMR (referred to as memory EMR in the rest) and present the analytical model on the spectrum patterns of memory EMR.

Generation of memory EMR. DDR memory is composed of a large number of basic memory units, each consisting of a transistor and a capacitor. The charging state of

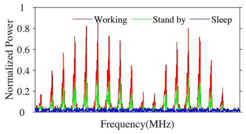


Figure 3: The spectrum mode of the memory EMR under different working conditions of the spy camera.

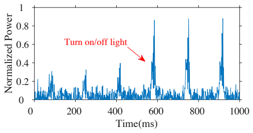


Figure 4: Distance-relevant RSS preservation.

the capacitor determines the memory unit's logic state, i.e., 1/0. When the DDR memory is read/ written, the memory clock immediately drives a large amount of switching voltage regulators' activities, by which the selected capacitors are charged/discharged by the transfer of electrons. Following the Faraday law of electromagnetic induction [20], the corresponding acceleration/deceleration of electrons produces EMR which resonates at the memory clock frequency [21].

If a certain read/write operation involves more memory units, the memory clock will drive more switching activities in proportion to the number of capacitors, resulting in stronger emanations at the memory clock frequency, and vice versa [21]. Therefore, the EMR amplitude is highly dependent on the memory workload.

Memory EMR analysis. The simplest form of a memory clock is a sine wave [15]. However, such a single-carrier clock can lead to excessive EMR intensity that may violate the regulatory requirement for electromagnetic compatibility. Modern clock generators instead adopt spread spectrum clock (SSC) techniques [22, 23], which distributes the EMR energy as $V_{ssc}(t) = A\cos 2\pi f_0 t + \frac{\Delta f}{f_m}\sin(2\pi f_m t)$, where f_0 is the center frequency of the memory clock, f_m and Δf are the modulation frequency and peak frequency offset, respectively. The descriptions of used symbols are listed in Appendix A. Due to stability requirements, the clock signal $V_{ssc}(t)$ has to go through a band-pass filter for harmonic components suppression. Then the energy of the memory EMR is non-zero only at frequencies f_{nz}, where the frequency of the i-th non-zero memory EMR can be expressed as

$$f_{nz}(i) = f_0 - i f_m. \tag{1}$$

Consequently, the memory EMR is composed of a series of *harmonic* components, where the frequency interval between consecutive harmonic components is f_m, and the first harmonic component is at f_0. Note that f_m is a constant and only depends on the memory hardware model. The theoretical analysis on memory EMR has been well explained in previous literature [24, 15].

Although existing techniques and regulations on electromagnetic interference and electromagnetic compatibility have made much effort to reduce the unintentional EMR leaked from DDR memory. However, it is still inevitable that DDR memory will produce EMR at the clock frequency when there are fluctuating currents.

2.4 Measurement of EMR Spectrum Patterns

We conduct empirical studies to verify the spectrum patterns of memory EMR. Our measurements are conducted with representative memory models, DDR2-800/ 1600, DDR3-1700/1866, and DDR4-2133/2400, which are widely equipped with spy cameras. We employ a USRP B210 with a log-periodic antenna to capture the EMR leakage. We configure the carrier frequency to be the memory center frequency f_0, set the sample rate to 2 Mhz, and take an FFT over 1s window of the captured signal to convert it into individual spectral components. Meanwhile, we create different workloads, i.e., write the memory intensively, put it in standby state with only periodical refresh operation, and put it to sleep state. We derive four observations from the measurement study.

First, from the analysis of spectrum patterns as illustrated in Figure 3, we observe that in both the working state and standby state, DDR memory operations do produce (leak) EMR signals, of which **the frequency spectrum features a series of energy peaks distributed over around 1 MHz near the memory clock frequency f_0 (e.g., 850MHz)**.

Second, an auto-correlation of the peak frequencies confirms that **the frequency interval (i.e., f_m) between consecutive peaks is constant**, agreeing well with the above analysis. For instance, f_m =31.25 kHz and 31.16 KHz, for DDR3-1700 and the DDR2-1600, respectively.

Third, the results also demonstrate that **the amplitude of memory EMR varies with the memory workload**. With the presence of intensive workload, the amplitudes of all energy peaks escalate significantly compared with the standby state. Yet the workload does not affect the locations of the frequency peaks, which corroborates the analysis in Eqn. (1).

Fourth, we further measure the correlation between scene change and the memory EMR. We find a **strong causal relationship between the stimulus of camera scene changes and the responsive EMR pattern**. By intentionally stimulating scene changes with turning on/off light, the spy camera immediately produces responsive EMR patterns as shown in Figure 4.

3 Threat Model

The attacker's goal is to surreptitiously record the video of privacy-sensitive environments with pre-deployed hidden spy cameras described in Figure 1. As a detector, ESauron tries to detect and quickly pinpoint the above hidden spy cameras.

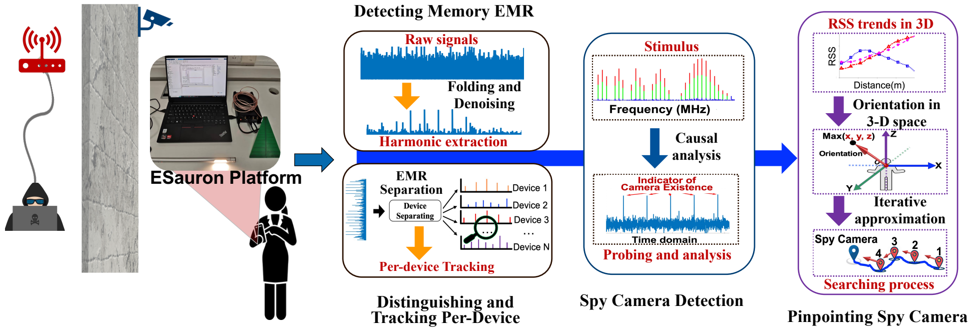


Figure 5: The architecture of ESauron, consisting of four major modules: Sensing of memory EMR, Separating and tracking per-device's EMRs, Stimulus-based spy camera detection, and Pinpointing of spy cameras.

3.1 Attack Model

We make the following assumptions about the attackers.

Concealment and deployment of the cameras. The attackers are able to hide the camera anywhere in the private space. For the sake of coverage, the attacker can install any number of spy cameras of the same or different brands.

Type of spy camera. The attackers can use different forms of spy cameras. Depending on the transferring methods of real-time video captures, the spy camera can be wired/ wireless to a remote receiver or stored in offline on-board memory. Moreover, we have no restrictions on the appearance and the manufacturer of the spy camera. In addition, the attacker may hack an innocuous camera (e.g., webcam or home security camera) used by the victim, and convert it into a spy cam.

Limited manipulation of camera hardware. We assume the attackers use consumer-grade cameras. They can change the camera configurations but are unable to modify the bottom memory system.

3.2 Capability of the Detection System

Equipment composition. The RF capturing device can capture the signal whose frequency band matches with that of the EMR signal of the spy cameras. A typical detection system of ESauron shall consist of the following devices: 1) broadband antenna that can capture the EMR signal, 2) software defined radio (SDR) that can down-convert and digitize the signal, 3) a plug-in miniature strobe light to bring continuous stimulus to trigger traceable and stable memory EMRs, and 4) PCs that can analyze the characteristics of EMR spectrum and run detection and localization algorithms.

Usage pattern of ESauron. The ESauron system can be carried by a user anywhere in a given privacy-sensitive room. To achieve a better detection performance, a user can request the other people at the scene to keep still or leave temporarily.

Environment. In a scene of privacy-sensitive environments, such as hotel, office and conference room, there will inevitably be various other electronic devices, especially some legal or authorized camera devices, which will also generate EMR and interfere with the detection. However, as a detector, the system should know the characteristics of the above-mentioned devices' EMRs to mitigate such interference and boost detection efficiency.

Detection range limitation. ESauron's detection capability is highly dependent to the strength of emitted EMR signals. A memory can be wrapped by a perfect electromagnetic shielding cage to prevent EMR leakage. ESauron's detection range depends on how well-established the electromagnetism shielding device (see Sec. 5.3.2).

4 ESauron Design

In this section, we presents the detail design of ESauron.

