Optical Air-Gap Attacks: Analysis and IoT Threat Implications

Lee, J., Yoo, J., Lee, J., Choi, Y., Yoo, S. & Song, J. S.

Original citation & hyperlink:
Lee, J, Yoo, J, Lee, J, Choi, Y, Yoo, S & Song, JS 2024, 'Optical Air-Gap Attacks: Analysis and IoT Threat Implications', IEEE Network, vol. (In-Press), pp. (In-Press).
https://dx.doi.org/10.1109/MNET.2024.3382969

DOI 10.1109/MNET.2024.3382969
ISSN 0890-8044
ESSN 1558-156X

Publisher: Institute of Electrical and Electronics Engineers

Optical Air-Gap Attacks: Analysis and IoT Threat Implications

Jieun Lee[§], JaeHoon Yoo[§], Jiho Lee[§], Yura Choi[§], Seong Ki Yoo[†] and JaeSeung Song[§]
† *Coventry University, United Kingdom; Email: ad3869@coventry.ac.uk*
§ *Sejong University, Korea; Email: love9ly, metoofire, twozio, dbfk1207@sju.ac.kr, jssong@sejong.ac.kr*

***Abstract*—Since 2008, the Korean government has instituted network separation technology, which physically isolates external internet networks from internal networks, aiming to thwart cyber-attacks. Consequently, the domestic financial sector was largely unaffected during global crises (2017 WannaCry ransomware outbreak and the 2021 Log4j vulnerability incident). However, there exist certain vulnerabilities owing to the presumption of their relative safety against cyber intrusions and the integration of cloud and Internet of Things (IoT) technologies in the current smart revolution. The existing network separation measures only mitigate one facet of potential cyber threats, rendering a comprehensive defense elusive. The rise of "air-gap" attacks, which exploit the isolated space between closed and external networks to illicitly transfer data and the existing research primarily substantiating the potential for data breaches from closed networks to their external counterparts are problems yet to be addressed. Thus, our study proposed a tangible optical air-gap attack methodology, harnessing readily available optical mediums within closed networks. Intricate measurement metrics that consider vital factors of the transmission environment were proposed. Moreover, acknowledging the proliferating integration of IoT devices, such as smart bulbs, to facilitate automation within closed networks, this study demonstrated the viability of optical air-gap attacks using these devices.**

***Index Terms*—Network Security, Air-Gap Attack, Closed Network, Network Separation, Optical Air-Gap Attack, IoT Air-Gap Attack.**

I. Introduction

A "closed network" is an internal business network segregated from the Internet to enhance security. It is comparatively insulated from cyber threats, such as malicious software (worms and viruses), distributed denial-of-service (DDoS) attacks, and data breaches. Historically, closed networks have been perceived as safer than external internet networks. Recognizing its robustness, countries have adopted network separation for their operational technology/industrial control systems (OT/ICS), which is pivotal in critical sectors such as power, gas, and energy, safeguarding them against external threats. Since 2008, the Korean government has pursued a network separation policy for national institutions and major companies, which resulted in affiliated financial and national institutions remaining unaffected during global cyber crises (WannaCry ransomware outbreak in 2017 and Log4j vulnerability in 2021) [1]. However, recent cyber incidents have emphasized that closed networks are not impregnable fortresses. Examples include Stuxnet's assault on Iran's nuclear infrastructure in 2010, the leakage of design data from Korea's KHNP nuclear power plant in 2014, and Industroyer2's attack on Ukraine's power grid amidst the Russia–Ukraine conflict [2].

Experts have unveiled novel methods for exfiltrating data from closed networks, even those not connected to the Internet, through "air-gap attacks". The air-gap is a physical void separating an external Internet network from a closed network devoid of any wired (e.g., Ethernet) or wireless (e.g., WiFi, Bluetooth) communication. An air-gap attack typically involves preinstalled malware collecting sensitive data across the air-gap using diverse transmission media (e.g., optics, electromagnetic waves, acoustics, electromagnetics, thermals, vibrations, and electricity) [3]. Receivers on external networks capture these signals and restore the intercepted sensitive data using equipment such as smartphone cameras and microphones. Traditional attacks require unintentional bridging by users; air-gap attacks bypass this necessity and remain unaffected by internal data loss prevention (DLP) systems. Existing research on air-gap attacks has focused on potential transmission media. However, executing an actual air-gap attack is challenging. A comprehensive understanding of the challenges is vital for their practical application. This paper presents the design principles, environmental specifications, and constraints essential for actualizing air-gap attacks using the most accessible optical medium in a closed network setting. The key factors influencing successful air-gap attacks using the hard disk light-emitting diode (LED) of a computer were identified, and concrete measurement metrics considering primary transmission factors (illumination, distance, and receiving equipment resolution) were proposed. Further, by simulating an air-gap attack via smart-bulb control and mirroring the increasing adoption of Internet of Things (IoT) devices, the feasibility of air-gap attacks using IoT optical devices was demonstrated.

Research contributions of this paper are listed as follows:

- Design and implementation of an optical medium-based air gap attack system.
- Identification of effective metrics for attack transmission based on detailed environmental parameters (light intensity, reception distance, reception equipment capability).
- Design and deployment of a novel air gap attack system incorporating IoT smart home devices.

First, the concept of an air-gap attack is introduced, highlighting its merits and limitations in Section II. Thereafter, various attack scenarios centered on transmission medium such as optics, electromagnetic waves, acoustics, electromag-

