Next-Generation Digital Forensics: Challenges and Future Paradigms

Reza Montasari
Department of Computing and Engineering
The University of Huddersfield
Huddersfield, U.K.
R.Montasari@hud.ac.uk

Richard Hill
Department of Computing and Engineering
The University of Huddersfield
Huddersfield, U.K.
R.Hill@hud.ac.uk

Abstract**—** In recent years, Information and Communications Technology (ICT) has rapidly advanced, bringing numerous benefits to the lives of many individuals and organisations. Technologies such as Internet of Things (IoT) solutions, Cloud-Based Services (CBSs), Cyber-Physical Systems (CPSs) and mobile devices have brought many benefits to technologically-advanced societies. As a result, commercial transactions and governmental services have rapidly grown, revolutionising the life styles of many individuals living in these societies. While technological advancements undoubtedly present many advantages, at the same time they pose new security threats. As a result, the number of cases that necessitate Digital Forensic Investigations (DFIs) are on the rise, culminating in the creation of a backlog of cases for law enforcement agencies (LEAs) worldwide. Therefore, it is of paramount importance that new research approaches be adopted to deal with these security threats. To this end, this paper evaluates the existing set of circumstances surrounding the field of Digital Forensics (DF). Our research study makes two important contributions to the field of DF. First, it analyses the most difficult technical challenges that need to be considered by both LEAs and Digital Forensic Experts (DFEs). Second, it proposes important specific future research directions, the undertaking of which can assist both LEAs and DFEs in adopting a new approach to combating cyber-attacks.

Keywords—*digital forensics, IoT forensics, cloud forensics, cybersecurity, digital investigation, encryption, anti-forensics*

I. Introduction

In recent years, we have witnessed rapid advancements in Information and Communication Technology (ICT) features. Technologies such as communication networks, mobile devices, Internet of Things (IoT) solutions, Cloud-Based Services (CBSs), Cyber-Physical Systems (CPSs) have brought many benefits to technologically-advanced societies [1, 2, 3]. As a result, commercial transactions and governmental services have rapidly grown, revolutionising the life styles of many individuals living in these societies. While technological advancements undoubtedly present many advantages, at the same time they pose new cybersecurity threats which have significant impacts on a variety of domains such as government systems, enterprises, ecommerce, online banking, and critical infrastructure. According to an official survey conducted by The Office for National Statistics [4], there were an estimated 3.6 million cases of fraud and two million computer misuse offences in a year. Although there is a variety of reasons for conducting cybercrimes, the motivation is often for financial gain. The fundamental issue associated with cybercrime consists of damage to reputation, monetary loss, in addition to impacts on the confidentiality, integrity and availability of data.

By exploiting technology, cybercriminals, for instance, will be able to turn IoT nodes into zombies (using malicious software), carry out distributed denial of service (DDoS) attacks (engineered through botnets), and create and distribute malware aimed at specific appliances (such as those affecting VoIP devices and smart vehicles) [1, 2], [5, 6, 7, 8, 9]. Other challenges resulting from such technological advancements include, but are not limited to: high volume of data, heterogeneous nature of digital devices, advanced hardware and software technologies, anti-forensic techniques, video and rich media, whole drive encryption, wireless, virtualisation, live response, distributed evidence, borderless cybercrime and dark web tools, lack of standardised tools and methods, usability and visualisation. The deployment of IP anonymity and the ease with which individuals can sign up for a cloud service with minimum information can also pose significant challenges in relation to identifying a perpetrator [2], [5], [8], [9, 10].

As a result, the number of cases that necessitate DFIs are on the rise, culminating in the creation of a backlog of cases for LEAs worldwide [11, 12]. Therefore, given the discussion above, it is of paramount importance that new research approaches be created to deal with the aforementioned security challenges. To this end, we evaluate the existing set of circumstances surrounding the field of DF. Our research study makes two important contributions to the field of DF. First, it analyses the most difficult mid and long-term challenges that need to be considered by both LEAs and DFEs. Second, it

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

proposes important specific future research directions, the undertaking of which can assist both LEAs and DFEs in adopting a new approach to combating cyber-attacks.

II. Challeneges

As the field of DF continues to evolve, its development is severely challenged by the growing popularity of digital devices and the heterogeneous hardware and software platforms being utilised [2], [13, 14]. For instance, the increasing variety of file formats and OSs hampers the development of standardised DF tools and processes [15]. Furthermore, the emergence of smartphones that increasingly utilise encryption renders the acquisition of digital evidence an intricate task. Additionally, advancements in cybercrime have culminated in the substantial challenge of business models, such as Crime as a Service (CaaS), which provides the attackers with easy access to the tools, programming frameworks, and services needed to conduct cyberattacks [2]. The following sub-sections analyse the key issues that pose significant challenges to the field of DF.

