Article Title: Industrial IoT Cross-Layer Forensic Investigation

Authors:

First author
Christopher M. Rondeau, 0000-0003-2315-3248, Department of Electrical and Computer Engineering, US Air Force Institute of Technology, Wright-Patterson AFB, Dayton, OH 45433, USA, Christopher.rondeau@afit.edu, No conflicts of interest.

Second author
*Michael A. Temple, 0000-0002-8016-3293, Department of Electrical and Computer Engineering, US Air Force Institute of Technology, Wright-Patterson AFB, Dayton, OH 45433, USA, michael.temple@afit.edu, No conflicts of interest.

Third author
Juan Lopez, Jr., 0000-0001-5083-8627, Global Security Directorate, Oak Ridge National Laboratory, Oak Ridge, TN, 37831, USA, lopezj@ornl.gov, No conflicts of interest.

Abstract

Cross-layer forensic investigation is addressed for Industrial Internet of Things (IIoT) device attacks in Critical Infrastructure (CI) applications. The operational motivation for cross-layer investigation is provided by the desire to directly correlate bit-level network anomaly detection with physical layer (PHY) device connectivity and/or status (normal, defective, attacked, etc.) at the time of attack. The technical motivation for developing cross-layer techniques is motivated by 1) having considerable capability in place for Higher-Layer Digital Forensic Information exploitation—real-time network cyberattack and post-attack analysis, 2) having considerably less capability in place for Lowest-Layer PHY Forensic Information exploitation—the PHY domain remains largely under exploited, and 3) considering cyber-physical integration as a means to jointly exploit higher-layer digital and lowest-layer PHY forensic information to maximize investigative benefit in IIoT cyber forensics. A delineation of higher-layer digital and lowest-layer PHY elements is provided for the standard network Open Systems Interconnection (OSI) model and the specific Perdue Enterprise Reference Architecture (PERA) commonly used in IIoT Industrial Control System/Supervisory Control and Data Acquisition (ICS/SCADA) applications. A forensics work summary is provided for each delineated area based on selected representative publications and provides the basis for presenting the envisioned cross-layer forensic investigation.

Introduction

This paper addresses the forensics of Internet of Things (IoT) devices with specific focus on the unique challenges associated with the Industrial IoT (IIoT) subset. Among the existing difficulties in the criminal justice system with respect to all cyber crime is the collection of evidence [1] and the fact that cyber crime cases are, by nature, some of the most difficult to investigate [2]. Operating under the definition of cyber forensics as "the practice of collecting, analysing, and reporting on digital information in a way that is legally admissible" [3], the difficulty in end-to-end operations is easily demonstrated. Whether or not the collection of forensic data supports criminal investigation or enables protecting exposed vulnerabilities, many cyber incidents and crimes are not reported because, among other reasons beyond the difficulty, there is a belief that there is little chance of successful prosecution and even fear of negative publicity [4]. Despite these factors, most consumer-level computing technology has increasingly asked its users to trust that technology will "function correctly and that information that [consumers] provide, or that is collected about them, will be adequately protected" [5] for a variety of daily uses (e.g., IoT devices, Google, Facebook, Instagram). Unfortunately, this trust benefits the criminals given that a "majority of the people in modern societies are not particularly concerned about their data as long as they can happily and securely use technologies and commercial products to [enhance] their social lives," [5] and this mindset shows no sign of changing any time soon.

There is clear concern that IoT and "smart devices can also be used to launch Distributed Denial of Service (DDoS) attacks against governments and corporate networks" [6]. However, the mindset that has emphasized ease of use has contributed to the explosion of IoT devices for consumers and industry alike, with some forecasts indicating that the number of internet-connected devices will reach 50 billion by 2020 [7] and pass $1.1 trillion ($USD) by 2021 [8]. This includes proliferation of the IIoT computing technology subset which adds to, and exacerbates, the existing challenges of cyber forensics. Many IoT devices are components of systems defined by the United States Department of Homeland Security [9] and the European Commission of Migration and Home Affairs [10] as part of Critical Infrastructure (CI) operations. The subordinate subset of IIoT devices provides complementary functionality in Industrial Control System/Supervisory Control and Data Acquisition (ICS/SCADA) applications.

The overlapping functionality and interdependencies of IoT and IIoT devices are illustrated Figure 1 along with the functional location of ICS/SCADA critical infrastructure elements. In preparing this review it became evident that the lack of cross-layer exploitation, and specifically the joint consideration of higher-layer digital and lowest-layer physical information, is prevalent across the broader IIoT domain and a decision was made to consider ICS/SCADA elements for illustration such as included in Figure 1. This decision was made given that the greatest sustained (documented, vocalized, etc.) concern over the lack of cross-layer exploitation has occurred within the ICS/SCADA community where researchers and practitioners alike have been “championing” the cause for cross-layer security and safety enhancements (Weiss, 2017; Weiss, 2018).

The interdependencies highlighted in Figure 1 expose the unique challenges associated with IoT and IIoT systems when compared to traditional Information Technology (IT) systems in terms of cyber forensics. These challenges can be generalized to six main points: 1) the unique quality of where forensic artefacts may exist is more holistic (i.e., unique evidence may exist in either/both the physical layer and the higher layers); 2) Commercial Off-The-Shelf (COTS) products have inherent vulnerabilities that have often been ignored and cybersecurity is often a retroactive measure placed on a system [11]; 3) many IIoT devices, and especially those used in SCADA/ICS applications, cannot be powered off to conduct forensic investigations; 4) forensic evidence is generally more volatile [3] in industrial applications; 5) forensic data for IIoT devices is generally an afterthought until such data are needed to investigate an attack and subsequently prevent a future attack—the “time to discover and unwind potential incidents can take weeks, if not months, of deep inspection by threat hunting experts” [12]; and 6) the specialization of specific systems and/or networks (e.g., SCADA) often requires a forensic specialist who “has to be an expert in such systems/networks … in order to identify where potential forensic evidence could be located” (Casey, 2011).

