Simulating SQL-Injection Cyber-Attacks Using GNS3

A. Mahrouqi, P. Tobin, S. Abdalla, and T. Kechadi

***Abstract*—Network Forensics is a subtopic of Digital Forensics wherein research on artificat investigations and intrusions evidence acquisition is addressed. Among many challenges in the field, the problem of losing data artifacts in the state of flux, (i.e., live volatile data), when network devices are suddenly non-operational remains a topic of interest to many investigators. The main objective of this article is to simulate an SQL injection attack scenarios in a complex network environment. We designed and simulated a typical demilitarized zone (DMZ) network environment using graphical network simulator (GNS3), Virtual Box and VMware workstation. Using this set-up we are now able to simulate specific network devices configuration, perform SQL injection attacks against victim machines and collect network logs. The main motivation of our work is to finally define an attack pathway prediction methodology that makes it possible to examine the network artifacts collected in case network attacks.**

***Index Terms*—Acquisition, anti-forensics, network forensics, SQL injection attack.**

I. Introduction

Digital forensics is a branch of forensics science and has been defined as "the use of scientifically proved methods towards the preservation, collection, validation, identification, analysis, interpretation, documentation, and presentation of digital evidence from digital source for the purpose of facilitating or furthering the reconstruction of events found to be criminal, or helping to anticipate unauthorised actions shown to be disruptive to planned operations [1].

Later researchers of this science have extended the digital forensics meaning to cover the forensics of each digital technology and its development. Digital forensics includes several sub-branches related to the investigation and acquisition of various types of media, devices and data artefacts, e.g., computer forensics, mobile device forensics, network forensics, forensics data analysis and database forensics. Digital forensics aims to extract digital evidence related to unauthorised actions happening in the target devices [2].

The rules to evaluate the admissibility of digital evidence are different from one country to another. For example, the United States use Federal Rules of Evidence whereas the United Kingdom use PACE and the Civil Evidence Acts. Digital forensics acts also differ from one country to another. For instance, in the US federal laws restrict seizures to items with only clear and obvious evidential value. However, in the UK digital investigators can seize any suspected evidence that has been found at a crime scene [3].

Digital investigators must be aware of two important issues while seizing and acquiring digital evidence – integrity and authenticity. Integrity ensures that the acquired digital evidence does not modify the original copy of the evidence. Whereas authenticity is the process of verifying the integrity of the acquired evidence [4]. The digital investigation should document the actions and evidence based on the chain of custody. This will ensure this evidence is admissible in the Court of Law. There must be enough evidence for extraction and examination without modification and bias. The link between evidence and criminal prosecution is potentially complicated because it relates to a series of interconnected events, depending on logical sequencing. Therefore, sufficient forensics evidence must be taken for analysis.

A network simulation tool allows end-users and professionals to emulate complex networks at low cost and consuming less time. GNS3 is an example of simulation tools and it refers to Graphical Network Simulators. GNS3 allows us to connect to Virtual Box virtual machines that are used to emulate different operating systems, e.g. Linux and Microsoft Windows. In addition, GNS3 allows the emulation of Cisco IOSs.

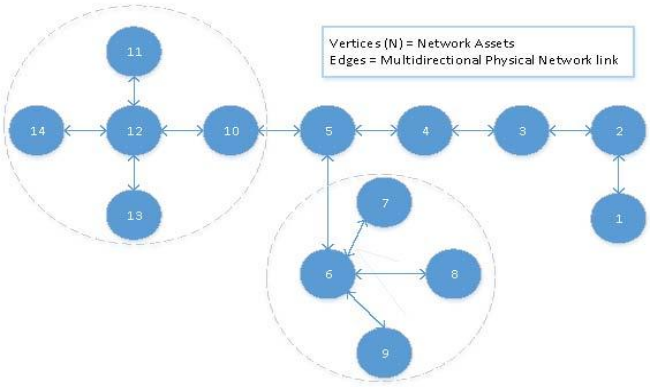

Fig. 1. Building node relationship.