A. *Cloud Forensics*

The cloud computing paradigm presents many benefits both to the organisations and individuals. One of such advantages relates to the manner in which data is managed by the cloud infrastructure. For instance, data is spread between various data centres to improve performance and facilitate load-balancing, scalability, and deduplication features. Because of this, data requires an efficient indexing so that retrieval and optimisation performance can take place to evade duplication that often contributes to the expansion of storage needs. As a result, evidence left by adversaries is more difficult to eliminate since it can be copied in various locations, rendering the acquisition of evidence and its examination easier to perform.

However, despite its many benefits, cloud computing poses significant challenges to the LEAs and DFEs from a forensic perspective. These include, but are not limited to, problems associated with the absence of standardisation amongst different CSPs, varying levels of data security and their Service Level Agreements [5], [16, 17], multiple ownerships, tenancies, and jurisdictions. Moreover, the distributed nature of cloud computing services presents a variety of challenges to LEAs as data often resides in a number of different jurisdictions. In contrast with traditional DF in which data is held on a single device, within cloud environments data is often spread over multiple different nodes. As a result, LEAs need to rely on local laws to be able to conduct digital evidence acquisition [1], [7], [18]. Therefore, the discrepancy in the legal systems of different jurisdictions combined with the lack of cooperation between CSPs also poses significant challenges from a DF perspective.

In addition, existing DF models, frameworks, methodologies and tools are mainly intended for off-line investigations, designed on the premise that data storage under investigation is within the LEAs' control [19]. However, performing DFIs within a cloud environment is increasingly challenging as digital evidence is often short-lived and stored on media beyond the control of DFEs [1]. Anonymising tools and distributed data storage in cloud services also enable criminals to cover their malicious activities more easily. Furthermore, the use of features such as IP anonymity and the ease with which one can sign up for a cloud service with minimal information make it almost impossible to identify criminals in cloud environments [1], [7, 8]. Another challenge for DF is the availability of different models for delivering cloud services (CSs). Specifically, investigating the data of an infrastructure-as-a-service (IaaS) user can be done without too many restrictions, but in the case of customers using software-as-a-service (SaaS) resources, access to information might be minimal or entirely absent.

Last, but not least, accessing a software application through a cloud computing system often leaves traces of evidence in various places on the OS, such as registry entries or temporary Internet files. However, evidence is lost once the user has exited the virtual environment as virtualisation sanitises traces of leftover artefacts. As a result, virtualisation limits the traditional examination of the leftover artefacts, rendering digital evidence traditionally stored on hard drives potentially unrecoverable [20, 21]. Therefore, cloud-based forensic investigations pose significant challenges related to the identification and extraction of evidential artefacts.

B. *Network Forensics*

A Network Forensic Investigation (NFI) pertains to the acquisition, storage and examination of network traffic (encapsulated in network packets) generated by a host, an intermediate node, or the whole portion of a network in order to establish the source of a security attack. Network traffic objects that require analysis consist of protocols used, IP addresses, port numbers, timestamps, malicious packets, transferred files, user-agents, application server versions, and operating system versions, etc. This data can be acquired from different types of traffic.

Similar to any other sub-fields of DF, NF poses various challenges to DFEs and LEAs. One of the challenges concerns traffic data sniffing. Contingent on the network set up and security measures where the sniffer is installed, the tool is likely not to capture all intended traffic data. However, this challenge can be addressed by utilising a span port on network devices in various places in the network. Another challenge for NF is that an attacker might be able to encrypt the traffic by utilising a SSL VPN connection. In this case, although the address and port will still be visible to DFEs, data stream will not be available. Therefore, additional analysis will need to be carried out so as to establish penetrated data.

Another challenge is determining the source of an attack since an attacker may use a zombie machine, an intermediate host to perform an attack, or simply use a remote proxy server. The deployment of such methods by an attacker makes it very difficult for DFEs to determine the source of the attack. However, this can be remedied by examining each packet only in a basic manner in memory and storing only certain data for future examination. Notwithstanding that this approach necessitates less amounts of storage, it often requires a faster processor to be able to manage the incoming traffic. To capture and analyse evidential network data, DFEs need to use a number of commercial and open-source security applications such as tcpdump and windump. Additionally, ensuring the privacy of legitimate end users is another challenging factor in NF as all packet data including that of the end user is captured during an investigation.

C. Internet of Things (IoT) Forensics

The Internet of Things (IoT) which is supported by the cloud, big data and mobile computing often connects anything and everything 'online'. The IoT represents the interconnection of uniquely identifiable embedded computing devices within the current Internet infrastructure. Some IoT devices are ordinary items with built-in Internet connectivity, whereas some are sensing devices developed specifically with IoT in mind. The IoT covers technologies, such as: unmanned aerial vehicles (UAVs), smart swarms, the smart grid, smart buildings and home appliances, autonomous cyber-physical and cyber-biological systems, wearables, embedded digital items, machine to machine communications, RFID sensors, and context-aware computing, etc. Each of these technologies has become a specific domain on their own merit. With the new types of devices constantly emerging, the IoT has almost reached its uttermost evolution. With an estimated number of 50 billion devices that will be networked by 2020 [20, 21], it is estimated that there will be 10 connected IoT devices for every person worldwide [22].