4.1 ESauron Overview

As depicted in Figure 5, ESauron consists of four major modules within its workflow: *(i) Detecting memory EMR:* ESauron employs a folding algorithm that leverages the unique spectrum characteristics of memory EMR to accurately capture EMR signals with high sensitivity; *(ii) Distinguishing and tracking per-device's EMRs:* ESauron then separates the memory EMRs of different devices, and tracks the EMR pattern of each potential spy camera based on their memory clock fingerprints; *(iii) Stimulus-based spy camera detection:* By intentionally inducing scene changes, ESauron excavates the causal relationship between the stimulation and responsive EMR pattern to assert the existence of a spy camera; *(iv) Pinpointing of spy cameras:* ESauron leverages both the sensitivity to scene changes and distance-relevant properties of memory EMR to quickly pinpoint the spy cameras' locations.

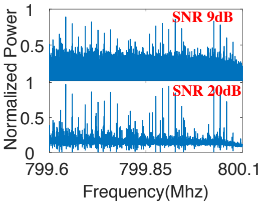


Figure 6: The spectrogram of original and denoised harmonic components.

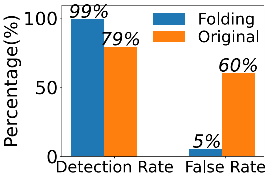


Figure 7: Folding-based memory EMRs extraction.

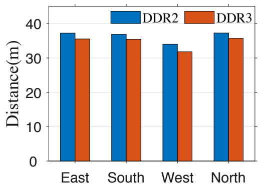


Figure 8: Folding-based memory EMRs detection range.

4.2 Detecting Memory EMR

We first present ESauron's EMR sensing module, which consists of folding-based memory EMR extraction and wavelet-based denoising.

4.2.1 Augmenting the strength of weak EMR signals

Recall that the memory EMR spectrum consists of a series of evenly separated harmonic components, of which the i-th harmonic component is located at the frequency of $f_0 - i \cdot f_m$. Based on this observation, we utilize the folding algorithm (see details in Appendix B), which is generally used for amplifying periodic astronomical signals, to search for possible peak frequency offset f_m to extract the harmonic components, and put all extracted EMR harmonic components with the same f_m into the same group.

To make the process computationally efficient, instead of performing a full spectrum scan, we investigated the typical center clock frequency ($f_0 \in G_{f_0}$) among the available DDR memory products in the market, and identify and checked the intervals between each visible harmonic components f_0 in the spectrum as the possible, referred to as $G_{f_0}^{f_m}$. In execution phase, ESauron iteratively selects a specific f_0 from G_{f_0}, and then scans the frequencies around f_0. Suppose $\mathcal{R}$ represents the series of N frequency samples in the memory EMR spectrum, and $\mathcal{R}[i][(\in [0,N])]$ is the amplitude of the i-th sample. The employing of folding is to search for and aggregate the energy from the EMR signals emitted from the same device, i.e., signals consisting of harmonic components separated by the same frequency samples offset, i.e., a possible candidate of $G_{f_0}^{f_m}$, denoted as f_m^c. To search for potential EMR harmonic components, the spectrum is first divided into small windows of samples with size f_m^c and then added in a window-wise fashion [25]

$$P_{f_m^c}[i] = \sum_{j=0}^{\lceil N/f_m^c \rceil - 1} \mathcal{R}[i + j \cdot f_m^c], \quad 0 \leqslant i < f_m^c. \tag{2}$$

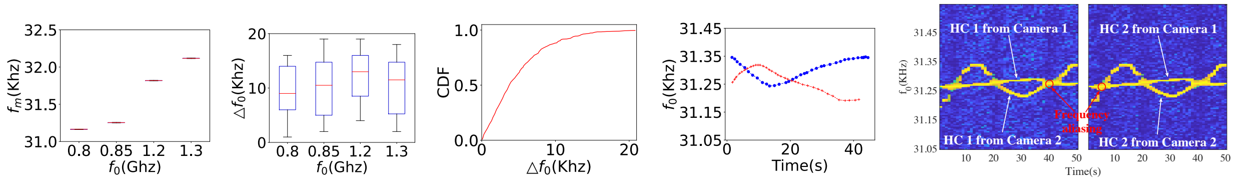
Through this folding mechanism, the energies of those harmonic components separated by f_m^c can be accumulated while the fused noise is likely smaller due to their non-periodicity in the frequency domain. The position of the folding peak, i.e., the i that maximizes $||P_T[i]||$, depends on the offset of memory's center clock frequency, i.e., Δf_0. For the application scenarios of ESauron, the actual f_m^c of potential spy cameras is unknown. So ESauron folds every possible f_m^c in $G_{f_0}^{f_m}$ to search for each harmonic component in the memory EMR spectrum. Even if some harmonic components may be completely overwhelmed by noise, ESauron still deduces the potential location of the harmonic component by Eqn. (1). In this way, ESauron can accurately extract all potential harmonic components sourcing from memory EMRs for follow-up processing. Those harmonic components with the same f_m are grouped into a set G_{f_m}. Note that these harmonic components could still originate from different devices with the same memory model. We will explain how to separate them in Sec. 4.3.

4.2.2 Wavelet based denoising

Even after the folding process, residual noise exists within the extracted harmonic components in G_{f_m}. The noise could come from various sources such as nearby electric devices and the camera's intrinsic noise. To eliminate such noise, we employ wavelet denoising [26] to eliminate the residual noise from the extracted harmonic components in G_{f_m}. The signal can be decomposed into detail coefficients (α_k) reflecting high-frequency information and approximate coefficients (β_k) reflecting low-frequency information through multi-layer wavelets. Through this multi-resolution decomposition, the peak frequency of memory EMRs and noise usually have different expressions at different layers of α_k and β_k. Finally, a level-dependent reconstruction is employed using all the coefficients as:

$$x_n = \sum_{k \in Z} \alpha_k^{(J)} \bar{g}_{n-2^J k}^{(J)} + \sum_{l=1}^{J} \sum_{k \in Z} \beta_k^{(l)} \bar{h}_{n-2^l k}^{(l)}$$

where $\bar{g}$ and $\bar{h}$ are rescaled discrete orthogonal functions. The spectrogram of a typical EMR harmonic component before and after the processing stage is shown in Figure 6, where the SNR is improved substantially, from 9 dB to 20 dB. The wavelet-based noise reduction is actually achieved by additional processing like thresholding or filtering the wavelet coefficients decomposition. By setting coefficients below certain amplitude thresholds to zero under the assumption



(a) Invariant f_m

(b) Δf_0

(c) CDF of Δf_0

(d) Clock drift

(e) Fluctuation trend of HCs

Figure 9: The device-specific static and dynamic traits. (a) f_m is only significant across different devices; (b) f_0 differs even across the identical devices; (c) f_0 is randomly distributed between 0 and 30 kHz; (d) Each camera device has a unique clock drift pattern; and (e) Fluctuation of harmonic components' (HC) drift pattern among two spy cameras (SP).

they manifest noise, crucial denoising is realized. Appropriate threshold selection is key to balancing signal distortion against noise removal.

To validate the effectiveness, we deploy different kinds of devices configured with a known DDR memory and then use both the folding algorithm and wavelet processing to extract the affiliated EMR harmonic components. As shown in Figure 7, compared with the energy threshold-based detection method, the folding algorithm can improve the detection rate from 79% to 99%, which means ESauron almost never misses any harmonic components. On the other hand, the false detection rate (i.e., mistaking other harmonic components for the target memory EMR) is decreased from 60% to around 5%. The residual error can be further reduced using the static/dynamic traits unique to each device (Sec. 4.3). Besides, the harmonic component energy after denoising can also significantly increase the SNR of received EMR and improve sensing distance. As illustrated in Figure 8, for both the DDR2-1600 and DDR3-1700, which are the most widely used memory models by spy cameras according to our investigation, the detection range can be up to 30 meters.

4.3 Distinguishing Per-Device's EMRs

ESauron further leverages both static and dynamic device-specific traits to differentiate different devices' EMRs in each group of EMR harmonic components (i.e., G_{f_m}).

