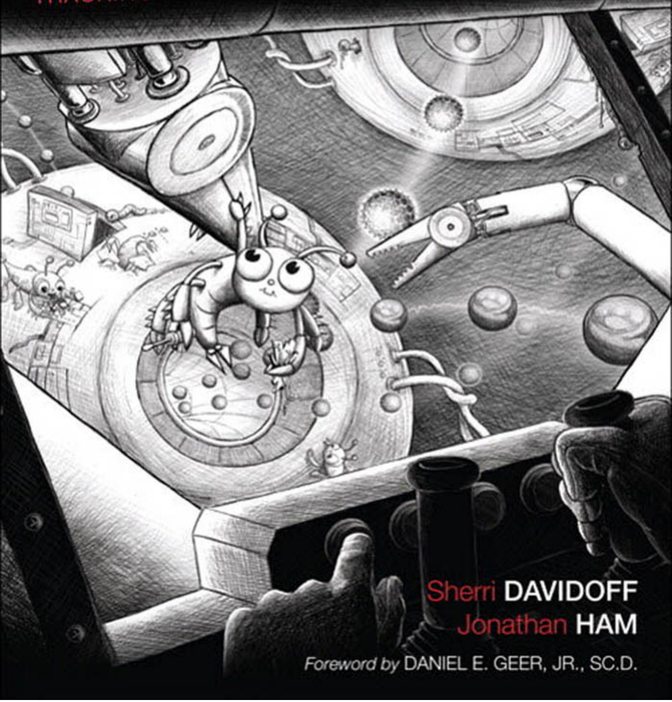
Network Forensics

TRACKING HACKERS THROUGH CYBERSPACE



Sherri DAVIDOFF
Jonathan HAM

Foreword by DANIEL E. GEER, JR., SC.D.

Network Forensics

This page intentionally left blank

Network Forensics

Tracking Hackers through Cyberspace

Sherri Davidoff
Jonathan Ham

PRENTICE HALL

Upper Saddle River, NJ • Boston • Indianapolis • San Francisco
New York • Toronto • Montreal • London • Munich • Paris • Madrid
Capetown • Sydney • Tokyo • Singapore • Mexico City

The publisher offers excellent discounts on this book when ordered in quantity for bulk purchases or special sales, which may include electronic versions and/or custom covers and content particular to your business, training goals, marketing focus, and branding interests. For more information, please contact:

U.S. Corporate and Government Sales
(800) 382-3419
corpsales@pearsontechgroup.com

For sales outside the United States please contact:

International Sales
international@pearson.com

Visit us on the Web: informit.com/ph

Library of Congress Cataloging-in-Publication Data

Davidoff, Sherri.
Network forensics : tracking hackers through cyberspace / Sherri Davidoff, Jonathan Ham.
p. cm.
Includes bibliographical references and index.
ISBN 0-13-256471-8 (hardcover : alk. paper)
1. Computer crimes–Investigation. 2. Computer hackers. 3. Forensic sciences. 4. Computer crimes—Investigation—Case studies. I. Ham, Jonathan. II. Title.
HV8079.C65D348 2012
363.25'968–dc23 2012014889

ISBN-13: 978-0-13-256471-7
ISBN-10: 0-13-256471-8
Text printed in the United States on recycled paper at Courier in Westford, Massachusetts.
Second printing, February 2013

To my mother, father, and sister.
Thank you for teaching me to follow my dreams
and for your endless encouragement and support.

SD

For Charlie and Violet,
and the rest of my family,
whose patience has made this possible.

JH

"May you find what you seek."

–An "ancient curse"; origin unknown.

Contents

This page intentionally left blank

Foreword

My great-grandfather was a furniture maker. I am writing this on his table, sitting in his chair. His world was one of craft, "the skilled practice of a practical occupation." [1] He made furniture late in life that was in superficial respects the same as that which he made earlier, but one can see his craft advance.

Cybersecurity's hallmark is its rate of change, both swift incremental change and the intermittent surprise. In the lingo of mathematics, the cybersecurity workfactor is the integral of a brisk flux of step functions punctuated by impulses. My ancestor refined his craft without having to address a change in walnut or steel or linseed. The refinement of craft in cybersecurity is not so easy.

Forensics might at first seem to be a simple effort to explain the past, and thus an affectation. It is not, and the reason is complexity. Complexity is cumulative and, as the authors say at the outset, enough has accumulated that it is impossible to know everything about even a *de minimus* network. Forensics' purpose, then, is to discover meaningful facts in and about the network and the infrastructure that were not previously known. Only after those facts are known is there any real opportunity to improve the future.

Forensics is a craft. Diligence can and does improve its practice. The process of forensic discovery is dominated by ruling out potential explanations for the events under study. Like sculpture, where the aim is to chip away all the stone that doesn't look like an elephant, forensics chips away all the ways in which what was observed didn't happen. In the terms popularized by EF Schumacher, forensics is a convergent problem where cybersecurity is a divergent one; in other words, as more effort is put into forensics, the solution set tends to converge to one answer, an outcome that does not obtain for the general cybersecurity problem.

Perhaps we should say that forensics is not a security discipline but rather an insecurity discipline. Security is about potential events, consistent with Peter Bernstein's definition: "Risk is simply that more things can happen than will." Forensics does not have to induce all the possibilities that accumulated complexity can concoct, but rather to deduce the path by which some part of the observable world came to be as it is. Whereas, in general, cybersecurity the offense has a permanent structural advantage, in forensics it is the defense that has superiority.

That forensics is a craft and that forensics holds an innate strategic advantage are factual generalities. For you, the current or potential practitioner, the challenge is to hone your craft to where that strategic advantage is yours—not just theoretically but in operational reality. For that you need this book.

It is the duty of teachers to be surpassed by their students, but it is also the duty of the student to surpass their teacher. The teachers you have before you are very good; surpassing

1. "WordNet Search—3.1," Princeton University, 2011. http://wordnetweb.princeton.edu/perl/webwn.

them will be nontrivial. In the end, a surpassing craft requires knowing what parts of your then current toolbox are eternal and which are subject to the obsolescence that comes with progress. It is likewise expeditious to know what it is that you don't know. For that, this book's breadth is directly useful.