The physical memory is the first concern of digital investigators. It contains critical and interesting volatile information about a computer and network device incident such as intruders IP addresses, information about running malicious programs, processes, worms, Trojans and so on [5]. In this paper, we simulated the network topology by using the open source network designing simulation tool called GNS3. There are many powerful open source tools designed for simulation and emulation of the data network like CLOONIX, CORE, GNS3, IMUNES, Marionnet, Mininet, Netkit, Psimulator2, Virtual square, VNX and VNUML. The main difference between network emulation and network simulation is that network emulation is a method

Manuscript received October 30, 2014; revised March 2, 2015. This work was supported in part by the Royal Court Affairs, Insight Centre for Data Analytics, University College Dublin. Simulating SQL-Injection Cyber attacks using GNS3.

The authors are with Insight Centre for Data Analytics, University College Dublin, Ireland (e-mail: aadil.al-mahrouqi@ucdconnect.ie, tahar.kechadi@ucd.ie, patrick.tobin@insight-centre.org, sameh@ucd.ie).

to simulate the properties of an existing, planned network by using emulation of specific network equipment like routers, switches, and computers. Whereas, network simulation defines the mathematical models of a data source, protocol and channels applied in the network topology.

II. Related Work

The purpose of this section is to provide an overview of the most relevant research done in network forensics simulation techniques and other different approaches. In addition, this section provides a platform for the case study which is the main focus of the research described in this paper. This literature review research is mainly focused on different primary areas: the simulation of the SQL injection attack in GNS3, network memory forensics analysis tools, investigation processes to detect interesting evidence by using Wire shark forensics tool and network anti-forensics techniques.

Digital Forensics uses software tools to get the results and gain data bit-by-bit from memory dumps. A prerequisite forensics examination is necessary and for this process we will be using a selection of tools such as EnCase and Forensic Toolkit (FTK) to extract and integrate information related to illegal activities using a network infrastructure [6].

Acquisition of data imaging from the target network is important and difficult [7]. Procedures and standards must be verified during acquisition: "acquisition of digital evidence, by its very nature, is fragile and can be altered, damaged, or destroyed by improper handling or examination" [8]. When data is identified from the network devices for acquisition purposes, the network forensics investigation should be conducted in a proper way in order to avoid any volatile information lost, network device locking or network power loss.

The acquisition stage of an investigation of a computer device is conducted in a digital forensics laboratory, but the acquisition process in network forensics is different. Network forensics investigators deal with live devices in most cases [9], and therefore investigators can't power off network devices and bring them to a laboratory.

Anti-forensics is defined as any attempt at compromising or destroying digital evidence according to the two forensics analysis methods [10]. Anti-forensics approaches are classified into several groups based on various techniques and tools. In [11] the authors proposed the most accepted subcategories of anti-forensics: data hiding, artefact wiping, trail obfuscation and attacks against the computer forensics processes and tools. After the illegal actions and attack activities, most of the professional attackers used anti-forensics to prevent proper digital forensics investigation processes that might be conducted. In addition, anti-forensics for networks has created major challenges. They use stealth and masking for hiding any digital evidence from the victim network device. The memory in most network devices like routers and switches, contain volatile information that requires continuous power.

The integrity of the data image must be examined during image processing of the victim devices. A hash function, such as md5sum or sha1sum for Linux, is the most common technique that is used to check the data integrity of the examined file. A hash function is a function that takes a relatively arbitrary input and produces an output of fixed size [12]. The hash will change if any modification occurs to the examined file. Consequently, hash functions are used to verify that no modifications have taken place when acquiring the data.

III. Attack Stimulation in GNS3

The case study scenario is selected based on the issues and problems that are faced in network forensics. In this research, the scenarios have been developed to demonstrate the results and to assist organisations and investigators in dealing with such attacks.

There are two attack scenarios that we can investigate, we have made certain assumptions about the attack strategies used in order to simplify and summarise an attack. One is an internal attack committed by a trusted person within the company, and the other is an external attack committed by an entity whose credentials are unknown to the company. These two scenarios present very different concerns for a company and support, to a point, two differing attack topologies. A third attack type is a hybrid of both of these attack types and can be described as a 'fuzzy' attacker. This is one where the attacker is external to the network, but establishes a presence within it by compromising a node, gaining a certain degree of control of a node from where he can launch an attack.