4.3.1 Device-specific EMR traits

Static traits: center clock frequency offset. For the static traits, the offset of f_m is invariable for the same devices, whereas sufficiently diverse across different devices as shown in Figure 9(a), which should be attributed to the different modulation frequencies of their clock generators. In comparison, due to the imperfect manufacturing process, there exists an intrinsic center clock frequency offset $\triangle f_0$ between DDR memory products even with the same model as shown in Figure 9(b).

To examine whether the static $\triangle f_0$ is sufficiently diverse, we select an arbitrary pair of identical DDR memory products, denoted as D_i and D_j. Due to the similar f_m, the separability of their EMR features depends only on $\triangle f_0$. If the $\triangle f_0$ is too small, the two devices' harmonic components will overlap in frequency. We observe that the bandwidth of all harmonic components is within 300 Hz, which is consistent with previous research [15] and significantly smaller than $\triangle f_0$, that randomly distributes between 0 and 20 kHz as shown in Figure 9(c), in other words, implying that the static traits of the memory EMR spectrum are device specific and can be exploited to identify and distinguish different devices. However, due to the randomness of the memory's center frequency f_0, when $\triangle f_0$ is no larger than the bandwidth of harmonic components, the static traits alone can not effectively separate the coexisting devices.

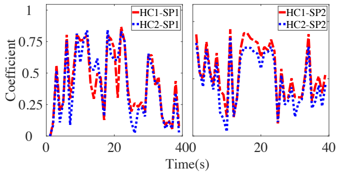


Figure 10: Fluctuation of the correlation coefficient of each harmonic component.

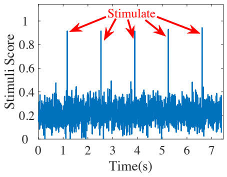


Figure 11: Causality analysis on stimulus-triggered EMRs patterns changes with t-test.

Dynamic traits: time-varying clock drift patterns. For the dynamic traits, due to the uncontrollable heating effect inside memory caused by read/write operation [27, 28], the memory clock f_0 experiences continuous change in amplitude and phase in the time domain as illustrated in Figure 9(d), leading to time-varying f_0 for a specific device. Due to the spread spectrum clocking, all affiliated harmonic components $V_{ssc}(t)$ on a device originating from the same clock source $V_{clk}(t)$, hence producing the same fluctuation trend for f_0. Figure 9(e) plots the traces of frequency drift of two devices with the same DDR memory. We observe completely different fluctuations, implying that even when two devices have identical static traits, namely (f_0, f_m, $\triangle f_0$), their EMR harmonic components can still be effectively separated by exploiting the dynamic traits. The highly diverse EMR patterns present device-specific traits for discriminating multiple potential spy cameras.

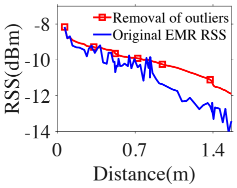


Figure 12: Distance-relevant RSS after Kalman filtering.

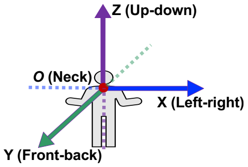


Figure 13: User-centered 3-dimensional reference system.

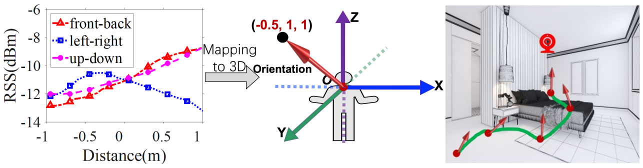


Figure 14: Example on the principle of mapping EMR RSS trend in three dimensions (left part) to an orientation vector in user-centered 3-D reference system (middle part); Iterative process (right part)

4.3.2 Distinguishing memory EMRs

We design an iterative process to leverage the above device-specific static and dynamic traits, to separate devices whose EMR harmonic components are clustered in the same group G_{f_m}. We first sort all harmonic components in G_{f_m} by center frequencies, i.e., $G_{f_m} = \{sc^0_{f_m}, sc^1_{f_m}, \cdots, sc^n_{f_m}\}$, where $sc^i_{f_m}$ denotes the i-th harmonic component in the group. We select a harmonic component with the highest peak (caused by a device that has the strongest memory EMR) in G_{f_m}, and also select all the affiliated harmonic components with the frequency offset $i \cdot f_m$, where i is an integer. Then we remove all the affiliated harmonic components from G_{f_m} and move to a temporary subgroup $\overline{G_{f_m}}$. Due to the spread spectrum clocking, the frequency drift of the same device's harmonic components must be with the same fluctuation trends. If these harmonic components in $\overline{G_{f_m}}$ are produced by the same devices, they can experience consistent fluctuations in time domain, such as the *harmonic component 1* and *harmonic component 2* of device 1 in Figure 9(e).

Thus, to quantify the consistency of all harmonic components in $\overline{G_{f_m}}$, we further calculate the correlation coefficient R_c of the consecutive samples of each harmonic component in the time domain. Although auto-correlations can effectively quantify the instantaneous match across harmonics at any given time, it is important to highlight that they are unable to measure the similarities in the long-term frequency drifting behavior of these harmonic bands. Thus we then employ Dynamic Time Warping (DTW) [29] to measure the similarity between the trends of R_c between different harmonic components in $\overline{G_{f_m}}$. Those harmonic components having similar trends in the time domain, as illustrated in Figure 10, can be exactly identified as the affiliated harmonic components belonging to the same device. If so, ESauron names the device as *Bob* and moves all the affiliated harmonic components in $\overline{G_{f_m}}$ to a device group $G_{f_m}(Bob)$. The above procedure is executed in a loop until the highest folding peak meets a predefined threshold. In this way, ESauron can separate all devices that are equipped with a DDR memory, which may include potential spy cameras.

In practice, due to the dynamic traits, the static traits (f_0, f_m, $\triangle f_0$) will be shifted over time. Hence, ESauron has to monitor the separated devices in real-time and leverages the limited frequency drift in a short period to keep track of each device's harmonic components in the time domain. It is worth noting that although the dynamic traits can also bring about the harmonic components' frequency aliasing of different devices, as illustrated in Figure 9(e), possibly mistaking the device track, it would not affect the follow-up spy camera detection and pinpointing as long as they can be separated.

4.4 Stimulus-based Spy Camera Detection

Note that the sudden change of light intensity can lead to significant scene change for a working camera, which in turn triggers a surge in camera memory access and processing to encode new visuals. In contrast, the activity of non-camera devices is independent of the change of light intensity. By intentionally inducing scene changes via turning on/off light, ESauron can excavate the causal relationship between the stimulation and responsive EMR pattern to assert the existence of spy cameras. Once the affiliated EMR harmonic components of device D are identified, we adopt a stimulation method to assert whether D is a spy camera.

Stimulation strategy. In the stimulation stage, we intentionally turn the room light on and off to create global scene changes, which can suddenly aggravate memory workload and hence trigger responsive EMR patterns from the cameras. Then, ESauron senses and extracts the affiliated EMR harmonic components of device D, and tracks the EMR pattern changes in response to the stimulus.

ESauron uses a method of t-test [30] to analyze the causal relationship between the intentional stimulus and responsive memory EMR pattern. Specifically, ESauron first samples the EMRs and calculates the variance of memory EMR pattern over a time window to determine whether it is in steady state. If so, ESauron instructs the user to start the lights on/off stimulation. Then, ESauron adopts t-test to perform causality analysis between the pre- and post-stimulus EMR patterns. A t-test takes the mean μ, variance σ, and the number of samples N of the two compared sets (i.e., the sets of samples captured in during pre- and post-stimulus), and then computes a t-score. Here the t-score represents the degree of difference between the two sets of samples. A large positive t-score is an evidence that the mean of samples in post-stimulus is significantly larger than that in pre-stimulus.

Existence assertion. Once the t-score is calculated, coupled with the sample size, the degree of the difference (referred to as “Stimulation score”) is normalized and determined by referring to the t-test distribution table [30]. Figure 11 shows the traces of intentional stimulus and stimulation score when probing the EMR signal of a spy camera. Even though normal video processing brings misleading stimulation scores, they are much lower compared with the case of intentional stimulation.