Because every forensics investigation is, in principle, different, the tools that will be needed for one study may well be a different set from those needed for another study. The best mechanics have all the specialized tools they can need, but may use a few tools far more than others. A collection of tools is only so good as your knowledge of it as a collection of tools, not necessarily that you've used each tool within the last week. Nicholas Taleb described the library of Umberto Eco as an anti-library that "...should contain as much of what you do not know as your financial means, mortgage rates, and the real-estate market allows you to put there."

You, dear reader, hold just such an anti-library of forensics in your hand. Be grateful, and study hard.

Daniel E. Geer, Jr., Sc.D.

Preface

Every day, more bits of data flow across the Internet than there are grains of sand on all the beaches in the world. According to the Cisco Visual Networking Index, the total global IP traffic for 2011 was forecast to be approximately 8.4 * 10^{18} bits per day. Meanwhile, mathematicians at the University of Hawaii have estimated the number of grains of sand on all the beaches in the world to be approximately 7.5 * 10^{18} grains. According to Cisco, global IP traffic is expected to increase at an annual growth rate of 32% per year, so by the time you read this, the number of bits of data flowing across the Internet *every day* may have *far* exceeded the estimated number of grains of sand on all the beaches in the world.[2, 3, 4]

Of course, these estimates are very rough, because in both cases the systems involved are far larger and more complex than humanity has the tools to quantify. The Internet has long since passed the point where we can fully analyze and comprehend its workings. We can understand bits and pieces of it and we can make broad generalizations; but the fact is that we humans have already created a monster far more powerful and complex than we can ever fully understand.

In this environment a new, endless field of study has emerged: network forensics. Forensics, in general, is "the application of scientific knowledge to legal problems, especially scientific analysis of physical evidence, as from a crime scene." Network forensics therefore refers to the scientific study of network-based evidence, commonly applied to legal questions. Of course, network forensics is a field of study independent of any specific legal case, and many of the scientific advances, tools, and techniques developed for the purposes of legal investigation can also be used for social study, historical analysis, and scientific exploration of network environments. In this book, we have endeavored to provide a technical foundation that will be practically useful not just for professional network forensic analysts conducting legal investigations, but also for students, independent researchers, and all those who are curious.

0.1 The Changing Landscape

The Internet is constantly changing. As new features are developed in hardware and software, new protocols are implemented to reflect those changes and old protocols are adapted and revised to suit the latest technology. In the past decade, we have seen the emergence of

2. Cisco estimates the total global IP traffic for 2011 at 28,023 petabytes per month. Dividing this by 30 days in one month, we get approximately 8.4 * 10^{18} bits each day.

3. "Networking Solutions," *Cisco*, http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360_ns827_Networking_Solutions_White_Paper.html.

4. Howard McAllister, "Grains of Sand on the World's Beaches," 21st Century Problem Solving, http://www.hawaii.edu/suremath/jsand.html.

protocols for distributed peer-to-peer video chat systems, protocols for conducting surgery on people from thousands of miles away, and protocols for driving robots halfway around the world.

Network forensics can seem daunting to investigators who are familiar with traditional filesystem forensics. There are a relatively small number of filesystem formats in widespread use, compared with hundreds of network protocols. On Windows systems, it is common to find FAT32 or NTFS filesystems. On UNIX/Linux systems, it is common to find ext2, 3, 4, ZFS, or HFS Plus filesystems. In contrast, on any given network you may find Ethernet, 802.11b/g/n, ARP, DHCP, IPv4, IPv6, TCP, UDP, ICMP, TLS/SSL, SMTP, IMAP, DNS, HTTP, SMB, FTP, RTP, and many other protocols.

On the Internet, there is no guarantee that any protocol you encounter will match the documented specifications, or that there are any published specifications in the first place. Furthermore, implementations of protocols change frequently. Manufacturers are not required to adhere to any standards, and so they implement the protocol as suits them best for any particular revision of software or hardware.

Sometimes protocols are developed before their time, and the applications that are built on top of them have not matured to the point where they can support all the features of the protocol. In the interim, the protocol or specific fields within it may go unused, or may be repurposed by vendors, standards committees, or hackers for other uses. Sometimes protocols are replaced because the environment has changed so much that the old protocols no longer work as intended. A perfect example of this is IPv4, which worked well in its original, relatively small environment. IPv4 is designed with 32-bit fields to store the source and destination addresses, which accomodates 2^{32}, or approximately 4.3 billion unique addresses. However, large segments of this address space were allocated to different organizations in the early years of the Internet, when there were relatively few users. Now that over a billion people have connected to the Internet, the 32-bit address space is too limited to accommodate the demand. As a result, IPv6 was developed with a far larger 128-bit address space (2^{128}, or 3.4 x 10^{38} unique addresses). With it have emerged even more protocols, such as Teredo (which is designed to tunnel IPv6 traffic over IPv4-only networks).

Forensics tools also go through changes and revisions as the protocols change. A tool made in 2010 may not correctly interpret a packet capture from 2002, or vice versa. Sometimes the errors can be very subtle, perhaps even undetectable. It is very important for investigators to understand how forensic tools work, and be capable of going down to the lowest layers to verify findings. Network forensics professionals must be highly skilled, highly motivated, and have a great deal of expertise because you can't always rely on tools that other people have written in order to correctly interpret results and perhaps even testify in court.

Compounding these issues is the overwhelming variety of network devices, including routers, switches, application servers, and more. Each system on any given network may have unique configuration, interface, and capabilities. It is not possible for investigators to be familiar with all network devices—or even a significant percentage—including current and past makes and models. Instead, network investigators must be prepared to research and learn about equipment on the fly, while at the same time managing an investigation and projecting an air of confidence. It's a fine balancing act.

Tracking down devices to examine them in the first place can be difficult or even impossible. Anonymity has been a hallmark of the Internet since its early days. While it may be possible to track an IP address down to a remote ISP, it is often impossible to wrangle any further identifying details out of a third party—particularly if they are located in a foreign country with lax information security laws. Even when the device is located inside your own organization, tracking it down depends on the quality of network documentation and logs, which are often not granular enough to suit investigative needs. With the rise of mobile networks, tracking down devices often feels like a game of hide-and-seek, where the mobile user (perhaps even unknowingly) has the upper hand.