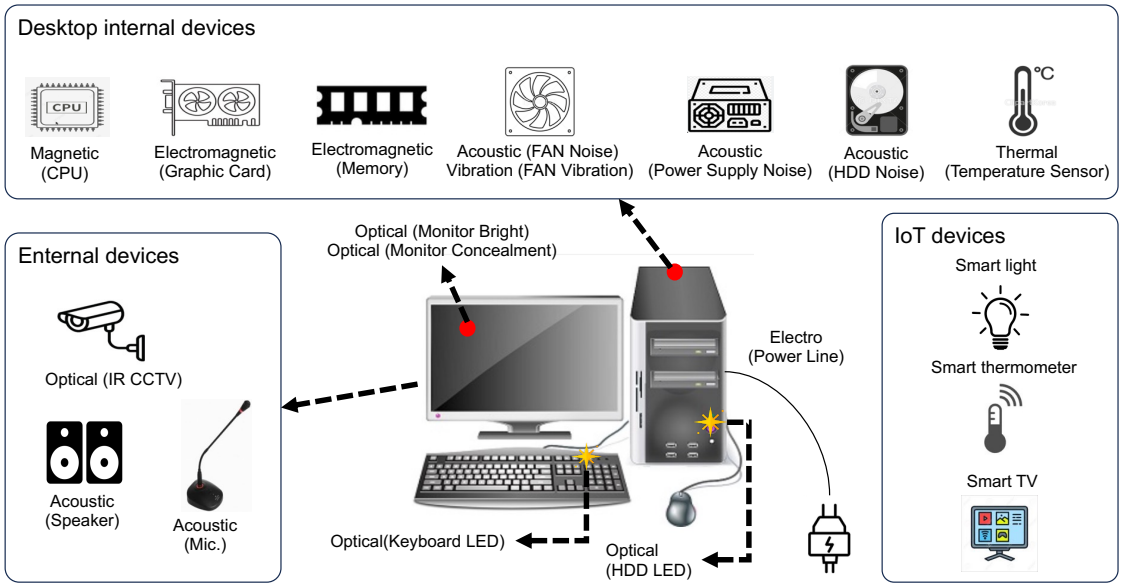


Fig. 1. Various mediums embedded in PCs and peripheral devices within closed networks that can be exploited for air-gap attacks.

netics, electricity, vibrations, and thermals are presented in Section III. Consequently, an air-gap attack leveraging optical devices connected to computers and peripherals and operational prerequisites and constraints are discussed in Section IV. Air-gap attacks under diverse settings using computer-attached LEDs within a closed network and their efficacy conditions are examined in Section V. Further, the creation of attack scenarios, environmental prerequisites, system architecture, and simulation outcomes for optical air-gap attacks using smart bulbs, mirroring the uptick in the incorporation of IoT devices, are discussed in Section VI. Finally, Section VIII concludes the study, summarizing findings and suggesting future research directions.

II. Air-Gap Attack Overview

Air-gap attacks discreetly gathers sensitive data within an isolated, network-separated environment, and conveys them to an external network utilizing optical, acoustic, and electromagnetic signals emanating from computer components and peripherals (Fig. 1). Such attacks transpire where the external Internet and the internal corporate network (closed network) are utterly distinct, and the void between these networks is an air-gap [3]. Within this closed network, malware equipped with air-gap attack capabilities alters the signals produced by computer components and peripherals such as hard disk drive (HDD) LEDs and speakers to exfiltrate data across the air-gap. Essentially, air-gap attacks eschew the need for network connectivity, with the internal computer functioning as a transmitter within this isolated environment and relaying sensitive data through various transmission media to the external network. The external receiver or attacker covertly positions themselves within the signal range of a closed network. By utilizing reception equipment such as smartphone cameras, Google Glasses, webcams, microphones, or drone cameras, the attacker interprets the signals and reconstructs the original data.

Considering attacks leveraging HDD LEDs, as developed by Ben-Gurion University, pre-installed malware within a closed network aggregates sensitive information, encodes it into digital data, and subsequently manipulates HDD LED flashes in line with binary signals [4]. An attacker employs a smartphone camera or an optical sensor within the line-of-sight to capture, interpret, and revert the signal from the closed network to its original data format. Such attacks deploy several transmission media with specific components and peripherals that vary based on the transmission medium and receiving equipment on the external network. Historical attacks on closed networks are reliant on insiders to smuggle malware-laden documents into closed networks to glean information, followed by covert extraction. However, an air-gap attack exploits the physical void (i.e., air-gap) between closed and external networks. As malware-infected computers and peripheral devices within a closed network serve as transmitters, such attacks can be executed independent of additional devices, i.e., universal serial buses (USB). However, the architecture and conditions both within and outside the closed network influence the efficacy of air-gap attacks, owing to stringent security policies and transmission environment constraints. Considering the limited transmission milieu of a closed network and requirement of clandestine transmission to elude users, administrators, and antivirus systems, the primary data targeted often comprises low-volume text (passwords, authentication keys, emails, and keystroke logs).

Having delineated the concept, operational mechanism, and merits and drawbacks of air gap attacks, Section III will spotlight prominent air gap attacks categorized by transmission mediums like optics, electromagnetic waves, and sound,

further discussing the distinct attributes of each technique.

III. Major Air-Gap Attacks

In environments with external Internet access restricted by the internal network, air-gap attacks utilize various transmission media to exfiltrate data to external networks. As illustrated in Fig. 1, multiple signals originate from the primary components and peripherals within a segregated network, enabling air-gap attacks. Malware tailored for these attacks can manipulate and transmit signals to external networks, deliberately modulate the medium through which signals are passed, and exfiltrate sensitive data (passwords, contacts, and conversations) to external receivers. Air-gap attacks have employed methods to relay data to external networks using diverse transmission media, which originate from internal computer components and peripherals. This section delves into air-gap attacks that employ electromagnetic waves, optics, and acoustics as their primary transmission mediums.

A. Electromagnetic Waves

Electromagnetic waves emanate from an antenna because of the perpendicular relationship between electric and magnetic fields. Air-gap attacks that utilize electromagnetic waves encode sensitive data obtained from air-gap malware within a segregated network into electromagnetic signals innately produced by internal computer components and peripherals; and then transmit them to attackers outside the network. Documented instances of attacks using electromagnetic waves include those exploiting frequency modulation (FM) signals from PC graphics cards [3], frequency-modulated attacks on global systems for mobile communications (GSM) bands from data exchanges on multichannel memory buses between central processing units (CPU) and random-access memory (RAM) [5], and attacks using radio frequency (RF) signals from USB device data buses [6].

Although these attacks are covert, generating the necessary electromagnetic waves by controlling HDDs requires sophisticated methods. A supplementary device is crucial for filtering noise from standard computer operations. In segregated networks with enhanced security protocols, the attack efficiency can be hampered by installing Faraday cages or devices designed to mitigate electromagnetic interference.

B. Optical

Computers use multiple sources of light. The chassis houses LED indicators (e.g., HDD operation and network status LEDs) and monitors that display user information. Furthermore, keyboards contain LEDs for Caps lock, Num lock, and language switching. Certain laptops feature touch bars for varied control and display functions. Several such light sources can be manipulated via firmware. For instance, repeatedly writing and reading a specific file can cause the HDD LED to blink synchronously with the file's access patterns.