4.5 Pinpointing Spy Camera’s Location

After determining the existence of a spy camera D, we further leverage both the sensitivity to scene changes and distance-relevant properties of the memory EMR to quickly pinpoint the spy camera’s location.

Unearthing distance-relevant information. ESauron leverages the relative change of RSS along the receiver’s trajectory to guide it approximately towards the EMR source, i.e., the camera. To this end, ESauron equips a plug-in miniature strobe light [31] on the receiver, to produce continuous stimulus and trigger traceable EMRs. Although the RSS of EMR from cameras situated indoors does not strictly adhere to the Friis free space path loss model given the rich multipath propagation, the general trend of RSS decrease over distance is invariant. So ESauron employs a simple Kalman filter to average out the multipath effect. We move the ESauron in the opposite direction of the spy camera. From Figure 12, we observe that although there are several solid or weak points in the middle, the RSS preservation after filtering continues to decline.

Orientation in 3-D space. We introduce a simple process to track the EMR RSS tendency along different directions, and then determine the orientation of spy camera relative to the ESauron receiver/user. Specifically, for each orientation, we construct a 3-D reference coordinate with the center of the user’s neck as the origin, as illustrated in Figure 13. The horizontally extended left and right arms are respectively marked as the negative and positive x-axis; and accordingly the y and z axes can be defined.

Then, the user needs to move the receiver along each axis. ESauron can analyze the EMR RSS tendency accordingly, as illustrated in the left part of Figure 14, where we normalize the RSS and the distance changes in one arm length. ESauron then selects the maximum RSS values along each direction, denoted as $(RSS^x_{max}, RSS^y_{max}, RSS^z_{max})$. ESauron leverages the EMR RSS tendency in different directions to generate an overall direction of spy camera in 3-D coordinate system relative to the ESauron user, as illustrated in the middle part of Figure 14.

Iterative-approximation process. The user iterates the above process of moving towards the RSS-maximizing direction, until the RSS reaches its maximum value, i.e., the user is closest to the camera. It is worth noting that this iterative-approximation process has great tolerance against the deviation from the optimal searching path, because the deviation does not accumulate over search steps, and the trend of RSS drop over distance remains unchanged. By holding ESauron in an apartment room where a spy camera was installed in a corner of the roof, ESauron can identify the RSS-maximizing direction at any location of the room. The right side of Figure 14 shows the iterative process to move towards the spy camera.

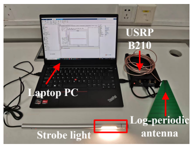


Figure 15: The prototype of ESauron.

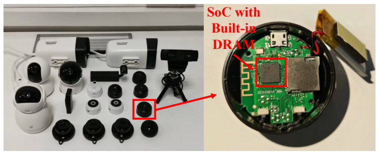


Figure 16: Some cameras used in our experiments.

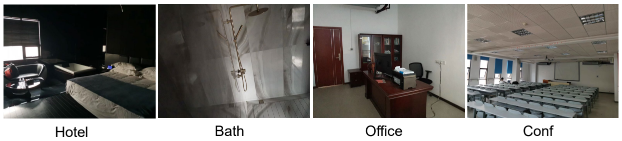


Figure 17: Snapshot of the four typical privacy-sensitive environments.

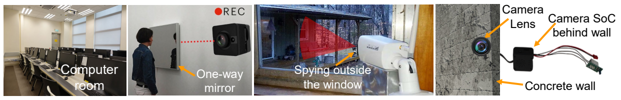


Figure 18: Four extreme scenarios.

5 Evaluation Results

5.1 Experimental Setup

Figure 15 shows the prototype platform of ESauron, composed of a laptop PC equipped with a plug-in miniature strobe light and a USRP B210 with a log-periodic antenna LP0965 (850 MHz-6.5 GHz, 5-6dBi, Size: 15*14cm). The USRP B210 is controlled by using Matlab in the laptop PC, which implements the signal processing mechanisms that constitute ESauron.

Testing cameras. We evaluated ESauron on 50 cameras of 10 typical brands on the market of both the US and China, as shown in Figure 16 and Table 1. All the cameras use H.264/H.265 codecs which are the most popular codecs used by spy cameras. Those cameras can be divided into 4 categories, including 13 miniature spy cameras, 10 surveillance cameras, 17 IP webcam in laptop/smartphone, and 10 home security cameras.

Testing scenarios. We experiment with ESauron in 4 typical privacy-sensitive environments (Hotel, Bathroom, Office

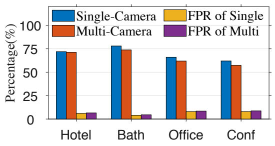


Figure 19: The detection rate and FNR of ESauron in a single stimulus round.

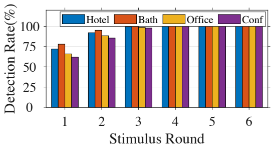


Figure 20: The single-camera detection rate of ESauron in the multiple stimulus rounds.

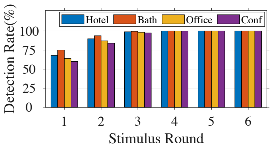


Figure 21: The multi-camera detection rate of ESauron in the multiple stimulus rounds.

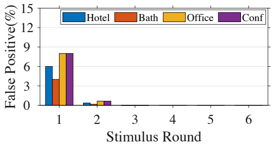


Figure 22: The false positive rate of ESauron in the multiple stimulus rounds.

and Conference room) as illustrated in Figure 17. The dimensions of the rooms are: Hotel ($5 \times 6m^2$), Bathroom ($2 \times 3m^2$), Office ($3 \times 4m^2$), and Conference room ($10 \times 12m^2$). For each environment, we hide the cameras inside furniture or appliances (e.g., a wardrobe or a bedside table) to imitate practical attack scenarios. In addition, we set up 4 extreme scenarios (Computer room with severe memory EMR interference, spying outside the window, spying with one-way mirror, and camera SoC behind concrete wall by camera re-factoring) that could possibly happen, as illustrated in Figure 18 .

Note that in all these experiment scenarios, spy cameras may coexist with other electronic devices, such as voice assistants, smart projectors, desktop computers, etc., that can produce memory EMRs to interfere with the ESauron detector. Also, we know the models of legal or authorized camera equipment in the environment. To ensure the objectivity, we conduct double-blind experimental procedure, which involves an attacker team for hiding spy cameras and an isolated defender team (the co-authors of this paper) for detecting the hidden cameras. Specifically, the defender team operated the ESauron prototype system from an isolated room with no information on the number, models or locations of the hidden cameras.

Comparing ESauron with Existing Work. We consider EarFisher [15] and CamRadar [18] to compare our work in terms of robustness: EarFisher relies on the network to actively import data volumes to stimulate devices to generate electromagnetic leakage. CamRadar detects spy cameras through electromagnetic leakage from ADCs.

Performance metrics. To evaluate the performance of ESauron, we use the following metrics:

- Detection rate. The ratio of the number of detected spy cameras to the total number of actually deployed spy cameras.

- False positive rate (FPR). The ratio of the number of devices that ESauron falsely regarded as spy cameras to the total number of devices that were detected as spy cameras.

- False negative rate (FNR). The ratio of the number of actual spy cameras failed to detect the presence to the total number of actually deployed spy cameras.

- Positioning distance efficiency (PDE). The ratio of the Euclidean distance between the initial location of ESauron and spy camera to the total length of walking route for finding out spy camera.

5.2 Spy Camera Detection Performance

We first deploy a single camera at eight fixed corners respectively at the room's top and middle wall, and calculate the average detection rate for all deployment locations. Figure 19 shows that, with a single stimulus (i.e., turn on and off the light once), the detection rate ranges from 62% to 78% across different room types. The detection rate tends to be higher in darker scenes. This is mainly because the stimulus significantly changes the scene, and accordingly the spy cam memory switches from the standby state to the high-load state, making the EMR signals more prominent.

Figure 20 further shows the detection rate of ESauron after the multiple stimuli. The interval between each stimulation is not less than 1 second. We observe that the detection rate improves quickly with the number of stimuli, e.g., from 78% to over 99.7% after 3 successive stimuli in the bathroom. Although the detection rate in the conference room is the worst, it increases rapidly to 97.7% after 3 stimuli.