The point is that the Internet functions as an ecosystem. It is not controlled by central forces or "designed" in the way one designs a car. When you examine network traffic, there is no telling what you may encounter or whether your tools will properly parse the specific versions of the protocols in your packet capture. When you gather evidence from network devices or reconfigure them, you may have to research the specific make and model to properly understand the interfaces and the sources of evidence. When you track down systems, you may have to wander all over creation chasing a mobile device, or call dozens of ISP contacts and law enforcement officials in multiple countries to pinpoint the source.

There is no specification to which manufacturers are universally bound to adhere, no set of rules that users around the globe must follow, and no manual that can tell you precisely how to conduct your investigation.

0.2 Organization

This book is designed to provide you with a broad overview of the most important topics in network forensics. It is divided into four parts: "Foundation," "Traffic Analysis," "Network Devices and Servers," and "Advanced Topics."

0.2.1 Part I, "Foundation"

Part I, "Foundation," covers the basic concepts of evidence handling, networking, and evidence acquisition. This provides a foundation for more advanced topics, which we cover later in the book. In addition to the topics in these chapters, we strongly recommend that all readers have a good understanding of TCP/IP networking. *TCP/IP Illustrated* by W. Richard Stevens is a fantastic book that we highly recommend as a reference.

Part I includes the following chapters:

- Chapter 1, "Practical Investigative Strategies," presents a myriad of challenges faced by network forensic investigators, introduces important concepts in digital evidence, and lays out a methodology for approaching network-based investigations.
- Chapter 2, "Technical Fundamentals," provides a technical overview of common networking components and their value for the forensic investigator, and presents the concept of protocols and the OSI model in the context of network forensic investigations.
- Chapter 3, "Evidence Acquisition," dives into passive and active evidence acquisition, including hardware and software used for sniffing traffic, as well as strategies for actively collecting evidence from network devices.

0.2.2 Part II, "Traffic Analysis"

Part II, "Traffic Analysis," discusses the many ways that investigators can analyze network traffic. We begin with packet analysis, from examination of protocol headers to payload extraction and reconstruction. Since flow record data retention is becoming commonplace, we subsequently devote a full chapter to statistical flow record analysis. This is followed by an in-depth look at wireless networks and the 802.11 protocol suite. Finally, we discuss network intrusion detection and prevention systems, which are designed to analyze traffic in real time, produce alerts, and in some cases capture packets on the fly.

Part II includes the following chapters:

- Chapter 4, "Packet Analysis," is a comprehensive study of protocols, packets, and flows, and methods for dissecting them.
- Chapter 5, "Statistical Flow Analysis," presents the increasingly important field of statistical flow record collection, aggregation, and analysis.
- Chapter 6, "Wireless: Network Forensics Unplugged," discusses evidence collection and analysis of wireless networks, specifically focusing on the IEEE 802.11 protocol suite.
- Chapter 7, "Network Intrusion Detection and Analysis," is a review of network intrusion prevention and detection systems, which are specifically designed to produce security alerts and supporting evidence.

0.2.3 Part III, "Network Devices and Servers"

Part III, "Network Devices and Servers," covers evidence acquisition and analysis from all kinds of network devices. We begin by discussing event log collection and examination, including challenges and benefits relating to different types of logging architectures. Next, we specifically talk about forensic investigation of switches, routers, and firewalls, which make up the backbone of our networks. Since web proxies have exploded in popularity, and often contain enormously valuable evidence, we close with a detailed discussion of web proxy evidence collection and analysis.

Part III includes the following chapters:

- Chapter 8, "Event Log Aggregation, Correlation, and Analysis," discusses collection and analysis of logs from various sources, including operating systems of servers and workstations (such as Windows, Linux, and UNIX), applications, network equipment, and physical devices.
- Chapter 9, "Switches, Routers, and Firewalls," studies the evidence that can be gathered from different types of network equipment and strategies for collecting it, depending on available interfaces and level of volatility.
- Chapter 10, "Web Proxies," reviews the explosion of web proxies and how investigators can leverage these devices to collect web surfing histories and even cached copies of web objects.

0.2.4 Part IV, "Advanced Topics"

Part IV, "Advanced Topics," includes a discussion of two of the most fascinating topics in network forensics: network tunneling and malware. We review legitimate and covert network tunnels and discuss investigative strategies for dealing with different types of tunnels. To close, we review a history of malware and the corresponding impact on forensic analysis, including the evolution of command-and-control channels, botnets, IDS/IPS evasion, and the advanced persistent threat (APT).

Part IV includes the following chapters:

- Chapter 11, "Network Tunneling," discusses both legitimate and covert network tunnels, methods for recognizing tunnels, and strategies for recovering evidence from tunneled traffic.
- Chapter 12, "Malware Forensics," is a condensed history of malware development, including the evolution of command-and-control channels, botnets, IDS/IPS evasion, and the advanced persistent threat. Along the way, we discuss how malware has changed—and been changed by—forensic investigations.

0.3 Tools

This book is designed to be accessible to a wide audience, and to teach you the fundamental principles and techniques of network forensics. There are many commercial, point-and-click tools that you can also use to reach the same answers, and we briefly touch on a few of those in this book. However, our focus is on tools that are freely available and that can be used to illustrate fundamental techniques. In this way, we hope to give you the ability to understand how forensic tools work at a low level, verify results gleaned from automated tools, and make educated decisions when selecting tools for your investigations.

0.4 Case Studies

Each of the chapters in Parts II, III, and IV includes a detailed case study designed to showcase the tools and techniques discussed in the chapter. You can download the evidence files and practice dissecting them on your own forensic workstation.

The case study evidence files are located here:

`http://lmgsecurity.com/nf/`

They are freely available for your personal use.

0.5 Errata

In any technical book of this size and density, there are bound to be a few errors. We will maintain a list of errata at:

`http://lmgsecurity.com/nf/errata`

If you find an error, we would like to know about it. Please email us at errata@lmgsecurity.com. Check the web site before emailing to make sure the error is not already listed.