Computers in segregated networks possess various light sources that can be manipulated by air-gaps to exfiltrate data for external attackers. Potential optical air-gap attack scenarios include:

- Keyboard LEDs: Control flashing through firmware [3].
- PC HDD LED: Indirectly manipulating LED blinking by accessing the HDD [4].
- Network status LEDs on routers or switches: Indirect LED control by generating network traffic [7].
- Monitor: Modulate brightness, refresh rates, embed discreet imagery, employ steganography, etc. [8], [9]

Advancements in camera technology (telephoto lenses) and the evolution of drones and mobile communication technologies (5G/6G), have allowed attackers to overcome previous limitations and access data from secure, remote locations. However, optical air-gap attacks have inherent limitations: the sender and receiver must have a clear line of sight; signal concealment is challenging (particularly during daylight hours); and overly rapid signal patterns can be detected using security cameras.

C. Acoustic

Typical human hearing spans frequencies ranging 20–20,000 Hz, with those beyond range being inaudible. Notably, PC speakers can produce both types of frequencies.

Air-gap attacks that use sound harness these inaudible frequencies for stealthy data exfiltration from isolated networks. Within a PC, the speaker can generate these sounds [10]. Further, a cooling fan designed to modulate the internal temperature of a PC can produce sounds within the audible range by manipulating the noise generated during operation [11]. Furthermore, data can be transmitted through various computer-generated noise sources. The HDD is equipped with an arm that adjusts the head position on the platter and an actuator that moves the arm; the consequent noise produced can be harnessed during air-gap attacks [12]. Moreover, attacks can exploit the noise generated by capacitors and transformers by adjusting the internal switching frequency of the computer power supply [13].

Acoustic air-gap attacks are stealthy as they use inaudible frequencies that are often undetected by humans. Moreover, such attacks can establish two-way communication channels for data transmission and command and control. However, high-security facilities may employ "Audio Gap" policies and disconnect speakers to prevent such vulnerabilities, thereby limiting air-gap attacks that exploit speaker-generated sounds in protected networks. In environments that adhere to this policy, air-gap attack techniques that harness noise signals from internal fans, HDDs, and power supplies are more appropriate. Conversely, in standard security settings, where speaker usage is permissible, attacks that leverage bidirectional speaker channels are most effective.

D. Other mediums

In addition to the aforementioned media, air-gap attacks can exploit vibrations, and thermal, electric, and magnetic channels. For example, PC fans can be manipulated to send vibration signals, whereas the heat from a CPU can be employed for data transmission to nearby PCs [14]. Variations in power consumption can be used to send signals through the power lines. The distance constraints for these methods

vary and an in-depth understanding is essential for effective exploitation.

Air-gap attacks that utilize heat exploit the warmth generated when electricity is supplied to and utilized by a PC CPUs [15]. They involve malware that increases power consumption across all components (GPU, HDD, RAM, etc.), increasing in the overall temperature of the PC. This is particularly true when the CPU workload is deliberately increased. Consequently, adjusting heating pattern based on the data bits captured by malware facilitates transmission of information to nearby PCs. An adjacent computer employs temperature sensors to interpret these data bits and decode them to recover the original information. To ensure effective communication, the transmitting and receiving PCs must be no more than 40 cm apart. The transmission rate for this method is restricted (maximum of 8 bits per hour), owing to the time required to modulate temperature. However, this approach provides the attacker with command and control capabilities.

In addition to air-gap attacks that leverage vibrations and heat, power, and electromagnetic energy are also employed. Malware designed for air-gap attacks can intentionally alter the CPU workload to adjust the system's power consumption. By encoding the essential data in the fluctuations of the current flow, the data can be transmitted externally through power lines. The recipient device measures this current, deciphers the signal, and relays it to an attacker via Wi-Fi [3]. Two strategies for power signal-based air-gap attacks are: line level power hammer (the attacker connects directly to a domestic power line) and phase-level power hammer (accessing the primary electrical service panel). In the latter, an attacker must physically access the main service panel. In both scenarios, the potential transmission distance is governed by the length of the power line. Electromagnetic wave-based air-gap attacks entail modulating the workload of the CPU core of a compromised PC and transmitting the accumulated data via a produced magnetic signal [3].

This section provides insights into diverse media for air-gap attacks, emphasizing primary transmission channels. Implementing a successful attack on a segregated network requires a comprehensive assessment of factors (transmission distance, environmental influences, and media limitations). For instance, in optical air-gap attacks, the achievable transmission distance for an optical signal can be influenced by ambient lighting, necessitating an assessment of the target network's illumination and light source positioning.

IV. Air-gap Attacks using Optical Mediums

As described in Section III-B, multiple light sources within an air-gap PC serve as potential transmission media for optical air-gap attacks. They include HDD LEDs, Power LEDs, Network Interface Card (NIC) LEDs, and keyboard LEDs. HDD LED are activated during user input/output actions and inherent operating system (OS) processes. The Power LED denotes the power status of the PC, whereas the NIC LED illuminates the PC during data transmission or reception over the network. In case of keyboard LEDs, the Caps Lock, Num Lock, and Scroll Lock illuminate for case conversion, keypad usage, and scroll function management, respectively.

Optical air-gap attacks involve the clandestine transfer of sensitive information by modulating inherent PC light sources (HDD and keyboard LEDs) that are detected using imaging devices (smartphone cameras, high-end cameras, smartwatches, and Google glasses), or through non-imaging optical sensors. Although non-imaging devices achieve faster data reception than imaging devices they are more prone to ambient light interference, typically exhibit reduced reception ranges, and are less portable.

We consider specific optical air-gap attack scenarios, focusing on those that are plausible even under restrictive closed-network conditions. Recognizing that closed networks commonly contain numerous PCs, we discuss attacks that exploit the ubiquitous LEDs. As shown in Fig. 2, an optical air-gap attack primarily comprises air-gap malware within a closed network and external receptors. This malware covertly gathers sensitive information, encodes it by following a protocol coordinated with an external receptor, and modulates the LEDs accordingly. The external receptor, locates the light source in the acquired footage (using imaging devices), decodes the data, retrieves the original information, and relays it to an adversary. We analyzed the roles and functionalities of air-gap transmitters and receptors.

