CURRENT CHALLENGES AND FUTURE RESEARCH AREAS FOR DIGITAL FORENSIC INVESTIGATION

David Lillis, Brett A. Becker, Tadhg O'Sullivan and Mark Scanlon

School of Computer Science,
University College Dublin, Ireland
{david.lillis, brett.becker, t.osullivan, mark.scanlon}@ucd.ie

ABSTRACT

Given the ever-increasing prevalence of technology in modern life, there is a corresponding increase in the likelihood of digital devices being pertinent to a criminal investigation or civil litigation. As a direct consequence, the number of investigations requiring digital forensic expertise is resulting in huge digital evidence backlogs being encountered by law enforcement agencies throughout the world. It can be anticipated that the number of cases requiring digital forensic analysis will greatly increase in the future. It is also likely that each case will require the analysis of an increasing number of devices including computers, smartphones, tablets, cloud-based services, Internet of Things devices, wearables, etc. The variety of new digital evidence sources poses new and challenging problems for the digital investigator from an identification, acquisition, storage, and analysis perspective. This paper explores the current challenges contributing to the backlog in digital forensics from a technical standpoint and outlines a number of future research topics that could greatly contribute to a more efficient digital forensic process.

Keywords: Digital Evidence Backlog, Digital Forensic Challenges, Future Research Topics

1. INTRODUCTION

The early 21st century has seen a dramatic increase in new and ever-evolving technologies available to consumers and industry alike. Generally, the consumer-level user base is now more adept and knowledgeable about what technologies they employ in their day-to-day lives. The number of cases where digital evidence is relevant to an investigation is ever-increasing and it is envisioned that the existing backlog for law enforcement will balloon in the coming years as the prevalence of digital devices increases. It is for these reasons that it is important to take stock of the current state of affairs in the field of digital forensics. Cloud-based services, Internet of Things devices, anti-forensic techniques, distributed and high capacity storage, and the sheer volume and heterogeneity of pertinent devices pose new and challenging problems for the acquisition, storage and analysis of this digital evidence.

Due to the sheer volume of data to be acquired, stored, analysed, and reported, combined with the level of expertise necessary to ensure the court admissibility of the resultant evidence, it was inevitable that a significant backlog in cases awaiting analysis would occur (Hitchcock et al., 2016). Three

particular aspects have contributed to this backlog (Quick and Choo, 2014):

1. An increase in the number of devices that are seized for analysis per case.
2. The number of cases whereby digital evidence is deemed pertinent is ever-increasing.
3. The volume of potentially evidence-rich data stored on each item seized is also increasing.

This backlog is having a significant impact on the ideal legal process. According to a report by the Garda Síochána Inspectorate [2015] (Irish National Police), delays of up to four years in conducting digital forensic investigations on seized devices have "seriously impacted on the timeliness of criminal investigations" in recent years. In some cases, these delays have resulted in prosecutions being dismissed in courts. This issue regarding the digital evidence backlog is further compounded due to the cross-border, intra-agency cooperation required by many forensic investigations. If a given country has an especially low digital investigative capacity, it can have a significant knock-on effect in an international context (James and Jang, 2014).

In this paper, we review relevant recent research literature to elucidate the developments and current challenges in the field. While much progress has been made in the digital forensic process in recent years, little work has made appreciable progress in tackling the evidence backlog in practice. While evidence is lying unanalysed in an evidence store, investigations are often left waiting for new leads to be discovered, which has serious consequences for following these new threads of investigation at a later date. A number of practical infrastructural improvements to the current forensic process are discussed including automation of device acquisition and analysis, Forensics-as-a-Service (FaaS), hardware-facilitated heterogeneous evidence processing, remote evidence acquisition, and cross-jurisdictional evidence sharing over the Internet. These infrastructural improvements will enable a number of both new and improved forensic processes. These may include 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. The aforementioned improvements should combine to aid law enforcement and private digital investigators to greatly expedite the current forensic process. It is envisioned that the future research areas presented as part of this paper will influence further research in the field.