0.6 Final Notes

This book is a labor of love. Each chapter required countless hours of research, discussion, debate, and writing. To create the case studies and corresponding packet captures, we first built a laboratory with the equivalent of a small business-sized network, configured and reconfigured it for each exercise, wrote each scenario, and then ran the scenarios over and over again until we got them all exactly right.

It's impossible to count all the late nights and early mornings, flipped circuit breakers and dead hard drives, warm beers and cold pizzas that went into this book. Even though this book is hundreds of pages, we feel that we have only scratched the surface of the very deep field of network forensics. We learned an enormous amount from all the effort, and we hope you do, too.

Acknowledgments

This book would not exist without the support of two widely respected security professionals: Rob Lee and Ed Skoudis. Three years ago, Rob Lee tapped us to create the network forensics curriculum for the SANS Institute. This was the first time that we pooled our joint knowledge on the subject, and actually gave that body of work a name. Since that time, Rob has continued to act as a mentor and friend, constantly pushing us to improve our work, incorporate feedback, and extend the limits of our knowledge. Rob: Thank you for your high standards, your open and honest feedback, and most of all for believing in us. We could not have accomplished this without you.

Ed, in turn, encouraged us to write this book and kindly took the time to introduce us to his editor. His advice on the process proved invaluable. Thank you, Ed, for all of your help and support. We are forever grateful.

Thanks to all the terrific staff at Pearson who put so much time and effort into producing this book, especially our wonderful editor, Chris Guzikowski, as well as Jessica Goldstein, Raina Chrobak, Julie Nahil, Olivia Basegio, Chris Zahn, Karen Gettman, Chuti Prasertsith, John Fuller, and Elizabeth Ryan. A very special thanks as well to the great team at Laserwords for producing this book, especially Patty Donovan for her patience and kindness.

Many thanks to Jonah Elgart for creating the awesome cover illustration. We deeply appreciate the work of Dr. Dan Geer, who kindly wrote the foreword for this book. We are very grateful to our friends and colleagues who conducted technical reviews of the book: Michael Ford, Erik Hjelmvik, Randy Marchany, Craig Wright, and Joshua Wright. Their perspectives and attention to detail made this book infinitely better.

We would like to give a shout-out to our fantastic crew at LMG Security, particularly: Eric Fulton, Jody Miller, Randi Price, Scott Fretheim, David Harrison, and Diane Byrne. Each of them devoted many hours to helping us with network forensics research and curriculum. Eric Fulton was instrumental in developing several of the ForensicsContest.com puzzles, which formed the basis for some of the case studies (in particular, "HackMe" and "Ann's Aurora"). Jody Miller became our He-Man when he swooped in and conquered the evil forces of Skeletor—*ahem*, we mean, formatted all of the footnotes for the book (nearly 500!).

Thanks to our friends, colleagues, and mentors who have taught us so much over the years: Shane Vannatta, Marsha and Bill Dahlgren, Pohl Longsine, Gary Longsine, John Strand, Michael P. Kenny, Gary and Pue Williams, the good folks at Modwest, Mike Poor, Kevin Johnson, Alan Ptak, Michael Grinberg, Sarah and Kerry Metlen, Anissa Schroeder, Bradley Coleman, Blake Brasher, Stephanie Henry, Nadia Madden and Jon McElroy, Clay Ward, the MIT Student Information Processing Board (SIPB), Wally Deschene, Steven and Linda Abate, Karl Reinhard, Brad Cool, Nick Lewis, Richard Souza, Paul Asadoorian, Larry Pesce, George Bakos, Johannes Ullrich, Paul A. Henry, Rick Smith, Guy Bruneau, Lenny Zeltser, Eric Cole, Judy Novak, Alan Tu, Fabienne van Cappel, Robert C de Baca, Mark Galassi, and Dan Starr.

Special thanks to the staff and faculty of the SANS Institute, in particular: Steven and Kathy Northcutt; Deb Jorgensen; Katherine Webb Calhoon; Lana Emery; Kate Marshall; Velda Lempka; Norris Campbell; and Lynn Lewis.

We would also like to thank everyone who contributed to ForensicsContest.com, whether by creating tools and writeups, submitting comments, or just playing the games. We have learned so much from all of you!

Thanks to our wonderful families for their encouragement and support while we were writing this book, especially Sheila Temkin Davidoff, E. Martin Davidoff, Philip and Lynda Ham, Barbara and Larry Oltmanns, Laura Davidoff, Michele, Kirk, and Naomi Robertson, Latisha, Mike, Makenna, and Braelyn Monnier, Chad, Amy, and Brady Rempel, Sheryl and Tommy Davidoff, Jonathan and Stefanie Davidoff, Jill and Jake Dinov, Jamie and Adam Levine, Annabelle Temkin, Norman and Eileen Shoenfeld, Brian and Marie Shoenfeld, and Debbie Shoenfeld.

Thanks to our little cat, Shark, who stayed curled up by our side for hundreds of hours as we wrote.

Most important of all: Thanks to our two daughters, Charlie and Violet. This book is for you.

About the Authors

Sherri Davidoff has over a decade of experience as an information security professional, specializing in penetration testing, forensics, social engineering testing, and web application assessments. She has consulted for a wide variety of industries, including banking, insurance, health care, transportation, manufacturing, academia, and government. Sherri is a course author for the SANS Institute and has published many articles on privacy and security. She is a GIAC-certified forensic examiner (GCFA) and penetration tester (GPEN), and holds her degree in computer science and electrical engineering from MIT.

Jonathan Ham specializes in large-scale enterprise security issues, from policy and procedure, to scalable prevention, detection, and response techniques. He's been commissioned to teach NCIS investigators how to use Snort, performed packet analysis from a facility more than 2,000 feet underground, taught intrusion analysis to the NSA, and chartered and trained the CIRT for one of the largest U.S. civilian Federal agencies. Jonathan has helped his clients achieve greater success for over fifteen years. He is a Certified Instructor with the SANS Institute, and regularly teaches courses on network security.

Sherri and Jonathan are principals of the security consulting firm, LMG Security. They are married and live in Montana, where they spend their days cracking jokes about TCP/IP and raising their two beautiful daughters, Charlie and Violet.

Source: Photo by Jonathan Qualben Photography, jonathanqualben.com.