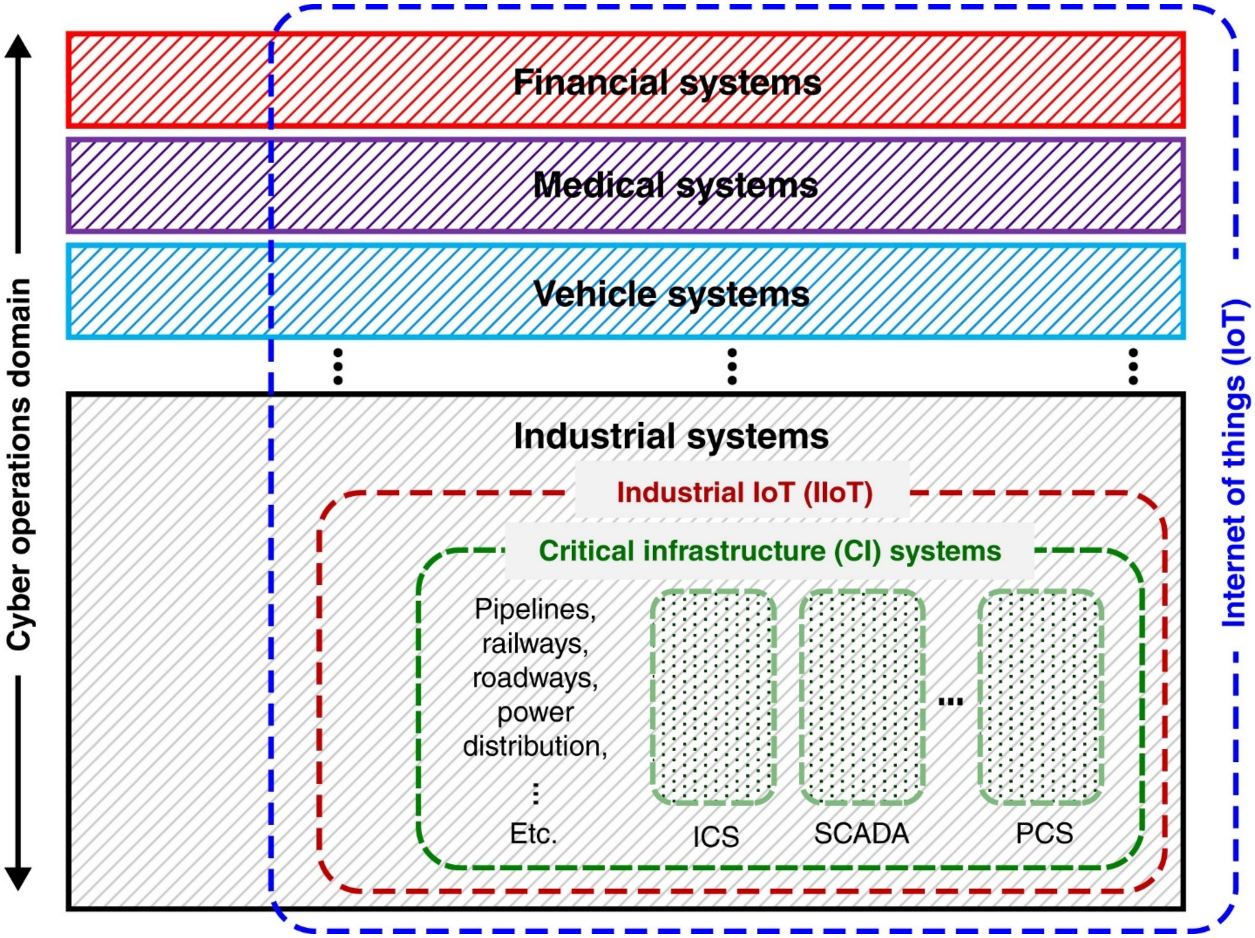

Figure 1: Functional overlap and interdependencies between broader IoT and the IIoT subset supporting cyber forensic physical systems and infrastructure exploitation using financial, medical and vehicular systems as examples (physical expanse of graphic elements not to scale).

This advanced review paper is not intended to represent an all-encompassing literature review in a subject area that is of great interest and which has yielded hundreds of related publications in just the past decade. Rather, a sufficient amount information is pulled from a selected number of cited references to 1) help the reader gain an appreciation for the breadth of global Research, Development, and Demonstration (RDD) activity that has occurred in IIoT forensics, and 2) provide motivation for researchers to stay the course and build upon a solid foundation while enhancing the effectiveness of forensic investigative tools and methods. The selected information is incorporated into presentation of an IIoT Forensic Investigative Framework (Section 1) that can be used for IIoT applications. The framework is currently supported by Higher-Layer Digital Forensic Information (Section 2) and Lower-Layer Physical Forensic Information (Section 3) , the integration of which is proposed for Cross-Layer Forensic Information (Section 4) exploitation—a largely under reported, under exploited mechanism that is envisioned as providing considerable benefit.

IIOT INVESTIGATIVE FRAMEWORK

Digital Forensic Frameworks

Any forensic technique must necessarily be part of an investigation framework to ensure that the evidence collected is useful for criminal prosecution and/or future vulnerability mitigation. The field and study of digital forensic science (addressed here in the High-Layer Digital Forensic Information section as a subordinate term to cyber forensics) dates back to approximately the late 1970's. Over the last 40 years, many forensic frameworks have been proposed and evolved based on changing technology, experience of evidence admissibility, ever-changing cybercrime attack methods, and government/public interest. It was during this time that the US Federal Bureau of Investigation (FBI) stood up their Computer Analysis and Response Team in 1984 [14] and the United Kingdom stood up their Computer Crime Unit in 1985 [15]. The forensic framework evolution has motivated governments around the world to evolve their criminal laws and law enforcement practices to accommodate the efficient collection of digital evidence. However, up to 2006 there were no US laws governing the collection of digital evidence [16] until an amendment was passed allowing for the collection of "electronically stored information" and its use in civil litigation cases [17].

The definition of cyber forensics adopted here for presentation follows the definition published by the US-based National Institute of Standards and Technology (NIST) and includes "the application of science to the identification, collection, examination, and analysis [four phases] of data while preserving the integrity of the information and maintaining a strict chain of custody for the data" [18] . The various definition elements are embodied in the digital forensic processes flow diagram in Figure 2 which includes selected elements and process flow attributed to (a) [19], (b) [20], (c) [21], and (d) [22] as indicated. The shaded box elements in Figure 2 include Proactive phase elements from [23] and are the basis for subsequent discussion in the paper—the unshaded box elements are among the Reactive phase elements noted in [23]. The US Department of Justice (DOJ) in [19] considers a more simplified framework that includes only four phases: assessment, acquisition, examination, documentation/reporting. More specific details of framework evolution during the 2005-2015 timeframe can be found in [20].

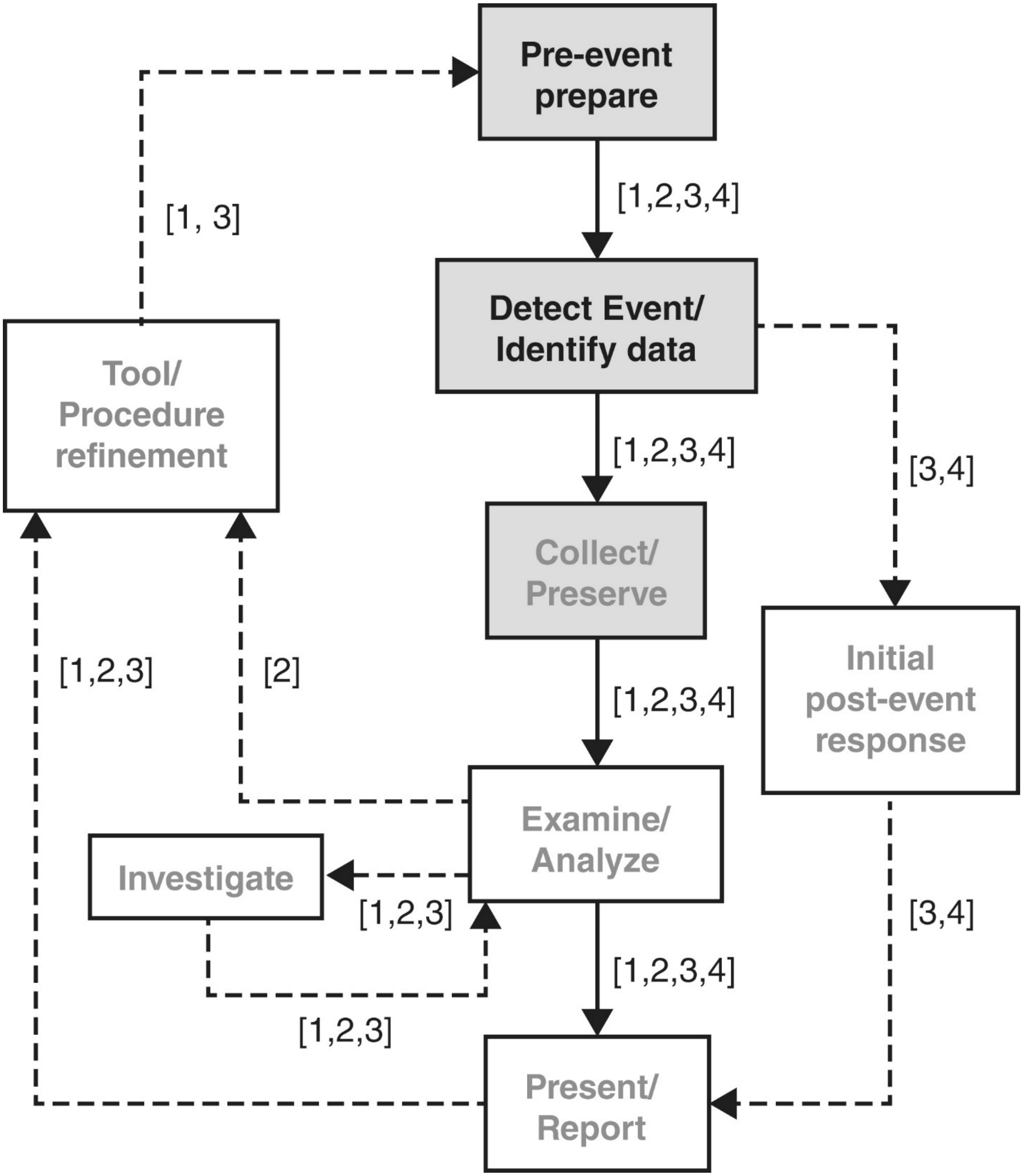

Figure 2: Digital forensic process flow diagram with flow attributed to the [19, 20, 21, 22] references as indicated. Shaded boxes denote Proactive Phase elements and non-shaded boxes denote Reactive Phase elements according to [23].