2. CURRENT CHALLENGES

Raghavan (2013) outlined five major challenge areas for digital forensics, gathered from a survey of research in the area:

1. The complexity problem, arising from data being acquired at the lowest (i.e. binary) format with increasing volume and heterogeneity, which calls for sophisticated data reduction techniques prior to analysis.
2. The diversity problem, resulting naturally from ever-increasing volumes of data, but also from a lack of standard techniques to examine and analyse the increasing numbers and types of sources, which bring a plurality of operating systems, file formats, etc. The lack of standardisation of digital evidence storage and the formatting of associated metadata also unnecessarily adds to the complexity of sharing digital evidence between national and international law enforcement agencies (Scanlon and Kechadi, 2014).

3. The consistency and correlation problem resulting from the fact that existing tools are designed to find fragments of evidence, but not to otherwise assist in investigations.
4. The volume problem, resulting from increased storage capacities and the number of devices that store information, and a lack of sufficient automation for analysis.
5. The unified time-lining problem, where multiple sources present different time zone references, timestamp interpretations, clock skew/drift issues, and the syntax aspects involved in generating a unified timeline.

Numerous other researchers have identified more specific challenges, which can generally be categorised according to Raghavan's above classification. Examples include Garfinkel (2010), Wazid et al. (2013), and Karie and Venter (2015).

It is widely agreed that the volume of data that is potentially relevant to investigations is growing rapidly. The amount of data per case at the FBI's 15 regional computer forensic laboratories has grown 6.65 times between 2003-2011, from 84GB to 559GB (Roussev et al., 2013). One cause of this is the growth in storage capacities that has occurred in recent years. Additionally, the increasing proliferation of mobile and Internet of Things devices adds to the number of devices that require examination in a given investigation. Beyond the magnitude of the data, the use of cloud services means that it may not be clear what data exists and where it is actually located.

As advanced mobile and wearable technologies have continued to become more ubiquitous amongst the general population, they also now play a more prevalent role in digital forensic investigations. Over the past decade the capabilities of these smart devices have reached a point where they can function at a level near to that of the average household computer and are currently only limited by processing power and storage capacity. This contributes to the diversity problem, where a greater variety of devices become candidates for digital forensic investigation (e.g. Baggili et al. [2015] has reported on forensics on smart watches). Mobile and IoT devices make use of a variety of operating systems, file formats and communication standards, all of which add to the complexity of digital investigations. In addition, embedded storage may not be easily removable from devices, unlike for traditional desktop and server computers, and in some cases, devices will lack persistent storage entirely, necessitating expensive RAM forensics.

Investigating multiple devices also contributes to the consistency and correlation problem, where evidence gathered from distinct sources must be correlated for temporal and logical consistency. This is often performed manually; a significant drain on investigators' resources. The requirements for RAM forensics also becomes pertinent in cases of anti-forensics, where a digital criminal takes measures to avoid evidence being acquired, including the creation of malware that resides in RAM alone. The increasing sophistication of digital criminals' activities is also a substantial challenge.

Other issues include limitations on bandwidth for transferring data for investigation, the volatility of evidence, the fact that digital media has a limited lifespan that may possibly result in evidence being lost, and the increasing ubiquity of encryption in modern communications and data storage.

The following sections concentrate on a number of important emerging trends in modern computing that contribute to the problems outlined above.

2.1 Internet-of-Things

The Internet-of-Things (IoT) refers to a vision of everyday items that are connected to a network and send data to one another. Juniper Research (2015) estimates that there are already 13.4bn IoT devices in existence 2015, and they expect this figure to reach 38.5bn by 2020. These IoT devices are typically deployed in two broad areas: in the consumer domain (smart home, connected vehicles, digital healthcare) and in the industrial domain (retail, connected buildings, agriculture). Some IoT devices are commonplace items that have Internet connectivity added (e.g. refrigerators, TVs), whereas others are newer sensing or actuation devices that have been developed with the IoT specifically in mind.

