Forensics for multi-stage cyber incidents: Survey and future directions

Antonia Nisioti[a,*], George Loukas[a], Alexios Mylonas[b], Emmanouil Panaousis[a]

---

1. Introduction

Modern adversaries, such as Advanced Persistent Threats (APTs) adversaries, are organised and sophisticated, having the time and computational and

monetary resources to achieve their goals. Their strategic nature includes comprehensive planning of the attacks they carry out, the use of anti-forensic techniques, as well as spreading their actions across large time periods to evade detection.

These characteristics have made cyber forensic investigations more complex, knowledge-demanding and time consuming. Investigators need to spend a significant amount of time to keep their knowledge and training up-to-date to be able to cope with the evolution of the current threat landscape. Consequently, the need has arisen for decision support methods that will allow investigators to increase their efficiency [1] and complete the analysis of sophisticated cyber incidents, involving a variety of Tactics, Techniques, and Procedures (TTPs), efficiently and in a timely manner.

Motivated by these challenges, we present a survey on methods that support cyber forensic investigations of multi-stage cyber incidents by strategic adversaries. We include approaches, both technical and non-technical, that can improve the efficiency of the investigation of sophisticated cyber security breaches. By comparing them against the current needs of the cyber forensic field, we aim to identify current support for practitioners and open areas and opportunities for future research. All works included are either focused or can be applied to support the investigation of multi-stage cyber incidents, such as APTs. Thus, works focusing on other types of cyber incidents or threats are not included in this survey. To the best of our knowledge there is no previous survey on this topic and thus in Section 2.3, we briefly summarise the surveys that more closely align with our work.

More specifically, the contributions of this paper are the following:

1. We provide a *comprehensive review* of the literature of works that support cyber forensic investigations and increase their efficiency. The review includes the current technical and non-technical approaches that can aid an investigator to cope with the ever-changing and sophisticated or even multi- staged attacks of the threat landscape.
2. These methods of the surveyed works are not only presented but also *compared* against a set of selected evaluation criteria, such as the type of

evaluation used, the dataset used. This allowed us to identify their strengths and weaknesses and thus draw recommendations for future works.

3. We *identify open issues* and discuss future directions that could allow systems to evolve and better support digital investigations in the threat landscape, e.g. by more realistic and representative modelling.

The rest of the paper is structured as follows. Section 2 provides the necessary background and presents previous surveys related to this work. Section 3 explains the collection and comparison criteria for the surveyed works. Section 4 provides a comprehensive review and comparison of those works. Finally, Section 5 discusses the identified open issues and future research directions, while Section 6 concludes the paper.

2. Background Information and Related works

This section provides the necessary background of concepts required for a sufficient understanding of the analysis and results of this review, as well as discusses related work.

2.1. Advanced Persistent Threats

Although originally the term APT was used to refer only to sophisticated threat groups that were sponsored by a third-party entity, currently it has a broader meaning [2]. Specifically, it represents strategic Attackers who are: (i) highly motivated and determined to achieve their goals, (ii) stealthy to avoid detection, (iii) deploy a variety of sophisticated TTPs, and (iv) perform long term or reiterated attack campaigns [3]. APT campaigns consist of multiple stages that range from the reconnaissance and planning of the attack to the achievement of the Attackers' final goals. There have been many similar lifecycle models proposed that represent these stages, such as the Unified Kill Chain [4]. The term *Tactics, Techniques, and Procedures* (TTPs) is used to describe the behaviour, aims and objectives of such adversaries, as well as the techniques they deploy to achieve them [5]. Currently, the most well-known and used knowledge base of TTPs is the MITRE ATT&CK[1] framework, which is constructed using real-world observations through the collection and processing of past incident reports.

2.2. Cyber Forensics

Digital forensics (DF) is the science of collecting, analysing and reporting evidence from data found on electronic media [6]. It combines computer science and investigative procedures to investigate [7] either: (i) criminal activity that involves but does not directly targets computer systems or any electronic device, or (ii) malicious activity that directly targets or involves computer systems and networks. The sub-field of DF that is focused on the latter, and thus overlaps with the field of cyber security, is called *cyber forensics*.

There are many similar or identical investigative process models for cyber forensics, but a typical one consists of the following stages: collection,

[1] https://attack.mitre.org/

examination, correlation, and reporting [8]. Each one of those stages consists of multiple technical or non-technical tasks. The collection stage includes the identification of the media or data that may be related and useful to the case, as well as their labelling, recording, acquisition, and integrity preservation according to the relevant guidelines. After the collection is complete, during the examination stage, the investigator needs to assess and extract relevant information from the collected data in an automated or manual way. This is essentially the technical analysis of the data collected and it may include a variety of tasks, such as examination of the registry, certain events from the collected logs, process analysis or malware extraction from the collected memory dumps or network analysis. It may also include tasks such as decryption, de-obfuscation, reverse engineering or bypassing of certain security features.

The correlation stage is where the uncovered information is connected and interpreted in an objective and logical way using different types of inference and reasoning methods. This may include guidelines and frameworks for good practices or/and tools and reasoner systems that deploy methods, such as casebased reasoning (CBR) [9], temporal aggregation [10] and forward and backward reasoning [11]. Finally, it is crucial that the whole investigative process has been completed in a forensically sound manner, following best practices and relevant legislation to ensure the integrity of the collected evidence, avoid mishandlings [12] or misinterpretations [6].

2.3. Related work

Saeed et al. [13] survey different game-theoretic approaches for modelling the interaction between an investigator and an Attacker using an anti-forensic method, such as a rootkit or steganography. Similarly, Conlan et al. [14] survey 308 anti-forensic tools and use their extracted capabilities to propose an extended taxonomy, as well as a comparison based on criteria, such as the country of origin and OS. Even though our work is interested in the use of anti-forensics by Attackers, we do not focus on specific anti-forensic methods and their analysis, e.g., rootkits and anti-rootkits. Instead, we focus on the decision support methods that are aware of the potential existence of anti-forensic techniques as part of the incident under investigation.

James and Gladyshev [15] present a survey on: (i) investigative processes and (ii) how the triage stage affects the decision to exclude or include piece of evidence from the full in-depth analysis, via interviews and experiments with forensic practitioners. While the authors focus on the high-level investigation process, we focus on the part of the technical analysis and its efficiency.

Quick and Choo [16] study how the large amount of produced data affects modern forensic investigations. Similarly to our work, the authors aim to identify research gaps that contribute to the efficiency of the investigation, but while they focus on data reduction methods, visualisation, data mining etc., we focus on decision support methods. One common finding that we also highlight as a research gap in the cyber forensic science is the absence of threat intelligence knowledge bases.

Lemay et al. [17] study publicly available reports regarding 40 well-known APT threat groups and present a comprehensive reference guide on their activities, targets, motives, and techniques. Likewise Ahmad et al. [2] study the definition of the term APT, focusing on the strategic nature of such threats, the role of human situation awareness and how it can be potentially used as a countermeasure. The authors use the results of their survey to develop and present an operational framework for the interpretation of APTs.

Finally, Alshamrani et al. [18] survey the different methodologies and TTPs used by APT groups, as well as detection, monitoring and mitigation methodologies from the past literature, which can be used by defenders. This work also focuses on strategic adversaries, but we study methods that support the investigator during the analysis of those adversarial activities.

In conclusion, although several surveys have been proposed focusing on digital forensics or APTs and strategic Attackers, none has focused on their intersection or the different levels and ways of supporting cyber forensic investigations. Contrary, in this survey we review the literature for works that aim to support cyber forensic investigations, i.e., the investigation of modern multi-stage incidents potentially performed by strategic Attackers (e.g., APTs), by increasing their efficiency. This could be achieved either on a technical or more abstract level, in a automated or manual way, by proposing tools, reasoning frameworks or decision support systems.

3. Methodology

3.1. Inclusion criteria

Our aim in this survey is to review works that can be used during cyber forensic investigations of modern sophisticated multi-stage attacks to increase the efficiency and quality of the investigative process. We note that the relevant literature is somewhat limited, therefore we included the most known or promising works in our list that directly or indirectly support any part of such an investigation.

Thus the inclusion criteria for this work can be summarised as: • Works that directly support and aim to make more efficient the technical or non technical phases of a forensic investigation in an automated or manual way.

• Works that although may have not been proposed as cyber investigation support methods, they can be used for this purpose. An example of this would be a work that has been proposed as a defence mechanism for event correlation, but it can obviously be applied to an investigation as well. • Works that directly or indirectly support the investigation of multi-stage attacks. Directly refers to the case where a paper takes the multi-stage nature of an incident in consideration in its modelling and methodology. On the opposite hand, indirectly refers to a work that although it is not created for a multi-stage incident, it can be applied to one.

With regards to our paper inclusion methodology, we initially created a first pool of papers based on searches on digital libraries, such as Google Scholar, IEEE Xplore and ACM Digital Library with keywords like “digital forensics optimisation”, “cyber forensics”, “forensic decision support”, “multi-stage attack forensic analysis”, etc. We also enumerated the proceedings of relevant top conferences and journals like ACM Symposium on Computer and Communications Security, Digital Forensic Research Workshop, IEEE Transactions on Information Forensics and Security, and Forensic Science International: Digital Investigation.