IIOT Forensic Information Domains

Contextualization of IIoT forensic information domains (or any other domain for that matter) requires consideration of specific implementation and operational details within the targeted application space. This includes consideration of the network protocol (communication rules, structure, etc.) being considered. There are two existing models shown in Figure 3 that collectively represent the major IIoT applications of interest here, including 1) the 7-layer Open Systems Interconnection (OSI) model [22] which represents a broad range of network applications, and 2) the 6-layer Perdue Enterprise Reference Architecture (PERA) [21] which is most common for ICS/SCADA applications. Figure 3 also shows the separation of cyber forensic information based on 1) evidence collected and analysed within higher layers (Digital Forensic Information), and 2) evidence collected and analysed at the lowest physical (PHY) layer (Physical Forensic Information). As detailed in subsequent sections of the paper, this functional separation is introduced to differentiate between higher-layer digital (bit-level) and lowest-layer PHY (waveform-level) forensic exploitation.

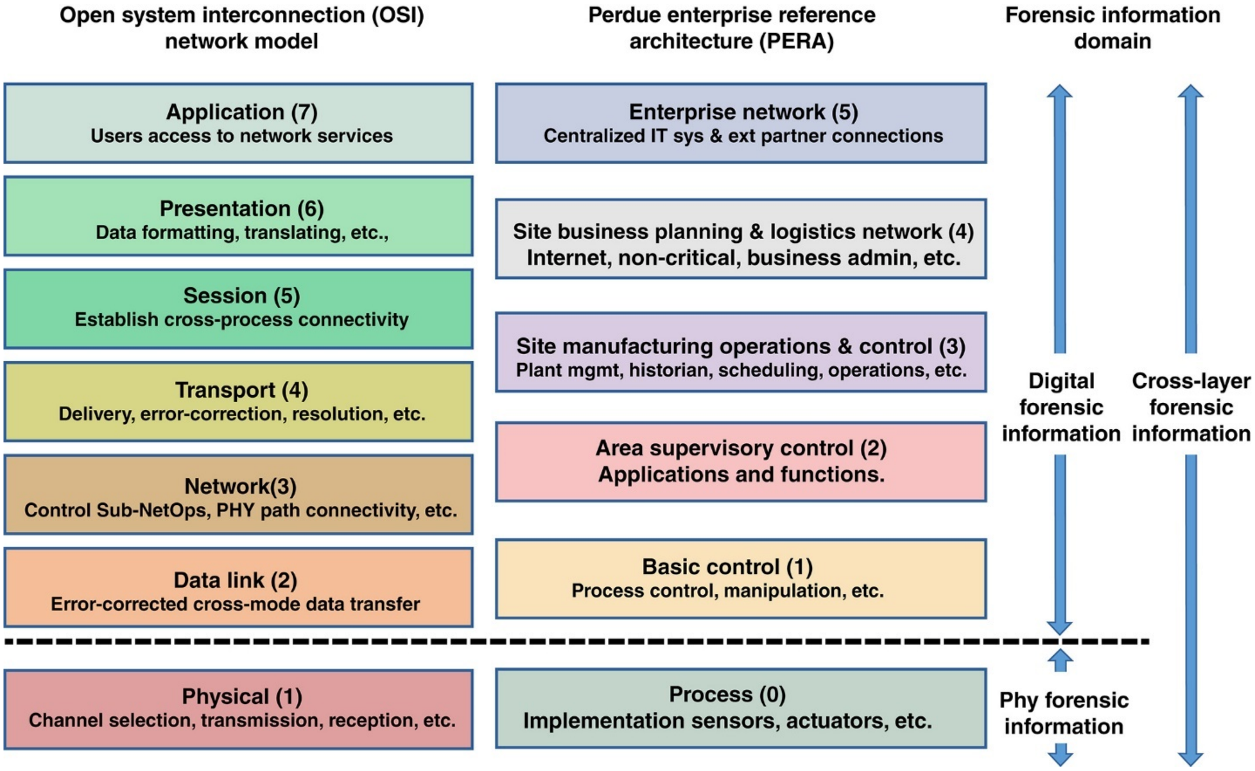
Open system interconnection (OSI) network model
Perdue enterprise reference architecture (PERA)
Forensic information domain
Application (7)
Users access to network services
Presentation (6)
Data formatting, translating, etc.,
Session (5)
Establish cross-process connectivity
Transport (4)
Delivery, error-correction, resolution, etc.
Network(3)
Control Sub-NetOps, PHY path connectivity, etc.
Data link (2)
Error-corrected cross-mode data transfer
Physical (1)
Channel selection, transmission, reception, etc.
Enterprise network (5)
Centralized IT sys & ext partner connections
Site business planning & logistics network (4)
Internet, non-critical, business admin, etc.
Site manufacturing operations & control (3)
Plant mgmt, historian, scheduling, operations, etc.
Area supervisory control (2)
Applications and functions.
Basic control (1)
Process control, manipulation, etc.
Process (0)
Implementation sensors, actuators, etc.
Digital forensic information
Cross-layer forensic information
Phy forensic information

Figure 3: Relationship between the Open Systems Interconnection (OSI) Seven-Layer Model [22] and six-layer Perdue Enterprise Reference Architecture (PERA) [21] showing layer association to Digital, Physical, and Cross-Layer Forensic Information domains.

Figure 4 shows a general IIoT network architecture consisting of different types of wired and wireless communication connections that may be employed—not all IIoT applications employ all of the indicated interconnectivity. Collectively, locations within the enterprise, plant, field site, vehicle, etc., where digital data may be stored and the required communication interconnectivity provide the opportunity for exploitation of higher-layer digital forensic information and lowest-layer physical forensic information. As addressed in the remaining major paper sections, there is 1) an overabundance of directly related work on exploiting Higher-Layer Digital Forensic Information for real-time network level protection, cyberattack detection/defence, and post-attack forensics analysis—this wealth of information is a result of the historical RDD emphasis in the IT domain, 2) a considerable amount of information related to Lowest-Layer Physical Forensic Information, especially in the area of Radio Frequency Fingerprinting (RFF)—a vast majority of this work provides little to no details on forensic exploitation, and 3) sufficient information to motivate consideration for Cross-Layer Forensic Information whereby digital and physical forensic information is jointly considered to support development of a corroborated forensic investigation.

