Internet of Things Software and Hardware Architectures and Their Impacts on Forensic Investigations: Current Approaches and Challenges

Abel Alex Boozer  
*University of Alabama, Huntsville, ab0166@uah.edu*

Arun John  
*University of Alabama, Huntsville, aj0126@uah.edu*

Tathagata Mukherjee  
*University of Alabama in Huntsville, tathagata.mukherjee@uah.edu*

Follow this and additional works at: [https://commons.erau.edu/jdfsl](https://commons.erau.edu/jdfsl)

Part of the Computer Law Commons, Information Security Commons, and the Other Computer Sciences Commons

**Recommended Citation**  
Boozer, Abel Alex; John, Arun; and Mukherjee, Tathagata (2021) "Internet of Things Software and Hardware Architectures and Their Impacts on Forensic Investigations: Current Approaches and Challenges," *Journal of Digital Forensics, Security and Law*. Vol. 16, Article 4. Available at: [https://commons.erau.edu/jdfsl/vol16/iss2/4](https://commons.erau.edu/jdfsl/vol16/iss2/4)
ABSTRACT

The never-before-seen proliferation of interconnected low-power computing devices, patently dubbed the Internet of Things (IoT), is revolutionizing how people, organizations, and malicious actors interact with one another and the Internet. Many of these devices collect data in different forms, be it audio, location data, or user commands. In civil or criminal nature investigations, the data collected can act as evidence for the prosecution or the defense. This data can also be used as a component of cybersecurity efforts. When data is extracted from these devices, investigators are expected to do so using proven methods. Still, unfortunately, given the heterogeneity in the types of devices that need to be examined, few widely agreed-upon standards exist. In this paper, we look at some of the architectures, current frameworks, and methods available to perform forensic analysis of IoT devices to provide a roadmap for investigators and researchers to form the basis of an investigation.

Keywords: Internet of things, IoT, forensics, architecture, tools

1. INTRODUCTION

The never-before-seen proliferation of interconnected low-power computing devices, patently dubbed the Internet of Things (IoT), is revolutionizing how people, organizations, and malicious actors interact with one another and the Internet. These devices are used in a variety of applications in homes and commercial fields like medicine, education, and transportation (Al-Fuqaha, Guizani, Mohammadi, Aledhari, & Ayyash, 2015). The IoT has introduced a new paradigm of machine interconnectedness by allowing IoT-capable devices to communicate with each other to share information. IoT devices provide a wealth of information about their surroundings, and their use in cyberattacks is of great interest to forensic investigators looking to learn more about individual attacks, the organizations involved, and their implications in crime scenes. Consequently, because these relatively simple devices collect large amounts of data, all of the processing that enables their functionality is performed by computers connected via the Internet. As a result, the simple hardware of the node devices, paired with the complex software of the destination computers, provides unique challenges for forensics teams (Yaqoob, Hashem, Ahmed, Kazmi, & Hong, 2019).

The heterogeneous nature of IoT nodes further contributes to the challenges of IoT forensics. These nodes are made by numerous global manufacturers and use varying software, hardware,
and distributed network architectures. This heterogeneity makes the invention of a universal tool or reference standard to aid in performing IoT forensic investigations highly unlikely (Al-Sadi, Chen, & Haddad, 2018; Guth, Breitenbucher, Falkenthal, Leymann, & Reinfurt, 2016). This unique issue is a significant challenge in the emerging field of IoT forensics. The lack of a standard framework or toolkit applicable to a variety of IoT endpoints creates a unique environment for investigators to navigate and requires experts that are highly specialized in the lowest level technical aspects of the IoT (Oriwoh, Jazani, & Epiphaniou, 2013). Previous works by Stoyanova et al. (Stoyanova, Nikoloudakis, Panagiotakis, Pallis, & Markakis, 2020) and Atlam et al. (Atlam, El-Din Hemdan, Alenezi, Allassafi, & Wills, 2020) focus on providing an overall review into the field of IoT Forensics.

This paper aims to guide the data discovery process for IoT forensics by examining and addressing the challenges of IoT forensic investigations from a practical standpoint. We examine the hardware and software architectures of specific IoT devices as well as developing paradigms of IoT devices in terms of their suitability for forensic investigations. Additionally, we consider what areas of the network relevant data may reside, challenges of accessing those areas, and methods and frameworks developed to assist in those investigations. This paper presents a series of research surveys involving IoT devices, corresponding frameworks, and methods aimed at forensic investigators. Thereby, we provide forensic investigators and researchers a roadmap to forensic approaches for standard IoT devices, challenges present in IoT forensics, and methods and frameworks for approaching current IoT forensics and new IoT paradigms.

The rest of this paper is organized as follows. Section 2 briefly discusses the concept of digital forensics and IoT forensics. Section 3 presents a generic overview of IoT architectures. Specific devices and families, their software-hardware and network architectures, as well as developing IoT paradigms and their respective challenges, are presented in Section 3. Section 4 provides a discussion of the available methods and frameworks and which device families they apply to, as well as modifications that may extend them to other device families. Section 5 summarizes lessons learned and concludes the study.

2. DIGITAL AND IOT FORENSICS

2.1 Overview of Digital Forensics

Digital forensics is the science of the identification and interpretation of information contained within digital devices, including IoT devices, in a way that preserves the integrity of the data. Like traditional forensics, the critical use of digital forensics is to prepare evidentiary documents for legal proceedings. Thus, digital forensics relies heavily on scientifically validated acquisition methods and tools in order to produce forensically sound results that can withstand the scrutiny of the judicial system (Chernyshev, Zeadally, Baig, & Woodward, 2018). This lack of standards poses an exciting challenge to forensic examiners working on newer, non-standard devices or data storage systems that cannot support entrenched tools, frameworks, and methodologies. In the modern era of computing, forensic examiners must often develop and validate these on their own in order to target specific systems (Watson & Dehghantanha, 2016).

