TOWARDS A THREAT INTELLIGENCE INFORMED DIGITAL FORENSICS READINESS FRAMEWORK

Nikolaos Serketzis, Vasilios Katos, Christos Ilioudis, Dimitrios Baltatzis, George Pangalos

Abstract

Digital Forensic Readiness (DFR) has received little attention by the research community, when compared to the core digital forensic investigation processes. DFR was primarily about logging of security events to be leveraged by the forensic analysis phase. However, the increasing number of security incidents and the overwhelming volumes of data produced mandate the development of more effective and efficient DFR approaches. We propose a DFR framework focusing on the prioritisation, triaging and selection of Indicators of Compromise (IoC) to be used in investigations of security incidents. A core component of the framework is the contextualisation of the IoCs to the underlying organisation, which can be achieved with the use of clustering and classification algoriihms and a local IoC database.

Keywords: Digital Forensic Readiness, Threat Intelligence, Indicators of Compromise..

1 Introduction and motivation

Digital forensics date over four decades. Unlike other forensic science disciplines, digital forensics faces the challenge to operate in a problem domain where the subject of study evolves in an intermittent, non-linear fashion; a routine, nightly update of the software or introduction of new hardware may substantially change the behaviour of the underlying system, requiring a significant revision of the digital forensics acquisition and analysis processes. Consider for example the case of evolution of traditional hard disks to solid state disk (SSD) technology. The way the latter operate invalidate many key assumptions under which forensic acquisition and investigation of disks is performed (Bednar and Katos 2011).

Moreover, the proliferation of heterogeneous networked devices and the amount of data they are capable of producing – as captured under the terms IoT and Big Data respectively – has exacerbated the problems and challenges of digital forensics. As such, digital forensic readiness (DFR) has become a critical function to the organisation's security processes and achieving effective DFR has become a high priority. However, research in digital forensics has primarily evolved through a responsive, practitioner-based attitude. The relevant literature on digital forensics is dominated by techniques and practical approaches for obtaining and analysing data in specific contexts and system configurations. When it comes to consider DFR approaches, the level of abstraction is high causing a void and eventually a disjoint between DFR and digital forensic investigations. Most DFR research publications are limited to describing high level and generic steps, whereas contextualisation is mostly absent. The aim of this work is to bridge the gap by proposing a framework for a closer coupling between DFR, forensics and incident response for addressing Advanced Persistent Threats.

The rest of the paper is structured as follows. Section 2 presents the relevant literature. In Section 3 our approach is developed. Section 4 outlines a representative APT scenario to be used as a vehicle to showcase our approach and section 5 summarises the conclusions.

2 Current research framework

In a seminal paper, Hutchins et al (Hutchins, Cloppert, and Amin 2011) proposed an approach for studying and improving incident response against APTs. They introduced a cyber kill chain which identifies a path comprised of 7 discrete and sequential phases an attacker follows in order to meet their adversarial goals. From a digital forensics perspective, the kill chain is particularly helpful in highlighting the following:

- Every successful (to the attacker) phase is a direct consequence of the respective security control failures.
- Detecting the security breach early in the chain infers low impact and potential damage.
- Late detection of the security breach implies that there are more security failures hence the scope of the digital forensic artefact collection is wider.

For the remainder of this section the relevant subtopics that will enable the key chain to leverage the proposed DFR framework are presented.

2.1 Threat Intelligence

It can easily become apparent from the current literature that there is limited consensus on a definition of threat intelligence. Threat intelligence has been defined for example as a product resulting from the collection, processing, integration, evaluation, analysis, and interpretation of available information concerning foreign nations, hostile or potentially hostile forces or elements, or areas of actual or potential operations (Sanders and Smith 2014). It can be therefore considered that threat intelligence is the elaborated information about threats targeting one or more organizations.

Threat intelligence can be produced both from internal (e.g. Firewall, IDS) and external sources, such as public or commercial threat and vulnerability repositories. Externally obtained intelligence is sought as being particularly beneficial to the organisation as this promotes cyber situational awareness.

Research on threat intelligence has highlighted the need for automated information exchange. To this extent, various standards and formats (openIoC, CybOX, STIX,) have been developed (MITRE n.d.) (MITRE 2017) (Mandiant Corporation 2013), with the most promising and publicly acceptable being CybOX, STIX and TAXII (Sauerwein et al. 2017), (Fransen, Smulders, and Kerkdijk 2015).