IoT-connected devices offer many benefits both individually and collectively. For instance, connected sensors can help farmers to monitor their crops and cattle so as to improve production, efficiency and track the health of their herds. Intelligent health-connected devices can save or significantly improve patients' lives through wearable devices. For instance, the wearable device developed by Intel can track symptoms of Parkinson's disease patients by passively collecting 300 observations per second from each wearer, tracking various activities and symptoms [23, 24].

However, despite its many benefits, IoT-connected devices pose significant privacy and security challenges as these devices and systems collect significant personal data about individuals. As an example of privacy challenge, employers can use their employees' security access cards to track where they are in the building to determine how much time the employees spend in their office or in the kitchen. Another example relates to smart meters that can determine when one is home and what electronics they use. This data is shared with other devices and stored in databases by companies. In relation to the security challenges, due to the constant emergence of new and diverse devices with varied OSs as well as the different networks and related protocols, IoT produces a wider security attack surface than that created by cloud computing. Examples of cyberattacks that can be carried out on IoT devices include: intercepting and hacking into cardiac devices such as pacemakers and patient monitoring systems, launching DDoS attacks using compromised IoT devices, hacking or intercepting In-Vehicle Infotainment (IVI) systems, and hacking various CCTV and IP cameras. Therefore, security is of paramount importance for the secure and reliable operation of IoT-connected devices.

Although IoT uses the same monitoring requirements similar to those utilised by cloud computing, it poses more security challenges resulting from issues such volume, variety and velocity. Furthermore, DFIs of IoT devices can be even more difficult than those of cloud-based investigations as more complex procedures are needed for investigation of these devices.

IoT Forensics must involve identification and extraction of evidential artefacts from smart devices and sensors, hardware and software which facilitate a communication between smart devices and the external world (such as computers, mobile, IPS, IDS and firewalls), and also hardware and software which are outside of the network being investigated (such as cloud, social networks, ISPs and mobile network providers, virtual online identities and the Internet). However, extracting evidential artefacts from IoT devices in a forensically-sound manner and then analysing them tend to be a complex process, if not impossible, from a DF perspective. This is due to a variety of reasons, including: the different proprietary hardware and software, data formats, protocols and physical interfaces, spread of data across multiple devices and platforms, change, modification, loss and overwriting of data, and jurisdiction and SLA (when data is stored in a cloud). Thus, determining where data resides and how to acquire data can pose many challenges to DFEs.

For instance, the DF analysis of IoT devices used in a business or home environment can be challenging in relation to establishing whom data belongs to since digital artefacts might be shared or transmitted across multiple devices. In addition, due to the fact that IoT devices utilise proprietary formats for data and communication protocols, understanding the links between artifacts in both time and space can be very complex. Another challenge related to the DFI of IoT devices concerns the chain of custody. In civil or criminal trial, collecting evidence in a forensically sound manner and preserving chain of custody are of paramount importance. However, ownership and preservation of evidence in an IoT setting could be difficult and can have a negative effect on a court's understanding that the evidence acquired is reliable.

Furthermore, existing DF tools and methods used to investigate IoT devices are designed mainly for traditional DF examining conventional computing devices such as PCs, laptops and other storage media and their networks. For instance, the current methods utilised to extract data from IoT devices include: obtaining a flash memory image, acquiring a memory dump through Linux dd command or netcat, and extracting firmware data via JTAG and UART techniques. Moreover, protocols such as Telnet, SSH, Bluetooth and Wi-Fi are deployed to access and interact with IoT devices. Likewise, tools such as FTK, EnCase, Cellebrite, X-Ways Forensic and WinHex, etc. and internal utilities such as Linux dd command (for IoT devices with OSs such as embedded Linux) are used to extract and analyse data from IoT devices. However, the forensic investigation of IoT devices necessitates specialised handling procedures, techniques, and understanding of various OSs and file systems. Additionally, by using conventional Computer Forensic tools to conduct IoT Forensics, it would be highly unlikely to maintain a chain of custody, the adherence to which is required by the Association of Chief Police Officers [25], concerning the collection of digital evidence.

Therefore, to deal with the aforementioned challenges posed by IoT-connected devices, cloud cybersecurity will need to be reviewed since each IoT device produces data that is stored in the cloud. Cloud cybersecurity policies must be blended with IoT infrastructure so as to provide timely responses for suspicious activities [20]. They must be reviewed in relation to evidence identification, data integrity, preservation, and accessibility. CSPs will need to ensure the integrity of the digital evidence acquired from cloud computing components in order to facilitate an unbiased investigation process in establishing the root cause of the cyberattack in IoT. Therefore, as the IoT paradigm is further developed, it becomes necessary to develop adaptive processes, accredited tools and dynamic solutions tailored to the IoT model.