Since there may be other electronic equipment in the environment, the FPR performance is also essential. Figure 22 shows that the FPR for all room types. The offices and conference rooms have the highest FPR because of the most number of electronic devices nearby. After the second round of stimulation, the FPR quickly dropped below 0.6% for all room types. After 3 rounds of stimulation, the FPR dropped to 0.

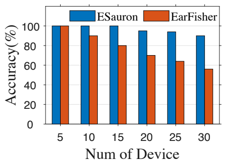


Figure 23: The upper limit of the number of ESauron separation devices.

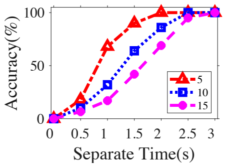


Figure 24: The time cost of the ESauron in separating the cameras.

5.3 Robustness Analysis

5.3.1 Multi-camera detection accuracy

We also conduct quantitative study to evaluate the maximum number of spy cameras that ESauron can successfully detect. Note that the detection of spy cameras consists of two steps, i.e., separating and classifying the harmonic components of different devices and then asserting whether they are spy cameras, and the former is decisive to the detection. Thus we first emulate a scenario where there are multiple devices equipped with the same DDR memory in a computer room with 30 personal computers (PCs). By starting different number of PCs, we evaluate how many devices can be successfully identified by separating and classifying their EMR harmonic components. Figure 23 shows that, when the number of devices increases to 30, the separation accuracy of ESauron remains at 90%, while the separation accuracy of EarFisher is only 56%. Because EarFisher does not consider the aliasing of harmonic components from different devices at all, the separation capability of EarFisher is limited to around 15. It is worth noting in practical applications, the number of devices that ESauron can accurately separate is much larger than 15, because of most devices in a deployment location usually use different DDR memory products, producing EMRs with different center frequency, such as DDR2-1600's 800 MHz and DDR3-1700's 850 MHz. ESauron can easily separate these devices by channel hopping.

Since the time cost for successfully separating multiple devices is also critical to ESauron's performance on spy camera detection. We define the time from the collecting of the first EMR signals to successfully separate all devices' harmonic components as "Separating time". Figure 24 shows when there are no more than 15 devices, ESauron can successfully separate 99% harmonic components within 3 second.

On that basis, we set 15 cameras at the same time to different locations in each of the deployment locations, and evaluate the average detection rate. Figure 19 shows that the multi-camera detection rate is similar to that of single-camera detection under a single stimulus. Figure 21 shows after only 4 stimuli in all deployment locations, all the deployed spy cameras (100%) can be successfully detected.

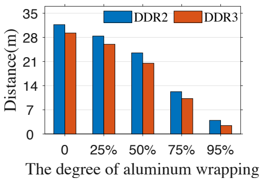


Figure 25: Impact of aluminum sheet's coverage on detection distance.

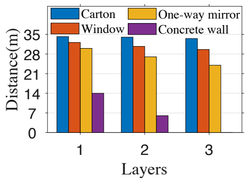


Figure 26: Influence of varying hiding conditions on detection distance.

5.3.2 Detection distance

We define the detection distance as the maximum distance at which ESauron can detect the camera over several rounds of stimulation. Considering the many methods an adversary may employ to prevent spy cameras from being spotted, we evaluate the EMR detection distance in several extreme situations. First, we emulate an scenario where the adversary packages the camera body to prevent EMR leakage. We cover the spy camera housing with an aluminum sheet, and gradually increase the coverage by changing the size of the sheet. The aluminum sheet in our experiments comes from the casing of a Coke can. Figure 25 shows the effect on the EMR detection distance. We observe that when the coverage is less than 50%, the detection distance decreases, and the variance increase. This is because the aluminum sheet reduces electromagnetic radiation in some directions while strengthening radiation in other directions. When the coverage is increased to 95%, the detection distance decreases significantly but remains above 2 meters. Note that 100% metal coverage is impossible since the lens area has to be exposed.

The spy camera can also be hidden and retrofitted in a variety of ways. For example, it can be hidden in a carton box. Spy cameras can also be deployed outside windows to monitor rooms. The processing unit can be decoupled from the lens unit (linked by a cable), and hidden inside a concrete wall to shield the EMR signals. In a bathroom, the camera can even be deployed behind a one-way mirror.

We evaluate the effects of cartons, outside windows, one-way mirrors, and concrete walls on the detection distance, respectively. Figure 26 shows that the carton (around 5 mm thick) does not affect the detection distance. As the number of layers of the windows (5 mm thick each) and the one-way mirror (12 mm thick each) increases, the detection distance decreases slightly. A concrete wall (25 cm thick) has the most significant impact on the detection distance, and can reduce the detection distance to 14 m. A double-layer concrete wall further reduces the distance to 6 m. Three layers of concrete walls can effectively mask the EMR signals.

Table 2: Detection distance (m).

Shell Type	ESauron	CamRadar
Plastic	30.6	0.7
Metal	11.2	0.2

Comparison with CamRadar. We have also conducted an evaluation of the detection distance for both ESauron and CamRadar using the same experimental setup. The detection distances for ESauron and CamRadar are listed in Table 2. Specifically, we consider two scenarios where the opponent encapsulates the camera body to prevent electromagnetic radiation (EMR) leakage, using either a plastic case or a metal case. In the case of a plastic housing, ESauron achieves a detection distance of 30.6m, while CamRadar only achieves a detection distance of 0.7m. When the housing is made of metal, ESauron has a detection distance of 11.2m, whereas CamRadar only reaches 0.2m.

It is important to note that CamRadar relies on detecting electromagnetic interference from analog-to-digital converters (ADCs) in cameras to identify hidden cameras. This limitation restricts CamRadar's detection range to less than 1 meter, as ADC emissions tend to be relatively weak. In contrast, our proposed solution focuses on capturing memory EMR from real-time encoding workloads, which exhibit significantly stronger emissions. This enables ESauron to achieve detection ranges beyond 10 meters in most cases.

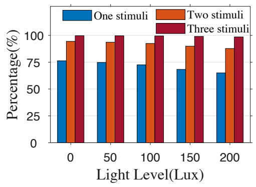


Figure 27: Detection performance under different light levels.

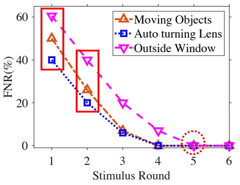


Figure 28: False negative rate in different conditions.

5.3.3 Impact of lighting conditions

We test ESauron at different times of the day with different ambient light intensities which are measured using a smartphone light sensor. Figure 27 shows the detection performance of ESauron at 5 ambient light levels. As the intensity of ambient light increases, the detection rate of ESauron decreases from 76% to 65% under a single stimulus. This is because strong ambient light can weaken the spy camera's sensitivity to the intentional stimulus by turning on/off room light. Even though, the increase of stimulus rounds can compensate for the impact of lighting conditions. Results show even with a light intensity of 200 Lux, the detection rate can be up to 98% after three rounds of stimulation.

5.3.4 False negative rate

When an intentional stimulus can not trigger an obvious responsive memory EMR pattern, it possibly causes the false negative rate (FNR). In our experiments, we observe three extreme situations can bring false negative detection: (1) Existing moving objects within spy camera's monitoring view; (2) Refactoring an automatic fine tuning mode for the camera lens by the adversary; and (3) Spy camera deployed outside a glass window that can be affected by the outdoor situation. Under these extreme situations, the spy camera's memory activities can possibly and coincidentally occur at the same time of EMR sensing, bringing non-causal EMRs pattern and leading to false negative results. By emulating the three scenarios, we repeat ESauron's spy camera detection mechanism more than 100 times under each scenario. Figure 28 shows the FNR of ESauron after different round of stimulus. All three extreme situations have high FNR (more than 40%) because of the non-causal EMRs pattern caused by other irrelevant activities. Even though, due to the picture estimation of advanced coding techniques, ESauron remains robust to these extreme situations after multiple stimulus rounds. As shown, after only 5 rounds of stimulus, the FNR is close to 0. The results demonstrate that ESauron has good robustness even in extreme situations. With the increase of stimulus round, ESauron can reliably detect spy cameras.