Cyber Observable eXpression (CybOX) is a standardized approach which leverages eXtensible Markup Language (XML) to encode and share information about observables in the operational cyber domain. CybOX can be used to describe almost any type of information. Typical examples include IP addresses, domain names, filenames, file content and any sort of text pattern.

Structured Threat Information eXpression (STIX) is another structured language which is used to specify, capture, characterize and communicate standardized cyber threat information. STIX represents a holistic approach to format threat intelligence, by incorporating a wide set of information like Indicators, Incidents, Tactics, Techniques and Procedures (TTP), Campaigns, Threat Actors, Exploit Targets, and Courses of Action (COA). As of version 3, CybOX has been integrated into the STIX schema (Barnum 2014).

Trusted Automated Exchange of Intelligence Information (TAXII) in turn, is a mechanism that facilitates the exchange of cyber threat information. TAXII is optimized to ensure the smooth exchange of information represented in STIX (OASIS Technical Committee n.d.).

It becomes clear that threat intelligence can assist in identifying an incident thus enhancing an organization's information security posture. On the other hand, absent or outdated information may considerably limit security personnel's awareness about an incident. Should this be the case, performing a comprehensive digital forensics investigation exercise could shed light to the root cause of the event.

2.2 Digital forensics

Digital forensics (DF) history dates back approximately forty years, but notable maturity took place post 1997 (Garfinkel 2010). DF encompasses a number of well defined steps, with the aim to assist an investigator to identify the source and the root cause of an event, thus answering six key questions; what, why, how, who, where and when (Ieong 2006).

Despite the continuous maturity and evolution of DF, its effectiveness is in debate, primarily due to the advances in IT industry (Garfinkel 2010). More specifically:

- The proliferation of portable devices such as smartphones, tablets, smart TVs, etc., resulted in significant increase in the information produced and the diversity of operating systems and data.
- The volume of data that need to be examined has been increased, making the investigations lengthier in time and effort, and more expensive.
- The broad adoption of cloud services fosters the perception that new approaches to digital forensics investigations need to be evolved.
- The expansive use of encryption both in commercial and personal devices deter the extraction of forensic artefacts.
- The sophistication in malware development prevents the production of permanent forensic evidence, as many malware variations write temporary data only in RAM.
- The dissimilarity among national legal frameworks and the absence of a unified international legal framework renders cross-border investigations a challenging task.

Furthermore, Garfinkel (2010) noted the upcoming crisis in modern digital forensics by identifying a number of challenges in both the approaches of building the specialised tools but also the forensic analyst's practices. Perhaps the most relevant and important highlight in Garfinkel's paper is the evidence-oriented design of the digital forensic tools where the emphasis is placed on detecting possession of evidence rather than the actual crime being committed. This approach essentially invalidates the relevant tools from conducting computer focused crime investigations. In addition, the silo and monolithic nature of digital forensic applications does not allow opportunities to integrate with digital forensic readiness processes.

2.3 Digital forensic readiness

Digital Forensic Readiness aims to maximize an organization's ability to collect credible evidence, whilst minimizing the cost of an investigation (Tan 2001). To date several approaches has been proposed.

Focusing at the policy dimension, Yasinsac and Manzano (Yasinsac and Manzano 2001) stated that a set of policies like information retention, planning of response, training, investigation acceleration, prevention of anonymous activities and protection of evidence could facilitate the digital forensics process. Rowlingson (Rowlingson 2004) stressed out the need for forensic readiness to be incorporated to an enterprise's forensic program. Proactive evidence identification, collection, secure storage and training are among the key priorities of their proposal. Grobler and Louwrens (Grobler and Louwrens 2007) underlined the overlap between information security and digital forensics and argued that digital forensic readiness must become a component of information security best practice. They also believe that the scope of DFR should be broadened to incorporate IS governance and augment the security program of the organization. Pangalos and Katos (Pangalos and Katos 2010) highlight that a relationship between Information Security and Digital Forensics exists. They identify the residual risk as the main reason that drives the need for digital forensics, and believe that a forensics-aware security strategy will manage to mitigate the impact of a security incident. Valjarevic and Venter (Valjarevic and Venter 2011) proposed a model following a holistic approach being comprised of 10 phases including scenario definition, Identification of possible sources, pre-incident collection, pre-incident analysis, incident detection, post-incident collection, post-incident analyses, definition of system architecture and assessment of implementation.

Approximately 15 years after Tan (Tan 2001) introduced the concept of forensic readiness, the emergence of ISO/IEC 27043 (International Organization for Standarization 2015) indicates a significant level of maturity in this field. In essence, this standard developed with the aim to provide guidelines for incident investigation principles and processes, but it also acknowledges the importance of digital forensic readiness and welcomes it as a special class within the model.