A. Definitions

To better understand what is being described it is best to define what an attacker and what a victim may be. In the context of these scenarios an attacker is any entity which maliciously attempts to access any data to which they are not allowed, within a network, without permission or authorisation. However that access is attempted, whether externally or internally, without proper authorisation. Malicious intent can include deliberate and knowing intent to control, copy, alter, destroy, obfuscate, insert, remove, corrupt, encrypt, retrieve or steal any information without proper authority, even if authority does exist to do any of these things to any other data, but not to those data accessed and includes any attempt to do one or more of those things.

A victim is any node which holds something of value and which is subject to an attack. A victim can be a server or any other node of the network. Evidence is anything of value, and may include data, emails, email addresses, log files, customer information or employee information.

For the purposes of this attack scenario it is safe to assume that an external attacker may not have the knowledge at his disposal that an internal attacker may have. This puts the external attacker at a distinct disadvantage in certain areas. He has to overcome the firewall before gaining access to the system and then has to find his target node in the system by probing the network to identify the target node. In a network with many nodes this can cause fingerprints to be left behind in log files on nodes probed. This can flag alerts to the network management of increased activity, perhaps triggering actions to curb the network activity of the attacker. There are however certain assets associated with being an external attacker. These can include anonymity and obfuscation of IP address and the advantage of being an

unknown entity. Anonymity and being an unknown entity are perhaps the two best advantages offered to an external attacker. All these contribute to presenting different attack strategies which can be difficult to defend against. Anonymity allows an attacker to use many different IP addresses, offering an attacker the advantage of being able to launch multiple attacks on a network without raising causing concern. Being an unknown entity presents network security with the difficulty of predicting possible attack patterns and preparing for and defending against those possible attacks. These difficulties are further added to by not knowing what the target or purpose of the attack is –is it to gain a foothold in the system, to attack or steal data, to disable servers or equipment, to compromise system security, etc.

The internal attacker already has access to and possesses knowledge of the network, its protocols, its design and security. He most likely knows his target, its network IP address, possibly its MAC address and may have knowledge of or access to user names and passwords. These can give the internal attacker a significant advantage, including being behind the externally facing firewall, The internal attacker can construct an attack with a high degree of precision, and by using the shortest known path to the victim can minimise his footprint on the network, reducing the likelihood of raising alerts and alarms, avoiding revealing his presence and circumventing access protocols, thereby concealing his activities. This knowledge gives the internal attacker a very distinct advantage relative to an external attacker.

We built the lab scenario by using open source tools; GNS3 and Virtual Box. After that, we simulated a real attack to the network core server. We ran experiment's to complete the tests, measurements and analysis of the given case study. In addition, we tried mathematical hypothesis analysing techniques in order to detect the source of the attack.