D. Big Data and Backlog of Digital Forensic Cases

Another key challenge that the field of DF is currently facing pertains to the substantial and continuing increase in the amount of data, i.e. big data – both structured and unstructured – acquired, stored and presented for forensic examination. This data is collected from a variety of sources such as digital devices, networks, cloud, IoT devices, social media, sensors or machine-to-machine data, etc. In particular, this challenge is relevant to live network analysis since DFEs are unlikely to acquire and store all the essential network traffic [2], [10]. This growth in data volume is the consequence of the ongoing advancement of storage technology such as growing storage capacity in devices and cloud storage services, and an increase in the number of devices seized per case. Consequently, this has resulted in an increase in the backlog of DF cases that are awaiting (often many months or years in some cases) investigations. The backlog of DF cases necessitating investigation has had a seriously adverse impact on the timeliness of criminal investigations and the legal process. The delays of up to 4 years in performing DFIs on seized digital devices have been reported to have significant effect on the timeliness of criminal investigations [5], [11], [26]. Due to such delays, some prosecutions have even been discharged in courts. This backlog of DF cases is predicted to increase due to the modern sources of evidence such as those of IoT devices and CBSs.

To address the aforementioned issues, i.e. the 3Vs of the big data, including: volume, variety and velocity, researchers have, in recent years, proposed various solutions ranging from data mining [27, 28, 29], data reduction and deduplication [27], [30, 31], triage [12], [32, 33, 34], increased processing power, distributed processing [35, 36], cross-drive analysis [31], artificial intelligence, and other advanced methods [30]. Despite the usefulness of these solutions, additional research studies are required to address the real-world relevance of the proposed methods to deal with the data volume that gravely challenges the field of DF. Therefore, it is of paramount importance to implement several practical infrastructural enhancements to the existing DF process. These augmentations should cover elements such as automation of device collection and examination, hardware-facilitated heterogeneous evidence processing, data visualisation, multi-device evidence and timeline resolution, data deduplication for storage and acquisition purposes, parallel or distributed investigations and process optimisation of existing techniques. Such enhancements should be integrated to assist both law enforcement and third-party providers of DF service to speed up the existing DF process. The implementation of the stated elements can significantly assist both new and augmented forensic processes.

E. Encryption

According to a survey conducted by the Forensic Focus [37], data encryption in addition to Cloud Forensics (discussed previously) are the most difficult challenges encountered by DFEs. Encryption is the fastest method used to prevent access to data held on a device. There exist numerous encryption methods that can be implemented on a system or its peripherals. Increase in storage devices has resulted in the creation of tools capable of encrypting the entire volume of a hard drive. Encryption can also be performed on an application, a folder, a cloud service, mobile devices, and data stored in a database or transmitted through email, etc. Concerning network-based data hiding, this can be facilitated through methods such as Virtual Private Network (VPN) tunnelling and the utilisation of proxy servers and terminal emulators. Regardless of data being stored in an unknown server in the cloud or on the perpetrator's computer's encrypted hard drive, encryption often makes it impossible for DFEs to acquire data essential for a DFI. Although such technologies are not unbeatable, they often necessitate large amount of time and luck to be bypassed [32], [38, 39].

Since many of the encryption schemes are implemented to resist brute-force attacks, it is, therefore, of paramount importance that researchers be able to design certain workarounds and exploits in order to be able to overcome encryption and acquire evidence from encrypted devices. Depending on the type of digital device involved, forensic challenges of encrypted devices differ. There are currently several exploits that DFEs can leverage to overcome encryption in DFIs. For instance, DFEs can decrypt a BitLocker volume by determining the correct Microsoft Account password. This can be achieved by recovering the matching escrow key directly from Microsoft Account. There are various tools and methods (the discussion of which is outside the scope of this paper) for retrieving the password. Another method of exploit used by the researchers is to conduct RAM Forensics (imaging the RAM) using a tool such as Belkasoft Live RAM Capturer and then draw out a binary decryption key from that RAM image. Using this method enables DFEs to bypass encryption and identify malware that is not placed in persistent storage. For instance, full-disk encryption on Windows desktop computers (BitLocker) can be attacked by imaging the RAM through a kernel-mode tool while the volume is mounted and examining that memory image to acquire the binary decryption key. This facilitates mounting BitLocker volumes in a short period of time.

However, the development of RAM Forensic tools as noted by Garfinkel [32] is more challenging than the creation of disk tools. Data stored in disks is persistent and intended to be read back in the future. However, data written to RAM can only be read by the running program. Garfinkel [32] argues that as a result there is less desire "for programmers to document data structures from one version of a program to another". Therefore, issues as such can complicate the tasks of tool developers.

F. *Limitations in DF Tools and Lack of Standardisation*

Existing DF tools and techniques are also limited in their functionality and are poorly appropriate to the task of identifying data which is “out-of-the-ordinary, out-of-place, or subtly modified” [32], [40]. Traditional DF tools, techniques and methods often lag behind new emerging technologies lacking adequate capabilities to address the resultant challenges presented by these technologies. Although current DF tools might be able to handle a case containing several terabytes of data, they are incapable of putting together terabytes of data into a succinct report. Furthermore, it is challenging to employ DF tools to recreate a unified timeline of past events or the activities of a culprit. Event and timeline reconstructions are often conducted manually during a given DFI. DF tools are also often slow to conduct data analysis. Furthermore, the task of creating digital documents which can be presented in courts has had an adverse effect on the production of DF methods that could process data that is not easily available [32], [41].