3 The proposed DFR framework

Comment [A1]: Δεν μπορούμε να μιλήσουμε ακόμη για methodology γιατί είναι ακόμα αρκετά high level η περιγραφή. Όπου μιλάμε για methodology στο κείμενο θα πρέπει να έχουμε αρκετή λεπτομέρεια.

As with most information security processes, DFR should be performed in a continuous and repeating fashion, rather than being a one-off process. Threat intelligence should be used to continuously inform and help prioritize the selection and collection of the necessary fields and features that would be used to support the digital forensic investigation in the event of a security incident. At this stage the distinction between a feature and an Indicator of Compromise, IoC, should be given:

Definition 1. A *feature* is an individual observable property capable of describing aspects of a state of a system.

Essentially a feature in this paper is meant to map to the concept of the feature as defined in the machine learning domain.

Definition 2. An *Indicator of Compromise* (IoC) is a specific instance or value of a particular feature.

From a machine learning perspective, an IoC can be seen as the labelling exercise, where specific tuples in a dataset are labelled. Labelling is needed in supervised or hybrid machine learning classification and clustering algorithms. For example identified features may include IP addresses, port numbers, file hashes, whereas an IoC would be the specific values, such as port:443, IP:61.12.13.14, hash:0x3e324ffd4e574639a0bc.

The proposed framework intends to provide a tool for prioritising – aka triaging – and identifying the stage of an attack in the cyber kill chain (Figure 1). Reflecting upon the work by Hutchins et al. (2010) it is assumed that an APT type of attack would involve attack patterns and malicious campaigns that may manifest in one or more organisations. By continuously receiving information on IoC from external sources, the information provided by and to the DFR would support correlation activities in order to answer the questions of forensic interest. A high level illustration of the framework is shown in Figure 2.

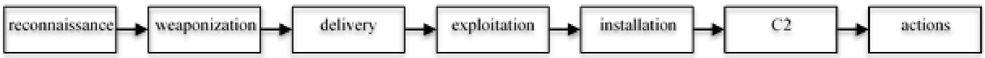

Figure 1. The cyber kill chain (adapted from: Hutchins et al., 2010)

As shown in Figure 2, the digital forensics investigation is triggered at time t_d, at the first instance of detecting a security control failure and a successful compromise. On the general model, there is an amount of delay between the security incident leading to the system compromise and its detection. This delay depends on a number of factors and is outside the scope of this paper. However, what is of the main interest and within the focus of the DFR is the efficiency by which the evidence is collected and prioritised. As such, efficient performance of forensic analysis would mean minimisation of t_a-t_d, that is, a reduced depth of attack, disruption of the malicious campaign and improvement of the intrusion detection and intrusion prevention layers.

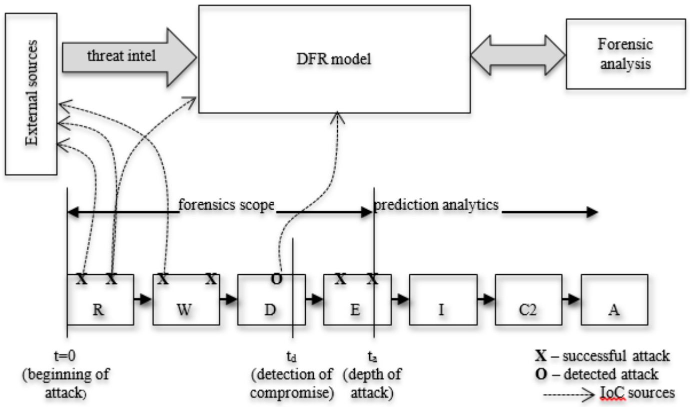

Figure 2. An integrated DFR framework

Another important aspect of the proposed approach is the continuous identification of the sources of IoCs. During an attack, not all information and IoCs will necessary be captured by the internal, in-house sensors, but some IOCs will be present in external repositories and sources. Consider for example shodan.io which captures and indexes the digital footprint of all contactable devices. During reconnaissance, an attacker may query the shodan servers to discover open ports for a specific IP range or organisations. This would be equivalent to a port scanning attack, but without even touching the actual servers; the victim organisation would be completely agnostic and oblivious of the port scanning activity since this would not be logged by their logging servers. Therefore, for every attack (denoted as **X** or **O** in Figure 2) the corresponding IoC could be located internally, externally, or in both places. Consequently, the forensic analysis process should tap into a DFR framework capable of integrating with both external threat intelligence feeds (Open Source, OSINT) as well as with internal, Security Information and Event Management, SIEM components. These requirements essentially transform a DFR from a logging facility to a fully blown process of clustering, classification of security incident features. This is shown in Figure 3.