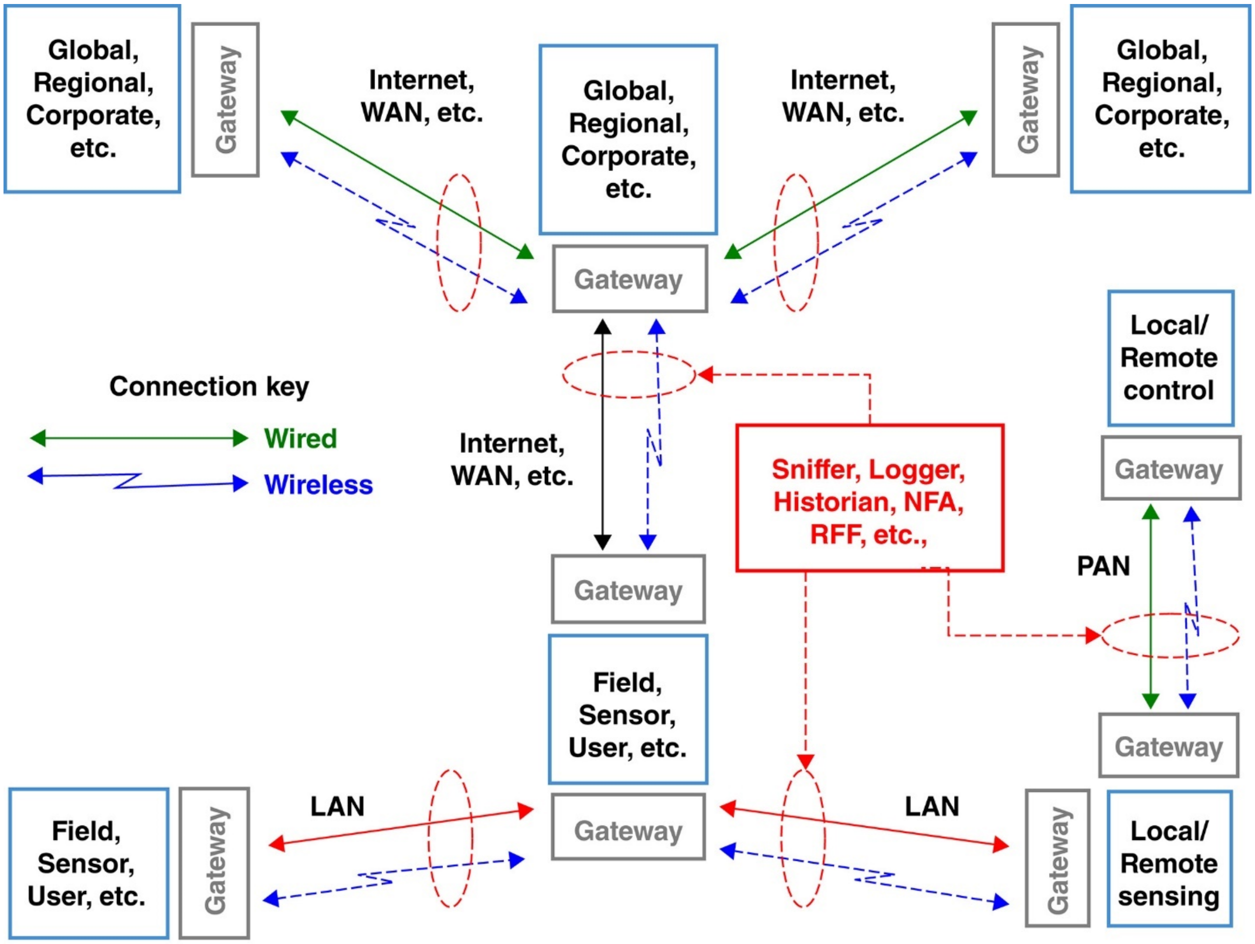

Figure 4. General IIoT network architecture consisting of Internet, Wide Area Network (WAN), Local Area Network (LAN), and Personal Area Network (PAN) connections. Wired and/or RF wireless interconnectivity provides access for bit-level data/packet sniffing, logging, historian, Network Forensic Analysis (NFA) and PHY-based Radio Frequency Fingerprinting (RFF) processes.

HIGHER-LAYER DIGITAL FORENSIC INFORMATION

Higher-layer digital forensic information may be viewed in terms of a hierarchy involving three general levels, including bit-level, network-level, and user-level forensic techniques—the distinguishing boundaries between these levels is somewhat arbitrary and may be blurred in some applications. The following review of higher-layer digital forensic techniques considers each of these levels independently. However, it should be noted that techniques applied at one level often, if not always, have application at other levels or can be used to corroborate findings at lower levels. IIoT forensic analysis utilizes similar forensic techniques as other applications, thus many surveys of applications that are not specific to IIoT are included here. In general, there is a lack of "tools and methodologies designed specifically to incorporate SCADA system[s], including their protocols and proprietary log formats" [11]. Therefore, the major difference between higher-layer digital forensics and those specifically applicable to IIoT is not necessarily in technique implementation itself but rather that specific system-level expertise is required to identify and collect useful data (Box 1).

Box 1. Insider Threats Present Unique IIoT Cyber Forensic Challenges: Forensic data collection for IIoT applications can be challenging in that the threat itself may very well emerge from among the very system experts charged with conducting forensic data collection. A 2017 report cites that 60% of industrial cyberattacks are from insider threats [23], with 44.5% of these having malicious intent [24]. When a forensic investigation requires system expertise for proprietary IIoT, ICS, and SCADA architectures that are decades old, there may be only one or a few qualified experts who can support the investigation. If these individuals are themselves complicit with the threat, soliciting there support will most likely be detrimental to the investigation. The risk of this occurring can be mitigated by 1) training additional personnel in the older IIoT, ICS, and/or SCADA system architectures, 2) updating plant/enterprise access and maintenance policies (e.g., two-person integrity access and control rule), and/or 3) modernizing selected system components without necessarily requiring an extensive and expensive plant-wide overhaul.

Bit-Level Forensic Analysis

The collection and analysis of bit-level information are reviewed here as being distinct from the Network Intrusion Detection (NID) and Network Forensic Analysis (NFA) discussed in the following section because many techniques developed for this level can be used or extrapolated to the network-level. Furthermore, bit-level information that is collected from

ICS/SCADA Programmable Logic Controller (PLC), Remote Terminal Unit (RTU), field sensor, communication gateway, etc., devices can be used to assist in post-attack analysis [25] while providing the most accurate view on how information is actually stored [26] and assisting in determining device authenticity.

Articles that specifically address bit-level forensics in the context discussed here are sparse, as are surveys and related reviews. However, as noted in [26] bit-level techniques are not only useful for copying or converting collected evidence but are also useful for detecting anomalies to allow investigators to reliably reconstruct data fragments or determine where bit insertion or deletion is of particular interest. In light of this concept, [27] provides a thorough review of statistical anomaly detection methods that are directly applicable to bit-level forensics, albeit their focus is on NID application. The authors in [28] illustrate and review works utilizing bit-level n-gram based forensics regarding authorship. This is followed by work in [29] which further demonstrates that bit-level classification techniques can help overcome potential deficiencies of stand-alone network-level forensic techniques, such as using such techniques in the absence of comments (very common in malicious code). Other types of collected evidence may be substantiated through bit-level techniques, e.g., passive digital video forensics using pixel-based feature extraction to detect spatio-temporal forgeries or copy-move attacks [30]. Combining the results of individual techniques from various plant-wide systems and connecting them to bit-level forensic information collected from IIoT devices builds a substantive post-attack narrative. However, the collection of such data is dependent on the data being stored long enough after an attack for investigators to have access when needed.

The amount of data collected at the bit-level and higher is continuing to increase to a point where previously utilized techniques need augmentation [31] in order to produce a timely contribution to a forensic investigation [32]. To this end, machine learning, data mining, and deep learning all hold promise for improving the efficiency of handling large amounts of data. However, as additional techniques such as these are developed, their use must be considered in relation to the admissibility and repeatability of forensic evidence provided by these methods [33].

Another consideration for bit-level forensic analysis is determining whether or not investigators have direct access to hardware components. For example, there has been a considerable amount of work addressing mobile device forensics where investigators have hardware-in-hand and can access the contents of 1) internal memory storage devices, and/or 2) removable storage devices containing user and service provider details. One hardware-in-hand example is SIM card evidence that is “forensically-retrieved from a mobile device in the form of call logs, contacts, and SMSs; a mobile forensic investigator should also be aware of the vast amount of user data and network information that are stored in the mobile SIM card such as Integrated Circuit Card Identifier (ICCID), International Mobile Subscriber Identity (IMSI), and Abbreviated Dialing Numbers (ADN)” [34]. This is more generally considered “mobile device forensics” [35] (which may be extended without loss of generality to many IIoT devices) whereby investigators acquire data from internal memory of handsets and Universal Integrated Circuit Cards (UICCs) or similar to extract features of cellular networks like logs of usage, geographic location, and other data.

As previously noted, bit-level techniques are not exclusive to hardware-in-hand or device level forensic analysis. The placement of a bit-level traffic historian somewhere within a network architecture to collect data for post-attack forensic analysis [36] is recommended, along with specific placement in the network so as to mitigate the chances of the historian itself becoming compromised. Similarly, packet sniffers, protocol analyzers, and similar devices enable the observation and/or collection of communications (packets, bits, etc.) flowing between computers—a partial survey and review of related network techniques is presented in [37]. In light of the bit-level techniques that bridge to the network-level, the authors in [38] contextualize the concept and also discuss specific challenges of data acquisition in SCADA systems. The bit-level to network-level forensic analysis connection becomes clear when successful bit-level forensic analysis properly attributes network-level activity to a given source.

Network-Level Forensic Analysis

The bit-level techniques in the previous section may be extended to apply to network-level forensic analysis whereby the nature of the attack and attack attribution begin to take place. As illustrated in Figure 3 and Figure 4, IIoT network-level forensics applications (PERA Layer 1 and higher) can utilize some network-level techniques of other communication applications (OSI Layer 2 and higher). The type of techniques that may be considered for transition between bit-level and network-level forensics generally include analysis of data files [18], disk image copy, log system security, transportation security, data security, data recovery, file system analysis, network traffic capture and analysis, network survey, and IP traceback [39].

Network-level forensics primarily focuses on network traffic analysis. Of note, the term network forensics is often used in the context of both post-attack forensic science as well as real-time monitoring and network security—the former has been defined as “strict network forensics” as opposed to “general network forensics” [39]. This paper focuses on techniques associated with strict network forensics (herein referred to simply as network forensics). Network forensics is defined as the use of scientifically proven techniques [40] to capture and analyse network traffic in order to discover the source of attacks, and also includes tasks such as reassembling transferred files, keyword searches, and parsing human communication-like emails or chat sessions [41].