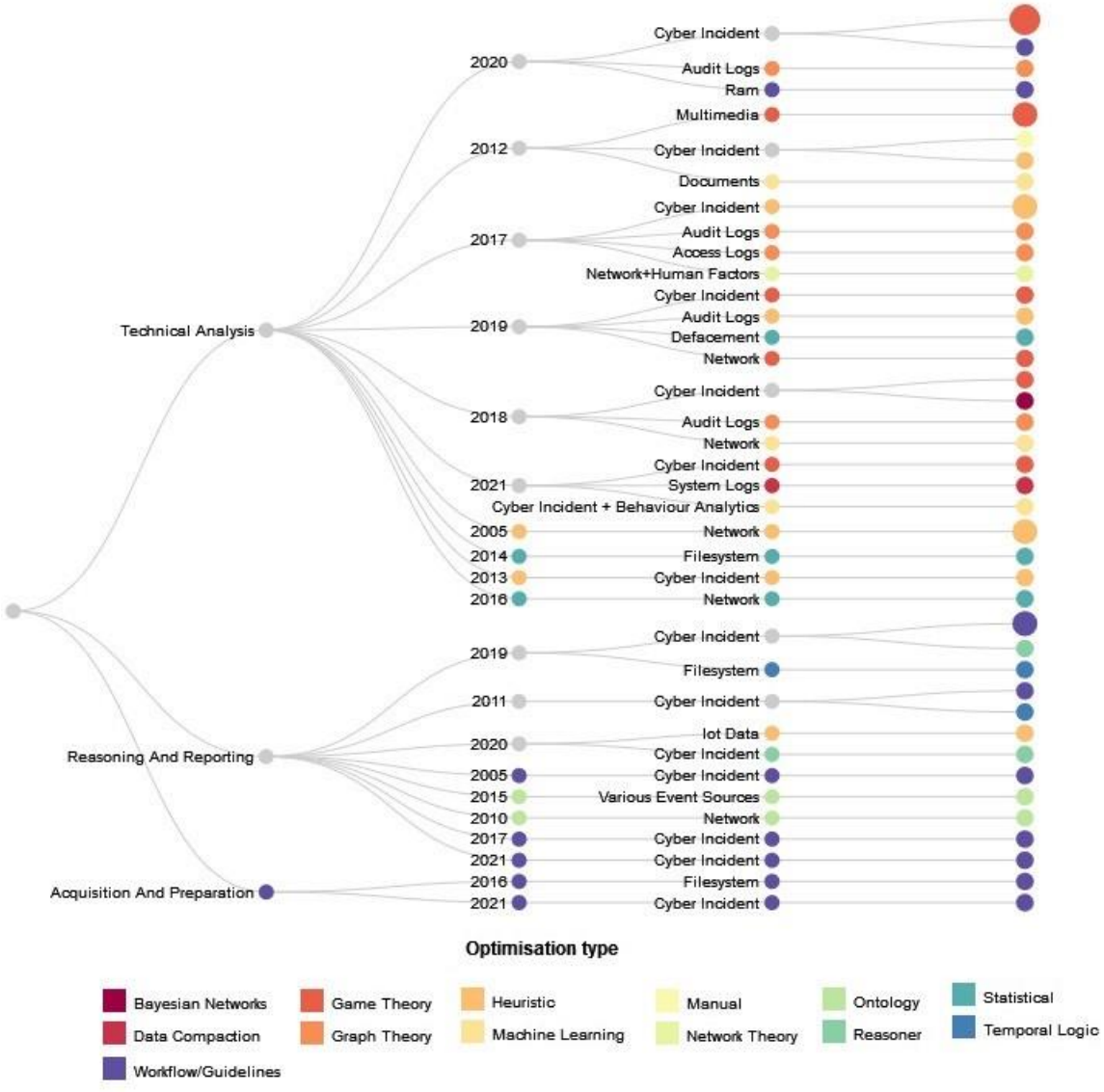


Figure 1: Linear Dendrogram depiction of surveyed works based on: (i) targeted part of the investigation, (ii) year of publication, (iii) data in use, and (iv) optimisation in use.

Our initial list was completed with recommendations of specific papers based on the authors' personal knowledge. Finally, we finalised this initial pool of papers with: a) relevant papers found in the reference list of the already selected papers in the initial pool and b) any additional papers in the proceeding of conferences or journals in which the selected papers were published, only if they have not been considered.

After the completion of this process we had accumulated a collection of 90 papers. Then we used our list of inclusion criteria to select a subset of these papers that met them. This resulted in a set of 49 papers that are presented and compared in this work.

The different categories that these works fall under are presented later in Section 4 and are summarised in Figure 1.

The figure organises the surveyed works based on: (i) part of the investigation studied, (ii) year of publication, (iii) data required for the proposed method, and (iv) optimisation in use. Different colours depict different optimisation methods, while the size of the circle depends on the number of works in the corresponding leaf of the tree. As shown in the presented dendrogram although 3 works included in this survey are published in 2005, the rest of the works have been published in the last 10 years with the majority been published in the last 5. Furthermore, as can be observed and will be discussed in Section 4, most of the earlier works used a guideline or workflow approach for the optimisation of the investigation and were focused on the reasoning and reporting phase. On the contrary, in the recent years most of the works utilise different non-manual methods of optimisation, most probably because of the increase of the data volume included in modern investigations, and focus on different parts of the technical analysis phase.

3.2. Comparison criteria

Finally, we have identified eleven evaluation and comparison criteria through the study of the literature, as well as the current challenges of cyber forensics on the industry.

These differ from the inclusion and selection criteria discussed in 3.1 as they are not used to select if a paper will be part of this survey, but rather to compare it with the rest of the review works. Moreover these will be used in Sections 4 and 5 for the comparison of the surveyed works, as well as the identification of research gaps and future directions:

1. **Investigation part:** refers to the part of the investigation that the surveyed work aims to support as presented in Section 2, e.g. collection, examination, correlation, and reporting. Please note that in many cases a work may be applicable in more than one investigation parts, e.g. examination and correlation as it aims to automatically extract evidence and identify the relations between them.
2. **Data:** refers to the data used by the method proposed in the work , e.g. network traffic, access or audit logs .If the proposed method is

independent of a specific data source and is applicable across the analysis of the whole incident, this criterion is assigned the value *incident*.

3. **Optimisation:** refers to the type of optimisation employed by the proposed method. For example, the value of this criterion may be gametheory, any heuristic approach or a guideline based approach.
4. **Modelling:** refers to the building block that is used by the proposed model or system, e.g., a graph-based system that used vulnerabilities or hosts as its nodes. This may not be applicable to all the surveyed papers.
5. **Evaluation:** refers to how the proposed method is evaluated. The evaluation method could be using a realistic simulated dataset or real data. In many cases the work may have not been evaluated but rather demonstrated.
6. **Knowledge base:** refers to the usage or not of an external well-recognised knowledge base, such as a threat intelligence knowledge base.
7. **Type of game:** this criterion applies only to works that use game theory and describes the type(s) of game used e.g. incomplete, zero sum etc.
8. **Parameters:** refers to the external parameters, e.g., cost or benefit, that are taken in consideration for the optimisation used by the proposed work.
9. **Parameter data:** refers to the type of data used for the aforementioned parameters during the evaluation section, i.e., real, realistic (through simulation), numerical (random) data.
10. **Anti-forensics:** this is a boolean criterion that is *true* when the proposed solution is formulated to be effective against Attackers that use anti-forensic techniques.
11. **Multi-stage:** this is a boolean criterion that is *true* only when the proposed solution explicitly takes in consideration the multi-stage nature of a cyber incident.

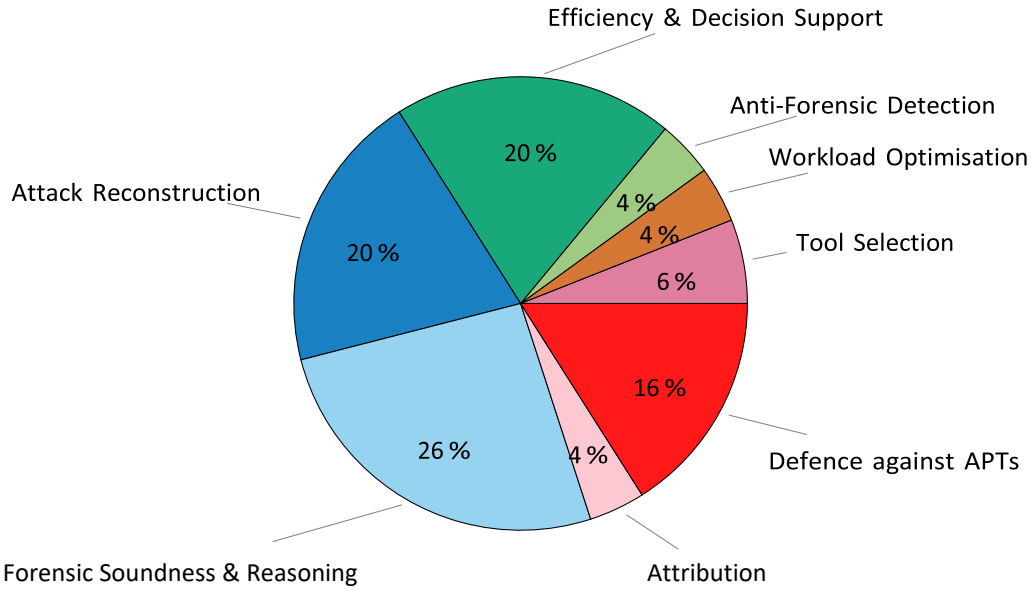
4. Literature Review

This section presents and compares a collection of papers from the literature that match the aforementioned collection criteria. Figure 2 presents a breakdown of the subcategories of papers surveyed in this work, while Tables 5 to 9 summarise the works against the aforementioned criteria.

We have divided the content of this subsection into the following categories based on the prime aim of each work, namely:

- **Tool Selection**: works that support the investigator on choosing the best tool to perform a certain task.
- **Workload Optimisation**: works that aim to optimise the use of limited resources of the investigator, such as the available time and computing power.
- **Anti-forensic Detection**: works that explicitly offer support against the use of anti-forensic techniques.

Figure 2: Breakdown of categories of the reviewed papers



- **Efficiency and Decision Support**: works that support the decision process of the investigator in various ways, technical or not, in order to increase the efficiency of the investigation thus decreasing the required resources.
- **Attack Reconstruction**: works that propose methods for the reconstruction of an attack, i.e., the identification of actions that belong in each individual phase and their connection if they belong to the same incident. This may be achieved either through an automated process or a reasoning framework.
- **Reasoning and Forensic Soundness**: works that aim to support and optimise the reasoning process of the investigator or the forensic soundness of the investigation.
- **Attribution**: works that focus on helping the investigator to attribute a cyber attack to a specific threat group or adversary.
- **Defence against APTs**: works that even though they have been proposed for defence against APTs, can also be utilised for cyber forensic investigations as they model multi-stage cyber incidents, APTs, their actions and goals.