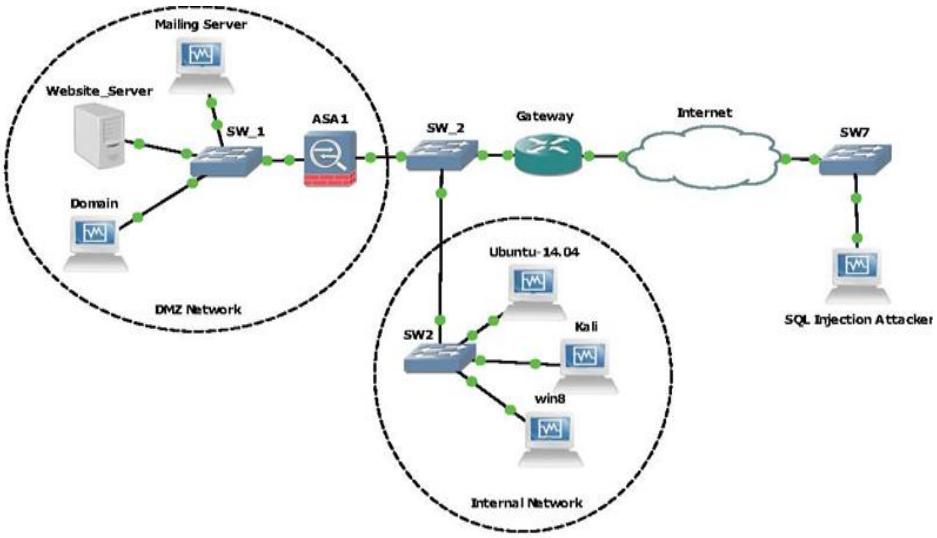

Fig. 2. Network topology of the proposed case study.

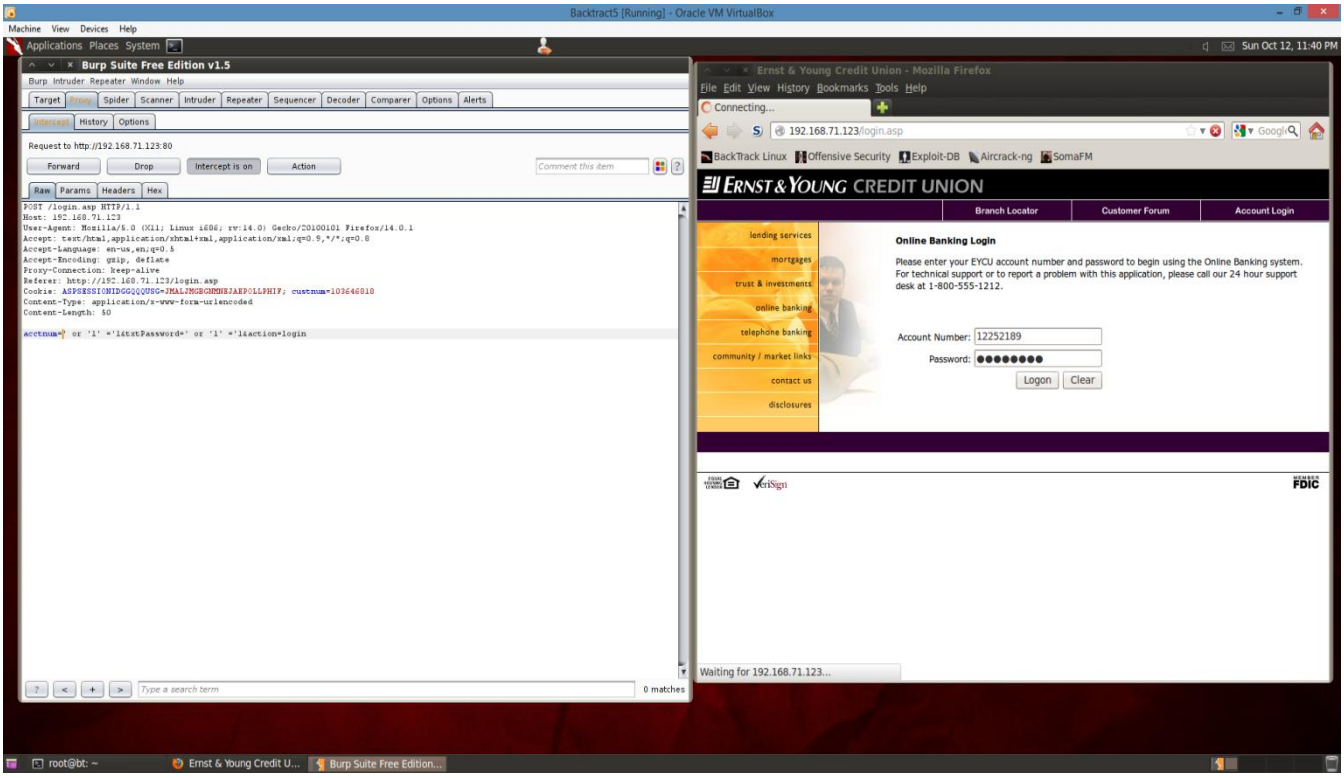

Fig. 3. Inserted SQL injections in the account number.