Given the nature of network layer data, and especially when considering IIoT devices, there is an assumed volatility and dynamic nature to the data and the storage of network traffic that may not exist at other layers. The works in [42], [43], and [20] collectively describe two network traffic classifications, including 1) "catch-it-as-you-can," and 2) "stop-look-and-listen." In "catch-it-as-you-can" systems all packets pass through a traffic point and are captured. This requires a large amount of data storage and analysis is performed after data capture, with storage capacity requirements increasing for Cloud computing [32] as part of the enterprise system. In "stop-look-and-listen" systems selected information is saved for future analysis following a trigger event.

There are Network Forensic Analysis Tools (NFATs) that capture network traffic, analyse the traffic, and allows administrators to identify, replay, and isolate anomalous network activity [44]. NFATs facilitate the organization of captured network traffic packets to be viewed as transport layer connections between machines. All NFATs exhibit three basic properties [45]: 1) data collection, 2) data preservation without alteration, and 3) data replay. The specific techniques associated with network forensics are tightly connected with the tools used in the process. This paper does not seek to catalogue or survey all available tools given the list is changing frequently. However, summary tables of selected NFATs are provided in [42] and [41], with [27] providing related survey and review information on NID forensics.

The bit-level and network-level forensic techniques discussed thus far and referenced herein are part of an overall investigation strategy that seeks to empirically describe the cyberattack and properly attribute it to the perpetrator(s). However, the evidence collected from the bit-level and network-level techniques may not be sufficient in and of itself when considered in isolation. Considering benefits of corroboration with other data sources, one final forensic layer within higher-level digital forensics is reviewed–the top-layers of the OSI and PERA models accounting for online data discoverable by forensic investigators (Box 2).

Box 2. Emergence of Operational Technology (OT) and IIoT Implications: The OT framework has only recently emerged as a means to discriminate between systems that use computing resources to support administrative operations (mainstream IT applications) versus industrial operations (mainstream ICS/SCADA applications) [46] - the Venn diagram in this reference shows ICS and SCADA functionality completely embodied within OT. Thus, the IT and OT distinction centres around different networking technologies (hardware and/or software) and different skill sets are required to properly maintain system operation. Whereas most businesses are familiar with IT career field specialization that includes a myriad of specialized certifications, the OT profession is currently much more limited and does not have widely accepted standards like the IT Security+ standard. Relative to this gap, "IT and OT have traditionally been developed and managed as two separate domains, the IIoT and Smart Manufacturing technologies have contributed to the blurring of lines between the two, and there are significant opportunities to be derived from aligning IT and OT" [47]. The need for doing so is evident when considering 1) a 2018 IIoT security survey that indicates most organizations globally are forecasting a 10% to 25% growth in connected IoT devices, with the potential for this rate to double every three to seven years [48], and 2) the fact that 32% of IIoT devices are connected directly to the internet and bypass traditional IT security layers, with only 40% of users reporting that they are applying routine patches and updates [48]. This increases the challenge for real-time IIoT device protection and cyber forensics as information from supposedly identical "witness" systems is collected, processed, etc.

User-Level Forensic Analysis

The user-level is defined here as data collected from the application-layer, data collected from Cloud storage, or any open-source intelligence (OSINT) application. Traditional forensic methods must be modified, or new methods created, in order to accommodate the large amounts of data required for IIoT forensic investigations [31]. Even when the amount of data collected is reduced, the amount of digital forensic information could still overwhelm an investigation. OSINT contributes a semi-automated process that extracts entity information and expands cross-device and cross-case analysis [49]. When combined with other higher-layer digital forensic information (previously discussed), or Physical Forensic Information (discussed in the next section), these data can synergistically produce the desired corroboration for forming a clear forensic picture. The benefits obtained from cross-device and cross-case analysis provide motivation for development of the concept of Cross-Layer Forensic Information presented in the final main section of the paper.

LOWEST-LAYER PHY

Many of the higher layer digital techniques addressed in the previous sections are generally not IIoT application specific. This section addresses lowest-layer physical information available for forensic analysis and having direct applicability to IIoT forensics. Many network and bit-level IIoT processes (Figure 3 PERA model Layer 1 and higher) rely on sensors for monitoring and controlling physical processes (pressure, temperature, speed, etc.). Thus, the IIoT PHY domain is rich in forensic information that may be exploited for a post-attack investigation. Outside of forensic applications there has been considerable RDD activity in single layer PHY-based Radio Frequency Fingerprinting (RFF) methods developed to support network defence of both wired and wireless communications [50], [51], [60], [52]–[59]. These works have predominantly addressed the discrimination of various hardware components (e.g., the gateway devices shown in Figure 4) used to establish communications. While some of these works suggest that the demonstrated methods support forensic analysis, the concept of forensic support is more of a theme used to motivate the readers and there are minimal details provided on actual forensic implementation. The nearly non-existent usage of RFF in PHY-based forensic applications is supported by an internet search using "radio frequency forensics" which yields an abundance of references addressing Radio Frequency Identification (RFID). The number of hits returned for this search (100,000s) is potentially misleading given that forensics in the RFID application space addresses the ability to "track and manage forensic evidence … streamline the capture, collection, and transfer of [digital] data to track assets and people" [61]. Thus, the use of a digital ID in RFID forensic applications is not at a true PHY-based forensic exploitation but rather is more consistent with the higher-level forensic exploitation methods addressed in the previous section.

While the cross-device uniqueness of electronic device fingerprints may not be totally on par with cross-human fingerprint uniqueness, results such as provided in [50], [51], [60], [52]–[59] routinely demonstrate near 100% discrimination for selected scenarios and have been sufficiently promising to sustain progressive RDD over the past 10 years. Collectively, these and other related RFF works have addressed nearly all common communication signalling schemes, including Bluetooth [53], Automation [50], [51], and ZigBee [60] Personal Area Networks (PANs); WiFi [52], [54], [55], [59] Wireless Local Area Network (WLANs), and WiMAX [57], [58] Wide Area Networks (WANs), to name a few. For the references provided, the unique RFF features have been reliably extracted from various signal domains, including 1) time [51]–[53], [55], [58], 2) frequency [50]–[52], 3) joint time-frequency [57], [59], and 4) constellation [54], [60]. A majority of RFF works available for forensic consideration are based on burst type communications, which when used for committing a cyberattack or electronic crime, may leave behind 1) only a single fingerprint—this may occur for a simple attack against a ZigBee control element that is designed to respond to a single command burst, or 2) 10s to 1000s of fingerprints—this may occur for a progressive multi-node WiFi network attack with the actual number of fingerprints “left behind” by the perpetrator(s) depending on the extent and duration of the attack. In the case of RFF-based forensics, anything “left behind” for analysis 1) must have been sensed, collected, and stored by a historian or similar device, and 2) be in the proper format to enable reliable extraction of features for the selected RFF method.

CROSS-LAYER FORENSIC INFORMATION

Inherent to any investigation that involves the use of cyber forensic artefacts is the integrated analysis of all artefacts in light of real-world conditions at the time of the attack. The biometric-based work in [62] describes this concept as a “cyber-physical integration” as it relates to supporting criminal investigations. There is a wealth of information on cross-layer attack and defence work in Cognitive Radio (CR) [63], [64], Metropolitan Area Network (MANET) [65], [66] and Wireless Sensor Network (WSN) [66], [67] applications, with a collective overview of these and other applications provided in [68]. Given the benefits realized in these application areas it is reasonable to infer that some of the defence mechanisms in these works may be adopted and exploited for IIoT forensic analysis and investigation benefit.

The cross-layer forensic investigative benefit becomes more evident by considering the related PHY-based ICS/SCADA example in (Lopez Jr. et al., 2018) which provides a generalization of IIoT attack assessments based on what are categorized as (a) Remote access attack (RAA) strategies that aim to alter the sensed/reported field device state being received and acted upon by the PLC and (b) physical access attack (PAA) strategies that aim to alter the physical device hardware, firmware, etc. The RAA and PAA assessments in (Lopez Jr. et al., 2018) were made using discriminating features extracted from communication links being used for exchanging control and sensing information between the PLC and field devices. Of note in these RAA and PAA assessments is that integrated cross-PERA layer and cross-SCADA zone processing was not implemented. Rather, the assessments are based on an inherent assumption that cross-layer only and within-zone only bit-level protection methods have been “fooled” and that 100% PHY-based device identity and state determination is desired. Using a normal-versus-anomalous (attacked) detection strategy and 12 assessment scenarios (six RAA and six PAA assessments using two devices from three manufacturers with each device operating at two distinct set points), the process in (Lopez Jr. et al., 2018) successfully yielded an anomaly detection rate (ADR) of ADR ≈ 91.3% and ADR ≈ 95.5% for RAA and PAA assessments, respectively. This PHY-based only performance suggests great promise for achieving ADR ≈ 100%, that is, 100% attack detection and/or post-attack attribution, in IIoT applications by including cross-layer methods similar to those noted in the previous paragraph.