Table 1: Comparative analysis of works for Tool Selection

Ref.	Data	Optimisation	Investigation part	Evaluation	Modelling	K.B.	A.F.	Multistage
[19]	cyber incident	game theory	examination	real data	tools	7	3 7	
[20]	cyber incident	game theory	examination	generated data	tools	7	7	7
[21]	cyber incident	game theory	examination	real data	tools	7	3 7	

4.1. Tool Selection

Hasanabadi et al. in [19] and[21] model the interaction between an Attacker who uses rootkits and an investigator who uses anti-rootkits as a non-zero-sum game and simulated as a Fictitious Play. They use a set of rootkits and antirootkits for Windows XP to extract characteristics for each software used to derive the benefit and cost values of each player's payoff function. The authors use the Nash Equilibrium (NE) of the game to identify future improvements that could enhance the performance of anti-rootkits.

The main difference between the two works is the performance improvement achieved in [21]. To overcome the re-simulation problem caused by the constant need of expansion of the action space, the authors propose the use of a memorybased mechanism. The mechanism takes as input the original strategy space as well as a list of strategies to be added. After solving the NE of the original game, it iterates through the list, adding one strategy at a time, finding the corresponding NE and observing the opponent's empirical frequencies.

Karabiyik and Karabiyik [20] propose a game theoretic model for the optimal selection of a tool during the file carving process of an investigation. Contrary to most of the literature, the Attacker is not simulated as a player, but instead the game is rather a cooperation between two investigators, who are working in the same case and can each select one tool to use during the investigation.

As summarised in Table 1, none of three works for tool selection take in consideration the multistage nature of modern incidents or utilise a knowledge base. While the former is expected based on the task they are aiming to support, the latter is a drawback due to high number of available tools and malware. Moreover, even though all three works utilise game theory to support the optimal tool selection by the investigator, [19] and [21] use a combination of benefit and cost parameters, while [20] use benefit, cost and effectiveness. Finally, all works have limitations on their evaluation. In [20] the authors generate disk images with hidden evidence and use two image carving tools, Photorec and Scalpel. The use of generated data for the disk images and the calculations of the parameters, and the limited number of carving tools decrease the applicability and significance of their results. On the other hand, while the authors in [19] and[21] use real rootkits and anti-rootkits for the evaluation, the quantity is still small and the chosen software could be more recent.

4.2. *Workload Optimisation*

Table 2: Comparative analysis of works for Workload Optimisation

Ref.	Data	Optimisation	Investigation part	Evaluation	Modelling	K.B.	A.F.	Multistage
[22]	cyber incident	game theory	examination	real & generated data	events	7	7	7
[23]	filesystem	workflow/ guidelines	collection examination	real data	N/A	7	7	7

Yan et al. [22] model the database auditing problem as a Stackelberg zerosum game between an auditor and an adversary. The authors assume different type of events that need to be audited using a limited budget, i.e., time, while considering the potential policy validations by strategic adversaries. The proposed method was shown to outperform traditional auditing policies. De Braekt et al. [23] propose a workflow management automation framework that aims to optimise the use of limited resources of hardware and software during the acquisition and preparation steps of forensic investigations. The authors propose three modules, each one of which consists of a workflow, which allows the automation of certain tasks during those steps so that more time can be spent on the analysis of evidence.

The study of the literature suggests (see Table 2) that the number of works that optimise the resources of investigations is limited. This does not align with the current industry needs, where companies are constantly trying to utilise their resources efficiently to meet service level agreements (SLAs) and decrease the respond time. Neither works utilise a knowledge base, which could decrease the active time required byt the investigator. Fortunately, both works present a complete evaluation using real data. The work in [22] uses real medical access logs and the Statlog (German Credit Data) dataset, while [23] is evaluated against a real human trafficking case. However, in the former the cost and benefit parameters were assumed, i.e., not generated or calculated based on real data. Thus, it is not clear how the variation of those parameters would affect the performance of the proposed policy.

4.3. *Anti-forensics Detection*

Stamm et al. [24], [25] model the interaction of an adversary who uses antiforensics to hide the manipulation of multimedia content and an investigator who uses forgery detectors as a zero-sum Nash game. By finding the NA of the proposed game, the authors identify the optimal trade-off strategy for each player. The adversary aims to balance the trade-off between the use of antiforensic techniques to hide evidence and the evidence generated by these techniques, while the investigator needs to balance the accuracy with which she

detects manipulated multimedia and the accuracy of detecting the use of antiforensics. Both works use a combination of real and generated data (Table 3).

Table 3: Comparative analysis of works for Anti-Forensic Detection

Ref.	Data	Optimisation	Investigation part	Evaluation	Modelling	K.B.	A.F.	Multistage
[25]	multimedia	game theory	examination	real & generated data	N/A	7	3	7
[24]	multimedia	game theory	examination	real & generated data	N/A	7	3	7

4.4. Efficiency and Decision Support

Horsman et al. [26] presents CBR-FT, a case reasoning-based method that aims to increase the efficiency of device triage. CBR-FT produces a list of file system paths with high probability of containing evidence. To do so it calculates a similarity rating of each path based on a knowledge base of past cases and input variables. Liu et al. in [27] propose the incorporation of attack graphs in forensic investigations to guide the decision process of the investigator. The authors also introduce to the graph the notion of anti-forensic steps, which may be taken by the adversary to hide malicious evidence, as parallel actions to the main nodes.

Later in [28] and [29], the authors presented a more optimised use of attack graphs in investigations, by proposing an algorithm that maps evidence graphs on attack graphs to identify missing evidence that the investigator needs to search for. In this way, a graph is not simply used as a manual guide for the investigator, but part of the reasoning process is automated to offer suggestions, which results in more efficient investigation.

Barrere et al. in [30] introduced core graphs, which are condensed versions of attack graphs. The authors propose an algorithm that transforms the structure of the original attack graph to allow for a more efficient exploration. The evaluation against three different simulated network topologies consisting of vulnerabilities, showed that the core graphs offer the same level of accuracy as traditional attack graphs, while significantly decreasing the network exploration rate. Thus, they have increased scalability and applicability in forensic investigations.

Nassif and Hruschka [31] propose the application of clustering techniques to tackle the problem of analysing large amounts of unstructured documents.

Table 4: Comparative analysis of works for Efficiency and Decision Support

Ref.	Data	Optimisation	Investigation part	Evaluation	Modelling	K.B.	A.F.	Multistage
[26]	filesystem	statistical	examination	real data	filesystem path	3	7	7
[27]	cyber incident	workflow/ guidelines	examination correlation	no evaluation	vulnera-bilities	7	3	3
[28]	cyber incident	heuristic	examination correlation	no evaluation	vulnera-bilities	7	7	3
[29]	cyber incident	heuristic	examination correlation	no evaluation	vulnera-bilities	7	7	3
[30]	cyber incident	heuristic	examination correlation	generated data	vulnera-bilities	7	7	3
[31]	documents	clustering	examination	real data	N/A	7	7	7
[32]	RAM	heuristic	examination	demonstration	N/A	7	7	7
[33]	cyber incident	bayesian networks	examination	demonstration	evidence	7	7	7
[34]	network & human factors	network theory	examination correlation	generated data	hosts	7	7	3
[35]	cyber incident	workflow/ guidelines	collection examination correlation	demonstration	N/A	7	7	3
[36]	cyber incident & human factors	machine learning	examination	generated data	N/A	7	7	3

Specifically, the authors test six clustering algorithms, namely K-means, Kmedoids, Single Link, Complete Link, Average Link, and CSPA. According to their results the Average Link and Complete Link algorithms outperform the rest with K-means and K-medoids following them if properly initialised.

Bohm et al. [32] utilise Visual Security Analytics (VSA) for decision support on Live Forensics. Specifically, they apply the Nested Blocks and Guidelines Model (NBGM) on the task of Memory Forensics to create a visual decision support system. The proposed system guides the investigator from the domain specific technical tasks and recovery of the relevant data to the required data operations and finally the encoding and visualisation.

Overill and Chow [33] use a Bayesian network to calculate a weight of each individual item of evidence of a case and thus support the decision process of the investigator through the triage stage. The authors create a Bayesian network that takes into consideration the cost-benefit ratio and the return-on-investment of

each evidential item, as well as a number of hypothesis and calculates a numerical estimate of weights for each of them.

Niu et al. [34] model an APT attack as a Targeted Complex Attack Network (TCAN) by combining dynamic attack graphs and network evolution. Specifically, the targeted network is modelled as a two-layered stochastic graph, where the first layer represents the human social interaction and the second layer the network topology. Nodes correspond to hosts, while edges represent the relationship between nodes. Finally, network theory is utilised to create rules that capture the time domain factor of the attacker.

Arshad et al. [35] propose FIMOSN, a model to semi-automate the forensic investigation of online social networks (OSN). FIMOSN is a formally defined process model that can be applied to the different stages of the investigation. It consists of two types of steps, automated, e.g. the formulation of the Information Extraction Zone, and manual, e.g. Data Examination. Contrary to other similar models FIMOSN includes the ability of hypothesis testing, as well as the ability to adapt to the size of the incident and its requirements.

Wei et al. [36] propose an unsupervised prediction framework for proactive forensic investigation of insider threats. The framework combines cascaded autoencoders (CAEs) and joint optimisation (CPJOS) for the detection of the anomalous activities through user behaviour features extracted via natural language processing.

Table 4 summarises the comparison of the works of this subsection. Most works of the works reviewed in this subsection directly target the investigation of multi-stage attacks, and many of them utilise some kind of graph to do so. However, 5 of these works use vulnerabilities or file paths as the building block for the modelling of the problem, which are limiting and out of date. Moreover, while [26] include a data-driven element, it has two main disadvantages: (i) it provides the same ranking regardless of the progress of the investigation, and (ii) the investigators need to populate the base at the end of each investigation.

Regarding the evaluation of the proposed method, 6 works ([27] [28], [29], [32], [33], [35]) present a simple demonstration or no evaluation or at all. Even though in some cases the demonstration includes real elements, such as the use of the Poweliks malware in [32] and a real BiTorrent case in [33], it is still does not allow the evaluation of the performance of the proposed methods.

4.5. Attack Reconstruction

Milajerdi et al. [37] propose HOLMES, a system that uses TTPs from MITRE ATT&CK and the typical APT life-cycle to correlate audit logs. Specifically, it uses tag propagation and predefined policies for backward and forward analysis to correlate audit logs to TTPs and then organise the TTPs in APT stages. In this way, HOLMES generated high-level graphs for the investigator and is able to reconstruct APT campaigns in a timely manner with high precision.