Research in digital forensics is essential to discover new methods to extract information and validate current digital forensics methods. Research is especially imperative in IoT forensics, where many devices are heterogeneous in design and often developed using proprietary computer organizations and architectures that will often not be available for access under standard modes of operation. In addition to on-device forensics, the requirements of IoT investigations may extend to the examination of cloud or network data. Thus, we can break down the field of IoT forensics into three subcategories based on the location of forensic artifacts: on-device, network-level, and cloud forensics (Chernyshev et al., 2018; Stoyanova et al., 2020). These three subdivisions provide varying types of forensic evidence contained in different artifacts depending
upon which layer of the IoT the extraction takes place. We expand upon this forensic model and generalize it into an architecture model in Section III of this paper.

### 2.2 What is IoT Forensics?

Due to the wide variety of IoT devices, it is hard to define IoT forensics using specific devices. Instead, digital forensic methods to extract information associated with or contained within IoT devices can be considered IoT forensics. The heterogeneity of IoT devices leads to multiple areas within digital forensics, such as network forensics, memory forensics, cloud forensics, hardware forensics, and many more.

### 2.3 Why Do IoT Forensics?

In addition to the wealth of information that can potentially be extracted from IoT devices and their component networks, the vast domain of applications of IoT devices is a primary reason those forensic examiners will want to pursue IoT forensic investigations. Unlike traditional digital forensics, which focuses primarily on in-device data from computers, laptops, portable storage devices, and other personal devices, IoT expands the scope of computing (and therefore digital forensics) to monitoring systems, vehicles, healthcare devices, surveillance systems, and intelligent home systems (Jahankhani & Ibarra 2019; Huang, Lu, & Choo 2017).

Furthermore, IoT forensics has many more applications than simply collecting data from a crime scene. Systems can be created by forensic examiners or cybersecurity professionals for the express purpose of collecting artifacts in a network in order to detect cyber attacks or to assist in forensically examining an attack post-execution (Zhang, Upton, Beebe, & Choo 2020; Chhabra, Singh, & Singh 2018; Widiyasono, Putra, Giriantari, & Sudarma 2019). These systems are most often found in cybersecurity applications on private networks as a cyber defense strategy. Additionally, systems can be developed with forensic-aware architectures that provide easier access to artifacts and exposes log files, network information, and on-device data more easily (Zawoad & Hasan 2015). It should be noted that both of these examples are only highly applicable in cyber defense contexts where privacy is a non-issue, whereas, in public investigations, this would be difficult to implement without infringing upon an individual’s right to privacy. Unfortunately, unlike traditional digital forensics and mobile forensics, these laws are somewhat ill-defined in the US code of law as far as IoT devices are specifically concerned (Weber 2010; Maras 2015; Peppet, 2014).

### 3. GENERIC OVERVIEW OF IOT DEVICE ARCHITECTURES

Attempting to design a generic architecture for IoT has been the subject of much research, though no apparent standard exists. This lack of standards can be partially attributed to the heterogeneous nature of IoT ecosystems and the lack of universal standardization indicating exactly how an IoT device should operate. Many authors have attempted a realization of a generic architecture and seem to have arrived, apparently independently, at a similar five-layer model (Zhong, Zhu, & Huang 2015; Mrabet, Belguith, Alhomoud, & Jemai 2020; Guth et al. 2016).

We present a version of this layered model in Table I as it is beneficial to concretely define areas of an IoT ecosystem when discussing varying devices, tools, frameworks, and implementations. Since, in certain implementations, components may have altered functions or be omitted entirely, it is convenient to keep this model abstract, not potentially to exclude any systems.

#### 3.1 Physical Interface Layer

The physical interface layer (Mrabet et al. 2020) is the fundamental element of an IoT device. On-device, this layer is composed of hardware sensors or actuators that allow the device to interface with the physical world around it. This interface may be gathering information and translating it into data in the case of sensors or physically manipulating something in its vicinity via mechanical action. An example of an IoT device that uses sensors would be an Amazon Alexa or Google Nest. In contrast, an example of an IoT
Table 1: Generic IoT device Architecture

<table>
<thead>
<tr>
<th>Layer</th>
<th>Components</th>
<th>Artifacts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical Interface Layer</td>
<td>Sensors, Actuators</td>
<td>Protocol packets, routing tables, device</td>
</tr>
<tr>
<td></td>
<td></td>
<td>identifiers, raw sensor or actuator data</td>
</tr>
<tr>
<td>Device Layer</td>
<td>Device-level hardware and software, Operating</td>
<td>Bytes from memory, logs, application data,</td>
</tr>
<tr>
<td></td>
<td>System</td>
<td>authentication data, containers</td>
</tr>
<tr>
<td>Network/Transport Layer</td>
<td>Network-specific protocols, TCP/IP, UDP/IP,</td>
<td>Packet traces, firewall alerts</td>
</tr>
<tr>
<td>Presentation Layer</td>
<td>Speakers, screens, user interfaces</td>
<td>User-level information, usage information</td>
</tr>
<tr>
<td>Integration Layer</td>
<td>Cloud services, IoT middleware, companion apps</td>
<td>Human-readable data, logs, usage history, user</td>
</tr>
</tbody>
</table>

A device that uses actuators would be a network-connected robotic arm (often found in industrial applications), an intelligent pacemaker, or an IoT oven. This layer will often require localized software to operate its components and may store artifacts of its sensor or actuator data locally, in the cloud, or on a companion device (such as a smartphone or tablet). Many non-cyber crime forensic investigations will focus on artifacts generated at this layer.