A. Air-Gap Malware

Air-gap malware penetrates a closed network through methods such as supply chain attacks (SCA) or social attacks, equipped with functionalities for internal data collection and deodorization. Consider that the air-gap malware has already infiltrated and infected the internal network computer, henceforth designated as the air-gap PC. The malware comprises four principal modules: data collection, preprocessing, data encoding, and LED controller:

- **Data Collection**: The Air-Gap Malware systematically collects text-based, low-capacity data (passwords, authentication keys, key logs, and emails), which is of primary interest to potential attackers.
- **Pre-processing**: The collected data are transformed into a binary format suitable for transmission. The processes involved are filtering the extraneous data, encryption, and modulation.
- **Data Encoding**: Herein, raw binary data are transmuted via a mutually agreed encoding method with the receiver. Conventional line coding methods (Morse code, Manchester coding, or Non-Return to Zero (NRZ) coding), traditionally used in wireless communication, can be repurposed for air-gap attacks.
- **LED Controller**: This module modulates the LED chosen for air-gap transmission. Considering the diverse LED control methods across various computer manufacturers and models, identifying the target device and integrating the appropriate control logic within the air-gap malware is essential.

Although the act of discreet data collection is a defining feature of air-gap malware, the primary focus is in the covert transmission of the accumulated data. The preprocessing function refines the collected data, emphasizing salient data

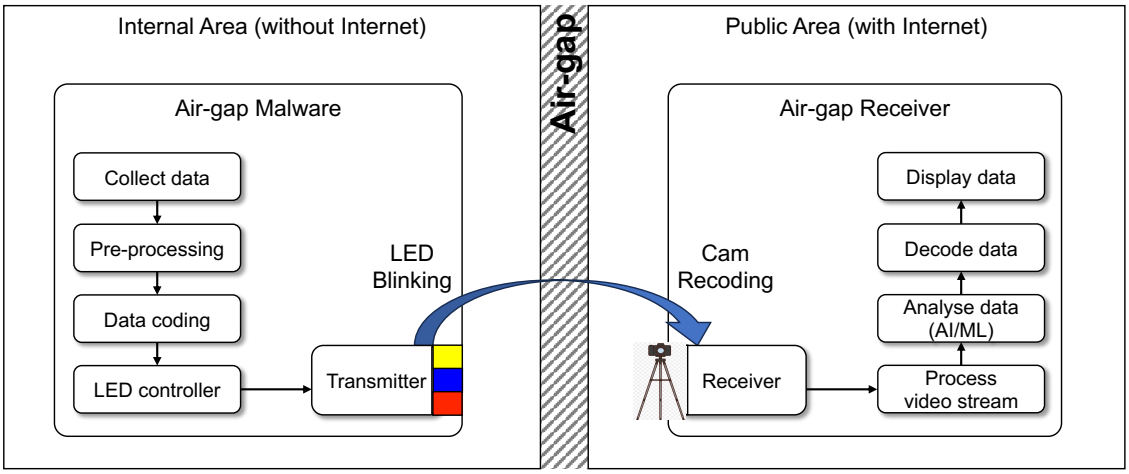


Fig. 2. Optical air-gap attack transmission/reception block diagram using optical devices (HDD LED and Status LED) of a computer

for external transmission. Considering its dependence on the attacker's specific requirements, further discussion of this topic is beyond the scope of this study.

Following preprocessing, the data encoding module converts the data intended for external transmission into binary digits. Both the computer and the receiver understand these digits, enabling optical signal modulation. The Morse code remains a fundamental communication tool that encodes characters through a combination of short (dot, `.`) and long (dash, `-`) signals (the letter '`A`' is represented as a dot followed by a dash). The duration of LED blinks can be manipulated to signify dots and dashes, necessitating prior agreement between the transmitter and receiver on the LED blink durations corresponding to these symbols.

Line coding, which entails the conversion of binary data into a digital signal, introduces functionalities such as e recovery and synchronization to ensure secure data transmission and reception. Synchronized methods are employed to minimize the transmission overhead in air-gap attacks. The data frame begins with a preamble character code for synchronization that, distinguishes it from a unique bit pattern. In this study, the preamble used the sequence '`10101010`' to indicate the frame's beginning. The payload comprises the original binary-converted data, whereas e the Cyclic Redundancy Check (CRC) detects transmission errors and is appended at the end of each frame.

The data encoding module employs various line-coding methods tailored to the transmission medium. Owing to its inherent characteristics, the on–off keying (OOK) modulation technique is suitable for LED-based data transmission. Further, Non-Return to Zero (NRZ), Return to Zero (RZ), and Manchester coding can be utilized for air-gap attack encoding. Selecting the most suitable line code for air-gap attacks requires a thorough evaluation of factors (DC Power, frequency bandwidth, and synchronization elements).

LED control on a PC requires diverse approaches based on the selected LED. While utilizing a HDD LED as an indicator of the storage device's operational status, air-gap malware cannot directly control the HDD LED. Instead, indirect control is realized through the system library provided by the operating system. This method capitalizes on the HDD LED activation whenever any process accesses the HDD. However, the desired data transmission can become intricate because of interruptions caused by the OS task noise signals or file I/O caching processes.

Certain inherent challenges to the creation of bespoke air-gap malware for individual target systems are differences in the processes governed by various operating systems, methodologies for executing read and write commands on HDDs, and distinct system libraries offered by each system.

B. Air Gap Attack Algorithms

In contrast, certain computer LEDs allow for direct manipulation. For instance, certain Network-Attached Storage (NAS) computer systems are equipped with status LED, which can be directly controlled via a dedicated controller. Considering the Synology NAS discussed herein, the target LED can be directly manipulated by sending a specific command bit to `/dev/ttyS1` — a unique controller's device file in the Command Line Interface (CLI) mode. For example, if the input bitstream value is "`1`", the command "`echo > /dev/ttyS1`" activates the status LED, glowing orange for a specified duration (T0). If the value is "`0`", the command "`echo 7> /dev/ttyS1`" deactivates the LED status for the same period. The data encoded by the internal network can be relayed to external networks using LED illumination as the signaling mechanism.

The algorithm shown in Fig. 3 presents the operational methodology for the air-gap malware transmission facet. The `LED_transmitter` integrates two variables: a frame composed of bits designated for transmission, and a specified time (`T`) defining the LED's duration for each bit.