5.4 Camera Location Inference Accuracy

We conducted localization experiments in the above 4 representative environments. In the hotel room, spy cameras are placed in five fixed locations: the upper left/corners of the wall, the headboard, the bedside table, and the opposite side of the bed. In the office and conference room, spy cameras are installed in the following locations: two sockets, on top of curtains, in the upper left/right corners of walls. In the bathroom, the cameras are hidden in the lampshades and the upper left/right corners of the walls. We tested 13 spy cameras in 10 trials for each situation. Positioning is determined to be successful when there is a spy camera within 0.5 meters of ESauron's final estimated location. For each trial, we calculate the localization time and the positioning distance efficiency for different rooms.

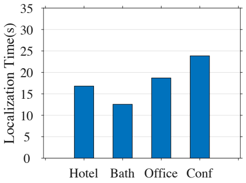


Figure 29: Localization time in different environments.

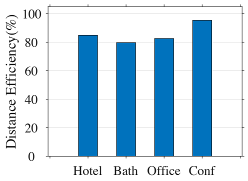


Figure 30: Positioning distance efficiency.

We successfully positioned the spy camera in all trials. From the results in Figure 29, we found that the bathroom has the short localization time, and the positioning task can be completed in as fast as 4.1 seconds. But due to multipath, it takes 16.7 seconds in the slowest bathroom case. In the conference room (the largest), ESauron still completes the task in 32.3 seconds. From Figure 30, we observe that the distance efficiency of the conference room is as high as 95.3%. In larger room types, ESauron is more efficient. In all room types, the efficiency of ESauron exceeds 79.7%.

6 Related Work

Detecting cameras by identifying light reflection and EMI. Prior research has shown that, by illuminating a suspicious place with laser or flashlight, hidden cameras can be pinpointed based on the tiny glint of the lens [32]. Similarly, a ToF sensor can reveal the location of lens [33], since the lens causes a sharp increase ToF along the depth direction. These approaches assume that the users are aware of the approximate location of the cameras and can shed light towards the lens, which may not be easily satisfied in practice. Besides, these methods are cumbersome to use and require significant user involvement. Alternatively, [34] employs a smartphone's magnetometer to sense the changes in electromotive force, and then infer the existence of spy cameras. However, all electronic devices can cause the change in electromotive force. Thus to search out a spy camera, the user must scrutinize the entire private space, which requires a meticulous scanning process while no guarantee that the detected electronic device is a spy camera.

Traffic analysis for wireless-connected cameras. Information leakage from wireless cameras has been investigated in prior research [35, 36, 13, 37, 14]. It is well established that scene variations (e.g., sharp change in lighting conditions) may modulate the video codec workload and hence the data traffic from a wireless camera. Exploiting such phenomenon, [12] further pinpoints the camera location. While these techniques are promising, they only work for wireless cameras, whereas a significant number of hidden cameras are either wired or store data to a local memory card. The key difference is that ESauron can accurately detect both wireless and wired spy cameras by leveraging the leakage of memory EMR signals and can quickly pinpoint cameras' locations.

EM side-channels. Recent research also leveraged the EM side channels of ADC, CPU and memory for attestation [38, 18], memory profiling [39, 15], and malware detection [40, 41, 42, 17]. Among them, EarFisher [15], Memscope [16], DeHiREC [17] and CamRadar [18] are the most relevant to ESauron. EarFisher and Memscope explored the possibility of exploiting memory EMR to detect wireless eavesdroppers, opening up new possibility of detecting stealthy electronic device with leaked EMR signals. Based on EarFisher, DeHiREC and CamRadar respectively leveraged the EMRs sourcing from the ADC to detect voice recorder and spy camera.

Although Memscope observed the negative impact of dynamic traits (frequency drift) of memory EMRs, rather than harnessing this feature to enhancing continuous tracking, it bypasses the dynamic traits by only leveraging stable harmonic peak interval to identify memory devices. Due to the ignorance of the critical dynamic traits of memory EMRs in camera devices, existing works are ineffective in distinguishing and continuously tracking the time-varying EMRs from even a single device, not to mention common scenarios deployed with multiple cameras probably of the same model. Moreover, due to the power of the ADC is much lower than the power of the memory, CamRadar and DeHiREC's detection range is limited within 1 meter, thus a user has to carry CamRadar and scrutinize the entire private space, requiring a slow and laborious sweep, bringing insurmountable hurdle for device identification and localization when spy devices are deployed in high places. Due to lack of thorough characterizing on memory EMR from camera devices, the inability to continuously differentiate devices of the same model, and lack of EMR-based positioning capability, existing works are still far from effective hidden camera detection in real scenarios.

Unlike these methods, ESauron features a new paradigm that actively stimulates the memory EMR of hidden cameras and pinpoint their locations. Moreover, based on the memory clock spectrum, ESauron innovates a signal processing chain to not only extract weak memory EMRs under poor signal conditions, but also distinguish and track individual memory EMRs when multiple devices coexist in a crowded environment.

Device Localization. Wireless indoor localization has been well explored in the past decade. Early work used signal strength fingerprinting or model-driven methods [43, 44, 45]. Angle of Arrival (AoA)[46, 47] or Time of Flight (ToF) [48] can further improve the spatial resolution. They require sophisticated receiver hardware or explicit synchronization with the transmitter. In comparison, ESauron adopts a received signal strength (RSS) based iterative-approximation search algorithm to heuristically direct the receiver towards the spy camera. Due to their relatively low frequency, the EMRs are less vulnerable to multipath and shadowing effects. On this basis, ESauron can continuously trigger the stimulus while the receiver moves towards the locations with an increasing RSS, until it approach the spy camera.

7 Discussions

ESauron's detection range is up to 20 meters, which applies to the majority of scenarios and outperforms the COTS camera detection methods, such as LAPD [33] 0.45-1.5m, E-Eye [37] 0.2m, and CamRadar [18] 1m. This significant advantage is mainly attributed to ESauron's unique techniques tailored for camera memory EMR, including the harmonic folding algorithm to extract weak EMR signals and the device fingerprinting to separate mixed EMRs. These innovations address the limitations of previous works that fail to consider memory EMR's critical traits.

Limitations. ESauron still has certain limitations: i) Stimulus restriction. ESauron relies on active stimulation which may not always be feasible, especially when the user has no control over the target environment. ii) Inapplicable to detect emissions from cameras in some smartphones. The latest smartphones employ low-power DDR techniques [49][50] and integrate Faraday cages internally to mitigate electromagnetic radiation (EMR) leakage [51]. The combination of the two measures significantly increase difficulty to detect smartphone cameras' memory EMRs using the current ESauron prototype. Based on ESauron, we will further study the characteristics of smartphone camera and try to detect its activities in the future. iii) Detection object limitation. Many low-power IoT devices do not use DRAM, hence the detection of such-kind IoT devices is beyond the capability of ESauron. iv) Deficiency in portability. As a prototype system, ESauron requires dedicated sensing hardware and is not easy to carry. The miniaturization of ESauron such as hosting our system on Raspberry Pi (PiSDR) to replace laptop and miniaturize ESauron system, and using a customized low-cost USRP (RTLSDR) to replace B210 for cost reduction, remains a work in progress. v) Antenna Limitations. The log-periodic antenna have some directionality, which could cause false negatives when the radiation direction is not aligned with the antenna orientation. By using ESaron to detect spy cameras, we can mitigate such errors by turning the antenna to change its direction to cover 360-degree. Furthermore, cameras emit EMR in complex electromagnetic patterns, not just in a single direction. Therefore, even if the main lobe is misaligned, other side lobes can still be captured.

Addressing these limitations would further enhance ESauron's applicability. On the whole, ESauron represents an important step towards protecting personal privacy against hidden spy cameras. The proposed techniques open up new capabilities to detect cameras in a generalized manner. The results clearly demonstrate the feasibility of exploiting inevitable memory EMR leakage for spy camera detection, which was not shown before.

Noting that compared with log-periodic antenna, Yagi antenna is with smaller size. Based on our study on the spectral characteristics of camera's memory EMRs in this work, we plan to work on miniaturize ESauron system in future work. We are likely to use a customized Yagi antenna in the miniaturized system to replace the log-periodic antenna.