Table 5: Comparative analysis of works for Attack Reconstruction

Ref.	Data	Optimisation	Investigation part	Evaluation	Modelling	K.B.	A.F.	Multistage
[37]	audit logs	heuristic	examination correlation	real data	TTPs	7 7 3		
[38]	audit logs	graph theory	examination correlation	realistic data	process	7 7 3		
[39]	audit logs	graph theory	examination correlation	realistic data	process	7 7 3		
[40]	audit logs	graph theory	examination correlation	realistic data	process	7 7 3		
[41]	network	clustering	examination correlation	realistic data	N/A	7 7 3		
[42]	network	statistical	examination correlation	realistic data	hosts	7 7 3		
[43]	network	heuristic	examination correlation	realistic data	hosts	7 7 3		
[44]	network	heuristic	examination correlation	realistic data	hosts	7 7 3		
[45]	access logs	graph theory	examination correlation	real data	events	7 7		7
[46]	cyber incident	heuristic	examination	realistic data	events	7 7 3		
[47]	cyber incident	data compaction	examination correlation	realistic data	events	3 7 3		

Hossain et al. present SLEUTH [38], an OS agnostic system for real-time attack reconstruction from COTS audit logs. The proposed system constructs a memory efficient dependence event graph and uses policy-based approach to correlate events that belong to the same attack. The alerts are then passed to a backward analysis algorithm to identify source of the attack. During the evaluation SLEUTH reconstructed 70% of the attacks in real time.

Later in [39] the authors propose two methods, namely full dependence preservation (FD) and source dependence preservation (SD), that allow the reduction of the size of dependence graphs, while preserving the accuracy of the results of the forensic analysis. Specifically, they prove that the proposed methods preserve forward and backward reachability and do not affect the results of backward and forward forensic analysis. The evaluation showed that FD and SD can achieve an average of 7x and 9.2x reduction correspondingly, while the LCD reduction algorithm proposed in [48] achieved only a 1.8x.

Finally, Hossain et al. [40] proposed MORSE, a system that uses the tag attenuation and tag decay propagation techniques, to help the analyst reconstruct attacks from large amounts of data. Similarly to their previous works, MORSE uses dependence graphs to identify events of the same attack. Contrary to SLEUTH, MORSE applies a conservative tag propagation policy to suspicious subjects and a more lenient to benign subjects, which allows it to produce a smaller number of false positives in real time.

Ghafir et al. [41] propose MLAPT, a machine learning based system for reconstruction of multi-stage APT attacks. MLAPT consists of three phases: (i) threat detection, (ii) alert correlation and (ii) prediction. At the first stage a combination of rules and blacklists are used to detect TTPs used by APTs from network traffic. The alerts from stage are fed to the second stage for correlation of the campaigns based on the APT lifecycle model. In the final stage, a classifier based prediction module calculates the probabilistic significance of each alert with the aim of minimising their potential damage. Although MLAPT does not outperform all the past works, it is able to offer a true positive rate of 82%, while minimising the false positive rate to 5.4% and providing decision support to the analysts through its prediction module.

Marchetti et al. [42] propose a framework for the reconstruction of the data exfiltration APT stage from large volumes of network flows. Contrary to previous works, the authors do not only use network wide statistics to identify malicious hosts, but focus on individual hosts by comparing their behaviour both to past data and other hosts in the network. To this end, they calculate a score from a set of three features consisting of three suspiciousness sub-scores. The proposed framework is evaluated using data from a real network containing 10K hosts injected with realistic exfiltration attacks and is shown to be able to detect burst and low-and-slow exfiltration attempts.

Wang and Daniels [43],[44] present a hierarchical reasoning framework for network forensics that correlates hosts that belong to the same attack using evidence graphs of IDS alerts. Specifically, a local reasoning component uses Rule-Based Fuzzy Cognitive Maps (RBFCM) to identify suspicious hosts, while a global reasoning component correlates hosts that belong to the same attack. However, the evaluation of the prototype was against a very small, simulated testbed and was not compared with any other method or a baseline.

Studiawan et al. [45] use an improved version of MajorClust to detect anomalies from graph structures of access log data. They utilise a calculated score and a dynamic threshold for the identifications of potential violations, which are then presented to the analyst in a graph-based visualisation for further observation. The score is based on a number of factors such as number of events per cluster, frequency of event, etc. The proposed method achieved a 83% accuracy, 82% specificity and 70% sensitivity.

Bryant and Saiedian [46] present a kill-chain based attack reconstruction framework that allows the fusion of multi-sensor data with alerts generated by cyber security tools, such as a Security Information and Event Management (SIEM), that represent attack actions across a network of sensors. The framework is able to cope with missing or incomplete data or data from disconnected

sensors. The proposed framework is evaluated using the Logrhythm SIEM and a virtual corporate network consisting of several Windows servers, such as domain controllers and mail servers, and workstations against realistic multi-stage attack scenarios. As shown by the evaluation results, the SIEM instance that was using the proposed framework achieved better accuracy and less false negatives and false positives than the original SIEM instance.

Zhu et al. [47] propose an OS agnostic system for real-time data compaction to enhance the investigation of APT attacks. The proposed system is based on two strategies of data compaction: (i) one that maintains the Global Semantics (GS) and removes redundant events, and (ii) one that is based on suspicious semantics (SS) and performs context analysis on the remaining events. In both cases, the system depends on the predetermined semantic rules. Through the evaluation the authors showcase how the proposed method outperforms the dependence preservation (FD) and source dependence preservation (SD). However, the evaluation also shows that the proposed method is not able to correlate network traffic events between more than two hosts.

As summarised in Table 5, most of the works of this subsection use a single source of data, such as audit logs ([37], [38] , [39], [40]) or network events ([41], [42], [43], [44]). While these are great sources of evidence, the use of single data source limits the applicability and visibility of the proposed method. Moreover, a none of the methods utilise any kind of external knowledge, which in the case of [37], [38] , [39] and [40] could be used to remove the need for predefined policies and tags that need to be updated and maintained.

Finally, most of the works uses realistic or real data for the evaluation, which of course increases the significance of their results. A summary of the used dataset, along with the dataset for the rest of the reviewed works can be found in Table 10. However in some cases ([45], [43], [44]) the size of the dataset was quite small and thus not representative.

4.6. Reasoning and Forensic Soundness

Horsman [49] proposes the Framework for Reliable Experimental Design (FRED), which aims to contribute their findings in a reliable and robust way. FRED supports researchers through the design, development, implementation but also testing and validation of their work, taking in consideration the SO/IEC 17025 requirements. This will allow investigators to incorporate current robust and reliable research findings and tools to their analysis, while fully understanding their functionalities. Later, the author [50] presents the Digital Evidence Reporting and Decision Support (DERDS) framework. DERDS is a workflowbased framework that acts as a guide for practitioners to allow them to make robust and safe interpretations of the uncovered evidence and reliable inferences, assumptions or conclusions. The framework consists of three main pathways, one based on previous cases, one based on published works and one that relies on validation via testing. Although such a framework can increase the credibility of the reporting findings, it could be improved by taking in consideration technical parameters, such as the different costs associated with the investigative process.

Table 6: Comparative analysis of works for Forensic Soundness and Reasoning

Ref.	Data	Optimisation	Investigation part	Evaluation	Modelling	K.B.	A.F.	Multistage
[49]	cyber incident	workflow/ guidelines	correlation	demonstration	N/A	7	7	7
[50]	cyber incident	workflow/ guidelines	reporting	demonstration	N/A	7	7	7
[51]	cyber incident	workflow/ guidelines	correlation	demonstration	N/A	7	7	7
[52]	cyber incident	workflow/ guidelines	correlation	No evaluation	N/A	7	7	7
[53]	cyber incident	workflow/ guidelines	correlation	demonstration	N/A	7	3 7	
[54]	cyber incident	temporal logic	correlation	demonstration	events	7	7 3	
[55]	events	ontology	examination correlation	No evaluation	events	7	7	7
[56]	filesystem	temporal logic	correlation	demonstration	events	7	7	7
[57]	network	ontology	correlation	demonstration	network objects	7	7	7
[58]	IoT data	heuristic	correlation	demonstration	users & devices	3	7	7
[59]	cyber incident	workflow/ guidelines & reasoner	examination correlation	demonstration	evidence	7	7	7
[60]	cyber incident	workflow/ guidelines	correlation	demonstration	N/A	3	7 3	

Finally, Horsman [51] presents the Device Evaluation and Prioritisation Scoresheet (DEPS). DEPS is a structure that assist responders to identify and collect the digital devices on the scene but also to assume and record their priority and importance on the current case. In this way DEPS does not only support the responder but also allows the investigator who receives the collected devices to optimise the workflow based on the provided formal decision record.

Similarly, Beebe and Clark [52] propose a multi-tier hierarchical framework for reasoning guidance during forensic investigations. The framework consists of two tiers. The first tier is a high level abstraction of a six phase investigation process that includes the following phases: preparation, incident response, data collection, data analysis, presentation of findings and incident closure. The second tier includes objective based sub-phases for each phase of tier one. In this way,

the proposed framework can be used as a guide for both high level and technical parts of the investigation, while being technologically neutral and easily extensible by the user community.

Rekhis and Boudriga [53] formally model the investigation process and propose the use of an inference system to detect the presence of anti-forensics. Specifically, the authors model the system under investigation, its deployed security controls, potential attacks and the corresponding evidence using statebased logic. The authors present a small case study to demonstrate the use of the proposed system. The authors also propose investigation-based Temporal Logic of Actions (I-TLA), a formal temporal logic for testing of potential multi-stage attack scenarios on a system [54]. I-TLA is a high-level specification language that allows an investigator to model different sources of evidence related to a case, define attack scenarios and hypothesis test them on the evidence.

Turnbull et al. [55] combine a multi-ontology representation of low level data and SPARQLer, a forward chain rule-based reasoning system to allow investigators to derive high level abstractions and triage the data more efficiently. First, the authors use multiple ontologies to represent system and user events from different sources. Then SPARQLer is used to correlate the events using predefined rules and forward chain logic.