The IoT has the potential to become a rich source of evidence from the physical world, and as such it poses its own unique set of challenges for digital forensic investigators (Hegarty et al., 2014). Compared to traditional digital forensics, there is less certainty in where data originated and where it is stored. Data persistence may be a problem. IoT devices typically have limited memory (and may have no persistent data storage). Thus any data that is stored for longer periods may be stored in some in-network hub, or sent to the cloud for more persistent storage. Therefore, this means that the challenges related to cloud forensics (as discussed below in Section 2.2) will likely apply in the IoT domain also.

Already, some efforts have begun to analyse IoT devices for forensics purposes (e.g. Sutherland et al. [2014] on smart TVs), however this work is in its early stages at present. The heterogeneous nature of IoT devices, including differences in operating systems, file systems and communication standards, adds significantly to the complexity, diversity, and correlation problems for forensic investigators.

Ukil et al. (2011) outline some security concerns of IoT researchers, which feed directly into the desires of forensic investigators, incorporating issues such as availability, authenticity, and non-repudiation, which are important for legally-sound use of the data. These are addressed using encryption technologies, which are easy to incorporate into computationally powerful devices that are connected to mains energy. However, it becomes more of a challenge for smaller, battery-operated, computationally constrained devices where such considerations may be sacrificed. This has inevitable consequences for the usefulness of the data in a legal context.

2.2 Emerging Cloud Computing or Cloud Forensic Challenges

Usage of cloud services such as Amazon Cloud Drive, Office 365, Google Drive, and Dropbox are now commonplace amongst the majority of Internet users. From a digital forensics point of view, these services present a number of unique challenges, as has been reported in the 2014 National Institute of Standards and Technology's draft report (NIST, 2014). Typically, data in the cloud is distributed over a number of distinct nodes, unlike more traditional forensic scenarios where data is stored on a single machine. Due to the distributed nature of cloud services, data can potentially reside in multiple legal jurisdictions, leading to investigators relying on local laws and regulations regarding the collection of evidence (Simou et al., 2014, Ruan et al., 2013). This can potentially increase the time, cost, and difficulty associated with a forensic investigation. From a technical standpoint, the fact that a single file can be split into a number of data blocks that are then stored on different remote nodes adds another layer of complexity thereby making traditional digital forensic tools redundant (Chen et al., 2015, Almulla et al., 2013).

Additionally, the Cloud Service Providers (CSP) and their user base must be taken into

consideration. Investigators are reliant on the willingness of CSPs to allow for the acquisition and reproduction of data. The lack of standardisation among the varying CSPs, differing levels of data security, and their Service Level Agreements are obstacles to both cloud forensic researchers and investigators (Almulla et al., 2013). The multi-tenancy of many cloud systems poses three significant challenges to digital forensic investigations. In the majority of cases the privacy and confidentiality of legitimate users must be taken into account by investigators due to the shared infrastructures that support cloud systems (Morioka and Sharbaf, 2015). The distributed nature of cloud systems, along with multi-tenancy, can require the acquisition of vast volumes of data leading to many of the challenges outlined below. Finally, the use of IP anonymity and the easy-to-use features of many cloud systems, such as requiring minimal information when signing up for a service, can lead to situations where identifying a criminal is near impossible (Chen et al., 2012, Ruan et al., 2013). Cloud forensics also face a number of challenges associated with traditional digital forensic investigations. Encryption and other antiforensic techniques are commonly used in cloud-based crimes. The limited time for which forensically-important data is available is also an issue with cloud-based systems. Due to the fact that said systems are continuously running data, can be overwritten at any time. Time of acquisition has also proved a challenging task in regard to cloud forensics. Thethi and Keane (2012) showed that commonly-used forensic tools such as the Linux dd command and Amazon's AWS Snapshot took a considerable amount of time to acquire 30Gb of data from a cloud service.

While advances continue with regard to the tools and techniques used in cloud forensics, the aforementioned challenges continue to impede investigations. Henry et al. (2013) produced results showing that investigations on cloud-based systems make up only a fraction of all digital forensic investigations. Many investigations are stalled beyond the point of a perpetrator's owned devices and rarely extend into the cloud-based services they use. Results such as these form a strong argument for continued research in this field.