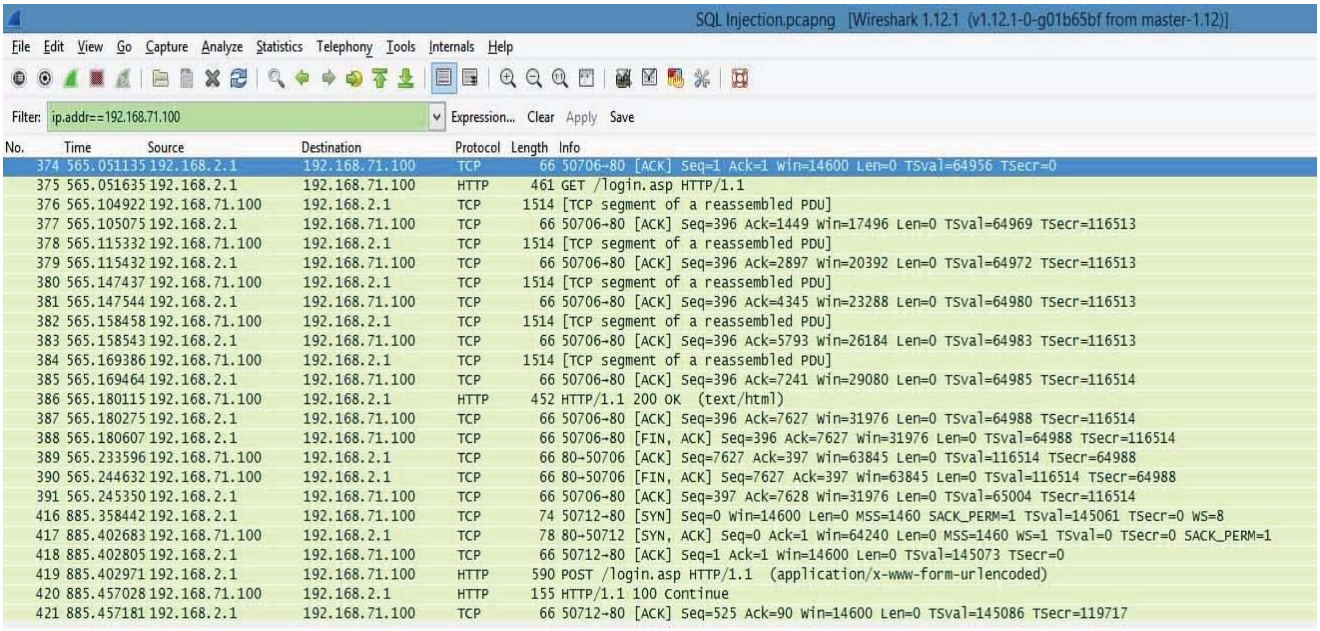

Fig. 4. Packets correlation by using expression features.

B. Incident Summary

Law enforcement received a report that Great International Bank's (http://192.168.71.129/) website has been compromised by an unknown attacker. Based on the initial investigation on the website the attacker used different techniques and tools to compromise the victim's website such as SQL injection, XSS, Broken caching, directory traversal and breaking the local authentication login to the server. Please see the (Fig. 2 depicts the Network topology).

C. SQL Injection Attack

The attacker used Burp Intercepting Proxy to intercept connection between his machine and the victim's server. The attacker sent the SQL injection code by using Burp Intercepting Proxy to gained access to the personal account of a Michael Nan Carrow (Fig. 3 shows the attack).

IV. Prediction Investigation Approach by Using ECPM

This approach gives information that can help Forensics Investigators about the network topology and network incidents. We can use the Prediction Investigation Approach to predict and trace the source of the attack or illegal activities in the computer network. The idea behind this approach is to identify the evidence collection path (ECP) by using evidence collection process model (ECPM). ECPM functions are divided based on the phases in the Cybercrime that has been discussed in [13].