As confirmed by the literature review conducted to prepare this advanced review, the overall conclusion in the earliest cited work remains valid today, i.e., “a majority of existing research on security issues in wireless networks mainly focuses on attack and defence in individual network layers” [63]. This includes a majority of single higher-layer approaches that tend to underutilize or entirely disregard PHY layer information. As a result, these approaches are not enable direct correlation of bit-level network anomaly detection (higher-layer digital forensic evidence) with PHY connectivity and/or hardware device status (attacked, defective, etc.). While likely a limitation across the broader IIoT domain, the greatest concern over this has been vocalized within the ICS/SCADA community where researchers and practitioners are “championing” the cause for security and safety improvement using cross-layer information [69], [70]. The envisioned IIoT cross-layer benefit could be demonstrated using a comparative performance assessment based on analysis using 1) only Digital Forensic Information per the Higher-Layer Digital Forensic Information section—comparable to conducting a criminal investigation using only witness information, 2) only PHY information per the Lowest-Layer Physical Forensic Information—comparable to conducting a criminal investigation using only physical fingerprint information, and 3) cyber-physical integration of both digital and physical forensic information—enabling generation of a more complete picture of events occurring during the crime.

The integration and exploitation of digital and PHY is conceptually illustrated using the SCADA architecture in Figure 5. This figure shows that the forensic investigative surface is comprised of multiple SCADA zones that include exploitable digital and physical information. The architecture in Figure 5 is based on the framework presented in (Eden et al., 2016), with the control zone intentionally expanded here to distinguish the PERA model Basic Control (a) and Physical (b) layer elements in Figure 3. As shown, there are interzone firewalls that enable forensic analysis using the typical firewall log data, event data, Internet Protocol (IP) addresses, and port log files (Eden et al., 2016). Additionally, there is a wealth of forensic information available throughout all SCADA zones that can be used to support analysis. Some details of this information and its location within each zone are provided in Table 1 which includes a representative sampling and summary of table entries provided in Eden et al. (2016).

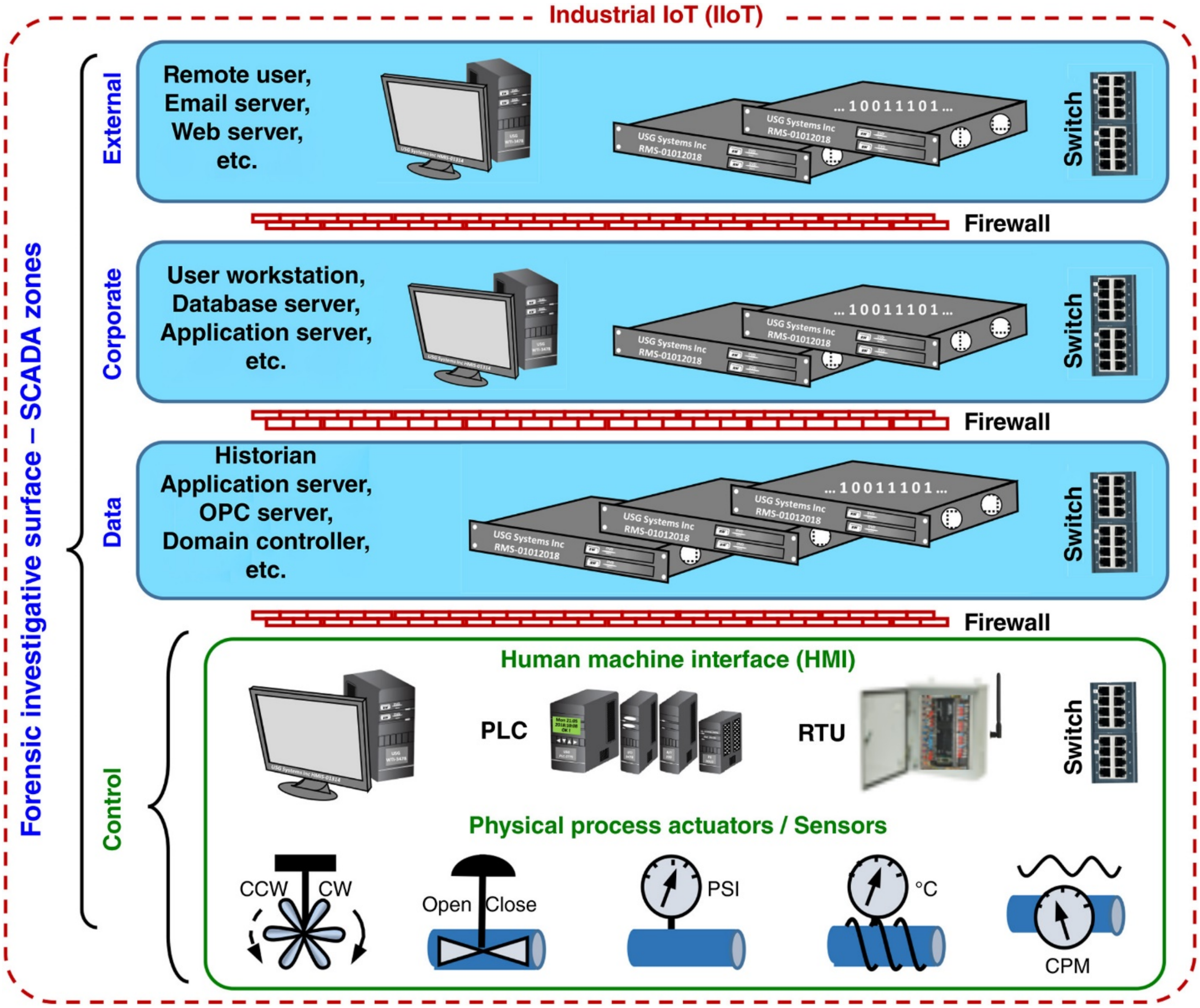

Figure 5: Forensic investigative surface comprised of SCADA zones with representative elements available for forensic analysis. The architecture is based on the framework presented in Eden et al. (2016) with the control zone expanded to highlight elements of higher-digital and lowest-physical elements.

Conclusion

This advanced review addresses cyber forensics of IoT devices with specific focus on the unique challenges associated with the IIoT subset. Given that the majority of IIoT devices are part of Critical Infrastructure (CI) systems, the necessity to have a robust forensic framework and well-defined, repeatable, and admissible techniques is paramount. This paper provides a representative sampling of information pulled from a selected number of cited references to 1) help the reader gain an appreciation for the breadth of global research, development, and demonstration (RDD) activity that has occurred in IIoT forensics, and 2) provide motivation for researchers to stay the course, build upon a solid foundation, and enhance the effectiveness of forensic investigative tools and methods.

A Forensic Investigative Framework is presented for use in IIoT applications, with primary support pillars including existing higher-layer digital and lowest-layer physical (PHY) forensic information methods. As supported by the summary information provided, the greatest degree of forensic exploitation occurs within the higher-layer digital domain(s), with the lowest-layer PHY domain remaining largely under exploited. This motivated the introduction of cyber-physical integration as a means for exploiting cross-layer forensic information to maximize investigative benefit in IIoT cyber forensics, as well as cyber forensics as a whole.

Acknowledgments

The views expressed in this paper are those of the authors and do not reflect the official policy or position of the Air Force Institute of Technology, the Department of the Air Force, the Department of Defense, or the US Government. This paper is approved for public release, Case No.: 88ABW-2018-5,312.

References

[1] R. Grimes, “Why it’s so hard to prosecute cyber criminals | CSO Online,” Chief Security Officer Online, 2018. [Online]. Available: https://www.csoonline.com/article/3147398/data-protection/why-its-so-hard-to-prosecute-cyber-criminals.html. [Accessed: 02-Jul-2018].

[2] S. S. Smith, “Roles and Responsibilities for Defending the Nation from Cyber Attack — FBI,” 2017. [Online]. Available: https://www.fbi.gov/news/testimony/cyber-roles-and-responsibilities. [Accessed: 02-Jul-2018].