3. FUTURE RESEARCH

3.1 Distributed Processing

Distributed Digital Forensics has been discussed for some time (Roussev and Richard III, 2004, Shanmugasundaram et al., 2003, Garfinkel et al., 2009, Beebe, 2009). However, there is more scope for it to be put into practice. Roussev et al. (2013) cite two main reasons that the processing speed of current generation digital forensic tools is inadequate for the average case: First, users have failed to formulate explicit performance requirements; second, developers have failed to put performance as a top-level concern in line with reliability and correctness. They proposed and validated a new approach to target acquisition that enables file-centric processing without disrupting optimal data throughput from the raw device. Their evaluation of core forensic processing functions with respect to processing rates shows intrinsic limitations in both desktop and server scenarios. Their results suggest that with current software, keeping up with a commodity SATA HDD at 120 MB/s requires between 120 and 200 cores.

3.2 HPC and Parallel Processing

Despite the bottleneck of many digital forensic operations being disk read-speed, there are steps in the process that are not limited by the physical read-speed of the storage device. For instance, the analysis phase can consume large amounts of time by computers and humans. High performance computing (HPC)

advantages should be employed wherever possible to reduce computation time, and in an effort to reduce the time required by humans. Traditional HPC techniques normally exploit some level of parallelism, and to date have been underexploited by the digital forensic community. There are many applications where HPC techniques and hardware could be employed, for instance, on expediting each part of the digital forensic process after the acquisition phase, i.e., preprocessing, storage, analysis, and reporting.

3.3 GPU-Powered Multi-threading

GPUs excel at "single instruction, multiple data" (SIMD) computations with large numbers of general-purpose stream processors that can execute massively-threaded algorithms for a number of applications and stand to do so for many digital forensics requirements in theory.

Marziale et al. (2007), noted that GPUs have traditionally been both difficult to program and targeted at very specific problems. More recently, multicore CPUs coupled with GPU accelerators have been widely used in high-performance computing due to better power efficiency and performance/price ratio (Zhong et al., 2012). In addition, there is now a multitude of integrated GPUs that are on the same silicon die as the CPU, bringing both easier programming models and greater efficiency.

With new heterogeneous architectures and programming models such as these, powerful and efficient computer systems can be found in workstations with transparent access to CPU virtual addresses and very low overhead for computation offloading, and Power et al. (2015) have shown such architectures to be advantageous in analytic processing. These seem very well suited for many digital forensics applications, particularly as technologies such as SSDs reduce the I/O bottleneck.

Nonetheless, the use of GPUs in digital forensics is largely absent from the literature and there are few standard digital forensic tools that utilise GPU acceleration. Marziale et al. (2007) measured the effectiveness of offloading processing typical to digital forensics tools (such as file carving) to GPUs and found significant performance gains compared to simple threading techniques on multicore CPUs. Although the programming of the GPUs was more complex, the authors found that the effort was worth the performance gains. Collange et al. (2009) researched the feasibility of employing GPUs to accelerate the detection of sectors from contraband files using sector-level hashes.

Their application was able to inspect several disk drives simultaneously and asynchronously from each other. In addition, disks from different computers can be inspected independently by the application. This approach indicated that the use of GPUs is viable.

However, Zha and Sahni (2011) employed multi-pattern search algorithms to reduce the time needed for file carving with Scalpel, showing that the limiting factor for performance is disk read time. The authors state there is no advantage to using GPUs, at least until mechanisms to read the disk faster are found. However, this conclusion assumes only one disk, and the traditional digital forensic model. In the new era of cloud forensics, SSDs, and other technological evolutions, this I/O bottleneck will be much less restrictive.

Iacob et al. (2015) have employed GPUs in information-retrieval cases where response time is of importance, similar to Digital Forensics. They demonstrate significant speed-up of two Bloom filter operations, which are used in

approximate matching forensic applications (Breitinger and Roussev, 2014).