With regards to the lack of standardisation in DF, although researchers in the field have made some attempts to agree on formats, schema, and ontologies on DF artefacts, very little progress have been made, if any [15], [42, 43, 44]. This is while analysis of advanced cyber-attacks often necessitates concerted efforts to deal with the processing of complex data. In most cases such cooperation does not exist amongst DFEs and DF researchers alike. As a result, the diversity problem arising from the absence of standardised methods and guidelines to detect, acquire, store, examine, analyse and present digital evidence also pose significant challenges for DFIs [45, 46]. The lack of formal and generic Digital Forensic Investigation Process Models (DFIPMs) also contribute to the intricacy of acquiring and analysing digital evidence in a forensically sound manner [42]. Therefore, it is essential that DF community engage in more collaborations to create effective standard formats and abstractions.

III. Research Directions

A. *IoT Forensics*

The Identification, Acquisition and Analysis (main phases of a conventional DFI) of digital evidence in IoT environments pose significant challenges to LEAs and DFEs. In relation to the identification of a particular user’s data, it would be difficult for investigators to determine how to conduct search and seizure when the location and provenance of data (representing potential digital evidence) cannot be determined. One of the ways to address this challenge is to integrate the IoT device data into Building Information Modelling. Thus, the research community can consider this as a research opportunity to be explored.

With regards to the problems of extracting a specific user’s data in IoT devices, the volatility of evidence in these devices is more complex than the evidence volatility in traditional devices. In IoT environments, data might be held locally by an IoT device. In this case, the lifespan of the data is very short before it is overwritten or compressed. Furthermore, digital evidence (data) from an IoT device might be shifted and used by another IoT device (or a local network of IoT-connected devices), or it might be moved to the cloud for aggregation and processing. As a result, the transmission and aggregation of evidence poses significant challenges for maintaining the chain of evidence. To deal with this challenge, we propose the development of new investigation methods that can track and filter the transfer of data across IoT-connected devices as supported by (Hegarty et al., 2014). Such methods can then pave the way for the acquisition of data that have been altered or deleted. Therefore, the creation of such techniques should be considered as a new research opportunity for further exploration

In terms of the challenges of the analysis process, IoT devices produce large amounts of data which are stored in large-scale distributed cloud environments. If this data requires Digital Forensic analysis, first it needs to be imaged in order to adhere to the principles of ‘forensically-sound investigations’. However, from a technical point of view, the imaging of such data (representing potential digital evidence) using the existing conventional DFI procedures is not a feasible acquisition process. This is due to the scale, distribution and remote nature of such data, generated by IoT and stored in the cloud. As a result, new research studies must be conducted to develop new distributed analysis techniques that could facilitate the examination of this kind of data, which is generated by IoT devices and stored in the cloud.

Last, but not least, we suggest the revision of standards in traditional DFIs against which digital evidence in IoT is assessed in order to accommodate the evolving nature of digital evidence in IoT environments.

B. *Big Forensic Data*

Analysing big forensic data (BFD) in both a timely and a forensically-sound manner poses significant challenges to LEAs and DFEs. However, there are a number of research directions that researcher can adopt to address these challenges. One of the research areas on which the researchers can focus is to alter the conventional principles (that ‘all data’ must be extracted in a ‘strict’ forensically-sound manner) and procedures. To do so, similar to our proposed research direction in the previous sub-section, the techniques related to the main phases of DF process (i.e. Identification, Acquisition, and Analysis) could be adapted to the context of BFD. For instance, concerning the Acquisition Phase, proper triage procedures can be developed (i.e. through the visualisation both for low-level file system analysis and higher level content analysis) to enable investigators to conduct prioritisation of data when conventional ‘bit-by-bit’ forensic image is not possible due to the sheer volume of data. This denotes that by using the new triage procedures, investigators should be able to scan ‘all’ data but only extract the parts applicable to the investigation. In these scenarios, investigators might need to access original source of evidence [12]. If this is the case, they must be able to justify and document their actions so as to adhere to the Principle two of the ACPO Guidelines, “In circumstances where a person finds it necessary to access original data, that person must be competent to do so and be able to give evidence explaining the relevance and the implications of their actions.” [25].