This page intentionally left blank

This page intentionally left blank

Chapter 1

Practical Investigative Strategies

> *"A victorious army first wins and then seeks battle; a defeated army first battles and then seeks victory." –Sun Tsu,* The Art of War[1]

Evidence scattered around the world. Not enough time. Not enough staff. Unrealistic expectations. Internal political conflicts. Gross underestimation of costs. Mishandling of evidence. Too many cooks in the kitchen. Network forensic investigations can be tricky. In addition to all the challenges faced by traditional investigators, network forensics investigators often need to work with unfamiliar people in different countries, learn to interact with obscure pieces of equipment, and capture evidence that exists only for fleeting moments. Laws surrounding evidence collection and admissibility are often vague, poorly understood, or nonexistent. Frequently, investigative teams find themselves in situations where it is not clear who is in charge, or what the team can accomplish.

An organized approach is key to a successful investigation. While this book is primarily designed to explore technical topics, in this chapter, we touch on the fundamentals of investigative management. This is particularly important because network forensics investigations often involve coordination between multiple groups and evidence that may be scattered around the globe.

We begin by presenting three cases from different industries in order to give you some examples of how network forensics is used to support investigations in the real world. Next, we explore the fundamentals of evidence collection and distinctions that are made between different types of evidence in many courts. We discuss the challenges specific to network-based evidence, such as locating evidence on a network and questions of admissibility. Finally, we present the OSCAR investigative methodology, which is designed to give you an easy-to-remember framework for approaching digital investigations.

1.1 Real-World Cases

How is network forensics used in real life? In this section, we present three cases:

- Hospital Laptop Goes Missing
- Catching a Corporate Pirate
- Hacked Government Server

1. Sun Tsu (Author), Lionel Giles (Translator), *The Art of War,* El Paso Norte Press, 2005.

These cases have been chosen to provide examples of common IT security incidents and illustrate how network forensics is frequently used. Although these cases are based on real-life experiences, they have been modified to protect the privacy of the organizations and individuals involved.

1.1.1 Hospital Laptop Goes Missing

A doctor reports that her laptop has been stolen from her office in a busy U.S. metropolitan hospital. The computer is password-protected, but the hard drive is not encrypted. Upon initial questioning, the doctor says that the laptop may contain copies of some patient lab results, additional protected health information (PHI) downloaded from email attachments, schedules that include patient names, birth dates, and IDs, notes regarding patient visits, and diagnoses.

1.1.1.1 Potential Ramifications

Since the hospital is regulated by the United States' Health Information Technology for Economic and Clinical Health (HITECH) Act and Health Insurance Portability and Accountability Act (HIPAA), it would be required to notify individuals whose PHI was breached.[2] If the breach is large enough, it would also be required to notify the media. This could cause significant damage to the hospital's reputation, and also cause substantial financial loss, particularly if the hospital were held liable for any damages caused due to the breach.

1.1.1.2 Questions

Important questions for the investigative team include:

1. Precisely when did the laptop go missing?
2. Can we track down the laptop and recover it?
3. Which patient data was on the laptop?
4. How many individuals' data was affected?
5. Did the thief leverage the doctor's credentials to gain any further access to the hospital network?

1.1.1.3 Technical Approach

Investigators began by working to determine the time when the laptop was stolen, or at least when the doctor last used it. This helped establish an outer bound on what data *could* have been stored on it. Establishing the time that the laptop was last in the doctor's possession also gave the investigative team a starting point for searching physical surveillance footage and access logs. The team also reviewed network access logs to determine whether the laptop was subsequently used to connect to the hospital network after the theft and, if so, the location that it connected from.

---

2. "HITECH Breach Notification Interim Final Rule," U.S. Department of Health and Human Services, http://www.hhs.gov/ocr/privacy/hipaa/understanding/coveredentities/breachnotificationifr.html.

There are several ways investigators could try to determine what time the laptop went missing. First, they could interview the doctor to establish the time that she last used it, and the time that she discovered it was missing. Investigators might also find evidence in wireless access point logs, Dynamic Host Control Protocol (DHCP) lease assignment logs, Active Directory events, web proxy logs (if there is an enterprise web proxy), and of course any sort of laptop tracking software (such as Lojack for Laptops) that might have been in use on the device.

Enterprise wireless access point (WAP) logs can be especially helpful for determining the physical location in the facility where a mobile device was most recently connected, and the last time it was connected. In order to ensure uniform availability of wireless networks, enterprises typically deploy a fleet of WAPs that all participate in the same network. Although they appear to the end user as a single wireless network, network operators can view which mobile devices were connected to specific access points throughout the building. Some enterprises even have commercial software that can graphically represent the movement of wirelessly connected devices as they traverse the physical facility. If the laptop was still connected to the hospital's wireless network as the thief exited the building, investigators might be able to use wireless access point logs to show the path that the thief navigated as he or she exited the building. This might also be correlated with video surveillance logs or door access control logs.

Once investigators established an approximate time of theft, they could narrow down the patient information that might have been stored on the system. Email logs could reveal when the doctor last checked her email, which would place an outer bound on the emails that could have been replicated to her laptop. These logs might also reveal which attachments were downloaded. More importantly, the hospital's email server would have copies of all of the doctor's emails, which would help investigators gather a list of patients likely to have been affected by the breach. Similarly, hospital applications that provide access to lab results and other PHI might contain access logs, which could help investigators compile a list of possible data breach victims.

There might be authentication logs from Active Directory domain controllers, VPN concentrators, and other sources that indicate the laptop was used to access hospital resources even after the theft. If so, this might help investigators track down the thief. Evidence of such activities could also indicate that additional information was compromised, and that the attacker's interests went beyond merely gaining a new laptop.

1.1.1.4 Results

Leveraging wireless access point logs, the investigative team was able to pinpoint the time of the theft and track the laptop through the facility out to a visitor parking garage. Parking garage cameras provided a low-fidelity image of the attacker, a tall man wearing scrubs, and investigators also correlated this with gate video of the car itself as it left the lot with two occupants. The video was handed to the police, who were able to track the license plate. The laptop was eventually recovered amongst a stack of stolen laptops.