Soltani and Seno [56] propose the use of temporal logic to allow the investigator to perform hypothesis testing for event reconstruction purposes. First, the system under investigation is modelled as a transition system both on its normal state (before the incident) and its current state. Then, the investigator is able to express the properties and hypothesis she wishes to evaluate using modal μ-calculus. Finally, a checking algorithm explores the state of the model and checks if the defined properties are present. Furthermore, the authors address the state space explosion problem, which is a common drawback on such systems. The proposed framework is demonstrated using a FAT file system.

Saad and Traore [57] present a framework for network forensics that consists of a method ontology and a reasoning module. The ontology has different classes to represent the network forensics domain objects and subjects, as well as their relations and it supports deductive, inductive, and abductive reasoning operations. Unfortunately, as with other ontology-based works, the construction and maintenance tasks require a lot of time and effort and could benefit from automation. Nieto [58] presented JSON Users and Devices analysis (JUDAS), as system for the extraction and correlation of heterogeneous forensic data from IoT devices. JUDAS extracts unique objects, such as users or devices, from case files and correlates them with additional internal sources, e.g., network traffic, and external sources, i.e., OSINT. The results are then presented using interactive graphs to the analyst. The proposed tool is validated using data from Amazon's Alexa from the DFRW 2017/18 challenge.

Amato et al. [59] present an extensible framework that utilises text processing techniques that allows the investigator to semantically represent and visualise evidence from different sources of the same case. Specifically, the authors propose an architecture that receives as input any type of multimedia and text

based evidence and allows the investigator to perform reasoning based on predefined rules, as well as query the data and visualise them. In this way, the proposed framework increases the efficiency of the correlation of evidence and enhances the investigator's reasoning capabilities.

Hettema [60] proposes a framework, which focuses on the application of belief revision theory on the tasks of incident handling, attribution, and creation of threat intelligence. The author focuses on the steps of belief expansion, contraction and revision using rationality constrains and argues how their application on the aforementioned tasks can support the decision process of the analysts. The proposed framework is demonstrated using the HAFNIUM exchange attack.

The comparison of the aforementioned works is summarised in Table 6. The most evident common disadvantage of all of the aforementioned works is the lack of evaluation. Even though in some cases a demonstration against a real case is presented, such as in [57] and [60], it does not offer any insights regarding their performance and accuracy. Moreover, all of the proposed works could benefit from some form of threat intelligence as well as automation as currently most of them are guideline based except [54], [56] and [58].

4.7. Attribution

Karafili et al. [61] developed an argumentation-based reasoner (ABR) for attribution of cyber attacks to certain threat actors. The proposed system consists of two parts: the reasoning rules and the background knowledge. The reasoning rules, which are created manually using past incident reports, as well as the evidence can be technical (such as an IP address), operational (e.g., the required capabilities), or strategic (such as who performed an attack).The investigator inputs technical and social evidence to ABR and then is able to submit queries that are processed using the reasoning rules and the background knowledge.

Han et al. [62] propose a data-driven, case-based reasoning framework for decision support on attribution of website defacement cases. The proposed framework allows investigators to compare their defacement case with past cases and identify the most similar ones. It consists of three components: a data collection and pre-processing module, a similarity module, and a clustering module. Initially, database of past defacement cases is created through a crawler and a pre-processing operation, which contains a case vector of 7 features for each past case. Then similarity module utilises this database and the characteristics of the current case to calculate a similarity score, which is then used by the clustering module.

In summary, the literature on attribution is still limited (see Table 7), even though is one of the most challenging tasks in cyber security. The main drawback of [61] compared to [62] is the use of a very high number of predefined rules, which is both expensive in terms of time and manual work, but also limits it to only previously known events. On the contrary the latter work uses a knowledge bases, which however would benefit from the addition of more data sources to achieve a higher accuracy. Both works need further evaluation as they both present an demonstration of the proposed systems.

Table 7: Comparative analysis of works for Attribution

Ref.	Data	Optimisation	Investigation part	Evaluation	Modelling	K.B.	A.F.	Multistage
[61]	cyber incident	argumentation & abductive reasoning	correlation	demonstration	N/A	3	7	7
[62]	defacement	statistical	examination correlation	demonstration	N/A	3	7	7

4.8. Defence against APTs

Zhu and Rass [63], [64] divide an APT campaign in three major phases and model it as a multi-phase, multi-stage (MPMS) game. The first phase is a Bayesian phishing game with two types: a phisher and a legitimate sender. The second phase is modelled as sequential game of N stages, each of which takes place on a firewall of the infrastructure. Finally, the third stage that corresponds to the final goal of the Attacker is zero-sum matrix game.

Rass et al. [65] model the interaction between a stealthy Attacker and a defender as a Bayesian game on a predefined attack graph. The authors allow different Attacker types to capture the uncertainty of the defender regarding the entry point used by the attacker. Moreover, this is the only work that does not require the two players to play in a synchronous way using the same rate. Specifically, the Attacker is allowed to move in continuous time using a probabilistic distribution, while the defender is moving in fixed intervals that represent working hours.

Huang and Zhu [66] formalise a multi-stage game between a defender and a deceptive Attacker as a Bayesian game. The authors propose an online game aimed for proactive defence against APTs. The defender utilises observations of previous actions to develop a belief regarding the Attacker and use it to enhance the infrastructure's defences and mitigate future incidents.

Yang et al. [67] tackle the APT repair problem, i.e., the efficient repair of the potentially hijacked hosts in a timely manner by the victim organisation, while taking in consideration the strategies of the Attacker. To this end, they model it as a differential Nash game in order to capture the factor of time and allow the evolution of the state of the network that depends: (i) on the exfiltration moves of the Attacker through the compromised hosts and (ii) on the repair moves of the Defender.

Min et al. [68] model the interaction between an APT Attacker aiming to steal data from a cloud storage system and a Defender as a Colonel Blotto game, i.e., a simultaneous game of limited resources. Both players choose how to allocate their Central Processing Units (CPUs) amongst the available storage devices of a cloud system. Two variations of the game are presented, one that allows the Defender to identify the optimal policy without the knowledge of the APT attack model, and one the enables the acceleration of the learning speed in cases of large number of devices and CPUs.

As summarised in Table 8, all the works offering defence support against APTs use game theory to model and solve the problem. This is easily explained given the strategic nature of APT actors and their motivations. A comparison between these works as well as the rest of the game theoretic methods reviewed in this survey can be found in Table 9.

Table 8: Comparative analysis of works for Defence Against APTs

Ref.	Data	Optimisation	Investigation part	Evaluation	Modelling	K.B.	A.F.	Multistage
[63],[64]	cyber incident	game theory	examination correlation	generated data	firewalls	7	7 3	
[65]	cyber incident	game theory	examination correlation	demonstration	vulnera-bilities	7	7 3	
[66]	cyber incident	game theory	examination correlation	demonstration	N/A	7	7 3	
[67]	network	game theory	examination	generated data	hosts	7	7 3	
[68]	cyber incident	game theory	examination	generated data	CPUs	7	7	7
[69]	cyber incident	game theory	examination	generated data	TTPs	7	7 3	

Most of the reviewed works presented an evaluation, except from [66] and [65]. The former presented a demonstration on a case study using the Tennessee Eastman (TE) process, while the second presented a numerical example. However, even though [63] and [64] evaluated their proposed method they do not use any real data. They also do not offer any insight to the meaning of the payoff values they use for each game and how they can be acquired. In the same way, [67] use random numeric values for the Attacker's benefit and the organisation's loss during the evaluation of the proposed method.

Interestingly, [65] was the only work that allowed an asynchronous movement of the defender and the attacker. Finally, another limitation for observed in many of the reviewed works of this subsection is the way they model the problem. Specifically, [63] and [64] use firewalls for the second phase of their game that corresponds to the movement of the Attacker from the initial access node to the final goal, which is unrealistic and limiting. Similarly, [65] use vulnerabilities which outdated as will be explained in the next section.

Table 9: Comparative analysis of works using Game Theory

	GameType				
Ref	1	2	3	Parameters	Parameters Data
[19]	non-zero sum	Fictitious game	incomplete & perfect	cost benefit	real data

[20]	non-zero sum	Nash	cooperative	cost, benefit & effectiveness	generated data
[22]	zero sum	Stackelberg	Nash	cost benefit	numerical
[25]	zero sum	Nash	N/A	payoff	numerical
[24]	zero sum	Nash	N/A	payoff	numerical
[63], [64]	zero sum	Bayesian	set of games	payoff	numerical
[65]	zero sum	Bayesian	N/A	payoff	numerical
[66]	zero sum	Bayesian	perfect	payoff	numerical
[21]	non-zero sum	Fictitious game	incomplete & perfect	payoff	numerical
[67]	non-zero sum	Nash	differential	loss & benefit	numerical
[68]	constant sum	Colonel Blotto	N/A	cost benefit	numerical
[69]	non-zero sum	Bayesian	sequential	payoff	numerical

Table 10: Overview of datasets used for the evaluation of surveyed works

Ref	**Dataset**	**Availability**
[19]	Set of 9 rootkits	Public
[20]	Non realistic generated data of disk images	Non Public
[22]	VUMC medical access logs and Statlog (German Credit Data) dataset	Mixed
[25], [24]	36 standard video test sequences and simulated motion compensated videos	Non Public
[26]	20 real DF fraud cases	Non Public
[23]	Real human trafficking case	Non Public
[30]	Naggen attack graph generation tool using with a random walk-based mechanism.	
[37]	3rd Transparent Computing Engagement by DARPA	Public
[39]	2nd Transparent Computing Engagement by DARPA	Non Public
[38]	2nd Transparent Computing Engagement by DARPA	Non Public
[40]	TRACE and CADETS from 5th Transparent Computing Engagement by DARPA	Public
[44], [43]	Realistic dataset generated by authors	Non Public
[45]	Public and open Security Repository (SecRepo)	Public
[31]	5 real-world investigation cases by the Brazilian Federal Police Department	Non Public
[41]	Realistic dataset generated by authors	Non Public
[42]	Realistic dataset generated by authors	Non Public
[21]	Set of 10 rootkits	Public
[62]	Website defacement dataset	Public
[34]	Realistic dataset generated by authors	Non Public
[67]	Syntethic dataset using Pajek software	Non Public

[68]	Realistic dataset generated by authors	Non Public
[69]	Evaluation using the Tennessee Eastman process	N/A
[46]	Realistic dataset generated by authors	Non Public
[36]	KDD CUP 99, Thyroid and Arrhythmia	Public

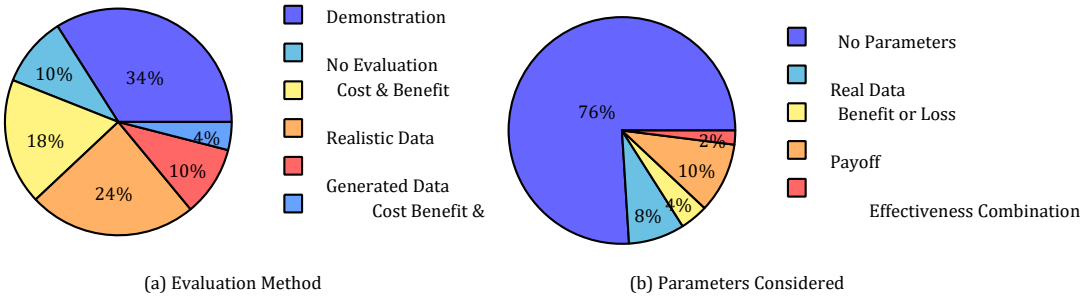
5. Discussion and Open Issues

In this section we draw useful insights regarding past works based on our literature review, which was presented in the previous section. We identify open issues and provide future directions that could enhance the decision support for cyber investigations and ultimately increase their efficiency and quality.