Throughout the transmission process, each bit s analyzed by modifying the LED state based on the bit value. The LED remains off for value "`0`" and on for value "`1`" (lines 4 and 7). The LED state can be altered using the NAS computer

Algorithm for LED controller

```
1  Procedure transmitter (frame, T)
2  for (bit in frame)
3      if (bit == '0') then
4          system('echo 7> /dev/ttyS1');
5          sleep(T);
6      if (bit == '1') then
7          system('echo :> /dev/ttyS1');
8          sleep(T);
9  end for
10 Return;
```

Algorithm for Air-gap Receiver

```
Input : Real-time camera screen
Output : String extracted from optical signal
1  x, y ← Pixel coordinates of optical signals in the camera screen
2  Frame ← On a real time basis, the frame image transformed to the HSV
3
4  while true
5      brightness ← frame[x,y] brightness value
6      threshold  ← The standard value deciding the ON/OFF of LED
7
8          if brightness > threshold
9              state On, array.append(1)
10         else if brightness <= threshold
11             state Off, array.append(0)
12         if state Off is maintained
13     break
14
15 Return morsecoding(array)
```

Fig. 3. Algorithms are provided for both the transmitter (which is part of the air-gap malware inside the closed network) and the receiver (located in the external network).

controller, which modulates the LED illumination by sending a designated byte. A single byte is relayed by the commands "`echo`" and 'redirect' (>) . The value "`7`" turns the LED off, whereas "`:`" activates it. As these values differ based on the target computer system, air-gap malware must be tailored to match the specifics of the system after a meticulous assessment of air-gap attacks. Additionally, to ensure precise identification of the LED state, its status (either on or off) is maintained for a predetermined duration (`T`) (lines 5 and 8).

C. Air-Gap Receiver

An air-gapped PC within a closed network communicates data using blinking LEDs, which are external network receivers with direct line-of-sight capture. This receiver is typically equipped with a camera, light reception sensor, and software to process the received images and signals. The camera was considered a part of the receiver setup. Algorithms that leverage pattern recognition or artificial intelligence can extract and digitize the LED blinking pattern from a captured image.

The operating procedure of the receiver, as in Fig. 3, begins by identifying the location of the LED within the captured image. The LED on/off signal within the frame is assessed, and its brightness value, after conversion to the Hue, Saturation, and Value (HSV) color spectrum, is extracted in real time (line 5). Depending on the brightness, the LED status is determined as 'On' or 'Off,' and the respective binary values (`1` to 'On' and `0` to 'Off') are added to the data array (lines 8–11). The collected binary stream is subsequently decoded synchronously using an algorithm with the transmitter encoding method to unveil the original data!(line 15).

In addition to the aforementioned setup and algorithm, the following factors also need to be considered when carrying out optical air-gap attacks in actual closed-network environments:

- The number of PCs having LED in the server room: If there are many PCs and peripheral devices in the server room, it can be difficult to identify the LED controlled by the air-gap malware.
- The location of the air-gapped devices: The air-gapped devices may be located in different places in the server room, which can affect the transmission and reception of optical signals.
- The ambient light conditions: The ambient light conditions in the server room can affect the sensitivity of the camera and the visibility of the optical signals.

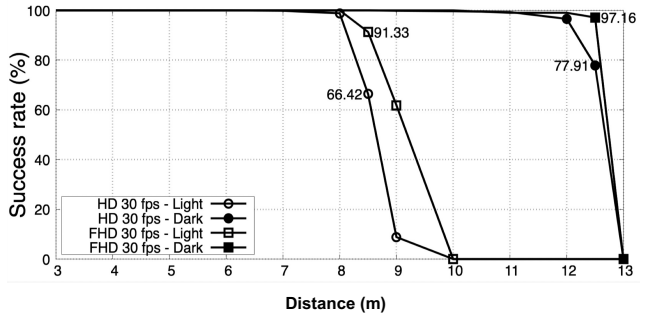
To mitigate the challenges associated with identifying the desired air-gapped PC and improving the success rate of optical air-gap attacks in real-world closed-network environments, several approaches can be considered. For example, employing a more sophisticated receiving device can significantly improve the detection of optical signals in noisy environments. Such devices often possess higher sensitivity and superior noise-filtering capabilities. Utilizing a more powerful light source can enhance the signal strength, making it easier for the receiver to detect. Also, implementing a directional light source can focus the light beam, reducing the amount of scattered light in the environment and potentially improving signal clarity.

However, these methods of improving the success rate or efficiency of the air-gap attack receiver require prior work, such as replacing computer components. Since the approach to achieving this varies depending on the specific closed network environment and the location of target internal PCs, significant customization is necessary for the receiver's design and configuration.

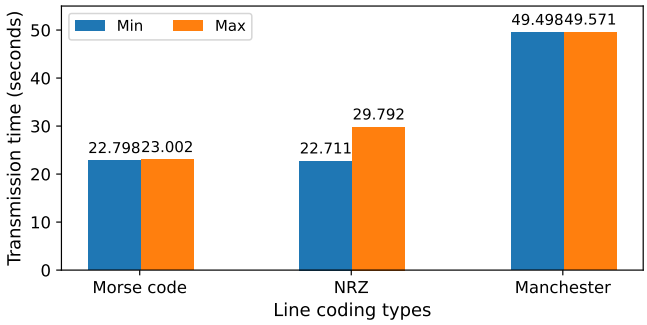
V. Optical Air-Gap Attack and Experiments

We conducted experiments and measurements to evaluate the effectiveness of air-gap attacks utilizing optical signals, specifically PC LEDs, in various isolated network environments with multiple constraints. We configured the experimental setup inside the server room containing multiple NAS devices and a single window facing the inner corridor. The NAS device, positioned within this isolated network and equipped with a status LED, was oriented toward this window, enabling observation by a receiver (Samsung Galaxy smartphone) located in the external network. The NAS device inside the isolated network was from Synology.

The air-gap malware controlled the LED by detecting the signal managing the Status LED in the Disk Station Management software provided by Synology. This detection was achieved through reverse engineering and the signal was subsequently sent to `/dev/ttyS1`, a specialized controller device, using a command-line interface (CLI). The metrics that assess the potential success of an attack in a genuinely isolated network scenario were identified. The variations in



(a)



(b)

Fig. 4. (a) The transmission success rate that varies depending on the resolution and illuminance of the video recorded at the receiver and the distance between the transmitter and receiver. (b) The transmission rate of data 'password' six characters encoded using three different line codings: Morse code, NRZ, and Manchester. Each coding employs a unique encoded data length and schemes: 65 bits for Morse code, 96 bits for NRZ, and 160 bits for Manchester.

the transmission reception rates affected by factors (ambient light, transmission distance, and resolution disparities) were examined.