The investigative team carefully reviewed VPN logs and operating system logs stored on the central logging server and found no evidence that the doctor's laptop was used to attempt any further access to hospital IT resources. Hard drive analysis of the recovered laptop showed no indication that the system had been turned on after the theft. After extensive consultation with legal counsel, hospital management concluded that patient data had ultimately not been leaked.

In response to the incident, the hospital implemented full-disk encryption for all laptop hard drives, and deployed physical laptop locking mechanisms.

1.1.2 Catching a Corporate Pirate

GlobalCorp, Inc., has a centrally managed intrusion detection system, which receives alerts from sites around the world. Central security staff notice an alert for peer-to-peer (P2P) file-sharing, and on closer inspection see filename references to movies that are still in theaters. Fearing legal ramifications of inaction, they investigate.

1.1.2.1 Potential Ramifications

Management at GlobalCorp, Inc., were highly concerned that an employee was using the company network for trafficking in pirated intellectual property. If this activity were detected, the owner of the intellectual property might sue the company. This case occurred in 2003, at the height of Digital Millennium Copyright Act (DMCA) fervor, and it was assumed that if an individual within the company was illicitly trading pirated music or movies, then it could place the company at risk of costly legal battles.

1.1.2.2 Questions

Important questions for the investigative team include:

1. Where is the source of the P2P traffic physically located?
2. Which user is initiating the P2P traffic?
3. Precisely what data is being shared?

1.1.2.3 Technical Approach

Using the IP address from the IDS alerts, investigators identified the physical site that was the source of the traffic. In order to specifically identify the client system, its location, and primary user, investigators worked with local network management staff.

Meanwhile, intrusion analysts in the central office began capturing all of the P2P-related packets involving the IP address in question. The local facility confirmed that this IP address was part of a local DHCP pool on the wired local area network (LAN). Intrusion analysts reviewed DHCP lease assignment logs for relevant time periods, and recovered the media access control (MAC) address associated with the suspicious activity. From the MAC address investigators identified the manufacturer of the network card (Dell, in this case).

In order to trace the IP address to a specific office, local networking staff logged into switches and gathered information mapping the IP address to a physical port. The physical port was wired to a cubicle occupied by an email system administrator. Investigators entered his office after hours one evening and recovered his desktop for forensic analysis.

Upon examination, however, it was clear that the confiscated desktop was not the source of the P2P activity. The MAC address of the network card in the confiscated system (a Hewlett-Packard desktop) was not consistent with the MAC address linked to the suspicious activity. Subsequent analysis of the company's email server produced evidence that the suspect, an email system administrator, had leveraged privileged access to read emails of key networking staff involved in the investigation.

Local networking staff took caution to communicate out-of-band while coordinating the remainder of the investigation. Investigators conducted a thorough search of the premises for a system with the MAC address implicated. The matching desktop was eventually found in the desktop staging area, buried in a pile of systems queued for reimaging.

1.1.2.4 Results

Network forensic analysts examined full packet captures grabbed by the IDS, and were ultimately able to carve out video files and reconstruct playable copyrighted movies that were still in theaters. Hard drive analysis of the correct desktop produced corroborating evidence that the movies in the packet capture had been resident on the hard drive. The hard drive also contained usernames and email addresses linking the hard drive and associated network traffic with the suspect.

Case closed!

1.1.3 Hacked Government Server

During a routine antivirus scan, a government system administrator was alerted to suspicious files on a server. The files appeared to be part of a well-known rootkit. The server did not host any confidential data other than password hashes, but there were several other systems on the local subnet that contained Social Security numbers and financial information of thousands of state residents who had filed for unemployment assistance. The administrative account usernames and passwords were the same for all servers on the local subnet.

1.1.3.1 Potential Ramifications

State laws required the government to notify any individuals whose Social Security numbers were breached. If the servers containing this sensitive information were hacked, the state might be required to spend large amounts of money to send out notifications, set up hotlines for affected individuals, and engage in any resulting lawsuits. In addition, disclosure of a breach might damage the careers of high-ranking elected state officials.

1.1.3.2 Questions

Important questions for the investigative team include:

- Was the server in question truly compromised?
- If so, how was the system exploited?
- Were any other systems on the local network compromised?
- Was any confidential information exported?

1.1.3.3 Technical Approach

The server in question appeared to contain files with names that fit the pattern for a well-known rootkit. Investigators began by examining these files and concluded that they were, indeed, malicious software. The rootkit files were found in the home directory of an old local administrator account that staff had forgotten even existed.

Investigators found that the local authentication logs had been deleted. Fortunately, all servers on the subnet were configured to send logs to a central logging server, so instead investigators reviewed Secure Shell (SSH) logs from the central logging server that were associated with the account. From the SSH logs, it was clear that the account had been the target of a brute-force password-guessing attack. Investigators used visualization tools to identify the times that there were major spikes in the volume of authentication attempts. A subsequent password audit revealed that the account's password was very weak.

The SSH logs showed that the source of the brute-force attack was a system located in Brazil. This was surprising to IT staff because according to network documentation the perimeter firewall was supposed to be configured to block external access to the SSH port of servers on the subnet under investigation. Investigators gathered copies of the current, active firewall configuration and found that it did not match the documented policy—in practice, the SSH port was directly accessible from the Internet. Subsequently, investigators analyzed firewall logs and found entries that corroborated the findings from the SSH logs.

When one system in the environment is compromised, there is a significant probability that the attacker may use credentials from that system to access other systems. IT staff were concerned that the attacker might have used the stolen account credentials to access other systems on the local subnet.

Fortunately, further analysis of the server hard drive indicated that the attacker's access was short-lived; the antivirus scan had alerted on the suspicious files shortly after they were created. Investigators conducted a detailed analysis of authentication logs for all systems on the local subnet, and found no other instances of suspicious access to the other servers. Furthermore, there were no records of logins using the hacked account on any other servers. Extensive analysis of the firewall logs showed no suspicious data exportation from any servers on the local subnet.

1.1.3.4 Results

Investigators concluded that the server under investigation was compromised but that no other systems on the local subnet had been exploited and no personal confidential information had been breached. To protect against future incidents, the state IT staff corrected the errors in the firewall configuration and implemented a policy in which firewall rules were audited at least twice per year. In addition, staff removed the old administrator account and established a policy of auditing all server accounts (including privileges and password strength) on a quarterly basis.