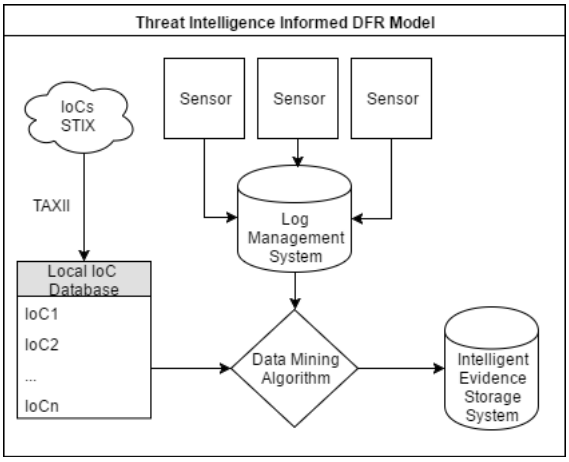

Figure 3. Threat intelligence informed DFR model.

More specifically, the proposed DFR methodology includes the following five steps:

3.1 Evidence Identification and Selection

Contemporary network equipment (routers, switches, etc.), security devices (Firewalls, IDSs, etc.), operating systems, and applications (web servers, mail servers, etc) offer logging data alongside their main operations. In a typical attack scenario, an adversary would attempt to discover the organizations' hardware and network assets, exploit their vulnerabilities, and attempt to install harmful applications in order to collect sensitive information or harm the systems themselves. In such a case, network sensors, operating systems or services themselves may collect useful data such as network connections, file changes etc. The organizational security policies are expected to define what data should be logged. Typically, such a selection of the data is the result of a formal risk assessment procedure. DFR in turn, can be used as a means to ameliorate the collection process, that is to further identify possible cases that require credible evidence collection. For instance, ISO/IEC 27043 incorporates the "scenario definition" process to describe how DFR assists in identifying the evidence required.

3.2 Evidence Collection

Different devices usually collect different types of data. The need to effectively elaborate these data requires a suitable level of centrality and uniformity. The former can be achieved relatively easy, by the employment of a central Log Management System. Storing this data into a central log management system is considered an effective approach from a management and security viewpoint (Elyas et al. 2014). Secure logging protocols can also be engaged to enforce the integrity and accountability of the collected evidence (Accorsi 2009).

On the other hand, the utilization of log parsers contributes to some extent, to the uniformity of the data captured. Unfortunately, it is not possible to achieve 100% homogeny in data, as they may describe various structures, like network traffic, connections, files, text, etc. Having that in mind, the authors acknowledge that the STIX language can be employed to effectively describe the data structures of the proposed framework.

Additionally, data storage mechanisms should be taken into consideration. While relational databases are considered stable and scalable solutions, the extensive employment of integrity procedures renders them inappropriate for managing log data. In contrast, NoSQL databases provide powerful query tools, but also demand hardware commitment, added programming, and administrative effort (Collins 2014). It is thus evident that choosing the suitable log management system borrows from the "scenario definition" phase.

3.3 Creation of the Local IoC Database

As stated above, IoCs can be produced internally, as a result of an incident analysis, or externally by third-party information security firms or individuals. Actually, more accurate IoCs are commonly produced externally, as they may be the result of extensive investigatory procedures, like malware analysis.

The model we propose is based on a separate, structured database containing only the appropriate IoCs. For efficiency and homogeneity reasons our proposed methodology uses the STIX language to describe the IoCs and TAXII to communicated them to the Local IoC Database.

For every IoC, their relevance to the organizations assets must also be considered before populating the Local IoC database, thus an initial IoC selection must be performed. This selection must take into account the results of the evidence identification phase. For example, it is worthless collecting IoCs that relate to operating systems an organization lacks. This contextualization process is a direct consequence of the threat intelligence and information sharing capabilities the DFR framework would possess.

Moreover, it is worth highlighting that all identifiable external IoCs would be subjected to the same risk assessment procedure applied for the Local IoC Database and subsequently decided whether it would be beneficial to include them.