3.2 Device Layer

The device layer (Mrabet et al. 2020) simply refers to the device itself and the associated hardware and software that composes it. The device is responsible for controlling the sensing and actuating at the physical interface layer, performing operations on that data, sending the data to the integration layer via the network/transport layer, receiving data from the integration layer, and finally translating and presenting that data at the presentation layer. Like any other modern computer, devices are controlled by operating systems with varying degrees of sophistication and integrate with the physical interface layer and the network/transport layers using driver software. Specific hardware may be implemented on a device that reduces the need for OS or driver overhead. However, it can generally be assumed that both of these software are needed to enhance reliability and compatibility through updates except in particular instances. This paper features a more in-depth discussion of specific devices at each layer in section IV.

3.3 Network/Transport Layer

The network layer (Mrabet et al. 2020) consists of any wired or wireless protocols supporting IoT networking. This layer includes Ethernet, Wi-Fi, Bluetooth, ZigBee, 5G, and associated transport protocols such as TCP/IP and UDP/IP. This layer is the layer at which device-to-device or device-to-gateway connections occur and may involve the sharing of data generated at the physical interface layer between devices or the routing of this data up to integration services at the integration layer. It also includes the flow of data down from the integration layer into the presentation layer. Artifacts at this layer include incoming and outgoing network packets or connection information. Further, all remote crimes that are committed leveraging IoT (botnets, cyberstalking) can be expected to generate artifacts on the network layer.

3.4 Presentation Layer

The presentation layer (Mrabet et al. 2020) is the non-sensing human-interface layer of the device. This layer provides end-user feedback either through its physical interface layer (on a speaker or screen) or through a connected command-and-control application on a companion device. This layer involves displaying data...
either generated by the device, by the integration layer, or jointly from both. Devices will often require interface software to provide this feedback.

3.5 Integration Layer

The integration layer (cloud services layer [Mrabet et al., 2020]) is the upstream non-local component of IoT and is responsible for doing complex computations, long-term storage of data, off-device command and control functions, and enhancing or enabling device functionality. This layer is the layer that enables IoT devices to be decentralized from their primary computing functions. This layer is enabled by cloud, middleware, and database technologies. It will often be located off-site except for in the cases of companion clients, though specific IoT architectures, often found in industrial or secure settings, may support local integration for security, confidentiality, or convenience. This layer will often contain valuable information for forensic investigations, particularly if the device in question has a comparatively small reserve of local memory for its function or exposes an interface to the end-user through a companion client. In the cloud or an integrated database, the primary issue at this layer is accessed since it will require permission, either voluntarily or compulsory through a subpoena, to access the data stores.

4. SPECIFIC DEVICE ARCHITECTURES & IOT PARADIGMS

As previously mentioned, the heterogeneity in the architecture of IoT devices makes any attempt at accessing the device correctly to conduct forensic analysis a challenge. Depending on the device, different forensic artifacts are generated, often in different formats and in different locations. This difference can even be observed in the same device family across generations. However, analyzing some of the architectures of typical IoT devices (like smart home speakers) has merit as a blueprint for future investigations of other devices. In this section, we look at a few of these devices and developing IoT paradigms, such as the Internet of Medical Things and the Internet of Industrial Things, examine potential sources of forensic artifacts, and discuss some of the related challenges.

4.1 Amazon Alexa Ecosystem

The Echo is a brand of home smart speaker systems developed by Amazon and first released in late 2014 [Newman, 2020]. It has various features that include playing music, voice interaction, and making lists. These various features are made possible by integrating Amazon Alexa, a cloud-based virtual assistant AI technology developed by Amazon. These devices require a Wi-Fi connection and a companion device (a mobile phone, laptop, tablet) for initial setup. The companion device is no longer needed after the initial setup is done. As an ever-increasing number of IoT devices are added for home automation purposes, the Echo can be connected with and used to control these smart home devices, functionally extending its physical interaction layer.

Multiple versions and generations of the Echo devices exist, giving rise to multiple hardware architectures. The software architecture for most of these devices remains the same, and a simplified version of the software architecture is given in Figure 1 (from [Pawlaszczyk et al., 2019]). The architecture allows for forensic artifacts to be extracted using the channels at the network/transport layer in between the companion...
ion clients, the echo devices themselves, or the cloud. It should be noted that since, in most cases, the forensic analysis is done after an incident, network analysis may not prove to be helpful to the investigation unless the real-time analysis is pursued. Echo devices themselves can be analyzed using a chip-off method outlined in Al-Sadi et al. to analyze forensic artifacts on device memory (Pawlaszczyk et al., 2019). This analysis requires the hybrid RAM and eMMC to be soldered off the device, the chip identified, and, using an appropriate adapter, a raw image of the memory chip can be acquired.

At the integration layer, analysis on the companion client can reveal forensic artifacts about when the Alexa app was last used, the user account associated with it, cards containing transcripts of what Alexa understood in each voice command, and its response to the voice command in JSON format (Chung, Park, & Lee, 2017). Audio recordings and their transcripts can be accessed on the cloud provided that the user has not deleted them, and each recording is stored with the date and time of the voice command. The location of artifacts on different operating systems of the companion devices has been provided in (Chung et al., 2017). User activity artifacts can be extracted using Amazon’s APIs as well, but these are not released to the public and would require the discovery of the APIs. Potential sources of forensic artifacts in the Amazon echo ecosystem can be summarized in Figure 2 (from Pawlaszczyk et al., 2019).