[3] J. Stirland, K. Jones, H. Janicke, and T. Wu, “Developing Cyber Forensics for SCADA Industrial Control Systems,” Proc. Int. Conf. Inf. Secur. Cyber Forensics, no. October 2014, pp. 98–111, 2014.

[4] D. Quick and K.-K. R. Choo, “Big forensic data reduction: digital forensic images and electronic evidence,” Cluster Comput., vol. 19, no. 2, pp. 723–740, Jun. 2016.

[5] C. S. Centre and A. Jones, “Information Security and Digital Forensics in the world of Cyber Physical Systems,” in International Conference on Digital Information Management, 2016, pp. 10–14.

[6] R. K.-K. Choo, “The cyber threat landscape : challenges and future research directions The Library,” Comput. Secur., vol. 30, no. 8, pp. 719–731, 2011.

[7] L. Yang et al., “System-level design solutions: Enabling the IoT explosion,” Proc. - 2015 IEEE 11th Int. Conf. ASIC, ASICON 2015, 2016.

[8] International Data Corporation, “IDC Forecasts Worldwide Spending on the Internet of Things to Reach $772 Billion in 2018,” Press Release, 2017. [Online]. Available: https://www.idc.com/getdoc.jsp?containerId=prUS43295217. [Accessed: 16-Jul-2018].

[9] United States Department of Homeland Security, “Critical Infrastructure Security | Homeland Security,” 2018. [Online]. Available: https://www.dhs.gov/topic/critical-infrastructure-security. [Accessed: 16-Jul-2018].

[10] European Commission - Migration and Home Affairs, “Critical infrastructure,” 2018. [Online]. Available: https://ec.europa.eu/home-affairs/what-we-do/policies/crisis-and-terrorism/critical-infrastructure_en. [Accessed: 16-Jul-2018].

[11] P. Eden, P. Burnap, A. Blyth, K. Jones, H. Soulsby, and Y. Cherdantseva, “A Forensic Taxonomy of SCADA Systems and Approach to Incident Response,” pp. 42–51, 2015.

[12] Dragos Inc. and OSIsoft, “The Case of the Inside Job,” 2018. [Online]. Available: https://dragos.com/media/Dragos-PI.pdf. [Accessed: 26-Apr-2018].

[13] T. Stirland, Joe; Jones, Kevin; Helge, Janicke; Wu, Digital evidence and computer crime : forensic science, computers and the Internet. Academic Press, 2011.

[14] M. G. Noblett, M. M. Pollitt, and L. A. Presley, “Recovering and examining computer forensic evidence,” Forensic Sci. Commun., vol. 2, no. 4, Oct. 2000.

[15] P. Sommer, “The future for the policing of cybercrime,” Comput. Fraud Secur., vol. 2004, no. 1, pp. 8–12, 2004.

[16] S. Zawoad and R. Hasan, “FAIoT: Towards Building a Forensics Aware Eco System for the Internet of Things,” Proc. - 2015 IEEE Int. Conf. Serv. Comput. SCC 2015, pp. 279–284, 2015.

[17] K. J. Withers, Electronically Stored Information: The December 2006 Amendments to the Federal Rules of Civil Procedure, vol. 4, no. 2. Northwestern University School of Law, 2006.

[18] K. Kent, S. Chevalier, T. Grance, and H. Dang, “Guide to integrating forensic techniques into incident response,” 2006.

[19] D. J. Daniels and S. V Hart, “Forensic Examination of Digital Evidence : A Guide for Law Enforcement,” U.S. Dep. Justice Off. Justice Programs Natl. Inst. Justice Spec., vol. 44, no. 2, pp. 634–111, 2004.

[20] G. Shrivastava, K. Sharma, and R. Kumari, “Network Forensics: Today and Tomorrow,” pp. 2234–2238, 2016.

[21] P. Didier et al., “Converged Plantwide Ethernet (CPwE) Design and Implementation Guide,” p. 564, 2011.

[22] P. Johnson, “An OSI Model for Cloud,” 2017. [Online]. Available:

https://blogs.cisco.com/cloud/an-osi-model-for-cloud. [Accessed: 19-Jul-2018].

[23] RSA FraudAction Research Labs, “Anatomy of an Attack,” Speaking of Security, 2017. [Online]. Available: https://blogs.rsa.com/anatomy-of-an-attack/.

[24] IBM X-Force Research, “2016 Cyber Security Intelligence Index,” 2016. [Online]. Available: https://www.autoindustrylawblog.com/wp-content/uploads/sites/8/2016/05/IBM_2016-cyber-security-intelligence-index.pdf. [Accessed: 03-Jul-2018].

[25] T. Spyridopoulos, T. Tryfonas, and J. May, “Incident Analysis & Digital Forensics in SCADA and Industrial Control Systems.” [Online]. Available: https://pdfs.semanticscholar.org/7e6b/4b5292ee5a53865c97f31ec8da5931c0fda3.pdf. [Accessed: 17-Jan-2018].

[26] C. Easttom, System Forensics, Investigation, and Response, 3rd ed. Burlington, MA: Jones & Bartlett Learning, 2019.

[27] P. Garcia-Teodoro, J. Diaz-Verdejo, G. Macia-Fernandez, and E. Vazquez, “Anomaly-based network intrusion detection: Techniques, systems and challenges,” Comput. Secur., vol. 28, no. 1, pp. 18–28, 2009.

[28] J. Peng, K. K. R. Choo, and H. Ashman, “Bit-level n-gram based forensic authorship analysis on social media: Identifying individuals from linguistic profiles,” J. Netw. Comput. Appl., vol. 70, pp. 171–182, 2016.

[29] G. Frantzeskou, E. Stamatatos, S. Gritzalis, and S. Katsikas, “Effective identification of source code authors using byte-level information,” Proceeding 28th Int. Conf. Softw. Eng. ICSE 06, pp. 893–896, 2006.

[30] S. Sharma and S. V. Dhavale, “A review of passive forensic techniques for detection of copy-move attacks on digital videos,” ICACCS 2016 - 3rd Int. Conf. Adv. Comput. Commun. Syst. Bringing to Table, Futur. Technol. from Arround Globe, 2016.

[31] S. Zawoad and R. Hasan, “Digital Forensics in the Age of Big Data: Challenges, Approaches, and Opportunities,” 2015 IEEE 17th Int. Conf. High Perform. Comput. Commun. 2015 IEEE 7th Int. Symp. Cybersp. Saf. Secur. 2015 IEEE 12th Int. Conf. Embed. Softw. Syst., no. August, pp. 1320–1325, 2015.

[32] D. Quick and K. K. R. Choo, “Impacts of increasing volume of digital forensic data: A survey and future research challenges,” Digit. Investig., vol. 11, no. 4, pp. 273–294, 2014.

[33] F. Iqbal, H. Binsalleeh, B. C. M. Fung, and M. Debbabi, “A unified data mining solution for authorship analysis in anonymous textual communications,” Inf. Sci. (Ny)., vol. 231, pp. 98–112, 2013.

[34] N. Ibrahim, N. Al Naqbi, and O. Alfandi, “SIM Card Forensics : Digital Evidence,” Annu. Conf. Digit. Forensics, Secur. Law Conf., no. c, p. 219 to 234, 2016.

[35] R. Ayers, S. Brothers, and W. Jansen, “Guidelines on mobile device forensics,” May-2014.

[Online]. Available: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-101r1.pdf. [Accessed: 24-Jul-2018].

[36] K. Stouffer, V. Pillitteri, S. Lightman, M. Abrams, and A. Hahn, “Guide to Industrial Control Systems (ICS) Security,” Jun-2015. [Online]. Available:

https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-82r2.pdf. [Accessed: 24-Jul-2018].

[37] S. Khan, A. Gani, A. W. A. Wahab, M. Shiraz, and I. Ahmad, “Network forensics: Review, taxonomy, and open challenges,” J. Netw. Comput. Appl., vol. 66, pp. 214–235, 2016.

[38] I. Ahmed, S. Obermeier, M. Naedele, and G. G. Richard, “SCADA systems: Challenges for forensic investigators,” Computer (Long. Beach. Calif)., vol. 45, no. 12, pp. 44–51, 2012.

[39] W. Ren and H. Jin, “Modeling the network forensics behaviors,” 2005 Work. 1st Int. Conf. Secur. Priv. Emerg. Areas Commun. Networks, pp. 1–8, 2005.

[40] G. Palmer, “the first Digital Forensic Research Workshop,” First Digit. Forensic Res. Work., no. 1, pp. 15–18, 2001.