Open Issue 1: As can be observed in Figure 3a more than 54% (Demonstration, No Evaluation and Generated Data) of the surveyed papers did not provide an adequate evaluation of the proposed method. Namely an assessment of the effectiveness of the proposed method, not a simple demonstration of its use, using real or realistic data and potentially its comparison against a baseline or past work to allow better understanding of the results. Specifically, 34% simply demonstrated the proposed method, i.e., showcased how it can be applied but did not evaluate its effectiveness or performance in any way, while 10% provided no evaluation at all. Moreover, 10% used generated data for the evaluation, i.e., unrealistic data that are randomly generated and not representative of the problem. These facts prohibit the rest of the community from drawing robust conclusions regarding the significance and contribution of those works.

The reliability and evaluation of a work is essential in any field but is especially critical in forensics given that the outcome of an investigation is presented either to: (i) the decision makers of the targeted organisation, or (ii) a court of law, where the investigator needs to be able to support their findings [12]. One of the reasons for this lack of evaluation could be the sensitivity of the data related to cyber investigations, which may have been part of ongoing criminal investigations, or cannot be shared for privacy reasons related to individuals involved in the case or organisations and companies. In many cases, sharing data from incidents can reveal private information regarding the infrastructure of an organisation, technologies in use and current defence mechanisms.

This can also be observed from Table 10, where most of the utilised datasets are not publicly available, which prohibits the comparison between the experimental results of the reviewed works.

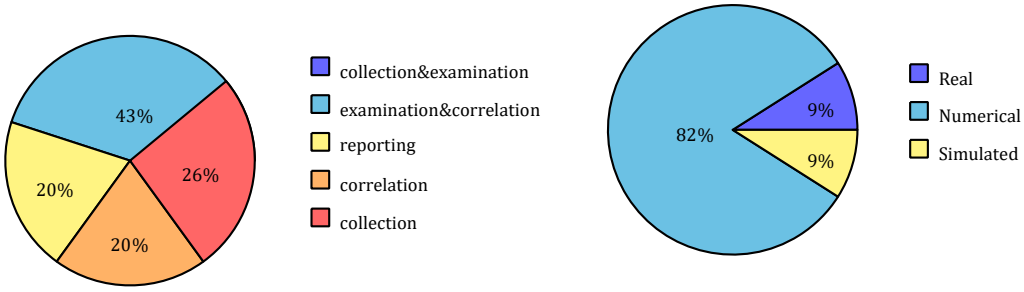


(a) Evaluation Method (b) Parameters Considered

Figure 3: Comparison of surveyed works based on the evaluation and parameters in use.

However, this obstacle can be overcome using the plethora of publicly available reports of past incidents, threat intelligence bases and tools that allow the creation of realistic datasets. Most of large cyber-security companies produce both individual incident reports, as well as quarterly or annual threat reports, where they present in detail the current TTPs used by Attackers on real attack campaigns, which they or their clients have faced. Moreover, organisations aggregate and process these reports, as well as other similar sources (blog posts, individual case reports, etc.) to create publicly available threat intelligence knowledge bases, such as the MITRE ATT&CK[2]. This information can be used by researchers to design and simulate realistic attack scenarios and datasets to develop, train and evaluate their works. In the same way, knowledge bases of past incidents can be used to produce probabilities or probability distributions needed for the proposed model. Naturally the same can be achieved using large quantities of past data, e.g., PCAPs, logs, etc., if they are available.

Finally, with regard to tools and malware used by Attackers, a large portion of them are: (i) open-source (or based on open-source versions), such as Empire[3], (ii) are OS native, such as PsExec [4], or (iii) can be found in online malware repositories, such as theZoo[5].



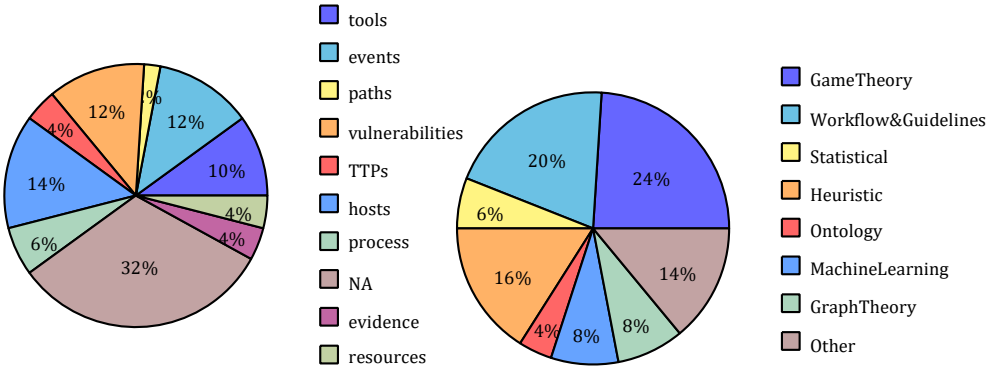
(a) Investigation Part (b)ParameterDataDuringEvaluation

[2] https://attack.mitre.org/
[3] https://www.powershellempire.com/
[4] https://attack.mitre.org/software/S0029/
[5] https://github.com/ytisf/theZoo [6]https://www.first.org/cvss/

Figure 4: Comparison of surveyed works based on the investigative phase on which they are applicable and type of evaluation data used for each optimisation parameter.

Similarly, 82% of the works that are based on game theory (Figure 4b) do not use real or realistic data for their parameters. On the contrary, they either generate numerical data or set specific values without providing any explanation or justification. Inevitably, the use of parameter values that do not reflect the reality limits the significance and credibility of the results. To overcome this, researchers could either gather data for their parameters through questionnaires and interviews with cyber-security practitioners, as well as utilise relevant public knowledge bases like the CVSS[6]. Moreover, in some cases (similarly to [19] and [21]), depending on the nature of those parameters, authors may be able to extract the required values from software simulations or inspection of an adequate amount of software (malicious or forensic).

Open Issue 2: The applicability and contribution of a work, as well as its produced results can be heavily decreased by the oversimplification of the modelling of the problem and the use of unrealistic assumptions. Almost half of the game theoretic approaches (Figure 3b) use a theoretical single payoff variable, without explaining what it represents in a real-world scenario and the meaning of its value.Instead, they could have used, for example, a payoff function that is formulated using parameters that represent actual characteristics of the problem, such as the cost (e.g., time or computational resources), benefit (e.g., impact of Attacker's actions on the system), etc.



(a) Modelling (b) Optimisation Method Deployed

Figure 5: Comparison of surveyed works based on their modelling and optimisation methods in use.

Similarly, only 38% of the game theoretic works (Figure 6) model the interaction between the Attacker and the Defender as incomplete games, which does not align with the current threat landscape, where the Defender is not able to observe all of the Attackers' past or future actions or be fully informed regarding their tactics. This is not an insignificant assumption as the *limited visibility* and *high uncertainty* of the Defender about the Attackers' actions, goals and resources are two of the main characteristics that can affect negatively her decision, decreasing the efficiency and quality of an investigation. In reality,

Investigators and Defenders in general have incomplete knowledge about their opponents' actions as they are not able to observe all their actions and they also do not know their specific capabilities or available resources. Another unrealistic assumption is the use of zero-sum games (46%, Figure 6) as they do not align with the motivations of players. The assumption that the payoffs of the players sum to zero, i.e., one player's gain is equivalent to the other player's loss, goes against the most simplistic model of an interaction between a Defender and an Attacker, where only one of the players may collect the benefit but they both suffer costs for their actions.

Open Issue 3: Only 2 works use TTPs for the modelling (Figure 7), while

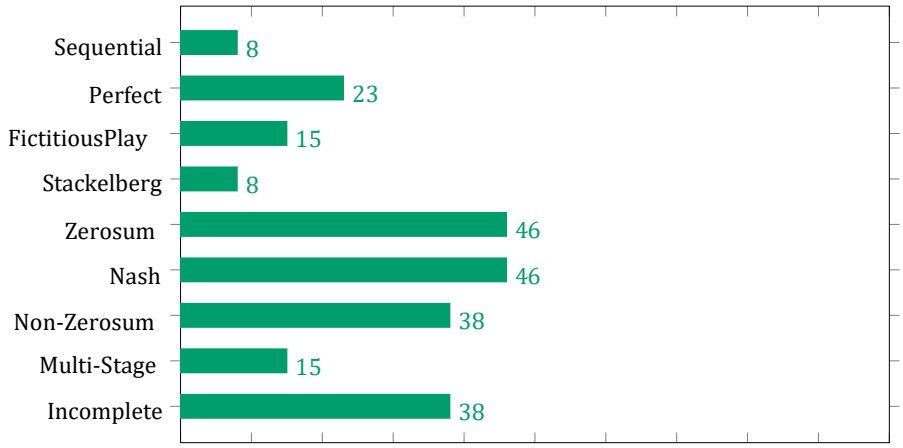


Figure 6: Comparison of surveyed works based game types in use (each work assumes more than one type) on a scale of %.