A. Network Assets Numbering

The first phase in the cybercrime chain is the proactive phase and its goal is to prepare the target network to automatically prevent and detect the attack or illegal activities before the network gets infected. Network Assets Numbering is that each network device in the target network should carry a unique number in order to identify this device forensically. There are many ways of numbering the nodes for example, numbering based on the device location in the network topology or numbering based on location plus the corresponding network layer for the given device, for example Cisco Router 3700 - its location in the network topology is 3so, the numbering for this device is (L3-3).

B. Learning Phase

Stage one is to build a Network as shown in topology Fig. 1, we built node relationships for the given case study scenario. After building the Network topology we will create an Adjacency Matrix based on the nodes relationship. The one's in the Adjacency Matrix shows the direct links between nodes while zero indicates no direct link between nodes in the network.

The second stage of this approach is to create a Network Union Matrix. Network Union Matrix will try to substitute all zero values in the Adjacency Matrix with distances between one node to another [Source - Distance (link(n))] for example, [7-1 (link(6))] and [1-2 (link(1))].The third stage of this approach is to create an Attack Pathway Detection. Sometimes the network investigators face difficulties in understanding the network infrastructure and the relationship between the victim node and others. The Attack Pathway Detection will try to utilise the Network Union Matrix as a road map for the investigation process to trace the source of the attack.

The Network Union Matrix will list all distances between the victim's node and other nodes. This will help examining all suspected nodes. The final step is recovering the remains from the suspected nodes which can be used as evidence (evidence collection).

V. The Investigation

For data acquisition of the SQL injection that has been generated by the attacker we used the Wire shark forensics tool for packets analysis. Fig. 4 shows all correlation packets between attacker and victim server. In Fig. 4, it's very clear that attacker has accessed to login.asp webpage after shake hand communications. After that, we inspect login.asp (webpage that has been infected by the SQL injection attack) as well as checking suspicious packets that we believe are the

results of the SQL injection attack. Wire shark forensics tool shows that the attacker opened the login.asp in his computer. In addition it shows that the attacker sent an SQL injection to the victim server.

VI. Conclusion and Future Work

In this paper we presented a simulation study network attack scenario. The main point of designing virtual network attack environments is to create a sandbox that allows us to perform such experiments from our real assets and at a low cost. The outcome of this experiment can be used as a recommendation in our real IT infrastructure. The core idea of the case study is to examine the Website that has been compromised by an SQL injection attack. To simulate this attack scenario we used many open source tools like Graphical Network Simulator (GNS3), Oracle VM Virtual Box and VMW are workstation.

We also used the Wire shark forensics tool to detect criminal activity from the network layer (Layer 3 in OSI model) and in addition, we also examined the victim and attacker's devices by using the Volatility Framework 2.4.

This paper will also act as a first step towards an attack simulation analysis. For future work, we will move on to simulate an attack in a cloud network. Furthermore, we will focus on mathematical modelling and algorithm for the evidence detection process. In addition, we will try to utilise criminology science to enhance the results and the efficiency of the investigative process in more complex case studies.

References

[1] P. Cisar and S. M. Cisar, "General direction of development in digital forensics," *Ata Techica Corviniensis-Bullentin of Engineening*, pp. 87-91, 2012.

[2] B. Carrier and E. H. Spafford, "Information assurance and security — CERIAS Purdue University," *International Journal of Digital Evidence*, pp. 2-20, Fall 2013.

[3] C. J. McGuinness, P. T. R. Clarke, S. F. McAuley *et al*., *Documentary and Electronic Evidence*, Dublin: Law Reform, 2009.

[4] I. C. N.-C. W. Group, *Anti-Cartel Enforcement Manual, Internation Competition Network*, 2010.

[5] L. Cai, J. Sha, and W. Qian, "Study on forensic analysis of physical memory," in *Proc. 2nd International Symposium on Computer, Communication, Control and Automation (3CA 2013)*, 2013.

[6] L. Aouad and M. T. Kechadi, "An open architecture for digital evidence acquisition on smart phones," in *Proc. the 8th Annual Workshop on Digital Forensics (IFIP'12)*, Pretoria, South Africa, January 3-5, 2012.