1.2 Footprints

When conducting network forensics, investigators often work with live systems that cannot be taken offline. These may include routers, switches, and other types of network devices, as well as critical servers. In hard drive forensics, investigators are taught to minimize system modification when conducting forensics. It is much easier to minimize system modification when working with an offline copy of a write-protected drive than with production network equipment and servers.

In network forensics, investigators also work to minimize system modification due to forensic activity. However, in these cases investigators often do not have the luxury of an offline copy. Moreover, network-based evidence is often highly volatile and must be collected through active means that inherently modify the system hosting the evidence. Even when investigators are able to sniff traffic using port monitoring or tapping a cable, there is always some impact on the environment, however small. This impact can sometimes be minimized through careful selection of acquisition techniques, but it can never be eliminated entirely.

Every interaction that an investigator has with a live system modifies it in some way, just as an investigator in real life modifies a crime scene simply by walking on the floor. We use the term "footprint" throughout this book to refer to the impact that an investigator has on the systems under examination.

You will always leave a footprint. Often, the size of the footprint required must be weighed against the need for expediency in data collection. Take the time to record your activities carefully so that you can demonstrate later that important evidence was not modified. Always be conscious of your footprint and tread lightly.

1.3 Concepts in Digital Evidence

What is evidence? The *Compact Oxford English Dictionary* defines "evidence" as:[3]

> evidence (noun)
>
> 1. information or signs indicating whether a belief or proposition is true or valid.
> 2. information used to establish facts in a legal investigation or admissible as testimony in a law court.

In this book, we are concerned with both of the above definitions. Our goal in many investigations is to compile a body of evidence suitable for presentation in court proceedings (even if we hope never to end up in court!). Both relevance to the case and admissibility are important, but the first goal is to ascertain the facts of the matter and understand truly and correctly what has transpired.

Consequently, we define "evidence" in the broadest sense as any observable and recordable event, or artifact of an event, that can be used to establish a true understanding of the cause and nature of an observed occurrence.

Of course, it's one thing to be able to reconstruct and understand the events that comprise an occurrence, and yet another to be able to demonstrate that in such a way that

3. "Oxford Dictionaries Online—English Dictionary and Language Reference," *Oxford Dictionaries*, http://www.askoxford.com/concise_oed/evidence?view=uk.

victims can be justly compensated and perpetrators justly punished within our legal framework. Within this system there are a few categories of evidence that have very specific meanings:

- Real
- Best
- Direct
- Circumstantial
- Hearsay
- Business Records

We'll take each of these in turn and discuss their nature and relative usefulness and importance. Due to the rising popularity of electronic communications systems, we also include the following general categories of evidence:

- Digital
- Network-Based Digital

In this book, our discussion of evidence is based on the United States common law system and the U.S. Federal Rules of Evidence (FRE).[4] Many of these concepts may be similar in your jurisdiction, although we also recommend that you familiarize yourself with the rules specific to your region of the world.

1.3.1 Real Evidence

What is "real" evidence? "Real evidence" is roughly defined as any physical, tangible object that played a relevant role in an event that is being adjudicated. It is the knife that was pulled from the victim's body. It is the gun that fired the bullet. It is the physical copy of the contract that was signed by both parties. In our realm it is also the physical hard drive from which data is recovered, and all the rest of the physical computer components involved.

Real evidence usually comprises the physicality of the event, and as such is often the most easily presented and understood element of a crime. Human beings understand tangible objects much more readily than abstract concepts, such as data comprised of ones and zeros (which are themselves comprised of the presence or absence of magnetization on microscopic bits of a spinning platter). Unless the hard drive was used as a blunt object in an assault, and as a consequence is covered in identifiable traces of blood and hair follicles (DNA is real evidence too), the judge or jury may have a difficult time envisioning the process through which the evidence reached its current state and was preserved.

4. Committee on the Judiciary House (US) and US House Committee on the Judiciary, *Federal Rules of Evidence (December 2011)* (Committee on the Judiciary, 2011), http://judiciary.house.gov/hearings/printers/112th/evidence2011.pdf.

Examples of "real evidence" can include:

- The murder weapon
- The fingerprint or footprint
- The signed paper contract
- The physical hard drive or USB device
- The computer itself—chassis, keyboard, and all

1.3.2 Best Evidence

"Best evidence" is roughly defined as the best evidence that can be produced in court. The FRE states, "To prove the content of a writing, recording, or photograph, the *original* writing, recording, or photograph is required, except as otherwise provided in these rules or by Act of Congress" [emphasis added].[5] If the original evidence is not available, then alternate evidence of its contents may be admitted under the "best evidence rule." For example, if an original signed contract was destroyed but a duplicate exists, then the duplicate may be admissible. However, if the original exists and could be admitted, then the duplicate would not suffice.

Our favorite illustration of the "best evidence rule" comes from Dr. Eric Cole, as presented in his SANS courses: Imagine that a helicopter and a tractor trailer collide on a bridge. Real evidence in this case would be the wreckage, but there is no hope of bringing the real evidence into depositions, much less in front of a jury. In such a case the photographs of the scene comprise the best records that can be brought to court. They will have to suffice, and most often do.

Forensic analysts, lawyers, and jurors have questioned what constitutes "original" evidence in the case of digital evidence. Fortunately, the FRE explicitly addresses this issue, as follows:[6]

> An "original" of a writing or recording means the writing or recording itself or any counterpart intended to have the same effect by the person who executed or issued it. For electronically stored information, "original" means any printout—or other output readable by sight—if it accurately reflects the information. An "original" of a photograph includes the negative or a print from it. (e) A "duplicate" means a counterpart produced by a mechanical, photographic, chemical, electronic, or other equivalent process or technique that accurately reproduces the original.

In other words, a printout from a computer hard drive that accurately reflects the data would normally be considered "original" evidence.

With network forensics, the bits and bytes being presented have been recorded and may be treated in the same way as a photograph of an event. It is as though we've photographed

5. Committee on the Judiciary House (US) and US House Committee on the Judiciary, *Federal Rules of Evidence*, 25.

6. *Ibid.*

the bullet as it traveled through the air. The difference is that network forensic investigators can often reconstruct a forensically identical copy of the entire bullet from the snapshot, rather than just presenting a grainy photograph from which legal teams hope to divine trajectories, masses, and the sending barrel's rifling.