4.2 Google Nest Ecosystem

Google has a range of intelligent speakers similar to Amazon’s Echo, part of its home automation range. Google Nest (previously known as Google Home, owned by Nest before Google acquired them) was Google’s answer to Amazon’s Echo range and was released in 2011. It is estimated to have sold around 52 million units since its release (D’Onfro, 2018). Other Google Nest products include a smart thermostat, cameras, alarm systems, doorbells, smart locks, and smoke alarms. All the devices require a Wi-Fi connection and can be controlled via a companion device like a laptop, mobile phone, or tablet.

Figure 3: Google Nest Ecosystem (Dorai et al., 2018)

The hardware architecture of the Nest devices varies by device type (smart speaker vs. thermostat) and across different generations of the device. The overall software architecture is given in Fig. 3 (from Dorai et al., 2018) and is similar to the Amazon Echo ecosystem. The IoT devices are connected to the cloud in tandem with a companion device registered to them. Different opportunities to collect forensic artifacts exist in the companion device, the cloud, the IoT device, and the IoT devices’ network.

Unlike Echo devices, Nest devices do not have persistent storage, due to which IoT device analysis may not prove valuable (Dorai et al., 2018). Thus, this illustrates that even among competing manufacturers of the same product type, significant differences exist in architectures. However, forensic investigations may necessitate examining the IoT device before it can be ruled out for not containing any relevant forensic artifacts.
Dorai et al. (Dorai et al., 2018) propose examining the companion devices to the Nest smart speaker system. An unencrypted logical backup of an iOS mobile device is taken and then examined to uncover information about user interactions with the device or the data collected by the IoT devices. This backup is used to analyze SQLite databases on the device relating to the Google Nest application, and an inference engine is built to analyze the data and produce a FEAAS (Forensic Evidence Acquisition and Analysis System) report. This report contains details about the device, the app’s account, geofence events, thermostat events, and camera events. Camera artifacts from the Nest device were recovered by analyzing Google Chrome’s cache on a companion client (laptop). After parsing the cache, links containing text files, image files, event clips, and a profile picture of the user can be extracted (Dorai et al., 2018).

4.3 Windows 10 IoT

Windows 10 IoT Core is the free version of an IoT operating system developed by Microsoft. This OS is optimized to run on ARM and x86/x64 devices such as Raspberry Pi, DragonBoard 410c, AAEON Up Squared, or MinnowBoard Turbot and supports applications developed in the Universal Windows Platform (UWP). Microsoft provides the Windows 10 IoT Dashboard application for Windows 10 computers to allow remote access to the Windows 10 IoT system. This OS features secure boot, BitLocker encryption, device guard, Bluetooth, Windows Update, and hardware connection capabilities for many physical interface layer attachments, including Wi-Fi adapters, Ethernet adapters, cameras, RFID, and other sensors.

Gmez et al. (Gmez et al., 2019) provide a method for conducting non-volatile memory analysis on a Windows 10 IoT system. This forensic analysis was done on a Raspberry Pi board with Windows 10 IoT Core installed. Non-volatile memory analysis was performed on the SD card from the Raspberry Pi, where the Windows IoT OS resides after installation. Conveniently, Windows 10 IoT systems are similar to traditional Windows systems in software architecture and require minimal adaptation of traditional forensic techniques. Examination of the file system can be performed using Autopsy, registry explorers and parsers, and master file table explorers and parsers. Windows system events can also be viewed via their associated log files (Gmez et al., 2019).

Windows 10 IoT has an app-based framework for programs, though it operates much the same as the desktop version of the OS. In the Windows 10 IoT ecosystem, apps are stored in a directory under \Programs\WindowsApp and the packages needed are stored in \ProgramData\Microsoft\Windows\AppRepository\Packages. User information from the app is stored in \Users\DefaultAccount\AppData\Local\Packages\ (Gmez et al., 2019). These apps may provide high-value forensic evidence to an examiner since they ultimately provide meaning and utility to the system for the user.

4.4 Smartwatches: Apple Watch and Fitbit Versa 2

Smartwatches are a class of IoT devices similar to smartphones in that they are often on the user’s person. They include devices like the Apple Watch and Fitbit Versa, though they are produced by many companies, including Samsung, Amazon, Garmin, and Fossil, to name a few. They provide a broad range of features like health and wellness, call and messaging, time-related features (multiple timezone display), alarms, calendars, and notifications.
These features and their implementations vary depending on the device model, device generations, and device maker and can be implemented with apps depending on the sensors that the device provides. Because these devices are typically on a person’s wrist for a significant portion of a day, they can store their location, heart rates, and fitness patterns. The architecture of a smartwatch ecosystem generally consists of the smartwatch device, a companion device that provides enhanced functionality and connectivity (a local server of sorts) for the smartwatch, and the cloud that data is backed up to and processed in.

The Apple Watch, Apple’s flagship smart wear device intended for use with iOS, was examined in Dorai and Houshmand et al. (Dorai et al., 2018). The watch requires a companion iPhone to be connected at all times to sync information, install apps, and change settings. The content from the Apple watch is constantly backed up onto the companion device and is available in the iTunes backup. The cache size for updates is small on the device, and any that would typically be data synced with the companion device may be lost if the companion device is disconnected for some time. The authors developed the Device Data and Forensic Analysis (DEFA) Model to extract artifacts from the watch and other devices, which uses an inference engine to extract relevant data from the activity logs. Given the general architecture of smartwatches, it is evident that the opportunities for collecting forensic artifacts are limited to within the watch itself, the companion device, or the cloud where the backups of the device are stored. However, the DEFA model was used only to extract data from the companion device, which, fortunately, is host to much of the forensically relevant data generated by various apps on the watch (Dorai, Houshmand, & Aggarwal, 2020).