Another research possibility to address the BFD is to develop new tools and techniques or adapt the existing ones. For instance, machine learning algorithms (MLAs) can be adapted for specific use in the unique context of DF for triage and analysis of big forensic data (such as disk images and network traffic dumps). Currently, there are only few DF tools that make use of MLAs for the triage and analysis of forensic data. On the other hand, the existing machine learning tools and libraries used in 'data science' are not fit or court-approved for use in the context of DF. Another example of adapting the existing tools to the context of DF is that of MapReduce, widely used in data science. As a research opportunity, MapReduce can be adapted to the task of processing the big data sets in DF with a parallel, distributed algorithm on a cluster. Similarly, Neural Networks can be extended to facilitate the complex patter recognition in various branches of DF such as Cloud Forensics and Network Forensics. Researchers can also build upon Natural Language Processing (NLP) techniques, for clustering and categorisation of unstructured DF data.

C. *Distributed Processing*

Although researchers have investigated Distributed Digital Forensics [30], [47], there is more scope for research in this area. The processing speed of existing tools is insufficient for the average case [13]. This is due to the fact that users have not been able to define clear performance requirements and that developers have not prioritised performance in accordance with reliability and accuracy [5], [13]. In their research paper, Roussev et al. [13] suggest a method for conducting data collection in such a way that facilities file-centric processing without disrupting optimal data throughput from the raw device [5]. Roussev et al.'s [13] assessment of core forensic processing functions in relation to processing rates demonstrate limits both in desktops and servers.

D. *Digital Forensic Data Abstraction*

In order for DFEs and researchers in the field to maintain DF capabilities, research studies in the field necessitates becoming more effective and being harmonised better. Due to the fact that DFEs often encounter a large amount of complex data, it is of paramount importance for them to create standards for data exchange. Furthermore, to enhance DF research, it is vital to implement standards for case data, data abstractions, and "composable models" for DF processing [32]. There are five broadly utilised abstractions including: disk images, packet capture files, files, file signatures and Extracted Named Entities. Due to the absence of standardised data abstractions and data formats, researchers are often made to implement more parts of a system prior to being able to create initial results. As a result, this hinders their progress. Therefore, new abstractions are needed to be developed in order to represent and compute with large amount of data. For instance, the researchers in the field can consider implementing the followings [32], [48]:

- Signature metrics for representing parts of files or whole files,
- File metadata JPEG EXIF information or geographical information,
- File system metadata and the physical location of files in a disk image,
- Application profiles, the Windows Registry or Macintosh plist information related to an application, document signatures, and network traffic signatures,
- User profiles, and
- Internet and social network information associated with the user, e.g. the acquisition of accounts accessed by the user, or user's Internet "imprint" or "footprint".

E. *Digital Forensics as a Service (DFaaS)*

Digital Forensics as a Service (DFaaS) is an extension of the traditional DF process. DFaaS can be used to reduce the backlog of DF cases. DFaaS solution can address issues such as the storage, automation, investigators' queries in the cases in which they are responsible. Furthermore, it facilitates efficient resource management, allows DFEs detectives to query data directly and enables easier teamwork amongst DFEs [5], [49]. Although DFaaS already provides multiple benefits, there are still many enhancements that can be made to the existing model in order to accelerate the existing process [5]. For instance, such improvements can be made in relation to DFaaS' functionality, indexing capabilities and identification of incriminating evidence during the Collection Phase in a DFI process [49]. However, it should be noted that DFaaS is not devoid of drawbacks, one of which pertains to latency concerning the online platform. Furthermore, DFaaS relies on the upload bandwidth available during the physical storage of data acquired through the Collection Phase in a DFI process.

F. *HPC and Parallel Processing*

The benefits of HPC should be considered to decrease computation time and the time needed by the users. HPC methods, which leverage a degree of parallelism, have not been adequately investigated by researchers in the field of DF. HPC methods and hardware could be used for various purposes such as accelerating each phase in a Digital Forensic Investigation Process following the Collection Stage, i.e., Storage, Examination, Even Reconstruction, and Presentation and Reporting Ps, etc. and reporting.

G. *Development of New Tools*

By default, the existing DFI tools are designed to run on the perpetrator's device. However, these tools provide restricted ability to examine complex cyberspace such as cloud sources [2], [32]. Therefore, many of DFIs tools are inappropriate to discover anomalies in an automatic manner [2]. As a result, one of the key problems that need to be addressed as future research relates to the development of new tools and methods to examine the volume of data and provide potential digital clue to the DFEs for additional examination. However, the design and implementation of such tools and techniques are a complex task due to the absence of standardisation and computational requirements. Auspiciously, DFI can take advantage of the element of cloud computing, for example, to reduce the most challenging processes of a DFI, such as log examination, data reduction, indexing and carving.

IV. Conclusion

The field of DF is facing various challenges that are often difficult to overcome. As the new technologies are constantly being developed, DFEs are presented with numerous challenges that can have substantial socioeconomic impact on both global enterprises and individuals [2], [6], [10]. Evidential data is no longer restricted to a single host but instead spread between different or virtual locations, including: online social networks, cloud resources, and personal network–attached storage devices. Furthermore, advances in technology and propagation of innovative services have led to a significant rise in the complexity of DFIs that DFEs must manage [2]. Hence, to mitigate these challenges, worldwide collaboration among LEAs, academic institutions and corporates becomes of paramount importance. Without a clear plan to facilitate research efforts that extend one another, forensic research will lag behind, tools will become outdated, and law enforcements' products will be incapable of relying on the results of DF analysis [32]. Thus, the aforementioned entities will need to converge regularly to discuss the future of the discipline and work out how to address the challenging aspects of the field. Likewise, more skills, tools and time are required to reconstruct digital evidence in a forensically sound manner. We believe that the future research directions outlined in this paper can have a positive impact on further research in the field of DF.