6 use vulnerabilities, which limits the applicability of the proposed work on a wide range of incidents, as not all of the Attacker's actions can be represented as vulnerabilities. This is because many actions do not include the exploitation of a vulnerability, but also because a vulnerability is not able to capture any behavioural or contextual information regarding the Attacker. Adversarial TTPs are more abstract than vulnerabilities or specific events and thus they are able to describe the adversarial behaviour, its motive and execution instead of simply a technical flaw that may be exploited. In this way, TTPs are applicable in a wider range of investigations and do not face the zero-day shortcoming that vulnerabilities do.

Open Issue 4: The inclusion of a diverse variety of data sources can significantly increase the visibility, applicability, and performance of a system. As shown in Figure 7 52% are not tied to a specific data source but are applicable, directly or indirectly, to the whole cyber incident under investigation. However, the rest of the works can only use a specific data type, host or network based, which adversely could decrease the visibility of the investigator and result in decreased accuracy, missed correlations or misinterpretations. A unique and potentially significant addition to the data sources is presented by [34] and [36],

which include human factor and behaviour analytics data respectively along with the computer related data. This approach could allow a better representation of the human element in the investigation and lead to increased performance.

Similarly, even though knowledge bases of past incidents are widely utilised in the cyber security industry, only 30% of the reviewed works use them to enrich their proposed models, methods, or frameworks (Figure 8c). Threat intelligence, i.e., data that is collected, processed, and analysed from past incidents, is able to represent a threat actor's motives, behaviour, and targets. Thus, its use can introduce to the methods a data-driven element that can significantly increase

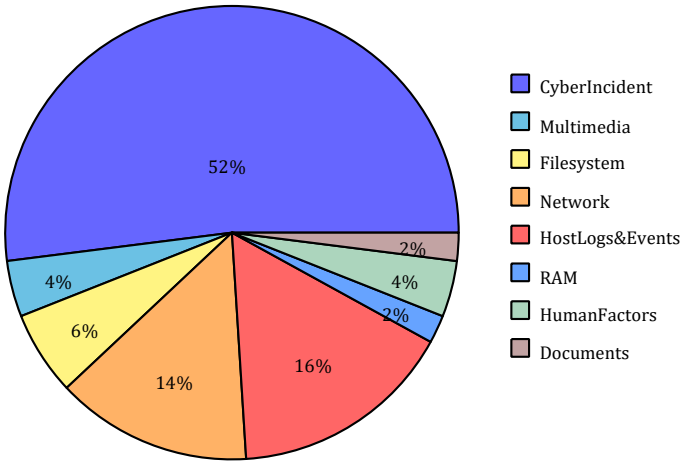
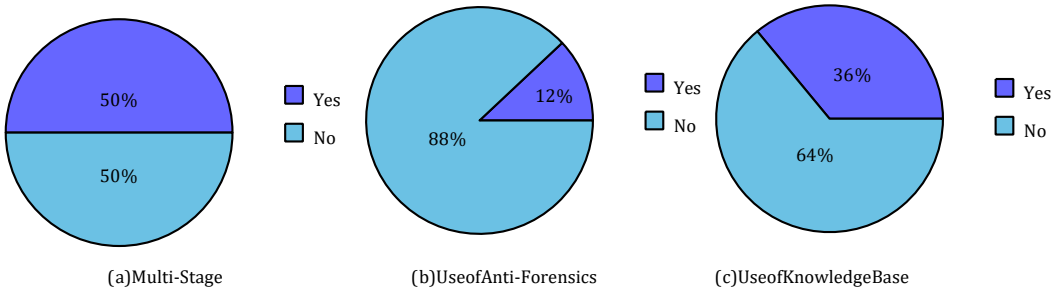


Figure 7: Comparison of surveyed works based on data in use.



(a)Multi-Stage (b)UseofAnti-Forensics (c)UseofKnowledgeBase

Figure 8: Comparison of surveyed works based on their consideration of multi-stage incidents and anti-forensics, and use of a knowledge base.

their performance. More importantly, the use of threat intelligence can provide a method and its user the ability to be up-to-date, in a fully automated way, about the current adversarial trends.

Open Issue 5: Only 12% (Figure 8b) of the works were aware of the potential anti-forensic capabilities of the Attacker and out of those two where specifically developed to tackle the anti-forensics problem. As the use of antiforensic or evasion techniques is on the rise [70], frameworks and systems that support the investigation of modern cyber incidents should be aware of such techniques and incorporate their existence to their modelling and development process. This is because anti-forensic techniques allow Attackers to conceal, delete or replace part of their traces. Thus, ignoring their existence may lead to wrong conclusions and interpretation of the evidence. This could be avoided by including on the proposed models or frameworks an element of uncertainty or trust level on the side of the Defender regarding the actions of the Attacker.

Open Issue 6: None of the survey works is implemented in a way that allows the feedback from the Investigator to be ingested and utilised for the readjustment of the output. The step-wise interaction between the proposed framework or method and the Investigator during the analysis, i.e., the production of suggestions by the framework to the Investigator and the submission of feedback by the Investigator to the framework, could allow the customisation of the produced output to increase its accuracy.

Open Issue 7: Finally, only one of the reviewed works included the element of time in their modelling, by using a differential Nash game [67]. However, time is an important factor for any investigation that can affect the rest of the variables, such as benefit and impact, of the investigation. For instance, the longer certain adversarial actions (e.g., active exfiltration channel) stay undisclosed the greater their impact on the infrastructure. Similarly, many types of evidence, such as IPs, are time sensitive and are no longer useful after a certain amount of time. Thus, the notion of time should be taken in consideration and potential methods that allow the definition of states and their transition, evolution or discount factors, which have been used in other security fields, could be applied in cyber forensics.

6. Conclusion

The increase of the sophistication and variety of TTPs used by strategic Attackers, such as APTs, has resulted in more challenging and time-consuming cyber forensic investigations. Thus, a need for decision support systems, which will increase the efficiency of those investigations and allow analysts to overcome problems like the ever-increasing variety of adversarial TTPs and the large amount of produced data, is evident.

In this paper, we presented a comprehensive review of works that aim to support cyber forensic investigations of multi-stage incidents on a practical or more abstract level through any manual or automated way. We compiled a list of evaluation criteria related to both the content of the works, as well as the requirements of modern cyber investigations.

Finally, we use them to compare the surveyed works and identify open issues that are not addressed by the current literature even though they align with the current threat landscape.

The most important issues highlighted by this work are: (i) inadequate evaluation and use of non realistic datasets, as well as (ii) oversimplified or outdated modelling by many of the past works in the field, which limits their applicability and significance.

References

[1] V. S. Harichandran, F. Breitinger, I. Baggili, A. Marrington, A cyber forensics needs analysis survey: Revisiting the domain's needs a decade later, Computers & Security 57 (2016) 1–13.

[2] A. Ahmad, J. Webb, K. C. Desouza, J. Boorman, Strategically-motivated advanced persistent threat: Definition, process, tactics and a disinformation model of counterattack, Computers & Security 86 (2019) 402–418.

[3] L. Martin, Cyber kill chain®, URL: http://cyber. lockheedmartin. com/hubfs/Gaining the Advantage Cyber Kill Chain. pdf (2014).

[4] P. Pols, J. van den Berg, The unified kill chain, CSA Thesis, Hague (2017) 1–104.

[5] C. Johnson, L. Badger, D. Waltermire, J. Snyder, C. Skorupka, et al., Guide to cyber threat information sharing, NIST special publication 800 (150) (2016).

[6] E. Casey, Handbook of digital forensics and investigation, Academic Press, 2009.

[7] J. R. Vacca, Computer and information security handbook, Newnes, 2012.

[8] K. Kent, S. Chevalier, T. Grance, H. Dang, Guide to integrating forensic techniques into incident response, NIST Special Publication 10 (14) (2006) 800–86.

[9] J. Kolodner, Case-based reasoning, Morgan Kaufmann, 2014.

[10] D. W. Gresty, D. Gan, G. Loukas, C. Ierotheou, Facilitating forensic examinations of multi-user computer environments through session-to-session analysis of internet history, Digital Investigation 16 (2016) S124–S133.

[11] T. Sharma, N. Tiwari, D. Kelkar, Study of difference between forward and backward reasoning, International Journal of Emerging Technology and Advanced Engineering 2 (10) (2012) 271–273.

[12] J. Williams, Acpo good practice guide for digital evidence, Metropolitan Police Service, Association of chief police officers, GB (2012).

[13] S. H. Saeed, H. L. Arash, A. A. Ghorbani, A survey and research challenges of anti-forensics: Evaluation of game-theoretic models in simulation of forensic agents' behaviour, Forensic Science International: Digital Investigation 35 (2020) 301024.

[14] K. Conlan, I. Baggili, F. Breitinger, Anti-forensics: Furthering digital forensic science through a new extended, granular taxonomy, Digital investigation 18 (2016) S66–S75.

[15] J. I. James, P. Gladyshev, A survey of digital forensic investigator decision processes and measurement of decisions based on enhanced preview, Digital Investigation 10 (2) (2013) 148–157.

[16] D. Quick, K.-K. R. Choo, Impacts of increasing volume of digital forensic data: A survey and future research challenges, Digital Investigation 11 (4) (2014) 273–294.

[17] A. Lemay, J. Calvet, F. Menet, J. M. Fernandez, Survey of publicly available reports on advanced persistent threat actors, Computers & Security 72 (2018) 26–59.

[18] A. Alshamrani, S. Myneni, A. Chowdhary, D. Huang, A survey on advanced persistent threats: Techniques, solutions, challenges, and research opportunities, IEEE Communications Surveys & Tutorials 21 (2) (2019) 1851–1877.

[19] S. S. Hasanabadi, A. H. Lashkari, A. A. Ghorbani, A game-theoretic defensive approach for forensic investigators against rootkits, Forensic Science International: Digital Investigation (2020) 200909.

[20] U. Karabiyik, T. Karabiyik, A game theoretic approach for digital forensic tool selection, Mathematics 8 (5) (2020) 774.