Potential problem of ESauron. While our system, ESauron, aims to empower users to detect hidden recording devices, we acknowledge that its detection capabilities could inadvertently result in privacy violations if misused. In fact, ESauron could potentially introduce a new privacy attack vector, as hackers could exploit it to infer users' activity routines within private locations.

8 Conclusion

Privacy protection in indoor environment has been an important but unsolved problem. In this paper we propose ESauron, which uses the leaked memory EMRs to detect spy cameras and pinpoint their location. We implemented and evaluated ESauron under various representative indoor scenarios, which demonstrates ESauron's effectiveness and robustness. We consider ESauron as a first exploration to detect and pinpoint all kinds of hidden spy cameras, including wireless/wire-connected and storage-based off-line devices.

Acknowledgments

We appreciate the shepherd's efforts and all the anonymous reviewers' insightful and constructive suggestions. This work was partially supported by National Key Research and Development Program of China under grant No. 2023YFB3001805, the NSFC under grants No. 62372166, No. 62072167, No. 62232007, No. 62202150, and No. U20A20181, Hunan Provincial Natural Science Foundation of China under grant No. 2023JJ30164, Excellent Youth Fund, and Fundamental Research Funds for the Central Universities.

References

[1] Time. Australian police charge tourist over spycam case at a bondi beach hostel, 2019.

[2] Nypost. Does your hotel or airbnb come with a hidden camera?, 2022.

[3] Ksat. Investigators uncover more than 2,100 images in growing hill country hidden camera case, 2022.

[4] Airbnb. Survey: Do airbnb guests trust their hosts?, 2019.

[5] Wizcase. Risk: Is this your webcam? you're being watched, 2019.

[6] CBC. We hired ethical hackers to hack a family's smart home — here's how it turned out, 2018.

[7] Consumerreports. How to protect yourself from camera and microphone hacking, 2019.

[8] Stanislaw Piasecki, Lachlan Urquhart, and Derek McAuley. Defence against the dark artefacts: Smart home cybercrimes and cybersecurity standards. *Comput. Law Secur. Rev.*, 42:105542, 2021.

[9] Sriram Sami, Sean Rui Xiang Tan, Bangjie Sun, and Jun Han. Lapd: Hidden spy camera detection using smartphone time-of-flight sensors. In *Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems*, SenSys '21, page 288–301, New York, NY, USA, 2021. ACM.

[10] Akash Deep Singh, Luis Garcia, Joseph Noor, and Mani B. Srivastava. I always feel like somebody's sensing me! A framework to detect, identify, and localize clandestine wireless sensors. In *30th USENIX Security Symposium, USENIX Security*, pages 1829–1846, Virtual Event, 2021. USENIX Association.

[11] Yushi Cheng, Xiaoyu Ji, Tianyang Lu, and Wenyuan Xu. Dewicam: Detecting hidden wireless cameras via smartphones. In *Proceedings of the 2018 on Asia Conference on Computer and Communications Security*, ASIACCS '18, page 1–13, New York, NY, USA, 2018. ACM.

[12] Yan He, Qiuye He, Song Fang, and Yao Liu. Motioncompass: Pinpointing wireless camera via motion-activated traffic. In *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services*, MobiSys '21, page 215–227, New York, NY, USA, 2021. ACM.

[13] Tian Liu, Ziyu Liu, Jun Huang, Rui Tan, and Zhen Tan. Detecting wireless spy cameras via stimulating and probing. In *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*, MobiSys '18, page 243–255, New York, NY, USA, 2018. ACM.

[14] Rahul Anand Sharma, Elahe Soltanaghaei, Anthony Rowe, and Vyas Sekar. Lumos: Identifying and localizing diverse hidden iot devices in an unfamiliar environment. In *31st USENIX Security Symposium, USENIX Security 2022, Boston, MA, USA, August 10-12, 2022*, pages 1095–1112. USENIX Association, 2022.

[15] Cheng Shen and Jun Huang. Earfisher: Detecting wireless eavesdroppers by stimulating and sensing memory EMR. In *18th USENIX Symposium on Networked Systems Design and Implementation, NSDI 2021, April 12-14, 2021*, pages 873–886. USENIX Association, 2021.

[16] Cheng Shen, Jun Huang, Guangyu Sun, and Jingshu Chen. Electromagnetic fingerprinting of memory heartbeats: System and applications. 6(3), sep 2022.

[17] Ruochen Zhou, Xiaoyu Ji, Chen Yan, Yi-Chao Chen, Chaohao Li, and Wenyuan Xu. Dehirec: Detecting hidden voice recorders via adc electromagnetic radiation. In *2023 IEEE Symposium on Security and Privacy*, SP '23, pages 658–673, San Francisco, CA, USA, 2023. IEEE.

[18] Ziwei Liu, Feng Lin, Chao Wang, Yijie Shen, Zhongjie Ba, Li Lu, Wenyao Xu, and Kui Ren. Camradar: Hidden camera detection leveraging amplitude-modulated sensor images embedded in electromagnetic emanations. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 6(4):1–25, 2023.

[19] AMD. Develop with amd, 2018.

[20] James Clerk Maxwell. *A treatise on electricity and magnetism*, volume 1. Clarendon press, 1873.

[21] Robert Callan, Alenka Zajić, and Milos Prvulovic. Fase: Finding amplitude-modulated side-channel emanations. In *Proceedings of the 42nd Annual International Symposium on Computer Architecture*, ISCA '15, page 592–603, New York, NY, USA, 2015. ACM.

[22] Takayuki Daimon, Hiroshi Sadamura, Takayuki Shindou, Haruo Kobayashi, Masashi Kono, Takao Myono, Tatsuya Suzuki, Shuhei Kawai, and Takashi Iijima. Spread-spectrum clocking in switching regulators for EMI reduction. *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, 86-A(2):381–386, 2003.

[23] Cornelis D. Hoekstra. Frequency modulation of system clocks for emi reduction. *Hewlett-Packard Journal*, 48(4):101–101, 1997.

[24] Boualem Boashash. Time-frequency signal analysis and processing: A comprehensive reference. *Signal Processing*, 2003.

[25] David H. Staelin. Fast folding algorithm for detection of periodic pulse trains. *Proceedings of the IEEE*, 57(4):724–725, 1969.

[26] Quan Pan, Lei Zhang, Guanzhong Dai, and Hongcai Zhang. Two denoising methods by wavelet transform. *IEEE Trans. Signal Process.*, 47(12):3401–3406, 1999.

[27] Jun-Young Park, Dae-Hwan Yun, Seong-Yeon Kim, and Yang-Kyu Choi. Suppression of self-heating effects in 3-d v-nand flash memory using a plugged pillar-shaped heat sink. *IEEE Electron Device Letters*, 40(2):212–215, 2019.

[28] Kentaro Nishimori, Keizo Cho, Yasushi Takatori, and Toshikazu Hori. Automatic calibration method using transmitting signals of an adaptive array for tdd systems. *IEEE Transactions on Vehicular Technology*, 50(6):1636–1640, 2001.

[29] Donald J. Berndt and James Clifford. Using dynamic time warping to find patterns in time series. In *Knowledge Discovery in Databases: Papers from the 1994 AAAI Workshop*, pages 359–370, Seattle, Washington, USA, 1994. AAAI Press.

[30] Sws Gosset. The probable error of a mean. *Biometrika*, 6(1):1–25, 1908.

[31] Shilin Zhu, Chi Zhang, and Xinyu Zhang. Automating visual privacy protection using a smart led. In *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking*, MobiCom '17, page 329–342, New York, NY, USA, 2017. ACM.

[32] LLC Logan Security Consulting. Pimall, 1993.

[33] Sriram Sami, Sean Rui Xiang Tan, Bangjie Sun, and Jun Han. Lapd: Hidden spy camera detection using smartphone time-of-flight sensors. In *Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems*, SenSys '21, page 288–301, New York, NY, USA, 2021. ACM.

[34] LLC Logan Security Consulting. Spy hidden camera detector, 2017.