References

[1] Montasari, R. (2017, a). An Overview of Cloud Forensics Strategy: Capabilities, Challenges, and Opportunities. In Strategic Engineering for Cloud Computing and Big Data Analytics, pp. 189-205. Springer, Cham.

[2] Caviglione, L., Wendzel, S. and Mazurczyk, W. (2017). The Future of Digital Forensics: Challenges and the Road Ahead. IEEE Security & Privacy, (6), pp.12-17.

[3] Pichan, A., Lazarescu, M. and Soh, S.T. (2015). Cloud forensics: Technical challenges, solutions and comparative analysis. Digital Investigation, 13, pp.38-57.

[4] BBC. (2017). 'Cybercrime and fraud scale revealed in annual figures'. Available at:

https://www.bbc.co.uk/news/uk-38675683 (Accessed: 21st September 2018).

[5] Lillis, D., Becker, B., O'Sullivan, T. and Scanlon, M. (2016). Current challenges and future research areas for digital forensic investigation. arXiv preprint arXiv:1604.03850.

[6] Jang-Jaccard, J. and Nepal, S. (2014). A survey of emerging threats in cybersecurity. Journal of Computer and System Sciences. 80(5), pp.973-993.

[7] Ruan, K., Carthy, J., Kechadi, T. and Baggili, I. (2013). Cloud forensics definitions and critical criteria for cloud forensic capability: An overview of survey results. Digital Investigation, 10(1), pp.34-43.

[8] Chen, G., Du, Y., Qin, P. and Du, J. (2012). Suggestions to Digital Forensics in Cloud computing ERA. The 3rd IEEE International Conference on Network Infrastructure and Digital Content (IC-NIDC), pp. 540-544.

[9] Ruan, K., Carthy, J., Kechadi, T. and Crosbie, M. (2011). 'Cloud Forensics'. International Conference on Digital Forensics, Springer Berlin Heidelberg, pp. 35-46.

[10] Cameron, L. (2018). 'Future of Digital Forensics Faces Six Security Challenges in Fighting Borderless Cybercrime and Dark Web Tools'. Available at:

https://publications.computer.org/security-and-privacy/tag/dark-web/

(Accessed: 19th September 2018).

[11] Montasari, R. (2016, a). The Comprehensive Digital Forensic Investigation Process Model (CDFIPM) for Digital Forensic Practice, PhD Thesis.

[12] Montasari, R. (2016, b). Formal Two Stage Triage Process Model (FTSTPM) for Digital Forensic Practice. International Journal of Computer Science and Security (IJCSS). 10(2), pp.69-87.

[13] Roussev, V., Quates, C. and Martell, R. (2013). Real-time Digital Forensics and Triage. Digital Investigation, 10(2), pp.158-167.

[14] Raghavan, S. (2013). Digital forensic research: current state of the art. CSI Transactions on ICT, 1(1), pp.91-114.

[15] Montasari, R. (2018). Testing the Comprehensive Digital Forensic Investigation Process Model (the CDFIPM). In Technology for Smart Futures, pp. 303-327. Springer, Cham.

[16] Morioka, E. and Sharbaf, M.S. (2015). Cloud computing: Digital forensic solutions. The 12th IEEE International Conference on Information Technology-New Generations (ITNG), pp. 589-594.

[17] Almulla, S., Iraqi, Y. and Jones, A. (2013). Cloud forensics: A research perspective. The 9th IEEE international conference on Innovations in information technology (IIT), pp. 66-71.

[18] Simou, S., Kalloniatis, C., Kavakli, E. and Gritzalis, S. (2014). Cloud forensics solutions: A review. International Conference on Advanced Information Systems Engineering, pp. 299-309. Springer, Cham.

[19] Grispos, G., Storer, T. and Glisson, W.B. (2012). Calm before the storm: The challenges of cloud computing in digital forensics. International Journal of Digital Crime and Forensics (IJDCF), 4(2), pp.28-48.

[20] MacDermott, A., Baker, T. and Shi, Q. (2018). IoT Forensics: Challenges for The IoA Era. The 9th IEEE IFIP International Conference on New Technologies, Mobility and Security (NTMS), pp. 1-5.

[21] Taylor, M., Haggerty, J., Gresty, D. and Lamb, D. (2011). Forensic investigation of cloud computing systems. Network Security, 2011(3), pp.4-10.

[22] Bojanova, I and Voas, J. (2015). 'Securing the Internet of Anything (IoA)'. Available at:

https://www.computer.org/web/computingnow/archive/securing-the-internet-of-anything-november-2015 (Accessed: 20th September 2018).