Illuminance was measured against a standard of 2 lux (lx) in the absence of sunlight and artificial light, and 240 lx under bright conditions. The distance between the transmitter and receiver was varied as 3– 13 m (1 m increments), providing insights into the impact on optical air-gap transmission. The smartphone camera's resolution was adjusted between High Definition (HD) (1280 × 720) at 30 frames per second (FPS) and Full HD (FHD) (1920 × 1080) at 30 FPS for different trials.

As shown in Fig. 4a, under bright illumination within the isolated network, the transmission reception rate received at HD resolution on a smartphone camera dramatically decreased from 100 to 8.8% over an 8–9 m distance. Beyond 10 m, no LED signals were detected. The 8–9 m range was further segmented (0.5 m intervals). This meticulous measurement revealed a reduction from 100 to 66.4% between 8–8.5 m and a drastic drop from 66.4 to 8.8% in 8.5– 9 m. Consequently, under brightly lit isolated network conditions, the optimal attack transmission range was obtained as 8 m for HD resolution. Under identical lighting, the signal at the FHD resolution declined from 100 to 61.8% between 8–9 m and a 61.8 to 0% between 9–10 m. A steady decreased of 100 to 91.3% for 8–8.5 m and down to 61.8% for 8.5–9 m was observed. This confirmed an optimal attack transmission range of 8 m for FHD resolution under bright conditions without any additional equipment such as a military telephoto lens.

Given the regular movements of personnel in a server room during daylight hours, optical air-gap attacks may be more feasible at night with less human activity. An additional experiment was conducted under dim lighting to simulate night-time conditions. Using the HD resolution, the reception rate decreased slightly from 100 to 96.5% between 12–13 m. Further analysis of 12–12.5 m revealed a drop from 96.5 to 77.9% between 12.5–13 m, rendering LED signals undetectable beyond 13 m. This established an optimal attack transmission range of 11 m for HD resolution in dim lighting. For the FHD resolution, the reception rate decreased progressively from 100 to 96.5% between 11–12 m and to 77.9% between 12–13 m. The LED signals were indiscernible beyond 12.5–13 m; Thus, the optimal attack transmission distance for FHD was 12 m under subdued lighting.

To further investigate the impact of line coding schemes on message transmission speed in optical air-gap attacks, we conducted an additional experiment. We encoded the same 6-character message ("password") using three different line coding schemes (Morse code, NRZ, and Manchester coding) and transmitted it to the receiver under varying illumination conditions. The time taken for transmission was measured 10 times for each coding scheme.

The results, presented in Fig. 4b, demonstrate that Morse code exhibited the fastest transmission speed, with an average time of 22.9 seconds. NRZ coding was slightly slower, with an average time of 26.25 seconds, while Manchester coding was the slowest, with an average time of 49.53 seconds. These findings align with our previous experiments, as illumination only impacts the attack's success rate, reinforcing the importance of balancing transmission speed and security when choosing a coding scheme for optical air-gap attacks.

VI. Air-Gap Attack Using Smart Bulbs

The increasing prevalence of IoT devices in smart homes and workplaces, designed for enhanced automation and convenience, has extended to isolated networks. Typical IoT devices include smart bulbs and household appliances, which can be manipulated for air-gap attacks by leveraging signals from built-in features such as LEDs, batteries, and fans. This study specifically explored the potential of optical air-gap attacks utilizing smart bulbs because of their energy efficiency and user-friendly features.

Smart bulbs, which are emblematic IoT devices, can produce a wide range of colors. Users can also fine-tune the desired color intensity and brightness using a separate controller. Considering their larger light-emission surfaces compared to computer peripherals, an overt flashing method to relay data may jeopardize the attack's discreetness in an isolated network. Hence, subtle adjustments in color or brightness are better suited for optical air-gap attacks using smart bulbs.

The proposed optical air-gap attack scenario using the brightness modulation of a smart bulb begins by scanning the network to detect and control smart bulbs within the

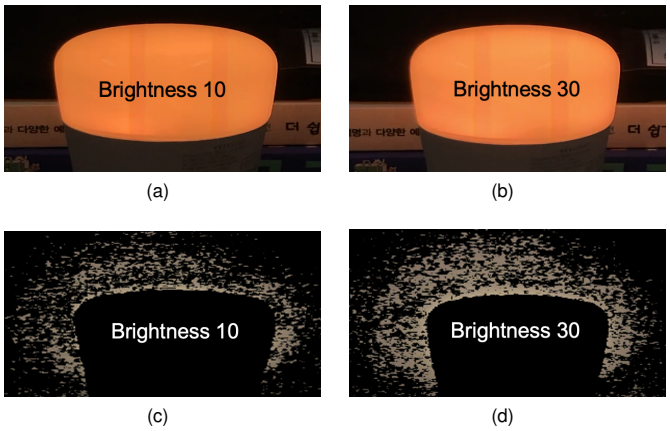


Fig. 5. Changes of smart bulb brightness: (a) Raw image with brightness level 10. (b) Raw image with brightness level 30. (c) Filtered image with brightness level 10. (d) Filtered image with brightness level 30. In the case of the filtered image, the size of the surrounding area varies based on changes in brightness.

same isolated network. Malwares can gain control of these bulbs by exploiting known vulnerabilities. The acquired data is transmitted to an external receiver by discreetly varying the brightness that is undetectable by nearby humans. A camera at the receiving end captures the brightness changes. Analyzing the captured images allows the detection of brightness variations, from which signals are extracted and the data are decoded.

In our experiment, we utilized a YeeLight smart bulb from Xiaomi, with a manufacturer set range of 1–100. When connected to the same internal Wi-Fi network as the smart bulb, we analyzed the network packets destined for the bulb to extract information about the color and brightness changes. Air-gap malware was designed to accurately adjust the brightness. A brightness level of 30 was associated with binary "1", whereas a level of 10 represented "0". For reference, the brightness levels of 10 and 30 were 80 and 240 lm, respectively in our experiments.

When transmitting signals through brightness modulation of smart bulbs, an external attacker requires a mechanism to detect these changes. Although using a camera for such detection appears intuitive, capturing minor brightness shifts in raw footage is not trivial. Fig. 5 (a) and (b) illustrate this challenge, showing minimal differences between brightness levels of 10 and 30. Thus, we leveraged the light-dispersion properties of bulbs as a function of brightness. Essentially, the extent of light emission varied with bulb brightness. Employing color filtering, which retains only specific hues in a frame, renders the differential brightness more evident. As shown in Fig. 5 (c) and (d), with brightness levels of 10 and 30, respectively, the latter emitted more radiant light dispersion. We used OpenCV's contour function, a technique for outlining regions of uniform color or intensity, which enabled us to distinguish light intensities based on the contour area.