[35] Christopher Wampler, A. Selcuk Uluagac, and Raheem A. Beyah. Information leakage in encrypted IP video traffic. In *2015 IEEE Global Communications Conference*, pages 1–7, San Diego, CA, USA, 2015. IEEE.

[36] Ben Nassi, Raz Ben-Netanel, Adi Shamir, and Yuval Elovici. Drones' cryptanalysis - smashing cryptography with a flicker. In *2019 IEEE Symposium on Security and Privacy*, pages 1397–1414, San Francisco, CA, USA, 2019. IEEE.

[37] Zhengxiong Li, Zhuolin Yang, Chen Song, Changzhi Li, Zhengyu Peng, and Wenyao Xu. E-eye: Hidden electronics recognition through mmwave nonlinear effects. In *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems*, SenSys '18, page 68–81, New York, NY, USA, 2018. ACM.

[38] Nader Sehatbakhsh, Alireza Nazari, Haider Khan, Alenka Zajic, and Milos Prvulovic. Emma: Hardware/software attestation framework for embedded systems using electromagnetic signals. In *Proceedings of the 52nd Annual IEEE/ACM International Symposium on Microarchitecture*, MICRO '52, page 983–995, New York, NY, USA, 2019. ACM.

[39] Nader Sehatbakhsh, Alireza Nazari, Alenka G. Zajic, and Milos Prvulovic. Spectral profiling: Observer-effect-free profiling by monitoring EM emanations. In *49th Annual IEEE/ACM International Symposium on Microarchitecture, MICRO 2016*, pages 59:1–59:11, Taipei, Taiwan, 2016. IEEE Computer Society.

[40] Yi Han, Sriharsha Etigowni, Hua Liu, Saman Zonouz, and Athina Petropulu. Watch me, but don't touch me! contactless control flow monitoring via electromagnetic emanations. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, CCS '17, page 1095–1108, New York, NY, USA, 2017. ACM.

[41] Alireza Nazari, Nader Sehatbakhsh, Monjur Alam, Alenka Zajic, and Milos Prvulovic. Eddie: Em-based detection of deviations in program execution. In *Proceedings of the 44th Annual International Symposium on Computer Architecture*, ISCA '17, page 333–346, New York, NY, USA, 2017. ACM.

[42] Zhenkai Zhang, Zihao Zhan, Daniel Balasubramanian, Bo Li, Péter Völgyesi, and Xenofon D. Koutsoukos. Leveraging EM side-channel information to detect rowhammer attacks. In *2020 IEEE Symposium on Security and Privacy*, S&P 2020, pages 729–746, San Francisco, CA, USA, 2020. IEEE.

[43] Rizanne Elbakly and Moustafa Youssef. A robust zero-calibration rf-based localization system for realistic environments. In *13th Annual IEEE International Conference on Sensing, Communication, and Networking*, SECON 2016, pages 1–9, London, United Kingdom, 2016. IEEE.

[44] Dian Zhang, Yunhuai Liu, Xiaonan Guo, Min Gao, and Lionel M. Ni. On distinguishing the multiple radio paths in rss-based ranging. In *Proceedings of the IEEE INFOCOM 2012*, pages 2201–2209, Orlando, FL, USA, 2012. IEEE.

[45] Zhijing Li, Zhujun Xiao, Yanzi Zhu, Irene Pattarachanyakul, Ben Y. Zhao, and Haitao Zheng. Adversarial localization against wireless cameras. In *Proceedings of the 19th International Workshop on Mobile Computing Systems &amp; Applications*, HotMobile '18, page 87–92, New York, NY, USA, 2018. ACM.

[46] Kun Qian, Chenshu Wu, Yi Zhang, Guidong Zhang, Zheng Yang, and Yunhao Liu. Widar2.0: Passive human tracking with a single wi-fi link. In *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*, MobiSys '18, page 350–361, New York, NY, USA, 2018. ACM.

[47] Yue Zheng, Yi Zhang, Kun Qian, Guidong Zhang, Yunhao Liu, Chenshu Wu, and Zheng Yang. Zero-effort cross-domain gesture recognition with wi-fi. In *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services*, MobiSys '19, page 313–325, New York, NY, USA, 2019. ACM.

[48] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. Predictable 802.11 packet delivery from wireless channel measurements. In *Proceedings of the ACM SIGCOMM 2010 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, pages 159–170, New Delhi, India,, 2010. ACM.

[49] JEDEC. Mobile memory: Lpddr, wide i/o, 2023.

[50] Micron. Micron and mediatek first to validate lpddr5x, 2021.

[51] Jungho Jin, Choongpyo Jeon, Byounggug Min, Heonsang Lim, and Jungki Kim. Effect of contact resistance on conformai shield package for mobile dram. In *2018 IEEE International Symposium on Electromagnetic Compatibility and 2018 IEEE Asia-Pacific Symposium on Electromagnetic Compatibility (EMC/APEMC)*, pages 341–344, 2018.

A Frequently Used Symbols

Table 3 lists the frequently used symbols and the corresponding descriptions in this paper.

Algorithm 1: ESauron's folding scheme.

```
Input  : $\mathcal{R}$, $G_{f_0}$, $G_{f_0}^{f_m}$
Output: $G_{f_m}$
for each $f_0 \in G_{f_0}$ do
    Initialize $P_{f_m^c}[i]$ ;
    for each $f_m^c \in G_{f_0}^{f_m}$ do
        $P_{f_m^c}[i] = \sum_{j=0}^{\lceil N/f_m^c \rceil - 1} \mathcal{R}[i + j \cdot f_m^c]$;
    end
    Find the max in $P_{f_m^c}[i]$, denote it as $f_m$;
    Put $f_m$ inte set $G_{f_m}$;
end
return $G_{f_m}$;
```

Table 3: Symbol Description.

Symbols	Description
$V_{clk}(t)$	The EMR energy.
$V_{ssc}(t)$	The EMR energy after spread spectrum clock (SSC) techniques.
f_0	The center frequency of clock.
f_m	The modulation frequency of clock.
Δf	The peak frequency offset of clock.
f_{nz}	The frequency of the *i*-th non-zero memory EMR.
G_{f_0}	The set of potential memory center frequency f_0.
G_{f_m}	The set of harmonic components with the feasible f_m.
$\mathcal{R}$	The series of N frequency samples.
f_m^c	The harmonic components separated by the same frequency samples offset.
$\|P_T[i]\|$	The position of the folding peak.
α_k	The detail coefficients.
β_k	The approximate coefficients.
$sc_{f_m}^i$	The *i*-th harmonic component in the group G_{f_m}.
μ	The mean of the two compared sets.
σ	The variance of the two compared sets.
RSS	The received signal strength.

B The Folding Algorithm

Algorithm 1 outlines the folding algorithm used to extract harmonic components from the electromagnetic radiation (EMR) spectrum. The algorithm takes three inputs: the series of frequency samples from the captured EMR signal ($\mathcal{R}$), the set of potential memory center clock frequencies (G_{f_0}), and the set of feasible modulation interval values ($G^{f_m}f_0$) derived from Gf_0. It produces the set of detected modulation intervals (G_{f_m}) that identify distinct memory emanation sources.

The algorithm proceeds by iterating through each candidate center frequency f_0 from the set G_{f_0} (Line 1). For each assumed f_0, it attempts to fold the EMR spectrum $\mathcal{R}$ using different modulation interval values f_m^c from the associated set $G^{f_m}f_0$ (Lines 3). The folding operation involves summing the amplitude samples within windows of width f_m^c (Lines 4), thereby aggregating harmonic components separated by that interval. The modulation interval f_m resulting in the maximum folded amplitude peak is selected as the preferred value for a detected device (Line 6). Subsequently, f_m is added to the output set Gf_m for further analysis (Line 7). Once all potential f_0 values have been scanned, the algorithm returns the final set G_{f_m} containing the extracted modulation intervals (Line 9).

Through iterative folding of the EMR spectrum using feasible intervals based on memory clock models, the algorithm proves efficient in extracting harmonic patterns from multiple devices even at low signal-to-noise ratios. The resulting set G_{f_m} forms the basis for distinguishing and tracking individual memory EMR sources.