GPUs, like many new technologies, present new considerations for digital forensics. Breß et al. (2013) researched the use of GPUs to process confidential/sensitive information and found that data in GPU RAM is retrievable by unauthorised users by creating a dump of device memory. However, this does not impede the use of GPUs for processing confidential information when the system itself is only accessible to authorised users

3.4 DFaaS

Digital Forensics as a Service (DFaaS) is a modern extension of the traditional digital forensic process. Since 2010, the Netherlands Forensic Institute (NFI) have implemented a DFaaS solution in order to combat the volume of backlogged cases (van Baar et al., 2014). This DFaaS solution takes care of much of the storage, automation, investigator enquiry in the cases it manages. Van Baar et al. (2014) describe the advantages of the current system including efficient resource management, enabling detectives to directly query the data, improving the turnaround time between forming a hypothesis in an investigation its confirmation based on the evidence, and facilitating easier collaboration between detectives working on the same case through annotation and shared knowledge.

While the aforementioned DFaaS system is a significant step in the right direction, many improvements to the current model could greatly expedite and improve upon the current process. This includes improving the functionality available to the case detectives, improving its current indexing capabilities and on-the-fly identification of incriminating evidence during the acquisition process (van Baar et al., 2014).

Seeing as the DFaaS model is a cloud-based, remote access model, two significant disadvantages to the model are potential latency in using the online platform and being dependent on the upload bandwidth available during the physical storage acquisition phase of the investigation. A deduplicated evidence storage system, such as that described by Watkins et al. (2009), would facilitate the faster acquisition with each unique file across a number of investigations only needing to be stored, indexed, analysed, and annotated once on the system. Eliminating non-pertinent, benign files during the acquisition phase of the investigation would greatly reduce the acquisition time (e.g., operating system, application, previously acquired non-incriminating files, etc.). This could greatly expedite pertinent information being available to the detectives working on the case as early as possible in the investigation. In order for any evidence to be court admissible, a forensically sound entire disk image would need to be reconstructible from the deduplicated data store, improving upon the system proposed by Watkins et al. (2009). Employing such a system would also facilitate a cloud-to-cloud based storage event monitoring of virtual systems as merely the changes of the virtual storage would need to be stored between each acquisition.

3.5 Field-programmable Gate Arrays

FPGAs are integrated circuits that can be configured after manufacture. FPGAs can implement any function that application-specific integrated circuits can, and offer several advantages over traditional CPUs. FPGAs can exploit inherent algorithmic parallelism (including low-level parallelism), and can often achieve results in fewer logic operations compared to traditional general purpose CPUs, resulting in faster processing times. FPGAs have recently found application in areas such as digital signal processing, imaging and video applications, and

cryptography. Despite demonstrating desirable traits for digital forensics researchers, they have yet to be exploited for non-I/O-bound facets of digital forensics. Furthermore, as SSDs and other technologies ease the I/O bottleneck, FPGAs stand to be more broadly applicable in digital forensics.

3.6 Applying Complementary Cutting Edge Research to Forensics

Current investigation practice involves the analysis of data on standalone workstations. As such, the sophistication of the techniques that can be practically employed are limited. Much research has been conducted in a variety of areas that have theoretical relevance to digital forensics, but also have been impractical to apply to date. A movement towards DFaaS and high-performance computing, as discussed above, offers advantages beyond merely expediting the techniques currently used in forensics investigations, which remain reliant on manual input. It also promises a situation where this complementary research may practically be brought to bear on digital forensic investigations.

One such research area is that of Information Retrieval (IR). Traditionally, IR is concerned with identifying documents within a corpus that help to satisfy a user's "information need." Traditionally, IR researchers have been faced with the trade-off between the competing goals of precision (retrieving only relevant documents) and recall (retrieving all the relevant documents), whereby improving on one of these metrics typically results in a reduction in the other. In IR for legal purposes, recall has long been acknowledged as being the more important metric, given that a single missing relevant document could have serious consequences for the prosecution of a criminal case, the enforcement of a contract, etc. However, focusing on recall frequently results in an investigator being required to manually sift through a large quantity of non-relevant documents. This is in contrast to web search, for example, where users typically do not require all of the relevant documents to be retrieved, of which there may possibly be millions. Instead, a web searcher wishes to avoid wasting time on non-relevant material.