3.4 The Data Mining Process

Information originating from the Log Management System and the Local IoC Database feed the Data Analysis System. Employing both unsupervised and supervised data mining algorithms, the Data Analysis System:

- firstly identifies whether an incident has taken place and
- thereafter correlates the information pertaining to this incident.

The outputs and results of this process are then forwarded to the Intelligent Evidence Storage System.

In detail, records entering the Data Analysis System are classified into categories according their type and sensor location, thus producing clusters of similar information. Data classification algorithms are then applied to every cluster record to further identify whether it relates to a security incident.

Data classification algorithms partition data sets into predefined classes. Such categorization is based on group identifiers of these classes that are commonly known as "class labels" (Aggarwal 2015). The proposed methodology employs indicators of compromise to define two class labels, "benign data" and "malicious data".

For example, if a record that comes into the Data Analysis System contains information relating to any IoC within the Local IoC Database, then this record is considered "malicious". Should this occur, that record is forwarded to the Intelligent Evidence Storage System, while a link analysis procedure is initiated for the discovery and association with similar records.

3.5 The Intelligent Evidence Storage System

The Intelligent Evidence Storage System is the last component of the proposed methodology. It is comprised of a central database that only stores information about incidents, but also includes links to related records. In an event of a security incident, it is more practical and time efficient for investigators to search for evidence within the Intelligent Evidence Storage System, than checking the whole logging inventory.

4 Pilot implementation - Example APT scenario

The following scenario which is used to demonstrate the advantages of the proposed DFR methodology, is based on a real case attack which is part of a popular malicious campaign. An adversary group targets a company, aiming to exfiltrate confidential data. Employing the sophisticated type of watering-hole attack and social engineering techniques, the offenders exploited a zero-day web browser vulnerability, and managed to install a custom-made malware on a PC which in turn initiated a tunnel connection to a command and control (C2) server listening on port 443. In this scenario we assume that information security devices like firewalls, IDSs, etc. are updated to the most recent versions of signatures.

On a first decomposition of the incident, we can note the following assumptions and observations. One of the employees visited a normal webpage, but this webpage has previously been tampered with a browser exploit. The network devices log the connections to the webpage. The browser also holds web history. The exploitation of the web browser allowed the installation and execution of the malware on the PC, and thus, the alteration of file system's and registry's records. The hash signature of the malware executable is not included in the antivirus database, so no alarm is raised. The destination IP address where the C2 server resides does not belong to any list of blocked IP addresses while port 443 maps to https protocol, thus TCP connections to this IP address are permitted, but logged as well.

A week later a private information security firm informs the company that its confidential records has been published on the internet. This firm has also identified that a new malware distribution campaign exists, analysed its characteristics and published the relevant indicators.

Following the traditional approach, should the company need to identify how the adversaries compromised its systems, a full forensic investigation is needed. In particular, all log files produced by the network devices, along with terminal equipment must be thoroughly examined. It is interesting to note that, following this approach, the investigator has no a priori knowledge of the way the incident took place, thus more time is needed to identify the indicators of compromise.

Our methodology employs the use of IoCs. In this scenario such indicators have already been created and communicated by the private security firm. Using these indicators to filter the whole logging repository and correlate the events, a subset database is produced. In this case, the database contains the hash value of the malware among other indicators. The correlated records within the Intelligent Evidence Storage System are able to suggest the investigator what the most possible reason of the compromise is, narrowing the timeframe need for performing a full forensic investigation, thus enhancing forensic readiness.

Initial results from the application of the proposed methodology on the above example APT scenario have produced encouraging results. More detailed experimentation is currently on the way and we are expecting that it will also confirm the correctness of our approach.

5 Conclusions and outlook

The volume, variety and velocity of data produced by contemporary networked devices challenges the efficiency of traditional digital forensics approaches. While DFR promises maximized collection of credible evidence and more cost effective investigations, most of its research publications are limited to describing high level and generic steps, whereas contextualisation is mostly absent. In addition we argue that a high volume of collected evidence may reach a point that would undermine the cost effectiveness and as such we recognise the need to employ This paper presented a methodology that incorporates the strengths of the threat intelligence domain into DFR with the aim to facilitate the DFR process.

References

Accorsi, Rafael. 2009. “Safe-Keeping Digital Evidence with Secure Logging Protocols: State of the Art and Challenges.” In *2009 Fifth International Conference on IT Security Incident Management and IT Forensics*, IEEE, 94–110. http://ieeexplore.ieee.org/document/5277863/.

Aggarwal, Charu C. 2015. *Data Mining*. Cham: Springer International Publishing. http://link.springer.com/10.1007/978-3-319-14142-8.