[7] M. T. Kechadi and L. Aouad, "Cloud-based mobile: What about digital forensics?" *eForensics Magazine Mobile*, vol. 1, no. 2, pp. 22-27, 2012.

[8] J. Ashcroft, D. J. Daniels, and S. V. Hart, *Forensic Examination of Digital Evidence: A Gude for Law Enforecement*, Washington, DC: U.S. Department of Justice office of Justice Program, 2012.

[9] ACPO E-Crime Working Group, "Good practice guide for computer-based electronic evidence," *7 Safe Information Security*, 1996.

[10] R. Harris, "Arriving at an anti-forensics consensus: Examining how to," *ELSEVIER and Digital Investigation 3S*, pp. S44-S49, 2006.

[11] M. Rogers, "Anti-forensics," presented at Conference in Lockheed Martin, San Diego, San Diego, 2005.

[12] The Law Reform Commission group, "Documentary and electronic evidence," *Law Reform Commission*, 2009.

[13] A. A. AlMahrouqi, S. Abdalla, and T. Kechadi, "Network forensics readiness and security awareness framework," in *Proc. International Conference on Embedded System in Telcommunications and Instrumentation*, Annaba, 2014.

A. Mahrouqi was born in Ibri, Oman, in 1981. He received the B.E. degree in computer hardware and networking from the Coventry University, UK, in 2010, and the M.Tech. degree in computer science from the University Banasthali Vidaypith, India, in 2012, and M.Sc. degree in digital investigation and forensics computing in Dublin, Ireland, in 2013, respectively.

In 2002, he joined the Sultan's Armed Forces Signal, Royal Army of Oman, as a data communication technician, and in 2007 until now, he is a network engineer in the General Department of Information Technology, Royal Court Affairs. His current research interests include cybercrime, network forensics, security readiness, anti-forensics, reverse engineering, network event logs. Mr. Aadil is a PhD researcher in Insight Centre for Data Analytics, University College Dublin, Ireland.

P. Tobin was born in Wexford, Ireland. He received a BSc. (Hons) degree in computer applications from Dublin City University and a MSc. degree in forensic computing and cybercrime investigation from University College Dublin in 2010. He has served 31 years as a garda (police officer) and is now a full-time PhD student researching evidence recovery from, and forensic examination of, virtual machines.

S. Abdalla received his master degree from the Universita Cattolicadel Sacro Cuore in Milan, Italy in November 2004. In 2009, Dr. Sameh received his PhD degree in computer science from the University of Trento, Italy. From November 2009 to April 2011, Dr. Sameh worked as a postdoctoral research fellow at Lille University of Science and Technology in France. He is a senior postdoctoral researcher at University College Dublin since 2011. Dr. Sameh has a solid industrial experience too. From 2005 to 2008, Dr. Sameh worked as a scientific researcher for ARS LOGICA SRL - a private sector ICT research laboratory in Trentino, Italy. In 2004, Dr. Sameh worked as a multimedia services analyst at the R&D labs of SIEMENS Mobile Communications, Milan, Italy. In addition, during his undergraduate studies, Dr. Sameh worked for four years at Orascom Telecom, Egypt.

T. Kechadi received his master and PhD degrees in computer science from University of Lille1, France. He was appointed as a lecturer at the Computer Science Department of Lille University. Subsequently he worked as a post-doctoral researcher under the TMR program at UCD. He joined UCD in 1999 as a permanent staff member of the School of Computer Science & Informatics (CSI). He is currently a professor of Computer Science at CSI, UCD. His research interests span the areas of data mining, distributed data mining, heterogeneous distributed systems, grid and cloud computing, digital forensics and cyber-crime investigations. Prof. Kechadi published over 210 research articles in refereed journals and conferences. He serves the scientific committees for a number of international conferences and he organised and hosted some of the leading conferences in his area. He is currently an editorial board member of the Journal of Future Generation of Computer Systems and of IST Transactions of Applied Mathematics-Modelling and Simulation. He is a member of the communication of the ACM journal and IEEE computer society.