A Fitbit Versa 2 was analyzed using NIST-approved tools Magnet AXIOM and MSAB XRY by Yoon et al. via its android companion device. The main focus of the analysis was to determine the types and sources of forensic artifacts on the rooted companion device. The authors provide relative paths on the smartphone to locate GPS location, heart rate, calories burned, web cookie, credit card information, and credit card image artifacts. No message data was stored on the Fitbit app even though message notifications were sent to the Fitbit device. Further, some user data such as OAuth refresh tokens in plaintext were recovered (Yoon & Karabiyik, 2020).

4.5 Vehicles

Modern vehicles often come equipped with IoT infotainment centers that connect with user’s smartphones via Bluetooth and may even have associated smartphone apps to enable additional features. At a very minimum, these centers often provide an interface for phone-to-car music streaming, GPS navigation, rear-view cameras, and hands-free phone calls, and sometimes control car features, such as seat actuation and climate control. Specific models may even expose cruise control, steering assistance, fuel economy, and other features through this central infotainment center. Occasionally, these IoT devices are connected, via buses, to other computerized units of the car (Lacroix, El-Khatib, & Akalu, 2016). This level of inter-connectedness creates an environment that may expose a wealth of forensic artifacts to an examiner.

The external architectures of these vehicular devices are often simple and typically only involve a singular connection with a user’s phone, though multiple phones may be registered to the vehicle. Figure 5 gives an example of a vehicle’s potential inner architecture, though it should be noted that some vehicles may not have all of these connections. More modern vehicles often include both an Event Data Recorder (EDR) and insurance black box that works with the telematics unit to provide data relevant to crash incidents (this may provide emergency call functionality instead or as well, like OnStar) (Mansor, Markantonakis, Akram, Mayes, & Gurruian, 2017). Both of these units provide several data logging opportunities for forensics investigators. They may contain artifacts such as a driver behavior profile and event information, though both of these units provide information inaccessible to the users and thus cannot be verified by them. Mansor et al. (Mansor et al., 2017) propose a forensics logging mobile application,
DiaLOG, that provides more excellent privacy features, transparency of data being transmitted from the vehicle, and an integrity-protected data logging feature aimed at forensic investigations. This application may also be used to alert owners of malicious intrusion by cyberattackers to cause a denial of service or control operations of the vehicle (Mansor et al., 2017).

For forensic approaches that do not require proaction, both Jacobs et al. (Jacobs, Choo, Kechadi, & Le-Khac, 2017), and Lacroix et al. (Lacroix et al., 2016) provide an in-depth examination of a 2012 Volkswagen Golf and a Ford F-150, respectively. Jacobs et al. (Jacobs et al., 2017) removed modules from the Volkswagen Golf in order to perform a hard-drive examination and chip-off of the multimedia device in the vehicle. From the device’s flash memory, they were able to retrieve the last-known GPS coordinates of the vehicle two partitions on the hard drive, a FAT32 partition, and a WindRiver Systems DosFs 2.0 partition. On the FAT32 partition, a Siemens AG 2.0.0 Europe Map version 5.0.5 was found, leading to the further examination of the disk revealing navigation cart data containing more locations in Europe used by the navigation system. Using Photorec for file carving, 7,431 files were carved from the WindRiver partition, of which 7,414 of them were mp3 files, while some were playlists in text file format (Jacobs et al., 2017).

Lacroix et al. (Lacroix et al., 2016) procured a logical dump, as well as two separate physical dumps, of the Ford-150’s SYNC infotainment file system and its associated content and files. Ford’s SYNC architecture is based upon Microsoft’s Windows CE automotive operating system. It is speculated that this system may interact with telematics and communications modules and an insurance companies’ black box for habit reporting, though this could not be confirmed without the device. Direct access to the data inside the system is complicated without a forensic toolkit or special forensics software since the encryption is employed on the data itself. Using Encase and Autopsy on the logical copy, the authors were able to retrieve a phone book containing device IDs, call names, call types, and
call times; Bluetooth connection attempt logs and potential security PIN artifacts in hexadecimal formats when authenticated; logs of USB device connections and respective file structures; last known AM/FM frequencies and Sirius radio related information (useful for localization); SQLite databases of fuel price listings, Wi-Fi network listings, and movie listings; GPS logs; climate state and configuration data (valid for localization); cryptographic seeds; mobile carrier information; and Internet profile information. It is noted that some of this information went unanalyzed, but it is apparent that the vehicle’s infotainment system stores a multitude of forensically valuable data (Lacroix et al., 2016).

4.6 Internet of Medical Things

Forensic analysis may be performed on medical devices for unique reasons, including autopsy reports, medical malpractice cases, or investigating ransomware attacks. As medical devices and data must comply with comprehensive government regulations, including the Health Insurance Portability and Accountability Act (HIPPA), forensic investigations may be accompanied by legal counsel. Additionally, as some medical devices are implanted into patients, retrieval and forensic analysis may require a patient to opt-in to surgical procedures (Jahankhani & Ibarra, 2019). New standards have been proposed for IoT devices operating within a clinical setting, as there are strict ethical and legal guidelines for medical devices in addition to storage and protection of patient data (Liu, Sasaki, & Uehara, 2020). Liu et al. (Liu et al., 2020) propose a holistic forensic investigation approach to comply with these standards that incorporate a "four-space model" in understanding the integration of medical devices with patients, institutions, and digital infrastructure: cyberspace; a "social space" incorporating legal knowledge and industry standards; a "physical space" where limitations of time, space, and biological conditions are included; and a "psychological" space where a patient’s behaviors are included (Liu et al., 2020).

