Advancing the Science of Digital Forensics

Gary C. Kessler, *Embry-Riddle Aeronautical University*

Digital forensics, the branch of forensic science that focuses on the recovery and investigation of digital data, has applications in many contexts outside the courtroom, including research, policy enforcement, and intelligence gathering.

Digital forensics combines methods from science, technology, and engineering to acquire and interpret information stored on digital devices for use in answering questions in court. Of course, these same methods allow for the acquisition of data for use in many contexts outside the courtroom, such as pure and applied research, policy enforcement, information security incident response, and intelligence gathering.

EARLY BEGINNINGS

My first foray into anything even remotely related to what we do today in computer forensics occurred in 1981. At that time, I was a programmer and coordinator of academic computing at a small college in Vermont, and our computer was an IBM System/34. Due to some catastrophic failure during shutdown the prior evening (yes, we shut the system down every night), the computer would not boot up the next morning. We found out later that this was the result of a corrupted volume table of contents (VTOC), the rough equivalent of today's file allocation table (FAT) or $Bitmap.

A consultant systems programmer came in to show us how to recover the files by reconstructing the VTOC based upon the prior morning's routine printout of the hard drive contents (yes, we made such a printout every day or two). Since we couldn't use the computer without overwriting the files in what was now, essentially, unallocated space and had no PC-class systems at the time, we did the hex conversions by hand—on paper. It took us three days to reconstruct the VTOC and get back online.

That was the beginning of the computer forensics process and that was our environment: using a hex editor to get down to the bare metal of the hard drive and file system. And that's how it was for most of the next 15 years—hackers (when the term was implicitly White Hat and, indeed, noble, before Black Hat hackers hijacked the term) with an interest in investigations, most often in the law enforcement community, building rudimentary tools for use in looking deep into the computer and its file system.

By the late 1990s, computer science departments began taking serious notice of computer forensics, and academic programs in digital forensics were introduced in the early 2000s. And yet, it was not until 2009 that the American Academy of Forensic Sciences adopted digital forensics as a science.

Forensic sciences are largely based on Locard's exchange principle: every contact leaves a trace—if one

0018-9162/12/$31.00 © 2012 IEEE Published by the IEEE Computer Society

person hits another on the head with a tree branch, part of the tree branch stays on the victim's head and part of the head stays on the tree branch. This is as true in cyberspace as it is in real space. The challenge with digital forensics is to find the traces, interpret them correctly—and place a person's fingers on the keyboard.

A primary difference between digital forensics and the other forensic sciences is that practitioners advanced the field before the computer science community generally got involved with research and education. Thus, although digital forensics has been around for several decades, it is still a young science, and the body of peer-reviewed, academic literature that is essential for every science is currently relatively small—but it is growing.

IN THIS ISSUE

The cover features in this special issue are not intended to provide a survey of the digital forensics field, but rather to offer a snapshot of four interesting, varied, and relevant areas of research activity: computer forensics, network forensics, control system vulnerabilities, and mobile device security.

Computer forensics

Many computer science applications use hashing to build a data structure for use in mapping one set of table entries to another, such as a variable name to an address in memory. For these applications, the hash values tend to be short, and hash collisions—that is, two different entries having the same hash value—are to be expected. Cryptographic hashing has a different function, namely attempting to provide data integrity and a unique identifier for a data item, such as a file on a hard drive.

Although hashes are not unique over the entire universe of possible files, hash collisions are rare in practice. Thus, hashes can be used as the basis for searching and filtering files in a computer forensics examination to identify known contraband and malware as well as known trusted files. While using hash sets to assist in identifying files of interest in an examination streamlines the process, this approach has severe limitations. If as little as a single bit in the file is altered due to system error or deliberate user action, the file's hash is very different than the expected value.

In "Distinct Sector Hashes for Target File Detection," Joel Young and his colleagues from the Naval Postgraduate School and Johns Hopkins University describe a method for employing hashes on a per-sector basis rather than per-file to identify known files and discuss the efficacy of using this approach with various file systems.

Network forensics

Today, it is unusual to find a computer that is not connected to the Internet. Just as investigators need to understand computer operating systems and file systems to get the most out of an examination of a computer, they also need knowledge of network applications and protocols when investigating a network.

While so-called hacker tools have become essential elements in a security officer's toolkit, these same tools can help in a network-based examination during a criminal or civil investigation, incident response analysis, or intelligence-gathering operation. Indeed, knowledge of network components, communication protocols, operating system utilities, the Transmission Control Protocol/Internet Protocol (TCP/IP) suite, the Internet, application software (including malware, browsers, and peer-to-peer clients), and cloud applications (including social networks and file-sharing sites) is essential to understanding the network artifacts found on computers.

A primary difference between digital forensics and the other forensic sciences is that practitioners advanced the field before the computer science community generally got involved with research and education.

Network forensics, a specialty within the digital forensics field, requires its own set of processes. "Network Forensics: An Analysis of Techniques, Tools, and Trends" by Ray Hunt of the University of Canterbury and Sherali Zeadally of the University of the District of Columbia provides an overview of the network forensics space, a review of state-of-the-art tools and methodologies, and a glimpse into the future.

Control system vulnerabilities

Supervisory control and data acquisition (SCADA) systems are employed to monitor and manage industrial control systems and processes. SCADA systems can be as simple as a temperature-sensing device used to turn a heater on and off or as complex as a radiological-sensing device that manages the position of control rods. Such systems are used extensively in critical infrastructures as varied as chemical plants, oil refineries, utility distribution systems, waterway and dam management systems, transportation systems, and manufacturing plants.

The information security vulnerabilities of SCADA systems have been studied extensively, and the vulnerable nature of these systems is well-known. But in the case of a security breach, what are the computer forensics ramifications? What tools and techniques are available to the investigator? How could the process be improved? Indeed, what training is available?

"SCADA Systems: Challenges for Forensic Investigators" by Irfan Ahmed and Golden G. Richard III from the University of New Orleans and Sebastian Obermeier and Martin Naedele from the ABB Corporate Research Center, Switzerland, explores these issues and describes the response from the digital forensics research community.

Mobile device security

During the past decade, mobile phones have evolved from being cool tech toys to become ubiquitous personal necessities. And, not surprisingly, cell phones—particularly, but not exclusively, those with cameras—are increasingly becoming the record keeper, victim, or instrument of criminal activity. Indeed, smartphones are essentially portable Internet terminals that, arguably, contain more probative data per byte examined than computers. At the same time, mobile phones have become a favored target of criminal hackers and a growing body of malware apps. Thus, mobile device forensics is a rapidly growing subspecialty of digital forensics.

In "Smartphone Security Challenges," Dakota State University researchers Yong Wang, Kevin Streff, and Sonell Raman describe some of the threats to smartphone security and suggest steps that users can take to further protect these intimate devices.

The authors and I thank *Computer* for contributing to the efforts to advance the science of digital forensics by publishing this special issue.

Gary C. Kessler *is an associate professor of homeland security at Embry-Riddle Aeronautical University, Daytona Beach, Florida; an adjunct associate professor at Edith Cowan University, Perth, Australia; and a member of the Northern Florida Internet Crimes Against Children (ICAC) Task Force. He received a PhD in computing technology in education from Nova Southeastern University, Fort Lauderdale, Florida. Contact him at gck@garykessler.net.*