Barnum, Sean. 2014. “Standardizing Cyber Threat Intelligence Information with the Structured Threat Information eXpression (STIX^{TM}).” *MITRE Corporation, July*: 1–20. http://blackberry8520.b277.doihaveamobilestrategy.com/http://stix.mitre.org/about/documents/STIX_Whitepaper_v1.0_(Draft).pdf.

Bednar, Peter, and Vasilis Katos. 2011. “SSD: New Challenges for Digital Forensics.” *ItAIS 2011, Proceedings of the 8th Conference of the Italian Chapter of the Association for Information Systems* (October 2011): 1–8.

Collins, Michael. 2014. *Network Security Through Data Analysis*. ed. MacDonald Andy,Oram Allyson. O’Reilly Media, Inc.

Elyas, Mohamed, Sean B Maynard, Atif Ahmad, and Andrew Lonie. 2014. “Towards A Systemic Framework for Digital Forensic Readiness.” *Journal of Computer Information Systems* 54(3): 97–105. http://www.scopus.com/inward/record.url?eid=2-s2.0-84900844255&partnerID=40&md5=56c48e92a269a383f51cadb7f07db87e.

Fransen, Frank, Andre Smulders, and Richard Kerkdijk. 2015. “Cyber Security Information Exchange to Gain Insight into the Effects of Cyber Threats and Incidents.” *Elektrotechnik & Informationstechnik* 18: 106–12. http://link.springer.com/10.1007/s00502-015-0289-2.

Garfinkel, Simson L. 2010. “Digital Forensics Research: The next 10 Years.” *Digital Investigation* 7: S64–73. http://www.sciencedirect.com/science/article/pii/S1742287610000368.

Grobler, C. P., and C. P. Louwrens. 2007. “Digital Forensic Readiness as a Component of Information Security Best Practice.” *IFIP International Federation for Information Processing* 232: 13–24.

Hutchins, Eric M., Michael J. Cloppert, and Rohan M. Amin. 2011. “Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains.” *6th Annual International Conference on Information Warfare and Security* (July 2005): 1–14. http://papers.rohanamin.com/wp-content/uploads/papers.rohanamin.com/2011/08/iciw2011.pdf%5Cnhttp://www.lockheedmartin.com/content/dam/lockheed/data/corporate/documents/LM-White-Paper-Intel-Driven-Defense.pdf.

Ieong, R. S C. 2006. “FORZA - Digital Forensics Investigation Framework That Incorporate Legal Issues.” *Digital Investigation* 3(SUPPL.): 29–36.

International Organization for Standarization. 2015. *ISO/IEC 27043:2015. Information Technology, Security Techniques, Incident Investigation Principles and Processes*. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=44407.

Mandiant Corporation. 2013. “OpenIOC.”

MITRE. 2017. “STIX: A Structured Language for Cyber Threat Intelligence.” https://oasis-open.github.io/cti-documentation/ (March 5, 2017).

———. “Cyber Observable eXpression (CybOX).” https://cybox.mitre.org/about/ (March 24, 2017).

OASIS Technical Committee. "Cyber Threat Intelligence Documentation." https://oasis-open.github.io/cti-documentation/ (March 24, 2017).

Pangalos, Georgios, and Vasilios Katos. 2010. "Information Assurance and Forensic Readiness." In *Next Generation Society. Technological and Legal*, , 181–88. http://www.springerlink.com/index/g60450l318835355.pdf.

Rowlingson, Robert. 2004. "A Ten Step Process for Forensic Readiness." *International Journal of Digital Evidence* 2(3): 1–28.

Sanders, Chris, and Jason Smith. 2014. *Applied Network Security Monitoring*. ed. David J. Bianco. Elsevier Inc.

Sauerwein, Clemens et al. 2017. "Threat Intelligence Sharing Platforms : An Exploratory Study of Software Vendors and Research Perspectives." In , 837–51.

Tan, John. 2001. "Forensic Readiness." https://isis.poly.edu/kulesh/forensics/forensic_readiness.pdf.

Valjarevic, Aleksandar, and H. S. Venter. 2011. "Towards a Digital Forensic Readiness Framework for Public Key Infrastructure Systems." *2011 Information Security for South Africa*: 1–10.

Yasinsac, Alec, and Yanet Manzano. 2001. "Policies to Enhance Computer and Network Forensics." In *Proceedings of the 2001 IEEE Workshop on Information Assurance and Security*, United States Military Academy, West Point, NY, 5–6.