Jahankhani and Ibarra note that Medical IoT devices are integrated into an information supply chain that links patients, physicians, providers, and Cloud Service Providers (CSPs) together. These can present unique security vulnerabilities as cyberattacks against digital medical infrastructure accounts for more than half of all cyberattacks. As devices, patients, providers, and CSPs have varying security standards; this information supply chain is especially vulnerable to malicious attacks. Additionally, the network is complicated by the need for a personal server (PS) that mediates access to IoT nodes in or on the body (Jahankhani & Ibarra, 2019). Smartphones are a prime candidate for personal servers, though they introduce new vulnerabilities to the network since they are used for applications other than mediating access to IoMT devices, acting as both monitoring devices and delivery devices in Figure 6. However, using a smartphone as a mediator for these devices provides rich opportunities for forensic investigators, as mobile forensics is a well-understood field that is more mature than IoT forensics and may impart less heterogeneity to IoMT forensics. IoMT is still a developing field, and what forensic artifacts may be recovered is not yet well understood, though we would surmise that if the PS acts as a monitoring tool, one could retrieve state information regarding the patient and possibly reports their health.

4.7 Internet of Industrial Things

The industrial Internet of things (IIoT), composed of both industrial control systems (ICS) and supervisory control and data acquisition systems, constitute the enabling technologies of many modern national and industrial infrastruc-
tures. These systems may be found in production facilities, power plants, nuclear facilities, and transportation networks, and their use is ever-increasing. These IIoT networks often expose an interface to critical systems and are thus prime attack vectors for malicious cyber actors, and because of the criticality of IIoT networks, they are the subject of much academic research in cybersecurity (Eden et al., 2017). However, examining their forensic value is an area of research that is still somewhat undeveloped even though they may contain artifacts that would enable forensic investigators to construct timelines of critical events, including cyberattacks, industrial-related fatalities, catastrophic failures, and general foul play. Typically, the architectures of these systems will be implementation-dependent. However, they all follow the 5-layer model outlined earlier in this paper.

Figure 7: IIoT Ecosystem (Awad et al., 2018)

Unfortunately, as is the case in all IoT forensics, the heterogeneity of these systems poses a significant challenge to investigators. However, the components of these systems are even less uniform and often more low-level than other types of devices (Awad et al., 2018). Fortunately, the control center is often composed of traditional computing systems, running a standard OS such as Linux or Windows. Provided that these terminals have serial or wireless interfaces, extracting data from them becomes a digital forensics issue and can be done using any number of standard toolkits, including Autopsy, Volatility, Rekall, and other automated analysis systems (Awad et al., 2018). The communication network also provides an opportunity for forensic artifact collection using familiar tools such as Wireshark. However, forensically interesting network traffic will often include packets whose contents are explicitly formatted for Programmable Logic Controllers (PLCs), Real-Time Automatic Controllers (RTACs), and Remote Terminal Units (RTUs) and will need to be translated depending upon the model and packet specification of the intended destination. Field devices must be assessed independently, but forensic investigations involving them may benefit from volatile memory analysis, chip-off methods, and side-channel methods (Awad et al., 2018).

5. FRAMEWORKS & METHODS

While investigators require tools and toolkits in order to conduct forensics examination of IoT devices, these have already been enumerated and evaluated extensively in Venkauskas et al. (Venkauskas, Toldinas, Grigalinas, Damaeviitis, & Jusas, 2015). However, many of these tools and toolkits are challenging to apply to IoT systems simply due to the heterogeneity of devices. Forensic investigators and researchers often develop frameworks and methodologies for pursuing an IoT-based forensic investigation to address this challenge. Many of these frameworks and methodologies focus on proactive forensics, while some provide responsive models. In general, most of the work centered primarily upon model-building only superficially discusses possible approaches to IoT forensics. Nevertheless, this section briefly surveys state-of-the-art research in IoT forensics, and each may potentially be achieved.

5.1 Frameworks

This subsection presents some generalized frameworks for the identification and analysis of forensic evidence on IoT systems. Currently, there exists no widely accepted framework for conducting IoT investigations. This fact not only complicates the extraction and analysis of evidence that would be admissible in a court of law, but
it also leaves these issues to the investigator to solve. Further, existing digital forensics frameworks may not be applicable, or only partially so, to IoT investigations even in research environments (Kebande & Ray, 2016).

Kebande & Ray et al. (Kebande & Ray, 2016) present an IoT framework based upon the ISO/IEC 27043: 2015, an international standard for security and incident investigation principles. The framework identifies four stages of IoT forensics: the proactive process, IoT forensics, the reactive process, and concurrent processes. The authors also compare their framework with other proposed frameworks and finally offer a critical evaluation of the framework (Kebande & Ray, 2016).

Kebande & Karie (Kebande et al., 2018), extend the research mentioned above into a multidimensional framework. This comprehensive framework specifies nine sub-processes to provide greater fidelity of the framework for forensic application and more closely match the ISO/IEC models. The new model takes into account the IoT network, readiness processes, management, policies, and standards (Kebande et al., 2018).