VII. Further Analysis and Discussion

In the previous sections, this paper has reviewed and analysed various air-gap attacks, identifying and presenting potential mediums that can be exploited for such attacks. and conducted experiments in a realistic closed-network environment, applying various constraints to derive concrete numerical values necessary for the successful execution of optical air-gap attacks. Moreover, the practicability of utilizing IoT devices for air-gap attacks within closed networks was demonstrated, reflecting upon recent technological advancements.

Attack mediums and restrictions: Table I [3] presents an additional analysis of the various air-gap attack mediums discussed earlier, showing the distance required for data transmission and the speed required to transmit 256 bits. Furthermore, a detailed examination of the restrictions inherent to each attack medium is provided. These insights can be instrumental in developing countermeasures against corresponding air-gap attacks. For instance, concerning the optical attack explored in this paper, neutralization of the air-gap attack can be achieved through the design of structures that obstruct the establishment of a direct line-of-sight or by employing an ancillary light source capable of manipulating the flickering of the initial light source. Additionally, in scenarios involving acoustic mediums, the application of sound-absorbing materials to structural elements can markedly diminish the probability of a successful air-gap attack employing acoustic mediums. Overall, Table I shows the existing body of knowledge on air-gap attacks by providing a comprehensive analysis of various attack mediums, their limitations, and potential countermeasures. The findings of this study can be used to improve the security of closed networks by implementing appropriate countermeasures against air-gap attacks.

Air-gap attacks using IoT: In this paper, we demonstrate that a smart light bulb IoT device located inside an air-gapped network can be exploited for an air-gap attack. The smart light bulb used in the experiment, as well as various other IoT devices, can be exploited for air-gap attacks. The following shows additional IoT devices that can be exploited for air-gap attacks, along with their characteristics and availability.

- Smart home camera LED: Controllable LEDs enable data transmission, ideal for covert operations due to in-room placement. However, low light intensity limits their range.
- MacBook Touch Bar: Malware installed on the target PC facilitates rapid data transfer via brightness changes or barcode displays. This offers high-speed transmission but requires compromising the target system.
- Smart humidifier: Discreet data encoding through humidity adjustments makes them stealthy. However, the analog output used for humidity control introduces noise, hindering reliable data transmission.

Practical considerations: Our study demonstrates that the feasibility of real-world air-gap attacks depends on the specific configuration of the targeted closed-network environment. For a successful air-gap attack, the following factors need to be considered:

- Physical configuration of the closed network: This includes information such as the specific floor or location of the closed network office, the presence or absence of windows connecting to the outside, and the structure of

TABLE I
CHARACTERISTICS AND LIMITATIONS OF VARIOUS AIR-GAP ATTACK MEDIUMS [3]

Type	Transmission mediums	Capabilities	Restrictions
Optical	HDD LED, Keyboard LED, Switch Status LED	Distance can be extended with tools such as telescope, High transmission rate (less than 2 sec to sent 256 bits)	• The transmitter and receiver must be positioned in a line-of-sight path without obstructions. • Interference may occur from factors such as lighting and sunlight within the closed network environment.
Electro-Magnetic	Graphic Card, USB Data Bus, SATA Cable	Distance (1 7 m), High transmission (less than 2 sec to transmit 256 bits)	• Operation is possible within a limited range where electromagnetic signals can reach. • Errors can occur due to factors such as signal reflection, fading, frequency interference, and Faraday cages on the transmission and reception routes. • Errors can occur if structures coated with electromagnetic/acoustic absorbing materials that hinder proper signal transmission are present within the closed network.
Acoustic	PC FAN, Speaker/Mic, HDD Noise	Distance (around 5 8 m), Speed is very depending on mediums (e.g., Speaker 2 20 sec, and FAN noise around 1,500 sec to send 256 bits)	• Operation is possible within a limited range where acoustic signals can reach. • Errors can occur if structures coated with electromagnetic/acoustic absorbing materials that hinder proper signal transmission are present.
Vibration	PC FAN	Distance (less than 2 m), Low transmission (more than 512 sec to send 256 bits)	• Errors can occur in conditions where road surface materials, cracks, etc., do not transmit vibration signals well.
Thermal	Thermal Sensor	Short distance, Very slow (32-256h to send 256 bits)	• A minimum separation distance must be maintained between adjacent PCs for optimized thermal signal transmission.
Electric	PC powerline cable	Length of powerline cable, Medium (30 300 sec to send 256 bits)	• Concealment is compromised due to the need for additional devices to extract power signals from power outlets. • Error can occur due to limited precision in power signal extraction.
Electronic signal	CPU	Very short distance (0 12 cm), High transmission (6 50 sec to send 265 bits)	• Operation is possible within a limited range where electromagnetic signals can reach. • Errors can occur due to Faraday cages within the closed network.

the closed network walls.

- Computers and peripheral devices located inside the closed network: This includes information such as general computers, keyboards, mice, NAS devices, routers, IoT devices, speakers, and so on.
- Location of the external receiver: This refers to the location of the receiver outside the closed network where it can effectively receive information transmitted by the air-gap malware.
- Infection location of the air-gap malware and system information of the infected host: This includes information such as the operating system, running server daemons, and versions of installed software.

By understanding these factors, attackers can tailor their air-gap attacks to the specific characteristics of the targeted closed network, increasing the likelihood of success. Additionally, defenders can use this information to strengthen the security of their closed networks by implementing appropriate countermeasures against the identified attack vectors.

Countermeasure: Countermeasures against air-gap attacks can vary depending on the type of attack medium. For instance, in the case of optical air-gap attacks utilizing LEDs on computer peripherals, the most straightforward countermeasure would be to disable the LED blinking function of a keyboard or to cover it with black tape. However, such methods restrict the usability of the original function. Additionally, CCTV monitoring devices can be used to observe LED blinking signals within the closed network. However, this necessitates continuous monitoring by the closed network security administrator and is inefficient due to the high possibility of false positives.