[41] R. Hunt and S. Zeadally, “Network Forensics: An Analysis of Techniques, Tools, and Trends,” Computer (Long. Beach. Calif)., vol. 45, no. 12, pp. 36–43, 2012.

[42] E. S. Pilli, R. C. Joshi, and R. Niyogi, “Network forensic frameworks: Survey and research challenges,” Digit. Investig., vol. 7, no. 1–2, pp. 14–27, 2010.

[43] M. Reith, C. Carr, and G. Gunsch, “An examination of digital forensic models,” Int. J. Digit. Evid., vol. 1, no. 3, pp. 1–12, 2002.

[44] R. Sira, “Network Forensics Analysis Tools: An Overview of an Emerging Technology,” no. Security 401, pp. 1–39, 2003.

[45] G. Shrivastava, “Network Forensics : Methodical Literature Review Network Forensics : Methodical Literature Review,” no. March, pp. 2203–2208, 2016.

[46] G. Williamson, “OT, ICS, SCADA – What’s the difference? - KuppingerCole,” Kuppinger Cole Analysts, 2015. [Online]. Available: https://www.kuppingercole.com/blog/williamson/ot-ics-scada-whats-the-difference. [Accessed: 31-Jul-2018].

[47] Cogent Industrial Technologies, “IT/OT Executive Series: Introduction - Cogent Industrial Technologies.” [Online]. Available: https://www.cogentind.com/it-ot-executive-series/.

[Accessed: 21-Dec-2017].

[48] SANS Institute, “Industrial IoT Security Survey Finds Practitioners Struggling to Identify Whats on Their Networks, How to Secure Things and Organizational Roles in IIoT Security,” Bethesda, MD, 2018.

[49] D. Quick and K. K. R. Choo, “Digital forensic intelligence: Data subsets and Open Source Intelligence (DFINT+OSINT): A timely and cohesive mix,” Futur. Gener. Comput. Syst., vol. 78, pp. 558–567, 2018.

[50] C. M. Talbot, M. A. Temple, T. J. Carbino, A. Betances, J. A. Betances, and A. Betances,

“Detecting Rogue Attacks on Commercial Wireless Insteon Home Automation Systems,” J. Comput. Secur., vol. Special Is, no. Internet and Cloud of Things, pp. 296–307, 2017.

[51] J. J. Lopez, N. C. Liefer, C. R. Busho, M. A. Temple, N. C. Liefer, and C. R. Busho, “Enhancing Critical Infrastructure and Key Resources (CIKR) Level-0 Physical Process Security Using Field Device Distinct Native Attribute Features,” IEEE Trans. Inf. Forenics Secur., vol. 13, no. 5, pp. 1215–1229, 2018.

[52] W. C. Suski, M. A. Temple, M. J. Mendenhall, and R. F. Mills, “Radio frequency fingerprinting commercial communication devices to enhance electronic security,” Int. J. Electron. Secur. Digit. Forensics, vol. 1, no. 3, p. 301, 2008.

[53] J. Hall, M. Barbeau, and E. Kranakis, “Detection of transient in radio frequency fingerprinting using signal phase,” Wirel. Opt. Commun., pp. 13–18, 2003.

[54] Y. Yuanling Huang and H. Hui Zheng, “Radio frequency fingerprinting based on the constellation errors,” in 2012 18th Asia-Pacific Conference on Communications (APCC), 2012, pp. 900–905.

[55] S. U. Rehman, S. Alam, and I. T. Ardekani, “An Overview of Radio Frequency Fingerprinting for Low-End Devices,” Int. J. Mob. Comput. Multimed. Commun., vol. 6, no. 3, pp. 1–21, Jul. 2014.

[56] P. Mirowski, D. Milioris, P. Whiting, and T. K. Ho, “Probabilistic Radio-Frequency Fingerprinting and Localization on the Run,” Bell Labs Tech. J., vol. 18, no. 4, pp. 111–133, 2014.

[57] D. R. Reising, M. A. Temple, and J. A. Jackson, “Authorized and Rogue Device Discrimination Using Dimensionally Reduced RF-DNA Fingerprints,” IEEE Trans. Inf. Forensics Secur., vol. 10, no. 6, pp. 1180–1192, 2015.

[58] S. Deng, Z. Huang, X. Wang, and G. Huang, “Radio Frequency Fingerprint Extraction Based on Multidimension Permutation Entropy,” Int. J. Antennas Propag., vol. 2017, pp. 1–6, Aug. 2017.

[59] F. Zhuo, Y. Huang, and J. Chen, “Radio Frequency Fingerprint Extraction of Radio Emitter Based on I/Q Imbalance,” Procedia Comput. Sci., vol. 107, no. Icict, pp. 472–477, 2017.

[60] C. M. Rondeau, J. A. Betances, and M. A. Temple, “Securing ZigBee Commercial Communications Using Constellation-Based Distinct Native Attribute (CB-DNA) Fingerprinting,” Secur. Commun. Networks.

[61] S. Williams, M. Taylor, J. Irland, and A. Mehta, “RFID Technology in Forensic Evidence Management :An Assessment of Barriers, Benefits, and Costs,” Gaithersburg, MD, Nov. 2014.

[62] T. Masahiro, “Special Issue on Cybersecurity Cyber-Physical Integrated Analysis Technology for Criminal Investigation Support,” NEC Tech. J., vol. 12, no. 2, 2018.

[63] W. Wang, Y. Sun, H. Li, and Z. Han, “Cross-Layer Attack and Defense in Cognitive Radio Networks,” in 2010 IEEE Global Telecommunications Conference GLOBECOM 2010, 2010, pp. 1–6.

[64] S. V Wankhede, M. N. Thakare, and S. R. Vaidya, “Design Approach for Cross Layer Attacks Defence in Cognitive Radio,” Int. Res. J. Eng. Technol., 2015.

[65] B. Premala, “International journal for scientific research et development : IJSRD.,” Int. J. Sci. Res. Dev., pp. 245–249, Oct. 2016.

[66] J. Tan, A. Liu, M. Zhao, H. Shen, and M. Ma, “Cross-layer design for reducing delay and maximizing lifetime in industrial wireless sensor networks.”

[67] J. Yu and X. Zhang, “A Cross-Layer Wireless Sensor Network Energy-Efficient Communication Protocol for Real-Time Monitoring of the Long-Distance Electric Transmission Lines,” J. Sensors, vol. 2015, 2015.

[68] G. Shrivastava, Handbook of research on network forensics and analysis techniques. IGI Global; 1 edition, 2018.

[69] J. Weiss, “Cyber Security of Sensors Are Not Being Addressed And Vulnerabilities Are Not Correlated to System Impacts,” Control - Unfettered Blog, 2018. [Online]. Available: https://www.controlglobal.com/blogs/unfettered/cyber-security-of-sensors-are-not-being-addressed-and-vulnerabilities-are-not-correlated-to-system-impacts. [Accessed: 08-Jan-2018].

[70] J. Weiss and J. Lopez, “The Gap in ICS Cyber Security and Safety – Level 0,1 Devices,” in 2018 ISA Power Industry Division (POWID) Conference, 2018.

Table 1: Location and type of forensic information available in SCADA implementation zones shown in Figure 5. This includes a representative sampling and summary of table entries provided in Eden et al. (2016) with the corresponding Figure 3 PERA layer added for completeness.

Zone	**Device Type**	**Forensic Data**	**PERA Layer**
Corp	Workstation(s)	Program/file execution, account/browser Usage, connected device logs	5
Corp	Email server	Server logs, email transactions	5
Corp	Web server	Server/event logs, IP addresses, session data	5
Corp	Database server	Data files, server/system event logs, trace files	4, 5
Corp	Routers/switches	Event logs, MAC addresses, IP/routing tables, CAM	3
Data	Workstation(s)	Program/file execution, account/browser usage, connected devices, logs	2, 3
Data	OPC server	Field device/communication logs	2, 3
Data	Application server	DHCP configuration log	2
Data	MTU	Logs, connected devices, PLC/RTU I/O data	1
Data	Domain controller	Logon/security events	1
Data	Routers/switches	Event log(s), IP table(s), CAM	1
Cont	Workstation	RAM, connected devices, PLC/HMI baseline images	1
Cont	HMI	Logs, issued commands, reports	1
Cont	Switches	Content addressable memory (CAM)	1
Cont	PLC/RTU	Logs, active processes, timestamps, ladder-logic program codes, SD card firmware versions	0