Kumar et al. (Kumar, Saha, Lal, & Conti, 2021), present an efficient blockchain-based IoT forensics framework while considering consortium blockchain to maintain the chain of custody in cross-border forensic investigations. The use of a Programmable Hash Function (PHF) in their approach allows for better blockchain security with a reasonable level of performance. This framework does not follow any ISO standards compared to (Kebande & Ray, 2016).

Hossain et al. (Hossain, Karim, & Hasan, 2018) propose using a public digital ledger (similar to bitcoin) based IoT forensic framework FIF-IoT. FIF-IOT can provide interfaces for evidence collection and use a tamper-proof scheme to maintain the integrity of evidence during a criminal investigation. Again, unlike (Kebande et al., 2018) does not follow any ISO standard explicitly.

5.2 Forensics Platforms

Some researchers have proposed platform-based IoT forensic models. These platform models provide a database of information for researchers and investigators to access and often involve the real-time acquisition of data over the Internet and the subsequent storage of that information for later examination or preprocessing by an Artificial Intelligence (AI) or Machine Learning (ML) model. These approaches can often essentially be qualified as big-data approaches because they are aggregating large amounts of data to be processed. These platforms may provide valuable information to security professionals and forensic researchers in the pre-investigation stages.

Torabi et al. (Torabi, Bou-Harb, Assi, & Debbabi, 2020) developed a system composed of an IoT data collection module, a darknet data collection module, an IoT threat repository, and an IoT traffic analysis module. The IoT data collection module would capture data "in the wild" to identify exploited devices and analyze data packets identified by the module. The darknet data collection module, which aggregates data from the UCSD real-time network telescope, would correlate dark data with the data collected by the IoT collection module to identify "suspicious IoT-generated activities". The IoT traffic analysis module would then identify compromised or exploited IoT devices and use the IoT threat repository to label malicious and compromised IoT devices. The model was tested on 4TB data set of network information and identified 27,849 compromised IoT devices that generated more than 300 million unsolicited packets (Torabi et al., 2020).

5.3 Real-Time Forensics

Real-Time analysis is a device-level technique that can prove invaluable for forensic investigators who are doing on-site analysis or otherwise have access to a device while it is running or in use. Though this will rarely be applicable in legal cases, the technique is helpful for malware and cyberattack investigation and cases in which access to the device is otherwise restricted. It is also helpful for investigations in which the device has no or little onboard memory to examine. Often, IoT devices operate as a black box from the viewpoint of investigators due to the lack of
standardization and publication of their internal mechanisms (Sayakkara, Le-Khac, & Scanlon, 2019). Real-time analysis is a method that assists in overcoming this challenge.

Zhou et al. (Zhou, Hu, & Makris, 2020) present an architecture-neutral non-intrusive real-time workload analysis framework for process tracing that leverages, in their case, ARM CoreSight. This framework requires the implementation of an on-device hardware tracing module. However, some architectures, such as those from ARM and Intel, already implement this module. The authors evaluate the framework on a Zedboard - a Zynq-7000 FPGA embedding an ARM Cortex-A9 core and evaluate the traces generated by the ARM CoreSight module using several machine-learning models (Zhou et al., 2020).

Sayakkara et al. (Sayakkara et al., 2019) examine the efficacy of electromagnetic side-channel analysis (EM-SCA) on a Raspberry Pi 3 B+ and an Arduino Leonardo. The authors utilize a HackRF software-defined radio (SDR) to acquire electromagnetic emissions from each device, apply a Fourier transform to the traces, average and normalize the results, and utilize a neural network to classify the results in order to detect and classify possible cryptographic algorithms executing on the devices. The authors note that both of these devices run heavyweight operating systems and posit that the method they supply could potentially achieve greater accuracy and recover a greater fidelity of information on simpler devices with fewer cores (Sayakkara et al., 2019).

5.4 Machine Learning and Artificial Intelligence Based Forensics

Artificial intelligence (AI) and machine learning (ML) is a well-developed and ever-maturing field of research with an extensive range of applications. These models in forensics have gained increasing interest in recent years and are already common in cybersecurity applications. The value of Machine learning and Artificial Intelligence-based forensic approaches is two-fold: they help preserve the privacy of the device owner, and they assist the investigator in determining what data is forensically valuable in a relatively short time (Kebande et al., 2020; Koroniotis, Moustafa, & Sitnikova, 2020). An exhaustive and up-to-date (at the time of writing) survey of ML and AI-based forensic models can be found in Kebande & Ikuesan et al. (Kebande et al., 2020).

Koroniotis et al. (Koroniotis et al., 2020) present a deep-learning framework for IoT network forensics called a particle deep framework (PDF), describing investigation phases for identifying and analyzing IoT attack behaviors. The author’s framework outlines the process of extracting data flows, implementing particle swarm optimization (PSO) algorithm to adapt the deep learning parameters, and developing a deep neural network based on the algorithm to uncover and trace abnormal events in the network. This framework is specifically targeted for use in intelligent home networks but is likely adaptable to other types of IoT networks as well (Koroniotis et al., 2020).

5.5 Blockchain-based Forensics

Many authors have proposed a blockchain model for IoT forensics that involves the insertion of a new forensics layer between the network layer and integration layer (Nakamoto, 2008; Le, Meng, Su, Yeo, & Thing, 2018; Yazdinejad, Parizi, Dehghantanha, Zhang, & Choo, 2020; Ryu, Sharma, Jo, & Park, 2019; Hossain et al., 2018). These proposed methods solve the problem of heterogeneity in IoT but require a change in paradigm that may not be widely implementable. The blockchain layer would require integration-provider buy-in and implementation or linking that makes this approach somewhat unpragmatic. However, it may have significant value in systems where command and control are provided and managed by the system’s user, such as in the industrial Internet of things, medical Internet of things, or military Internet of things.