We believe that artificial intelligence (AI) and machine learning (ML) technologies can be effectively utilized to detect and mitigate air-gap attacks. Our proposed method involves identifying light sources within the closed network that could be exploited for optical air-gap attacks (e.g., LEDs, monitors, lights) and training an AI model to recognize their normal blinking signal patterns. Since a light source controlled by air-gap malware would generate abnormal blinking patterns for the purpose of information transmission, the pre-trained optical air-gap countermeasure model is expected to effectively detect such abnormal blinking signals. This artificial intelligence-based air-gap attack countermeasure technique holds the potential to be applied not only to optical attacks but also to other air-gap attack mediums.

VIII. CONCLUSION

The closed-network policy, which distinctly separates external from internal networks, remains widely used. However, as the proliferation of interconnected devices has increased, attackers have persistently devised new methods for bridging these isolated networks. One method is the optical air-gap attack discussed in this study.

This study explores diverse media that can facilitate air-gap attacks and analyzes attack scenarios tailored to each medium. We developed an optical air-gap attack system using various LEDs attached to a computer. Our experiments were conducted using different closed network environmental conditions (illumination, reception distance, receiver performance, etc.), aimed at discerning the conditions favoring the success of air-gap attacks. Our results indicated that under bright interior illumination of the closed network, the effective transmission range was approximately 8 m. Under dimmer conditions, this

range extended to approximately 12 m. The resolution performance of the receiving device exerted minimal influence on the transmission reception rate or the expansion of the effective transmission distance; whereas the interior illumination of the closed network exerted heavy influence on these factors in optical air-gap attacks using LEDs.

Thus, launching an attack is more advantageous during nighttime or early morning hours, when lighting conditions are subdued. For the highest likelihood of success, an attacker should position themselves within 8 m of the transmitter in brightly lit conditions of the closed network and within 12 m when the illumination is lower. Moreover, using a telephoto lens into a receiver can substantially extend its transmission range. However, this tactic hinges on the requirement that the window of the closed network is faced externally.

Historically, the defense strategies proposed for air-gap attacks have been somewhat limited, often restricted to the deactivation of specific functionalities of closed-network PCs, establishing isolated spaces to deter unauthorized access, or the introduction of electronic devices. However, these strategies do not offer comprehensive solutions. In the future, we intend to focus on the potential of machine-learning software-based defense mechanisms against air-gap attacks, aiming to bolster the defense of closed networks against a wide range of transmission mediums.

ACKNOWLEDGMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2022-00165231) and Agency for Defense Development (ADD) and Defense Acquisition Program Administration (DAPA). (No. UD230020TD) Professor Song is the corresponding author.

REFERENCES

[1] Q. Chen and R. A. Bridges, "Automated behavioral analysis of malware: A case study of wannacry ransomware," in *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2017, pp. 454–460.

[2] L. M. Castiglione and E. C. Lupu, "Which attacks lead to hazards? combining safety and security analysis for cyber-physical systems," *IEEE Transactions on Dependable and Secure Computing*, pp. 1–16, 2023.

[3] J. Park, J. Yoo, J. Yu, J. Lee, and J. Song, "A survey on air-gap attacks: Fundamentals, transport means, attack scenarios and challenges," *Sensors*, vol. 23, no. 6, 2023. [Online]. Available: https://www.mdpi.com/1424-8220/23/6/3215

[4] M. Guri, B. Zadov, and Y. Elovici, "Led-it-go: Leaking (a lot of) data from air-gapped computers via the (small) hard drive led," in *Detection of Intrusions and Malware, and Vulnerability Assessment*, M. Polychronakis and M. Meier, Eds. Cham: Springer International Publishing, 2017, pp. 161–184.

[5] M. Guri, A. Kachlon, O. Hasson, G. Kedma, Y. Mirsky, and Y. Elovici, "GSMem: Data exfiltration from Air-Gapped computers over GSM frequencies," in *24th USENIX Security Symposium (USENIX Security 15)*. Washington, D.C.: USENIX Association, Aug. 2015, pp. 849–864. [Online]. Available: https://www.usenix.org/conference/usenixsecurity15/technical-sessions/presentation/guri

[6] M. Guri, M. Monitz, and Y. Elovici, "Usbee: Air-gap covert-channel via electromagnetic emission from usb," in *2016 14th Annual Conference on Privacy, Security and Trust (PST)*, 2016, pp. 264–268.

[7] M. Guri, B. Zadov, A. Daidakulov, and Y. Elovici, "xled: Covert data exfiltration from air-gapped networks via switch and router leds," in *2018 16th Annual Conference on Privacy, Security and Trust (PST)*, 2018, pp. 1–12.

[8] M. Guri, O. Hasson, G. Kedma, and Y. Elovici, "Visisploit:an optical covert-channel to leak data through an air-gap," in *2016 14th Annual Conference on Privacy, Security and Trust (PST)*, 2016, pp. 642–649.

[9] M. Guri, D. Bykhovsky, and Y. Elovici, "Brightness: Leaking sensitive data from air-gapped workstations via screen brightness," in *2019 12th CMI Conference on Cybersecurity and Privacy (CMI)*, 2019, pp. 1–6.

[10] M. Guri, Y. Solewicz, and Y. Elovici, "Mosquito: Covert ultrasonic transmissions between two air-gapped computers using speaker-to-speaker communication," in *2018 IEEE Conference on Dependable and Secure Computing (DSC)*, 2018, pp. 1–8.

[11] M. Guri, Y. Solewicz, and Y. Elovici, "Fansmitter: Acoustic data exfiltration from air-gapped computers via fans noise," *Comput. Secur.*, vol. 91, no. C, apr 2020.

[12] M. Guri, Y. Solewicz, A. Daidakulov, and Y. Elovici, "Acoustic data exfiltration from speakerless air-gapped computers via covert hard-drive noise ('diskfiltration')," in *Computer Security – ESORICS 2017*. Springer International Publishing, 2017, pp. 98–115.

[13] M. Guri, "Power-supplay: Leaking sensitive data from air-gapped, audio-gapped systems by turning the power supplies into speakers," *IEEE Transactions on Dependable and Secure Computing*, vol. 20, no. 1, pp. 313–330, 2023.

[14] M. Guri, "Exfiltrating data from air-gapped computers via vibrations," *Future Generation Computer Systems*, vol. 122, pp. 69–81, 2021.

[15] M. Guri, M. Monitz, Y. Mirski, and Y. Elovici, "Bitwhisper: Covert signaling channel between air-gapped computers using thermal manipulations," in *2015 IEEE 28th Computer Security Foundations Symposium*, 2015, pp. 276–289.