[21] S. S. Hasanabadi, A. H. Lashkari, A. A. Ghorbani, A memory-based gametheoretic defensive approach for digital forensic investigators, Forensic Science International: Digital Investigation 38 (2021) 301214.

[22] C. Yan, B. Li, Y. Vorobeychik, A. Laszka, D. Fabbri, B. Malin, Database audit workload prioritization via game theory, ACM Transactions on Privacy and Security (TOPS) 22 (3) (2019) 1–21.

[23] R. I. de Braekt, N.-A. Le-Khac, J. Farina, M. Scanlon, T. Kechadi, Increasing digital investigator availability through efficient workflow management and automation, in: 2016 4th International Symposium on Digital Forensic and Security (ISDFS), IEEE, 2016, pp. 68–73.

[24] M. C. Stamm, W. S. Lin, K. R. Liu, Forensics vs. anti-forensics: A decision and game theoretic framework, in: 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2012, pp. 1749–1752.

[25] M. C. Stamm, W. S. Lin, K. R. Liu, Temporal forensics and anti-forensics for motion compensated video, IEEE Transactions on Information Forensics and Security 7 (4) (2012) 1315–1329.

[26] G. Horsman, C. Laing, P. Vickers, A case-based reasoning method for locating evidence during digital forensic device triage, Decision Support Systems 61 (2014) 69–78.

[27] C. Liu, A. Singhal, D. Wijesekera, Using attack graphs in forensic examinations, in: 2012 Seventh International Conference on Availability, Reliability and Security, IEEE, 2012, pp. 596–603.

[28] C. Liu, A. Singhal, D. Wijesekera, Mapping evidence graphs to attack graphs, in: 2012 IEEE International Workshop on Information Forensics and Security (WIFS), IEEE, 2012, pp. 121–126.

[29] C. Liu, A. Singhal, D. Wijesekera, Creating integrated evidence graphs for network forensics, in: IFIP International Conference on Digital Forensics, Springer, 2013, pp. 227–241.

[30] M. Barr`ere, R. V. Steiner, R. Mohsen, E. C. Lupu, Tracking the bad guys: An efficient forensic methodology to trace multi-step attacks using core attack graphs, in: 2017 13th International Conference on Network and Service Management (CNSM), IEEE, 2017, pp. 1–7.

[31] L. F. da Cruz Nassif, E. R. Hruschka, Document clustering for forensic analysis: An approach for improving computer inspection, IEEE transactions on information forensics and security 8 (1) (2012) 46–54.

[32] F. B¨ohm, L. Englbrecht, G. Pernul, Designing a decision-support visualization for live digital forensic investigations, in: IFIP Annual Conference on Data and Applications Security and Privacy, Springer, 2020, pp. 223–240.

[33] R. Overill, K.-P. Chow, Measuring evidential weight in digital forensic investigations, in: IFIP International Conference on Digital Forensics, Springer, 2018, pp. 3–10.

[34] W. Niu, X. Zhang, G. Yang, R. Chen, D. Wang, Modeling attack process of advanced persistent threat using network evolution, IEICE TRANSACTIONS on Information and Systems 100 (10) (2017) 2275–2286.

[35] H. Arshad, E. Omlara, I. O. Abiodun, A. Aminu, A semi-automated forensic investigation model for online social networks, Computers & Security 97 (2020) 101946.

[36] Y. Wei, K.-P. Chow, S.-M. Yiu, Insider threat prediction based on unsupervised anomaly detection scheme for proactive forensic investigation, Forensic Science International: Digital Investigation 38 (2021) 301126.

[37] S. M. Milajerdi, R. Gjomemo, B. Eshete, R. Sekar, V. Venkatakrishnan, Holmes: real-time APT detection through correlation of suspicious information flows,

in: 2019 IEEE Symposium on Security and Privacy (SP), IEEE, 2019, pp. 1137–1152.

[38] M. N. Hossain, S. M. Milajerdi, J. Wang, B. Eshete, R. Gjomemo, R. Sekar, S. Stoller, V. Venkatakrishnan, SLEUTH: Real-time attack scenario reconstruction from COTS audit data, in: 26th USENIX Security Symposium (USENIX Security 17), 2017, pp. 487–504.

[39] M. N. Hossain, J. Wang, O. Weisse, R. Sekar, D. Genkin, B. He, S. D. Stoller, G. Fang, F. Piessens, E. Downing, et al., Dependence-preserving data compaction for scalable forensic analysis, in: 27th USENIX Security Symposium (USENIX Security 18), 2018, pp. 1723–1740.

[40] M. N. Hossain, S. Sheikhi, R. Sekar, Combating dependence explosion in forensic analysis using alternative tag propagation semantics, in: 2020 IEEE Symposium on Security and Privacy (SP), IEEE, 2020, pp. 1139–1155.

[41] I. Ghafir, M. Hammoudeh, V. Prenosil, L. Han, R. Hegarty, K. Rabie, F. J. Aparicio-Navarro, Detection of advanced persistent threat using machinelearning correlation analysis, Future Generation Computer Systems 89 (2018) 349–359.

[42] M. Marchetti, F. Pierazzi, M. Colajanni, A. Guido, Analysis of high volumes of network traffic for advanced persistent threat detection, Computer Networks 109 (2016) 127–141.

[43] W. Wang, T. E. Daniels, Building evidence graphs for network forensics analysis, in: 21st Annual Computer Security Applications Conference (ACSAC'05), IEEE, 2005, pp. 11–pp.

[44] W. Wang, T. E. Daniels, Network forensics analysis with evidence graphs (demo proposal), in: Proceedings of the digital forensic research workshop, 2005.

[45] H. Studiawan, C. Payne, F. Sohel, Graph clustering and anomaly detection of access control log for forensic purposes, Digital Investigation 21 (2017) 76–87.

[46] B. D. Bryant, H. Saiedian, A novel kill-chain framework for remote security log analysis with siem software, computers & security 67 (2017) 198–210.

[47] T. Zhu, J. Wang, L. Ruan, C. Xiong, J. Yu, Y. Li, Y. Chen, M. Lv, T. Chen, General, efficient, and real-time data compaction strategy for apt forensic analysis, IEEE Transactions on Information Forensics and Security (2021).

[48] Z. Xu, Z. Wu, Z. Li, K. Jee, J. Rhee, X. Xiao, F. Xu, H. Wang, G. Jiang, High fidelity data reduction for big data security dependency analyses, in: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, 2016, pp. 504–516.

[49] G. Horsman, Framework for reliable experimental design (fred): A research framework to ensure the dependable interpretation of digital data for digital forensics, Computers & Security 73 (2018) 294–306.

[50] G. Horsman, Formalising investigative decision making in digital forensics: Proposing the digital evidence reporting and decision support (DERDS) framework, Digital Investigation 28 (2019) 146–151.

[51] G. Horsman, Decision support for first responders and digital device prioritisation, Forensic Science International: Digital Investigation 38 (2021) 301219.

[52] N. L. Beebe, J. G. Clark, A hierarchical, objectives-based framework for the digital investigations process, Digital Investigation 2 (2) (2005) 147–167.

[53] S. Rekhis, N. Boudriga, A system for formal digital forensic investigation aware of anti-forensic attacks, IEEE transactions on information forensics and security 7 (2) (2011) 635–650.

[54] S. Rekhis, N. Boudriga, Logic-based approach for digital forensic investigation in communication networks, Computers & Security 30 (6-7) (2011) 376–396.

[55] B. Turnbull, S. Randhawa, Automated event and social network extraction from digital evidence sources with ontological mapping, Digital Investigation 13 (2015) 94–106.

[56] S. Soltani, S. A. H. Seno, A formal model for event reconstruction in digital forensic investigation, Digital Investigation 30 (2019) 148–160.

[57] S. Saad, I. Traore, Method ontology for intelligent network forensics analysis, in: 2010 Eighth International Conference on Privacy, Security and Trust, IEEE, 2010, pp. 7–14.

[58] A. Nieto, Becoming judas: Correlating users and devices during a digital investigation, IEEE Transactions on Information Forensics and Security 15 (2020) 3325–3334.

[59] F. Amato, G. Cozzolino, V. Moscato, F. Moscato, Analyse digital forensic evidences through a semantic-based methodology and nlp techniques, Future Generation Computer Systems 98 (2019) 297–307.

[60] H. Hettema, Rationality constraints in cyber defense: Incident handling, attribution and cyber threat intelligence, Computers & Security 109 (2021) 102396.

[61] E. Karafili, L. Wang, E. C. Lupu, An argumentation-based reasoner to assist digital investigation and attribution of cyber-attacks, Forensic Science International: Digital Investigation 32 (2020) 300925.

[62] M. L. Han, B. I. Kwak, H. K. Kim, Cbr-based decision support methodology for cybercrime investigation: Focused on the data-driven website defacement analysis, Security and Communication Networks 2019 (2019).

[63] Q. Zhu, S. Rass, Game theory meets network security: A tutorial, in: Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, 2018, pp. 2163–2165.

[64] Q. Zhu, S. Rass, On multi-phase and multi-stage game-theoretic modeling of advanced persistent threats, IEEE Access 6 (2018) 13958–13971.

[65] S. Rass, S. K¨onig, E. Panaousis, Cut-the-rope: a game of stealthy intrusion, in: International Conference on Decision and Game Theory for Security, Springer, 2019, pp. 404–416.

[66] L. Huang, Q. Zhu, Adaptive strategic cyber defense for advanced persistent threats in critical infrastructure networks, ACM SIGMETRICS Performance Evaluation Review 46 (2) (2019) 52–56.

[67] L.-X. Yang, P. Li, Y. Zhang, X. Yang, Y. Xiang, W. Zhou, Effective repair strategy against advanced persistent threat: A differential game approach, IEEE Transactions on Information Forensics and Security 14 (7) (2018) 1713–1728.

[68] M. Min, L. Xiao, C. Xie, M. Hajimirsadeghi, N. B. Mandayam, Defense against advanced persistent threats in dynamic cloud storage: A colonel blotto game approach, IEEE Internet of Things Journal 5 (6) (2018) 4250– 4261.

[69] L. Huang, Q. Zhu, A dynamic games approach to proactive defense strategies against advanced persistent threats in cyber-physical systems, Computers & Security 89 (2020) 101660.

[70] F. Mandiant, Special report: M-trends 2021 (2021).