The concept of the blockchain was introduced by Satoshi Nakamoto in 2008, concurrent with the creation of Bitcoin (Nakamoto, 2008). It consists of chains of digital ledgers called blocks that are managed jointly by all hosts on a peer-to-peer network with the intent of transparency and ver-
ification from all participating hosts (Nakamoto, 2008). Since its introduction to the world, blockchain has been applied in almost every sector from finance to agriculture (Yazdinejad et al., 2020).

Yazdinejad et al. (Yazdinejad et al., 2020) propose a blockchain organization in which the network is divided into software-defined network (SDN) clusters, each with an SDN controller that acts as a cluster head. Both public and private blockchain layers are inserted between these clusters to act as verified ledgers of forensic evidence. For SDN to SDN connections, a public blockchain network is maintained that new SDNs are free to join and participate in. Between the SDN head and individual IoT devices, a private blockchain is maintained that requires validation by the network starter or set rules dictated by the network starter. The authors tested this architecture using the Pyethereum test tool from the Ethereum platform (Le et al., 2018).

Le et al. (Yazdinejad et al., 2020) propose a framework that implements a law-enforcement-managed blockchain that defines a device from which evidence is generated as a digital witness (DW). A law enforcement agency (LEA) designated a digital custodian (DC) to examine the evidence. The LEA acts as a provider of the blockchain platform and is the only entity allowed to verify each transaction on the network and write to ledgers, and all other entities can read and write transactions. This network uses a Byzantine Fault Tolerance (BFT) consensus algorithm where unverified transactions are collected by the LEA, formed into a block, and broadcast back to the network for community verification. The author’s framework has not been tested, and an analysis of its effectiveness and soundness on an IoT network would be needed to evaluate its effectiveness. Further, it has not been examined in a legal sense and allowing LEAs sole access to the blockchain platform may likely constitute privacy violations and cause a considerable amount of legal trouble to citeb47.

5.6 Fog-based Forensics

Some authors have proposed a fog solution to IoT forensics that involves migrating computation power and data storage closer to the IoT implementation (Al-Masri, Bai, & Li, 2018; Huang et al., 2017). These solutions involve adding nodes to the network that act as mediators to the upstream integration-layer services and perform some of the integration-layer work on-site instead of in the cloud or on a backend. This solution solves some forensic issues, such as cyberattacks, by aggregating and interpreting forensic data on a single node. However, this potentially exposes forensic evidence to physical tampering and only partially solves heterogeneity since each device family would need a separate fog computer. Additionally, like blockchain-based methods, this requires integration service provider buy-in and restructuring of the larger IoT ecosystem as a whole, making it unrealistic for consumer-level devices but plausible for the industrial Internet of things, Internet of medical things, and Internet of military things.

Al-Masri et al. (Al-Masri et al., 2018) present the concept of a Fog-Based IoT Forensic Framework (FoBI) that is implemented on a generic fog node. FoBI requires the use of in-built ML algorithms to determine suspicious activity on the end devices. This essentially constitutes user entity and behavior analytics (UEBA) to construct profiles of device users and compare device use in real-time with the profile using log files and network packet aggregation. Suppose a mismatch between behavior and the profile occurs past a threshold. In that case, the fog node notifies other IoT devices on the network to stop executing instructions via message queuing telemetry transport (MQTT) until further analysis determines there is no longer a present threat or the system owner is notified. FoBI also includes using an evidence collection module that creates a forensic image of all data residing on the IoT nodes using bit-stream imaging and process examination and generates reports. This model has not been tested (Al-Masri et al., 2018).

Huang et al. (Huang et al., 2017) suggest a fog computing framework for vehicles using road-
side fog nodes for data collection and processing for both forensics and real-time traffic control. These fog nodes connect to a higher-level cloud system responsible for aggregating and storing data and providing larger-scale traffic control. This model, of course, relies upon the buy-in of local government and a yet-to-be-developed data fusion algorithm to account for the different formats of vehicle manufacturers. However, the authors also conduct a theoretical analysis of cyberattacks on these systems and propose an evidence-based system forensics approach that relies upon adjacent nodes and intelligent vehicles to determine a given fog node’s validity (compromised or not-compromised). These countermeasure approaches may have useful applications in other fog-based systems (Huang et al. 2017).

6. SUMMARY & CONCLUSIONS

With the increasing prevalence of IoT systems for every application and the increasing practice of leveraging them for cybercrime, there is a growing need for forensic investigators and security practitioners to utilize digital forensics to investigate these systems. However, because of these systems’ heterogeneity, there are no standardized frameworks, models, or methods for the extraction and handling of forensic evidence from them compared to the realm of digital forensics. This lack of standardization represents a gap between the field of digital forensics and the implementation of IoT. In order to assist investigators and researchers in filling this gap, we have provided a roadmap to some common existing architectures, developing IoT paradigms and both generalized and specific approaches to confronting the challenges of performing IoT forensics. Additionally, we have provided a general model for discussing and classifying different components of an IoT architecture. As the field progresses, we suggest that researchers continue to focus on developing generalized frameworks, methods, and tools for IoT devices while progressing beyond developing high-level models - an endeavor that has already had much effort placed into it. Future work involves further research into general frameworks and tools for IoT forensics while discovering sources for evidence collection in IoT devices.

REFERENCES


Retrieved from https://doi.org/10.1109/mitp.2018.032501747


doi: 10.26438/ijcse/v7i11.2029


doi: 10.18638/rcitd.2015.3.1.67