IR for digital forensics is often seen as a typical example of legal information retrieval (e.g. by Beebe and Clark [2007]). Although, this is certainly true at the point a case is being built for court, it could be argued that the level of recall required at the triage stage can be sacrificed somewhat for greater precision in order to allow investigators to make speedy decisions about whether a given device should be investigated fully. Thus, there is the potential for configurable IR systems to be utilised in forensics investigations, whose focus will change depending on the stage of the investigation.

The primary advantage of applying IR techniques to digital investigations is that once the initial preprocessing stage has been completed, searches can be conducted extremely quickly. Furnas et al. (1987) has shown that less than 20% of searchers choose the same keywords for topics they are interested in. This suggests that many queries must be run to achieve full recall, and also suggests that standard IR techniques such as query expansion and synonym matching could also be applied to increase recall.

However, increasing recall typically reduces precision by also retrieving non-relevant documents as false positives. There are a number of ways in which this problem can be alleviated. The use of the aforementioned data deduplication techniques would eliminate standard system files from consideration (Beebe and Dietrich [2007] note that the word "kill" appears as a command in many system

files). Additionally, common visualisation approaches such as ranking (Beebe and Liu, 2014) and clustering (Beebe et al., 2011) are likely to help investigators in their manual search of retrieved documents.

Another consideration is that event timeline reconstruction is extremely important in a criminal investigation (Chabot et al., 2014). When constructing a timeline from digital evidence, some temporal data is readily available (e.g. chat logs, file modification times, email timestamps, etc.), although it should be acknowledged that even this is not without its own challenges. Within the IR community, much research has been conducted into the extraction of temporal information from unstructured text (Campos et al., 2014). This can be used to dramatically reduce the manual load for investigators in this area.

4. CONCLUSION

In this paper a number of current challenges in the field of digital forensics are discussed. Each of these challenges in isolation can hamper the discovery of pertinent information for digital investigators and detectives involved in a multitude of different cases requiring digital forensic analysis. Combined, the negative effect of these challenges is amplified. The digital evidence backlog is currently in the order of years for many law enforcement agencies worldwide. The predicted ballooning of case volume in the near future will serve to further compound the backlog problem – particularly as the volume of evidence from cloud-based and Internet-of-Things sources continue to increase. In terms of research directions, practices already in place in many Computer Science sub-disciplines hold promise for addressing these challenges, including those in distributed, parallel, GPU and FPGA processing, as well as information retrieval techniques. These research directions can be applied to digital forensics requirements to help combat the backlog through more efficient allocation of precious digital forensic expert time through the improvement and expedition of the digital forensic process itself.

REFERENCES

Alexandru Iacob, Lucian Itu, Lucian Sasu, Florin Moldoveanu, and Constantin Suciu. Gpu accelerated information retrieval using bloom filters. In System Theory, Control and Computing (ICSTCC), 2015 19th International Conference on, pages 872–876. IEEE, 2015.

Arijit Ukil, Jaydip Sen, and Sripad Koilakonda.

Ben Hitchcock, Nhien-An Le-Khac, and Mark Scanlon. Tiered forensic methodology model for digital field triage by non-digital evidence specialists. Digital Investigation, 13(S1), 03 2016. Proceedings of the Third Annual DFRWS Europe.

Darren Quick and Kim-Kwang Raymond Choo. Impacts of increasing volume of digital forensic data: A survey and future research challenges. Digital Investigation, 11(4): 273–294, 2014.

E. Morioka and M.S. Sharbaf. Cloud computing: Digital forensic solutions. In Information Technology - New Generations (ITNG), 2015 12th International Conference on, pages 589–594, April 2015.

Embedded security for Internet of Things. In 2011 2nd National Conference on Emerging Trends and Applications in Computer Science, pages 1–6. IEEE, mar 2011. ISBN 978-1-4244-9578-8.