[23] Kobie, N. (2015). 'What is the internet of things?'. Available at:

https://www.theguardian.com/technology/2015/may/06/what-is-the-internet-of-things-google

(Accessed: 20th September 2018).

[24] McCallion, J. (2014). 'Parkinson's disease to be tracked by wearables'. Available at:

http://www.alphr.com/news/390259/parkinsons-disease-to-be-tracked-by-wearables

(Accessed: 20th September 2018).

[25] ACPO. (2012). 'ACPO Good Practice Guide ACPO Good Practice Guide for Digital Evidence'. Available at:

http://library.college.police.uk/docs/acpo/digital-evidence-2012.pdf (Accessed: 21st September 2018).

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

[27] Quick, D. and Choo, K.K.R. (2014, b). Data reduction and data mining framework for digital forensic evidence: storage, intelligence, review and archive.

[28] Beebe, N. and Clark, J. (2005). Dealing with terabyte data sets in digital investigations. IFIP International Conference on Digital Forensics, pp. 3-16. Springer, Boston, MA.

[29] Palmer, G. (2001). A road map for digital forensic research. First Digital Forensic Research Workshop, pp. 27-30, Utica, New York.

[30] Beebe, N. (2009). Digital forensic research: The good, the bad and the unaddressed. IFIP International Conference on Digital Forensics, pp. 17-36. Springer, Berlin, Heidelberg.

[31] Garfinkel, S.L. (2006). Forensic feature extraction and cross-drive analysis. Digital Investigation, 3, pp.71-81.

[32] Garfinkel, S.L. (2010). Digital forensics research: The next 10 years. Digital Investigation. 7 (Supplement). pp. S64-S73.

[33] Mislan, R.P., Casey, E. and Kessler, G.C. (2010). The growing need for on-scene triage of mobile devices. Digital Investigation, 6(3-4), pp.112-124.

[34] Casey, E., Ferraro, M. and Nguyen, L. (2009). Investigation delayed is justice denied: proposals for expediting forensic examinations of digital evidence. Journal of Forensic Sciences, 54(6), pp.1353-1364.

[35] Richard III, G.G. and Roussev, V. (2006). Next-generation digital forensics. Communications of the ACM, 49(2), pp.76-80.

[36] Roussev, V. and Richard III, G.G. (2004). Breaking the performance wall: The case for distributed digital forensics. Proceedings of the digital forensics research workshop, 94, pp. 1-16.

[37] Forensic Focus. (2016). Current Challenges in Digital Forensics. Available at:

https://articles.forensicfocus.com/2016/05/11/current-challenges-in-digital-forensics/

(Accessed: 19th September 2018).

[38] Grispos, G., Glisson, W.B. and Storer, T. (2013). Using smartphones as a proxy for forensic evidence contained in cloud storage services. The 46th IEEE Hawaii International Conference on System Sciences (HICSS), pp. 4910-4919.

[39] Casey, E. and Stellatos, G.J. (2008). The impact of full disk encryption on digital forensics. ACM SIGOPS Operating Systems Review. 42(3), pp.93-98.

[40] Scanlon, M. (2016). Battling the digital forensic backlog through data deduplication. The 6th IEEE International Conference on Innovative Computing Technology (INTECH). pp.10-14.

[41] Sencar, H.T. and Memon, N. (2009). Identification and recovery of JPEG files with missing fragments. Digital Investigation, 6, pp. S88-S98.

[42] Montasari, R. (2017, b). A Standardised Data Acquisition Process Model for Digital Forensic Investigations. International Journal of Information and Computer Security, 9(3), pp.229-249.

[43] Montasari, R. (2016, c). A Comprehensive Digital Forensic Investigation Process Model. International Journal of Electronic Security and Digital Forensics, 8(4), pp.285-302.

[44] Montasari, R., Peltola, P. and Evans, D. (2015). Integrated Computer Forensics Investigation Process Model (ICFIPM) for Computer Crime Investigations. International Conference on Global Security, Safety, and Sustainability, pp. 83-95. Springer, Cham.

[45] Montasari, R. (2016, d). An Ad Hoc Detailed Review of Digital Forensic Investigation Process Models. International Journal of Electronic Security and Digital Forensics, 8(3), pp.205-223.

[46] Montasari, R. (2016, e). Review and Assessment of the Existing Digital Forensic Investigation Process Models. International Journal of Computer Applications, 147(7), pp. 1-9.

[47] Garfinkel, S., Farrell, P., Roussev, V. and Dinolt, G. (2009). Bringing science to digital forensics with standardized forensic corpora. Digital Investigation, 6, pp. S2-S11.

[48] Garfinkel, S. and Cox, D. (2009). Finding and archiving the internet footprint. Naval Postgraduate School Monterey CA.

[49] Van Baar, R.B., Van Beek, H.M.A. and van Eijk, E.J. (2014). Digital Forensics as a Service: A game changer. Digital Investigation, 11, pp. S54-S62.