Frank Breitinger and Vassil Roussev. Automated evaluation of approximate matching algorithms on real data. Digital Investigation, 11:S10–S17, 2014.

Garda Síochána Inspectorate. Changing Policing in Ireland, November 2015.

George W. Furnas, Thomas K. Landauer, Louis M. Gomez, and Susan T. Dumais. The vocabulary problem in human-system communication. Communications of the ACM, 30(11):964–971, 1987.

Guangxuan Chen, Yanhui Du, Panke Qin, and Jin Du. Suggestions to digital forensics in cloud computing era. In Network Infrastructure and Digital Content (IC-NIDC), 2012 3rd IEEE International Conference on, pages 540–544, Sept 2012.

Iain Sutherland, Huw Read, and Konstantinos Xynos. Forensic analysis of smart TV: A current issue and call to arms. Digital Investigation, 11(3):175–178, sep 2014.

Ibrahim Baggili, Jeff Oduro, Kyle Anthony, Frank Breitinger, and Glenn McGee. Watch What You Wear: Preliminary Forensic Analysis of Smart Watches. In 2015 10th International Conference on Availability, Reliability and Security, pages 303–311. IEEE, aug 2015. ISBN 978-1-4673-6590-1.

Jason Power, Yinan Li, Mark D Hill, Jignesh M Patel, and David A Wood. Toward gpus being mainstream in analytic processing. 2015.

Joshua I James and Yunsik Jake Jang. Measuring digital crime investigation capacity to guide international crime prevention strategies. In Future Information Technology, pages 361–366. Springer, 2014.

Juniper Research. The Internet of Things: Consumer, Industrial & Public Services 2015-2020, July 2015.

Kathryn Watkins, Mike McWhorte, Jeff Long, and Bill Hill. Teleporter: An analytically and forensically sound duplicate transfer system. Digital investigation, 6:S43–S47, 2009.

Keyun Ruan, Joe Carthy, Tahar Kechadi, and Ibrahim Baggili. Cloud forensics definitions and critical criteria for cloud forensic capability: An overview of survey results. Digital Investigation, 10(1):34 – 43, 2013.

Kulesh Shanmugasundaram, Nasir Memon, Anubhav Savant, and Herve Bronnimann. Fornet: A distributed forensics network. In Computer Network Security, pages 1–16. Springer, 2003.

Lei Chen, Lanchuan Xu, Xiaohui Yuan, and N. Shashidhar. Digital forensics in social networks and the cloud: Process, approaches, methods, tools, and challenges. In Computing, Networking and Communications (ICNC), 2015 International Conference on, pages 1132–1136, Feb 2015.

Lodovico Marziale, Golden G Richard, and Vassil Roussev. Massive threading: Using gpus to increase the performance of digital forensics tools. digital investigation, 4:73–81, 2007.

Mark Scanlon and M-Tahar Kechadi. Digital evidence bag selection for p2p network investigation. In Proceedings of the 7th International Symposium on Digital Forensics and Information Security (DFIS-2013), pages 307–314. Springer, Gwangju, South Korea, 2014.

Mohammad Wazid, Avita Katal, RH Goudar, and Smitha Rao. Hacktivism trends, digital forensic tools and challenges: A survey. In Information & Communication Technologies (ICT), 2013 IEEE Conference on, pages 138–144. IEEE, 2013.

Neha Thethi and Anthony Keane. Digital forensics investigations in the cloud. In IEEE International Advance Computing Conference (IACC), Sept 2012.

Nickson M Karie and Hein S Venter. Taxonomy of challenges for digital forensics. Journal of forensic sciences, 60(4):885–893, 2015.

Nicole Beebe and Glenn Dietrich. A new process model for text string searching. In Advances in Digital Forensics III, pages 179–191. Springer, 2007.

Nicole Beebe. Digital forensic research: The good, the bad and the unaddressed. In Advances in digital forensics V, pages 17–36. Springer, 2009.

Nicole Lang Beebe and Jan Guynes Clark. Digital forensic text string searching: Improving information retrieval effectiveness by thematically clustering search results. Digital Investigation, 4(SUPPL.):49–54, 2007.

Nicole Lang Beebe and Lishu Liu. Ranking algorithms for digital forensic string search hits. Digital Investigation, 11(SUPPL. 2): 314–322, 2014.

Nicole Lang Beebe, Jan Guynes Clark, Glenn B. Dietrich, Myung S. Ko, and Daijin Ko. Post-retrieval search hit clustering to improve information retrieval effectiveness: Two digital forensics case studies. Decision Support Systems, 51(4):732–744, 2011.

NIST. NIST cloud computing forensic science challenges. 2014.

Paul Henry, Jacob Williams, and Benjamin Wright. The sans survey of digital forensics and incident response. In Tech Rep, July 2013.

RB van Baar, HMA van Beek, and EJ van Eijk. Digital forensics as a service: A game

changer. Digital Investigation, 11:S54–S62, 2014.

Ricardo Campos, Gaël Dias, Alípio M Jorge, and Adam Jatowt. Survey of temporal information retrieval and related applications. ACM Computing Surveys (CSUR), 47(2):15, 2014.

Robert C. Hegarty, David J. Lamb, and Andrew Attwood. Interoperability Challenges in the Internet of Things. In Paul Dowland, Steven Furnell, and Bogdan Ghita, editors, Proceedings of the Tenth International Network Conference (INC 2014), pages 163–172. Plymouth University, 2014.

S. Almulla, Y. Iraqi, and A. Jones. Cloud forensics: A research perspective. In Innovations in Information Technology (IIT), 2013 9th International Conference on, pages 66–71, March 2013.

Sebastian Breß, Stefan Kiltz, and Martin Schäler. Forensics on gpu coprocessing in databases–research challenges, first experiments, and countermeasures. In BTW Workshops, pages 115–129. Citeseer, 2013.

Simson Garfinkel, Paul Farrell, Vassil Roussev, and George Dinolt. Bringing science to digital forensics with standardized forensic corpora. Digital investigation, 6:S2–S11, 2009.

Simson L Garfinkel. Digital forensics research: The next 10 years. digital investigation, 7: S64–S73, 2010.

Sriram Raghavan. Digital forensic research: current state of the art. CSI Transactions on ICT, 1(1):91–114, 2013.

Stavros Simou, Christos Kalloniatis, Evangelia Kavakli, and Stefanos Gritzalis. Cloud forensics solutions: A review. In Lazaros Iliadis, Michael Papazoglou, and Klaus Pohl, editors, Advanced Information Systems Engineering Workshops, volume 178 of Lecture Notes in Business Information Processing, pages 299–309. Springer International Publishing, 2014. ISBN 978-3-319-07868-7.

Sylvain Collange, Yoginder S Dandass, Marc Daumas, and David Defour. Using graphics processors for parallelizing hash-based data carving. In System Sciences, 2009. HICSS'09. 42nd Hawaii International Conference on, pages 1–10. IEEE, 2009.

Vassil Roussev and Golden G Richard III. Breaking the performance wall: The case for distributed digital forensics. In Proceedings of the 2004 digital forensics research workshop, volume 94, 2004.

Vassil Roussev, Candice Quates, and Robert Martell. Real-time digital forensics and triage. Digital Investigation, 10(2):158–167, 2013.

Xinyan Zha and Sartaj Sahni. Fast in-place file carving for digital forensics. In Forensics in Telecommunications, Information, and Multimedia, pages 141–158. Springer, 2011.

Yoan Chabot, Aurélie Bertaux, Tahar Kechadi, and Christophe Nicolle. Event reconstruction: A state of the art. Handbook of Research on Digital Crime, Cyberspace Security, and Information Assurance, page 15, 2014.

Ziming Zhong, Vladimir Rychkov, and Alexey Lastovetsky. Data partitioning on heterogeneous multicore and multi-gpu systems using functional performance models of data-parallel applications. In Cluster Computing (CLUSTER), 2012 IEEE International Conference on, pages 191–199. IEEE, 2012.