Examples of "best evidence" include:

- A photo of the crime scene
- A copy of the signed contract
- A file recovered from the hard drive
- A bit-for-bit snapshot of a network transaction

1.3.3 Direct Evidence

"Direct evidence" is the testimony offered by a direct witness of the act or acts in question. There are lots of ways that events can be observed, captured, and recorded in the real world, and our court systems try to accommodate most of these when there is relevant evidence in question. Of course, the oldest method is the reportable observation of a fellow human being. This human testimony is classified as "direct evidence," and it remains some of the most utilized forms of evidence, even if it is often disputed and unreliable.

Direct evidence is usually admissible, so long as it's relevant. What other people witnessed can have a great impact on a case.

Examples of "direct evidence" can include:

- "I saw him stab that guy."
- "She showed me an inappropriate video."
- "I watched him crack passwords using John the Ripper and a password file he shouldn't have."
- "I saw him with that USB device."

1.3.4 Circumstantial Evidence

In contrast to "direct evidence," "circumstantial evidence" is evidence that does not directly support a specific conclusion. Rather, circumstantial evidence may be linked together with other evidence and used to deduce a conclusion.

Circumstantial evidence is important for cases involving network forensics because it is "the primary mechanism used to link electronic evidence and its creator."[7] Often, circumstantial evidence is used to establish the author of emails, chat logs, or other digital evidence. In turn, authorship verification is necessary to establish authenticity, which is required for evidence to be admissible in court. The DoJ elaborates:[8]

---

7. Scott M. Giordano, "Electronic Evidence and the Law," *Information Systems Frontiers* 6, no. 2 (June 1, 2004): 165.

8. H. Marshall Jarrett, Director, EOUSA, "Searching and Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations," 203.

> [D]istinctive characteristics like email addresses, nicknames, signature blocks, and message contents can prove authorship, at least sufficiently to meet the threshold for authenticity . . . For example, in United States v. Simpson, 152 F.3d 1241 (10th Cir. 1998), prosecutors sought to show that the defendant had conversed with an undercover FBI agent in an Internet chat room devoted to child pornography. The government offered a printout of an Internet chat conversation between the agent and an individual identified as 'Stavron' and sought to show that 'Stavron' was the defendant . . . 'Stavron' had told the undercover agent that his real name was 'B. Simpson,' gave a home address that matched Simpson's, and appeared to be accessing the Internet from an account registered to Simpson. Further, the police found records in Simpson's home that listed the name, address, and phone number that the undercover agent had sent to 'Stavron.' Accordingly, the government had provided evidence sufficient to support a finding that the defendant was 'Stavron,' and the printout was properly authenticated.

Examples of "circumstantial evidence" can include:

- An email signature
- A file containing password hashes on the defendant's computer
- The serial number of the USB device

1.3.5 Hearsay

"Hearsay" is the label given to testimony offered second-hand by someone who was not a direct witness of the act or acts in question. It is formally defined by the FRE as "a statement, other than one made by the declarant while testifying at the trial or hearing, offered in evidence to prove the truth of the matter asserted." This includes the comments of someone who may have direct knowledge of an occurrence, but who is unable or unwilling to deliver them directly to the court. Hearsay is generally not ruled admissible, unless it falls into the category of an exception as listed in the FRE (Rules 803 and 804).

Digital evidence can be classified as hearsay if it contains assertions created by people. The U.S. Department of Justice cites "a personal letter; a memo; bookkeeping records; and records of business transactions inputted by persons" as examples of digital evidence that would be classified as hearsay.[9]

However, digital evidence that is generated by a fully automated process with no human intervention is generally *not* considered heresay. The Department of Justice explains:[10]

> Computer-generated records that do not contain statements of persons therefore do not implicate the hearsay rules. This principle applies both to records generated by a computer without the involvement of a person (e.g., GPS tracking records) and to computer records that are the result of human conduct other than assertions (e.g., dialing a phone number or punching in a PIN at an ATM).

9. H. Marshall Jarrett, Director, EOUSA, "Searching and Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations," 192.

10. H. Marshall Jarrett, Director, EOUSA, "Searching and Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations," 193.

In some cases, courts have admitted digital evidence using the "business records" exception of the hearsay rule, which we discuss further in the next section. However, the Department of Justice points out that in these cases, the courts overlooked the question of whether the digital evidence should have been classified as hearsay in the first place. "Increasingly ... courts have recognized that many computer records result from a process and are not statements of persons—they are thus not hearsay at all."[11]

Examples of "hearsay" can include:

- "The guy told me he did it."
- "He said he knew who did it, and could testify."
- "I saw a recording of the whole thing go down."
- A text file containing a personal letter

1.3.6 Business Records

Business records can include any documentation that an enterprise routinely generates and retains as a result of normal business processes, and that is deemed accurate enough to be used as a basis for managerial decisions. The FRE specifically exempts business records from the rule that hearsay is inadmissible, stating that:[12]

> The following are not excluded by the rule against hearsay, regardless of whether the declarant is available as a witness: [...] A record of an act, event, condition, opinion, or diagnosis if ... the record was kept in the course of a regularly conducted activity of a business, organization, occupation, or calling, whether or not for profit ...

This can include everything from email and memos to access logs and intrusion detection system (IDS) reports. There may be legally mandated retention periods for some of this data. Other records may be subject to internal retention and/or destruction policies. The bottom line is that if the records are seen as accurate enough by the enterprise that they are the basis for managerial decision making, then the courts usually deem them reliable enough for a proceeding.

Digital evidence has been admitted under the "business records" exception to hearsay many times, although in some cases this was erroneous. The Department of Justice points out that "courts have mistakenly assumed that computer-generated records are hearsay without recognizing that they do not contain the statement of a person."

Examples of "business records" can include:

- Contracts and other employment agreements
- Invoices and records of payment received

---

11. H. Marshall Jarrett, Director, EOUSA, "Searching and Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations," 191.

12. Committee on the Judiciary House (US) and US House Committee on the Judiciary, *Federal Rules of Evidence*, 17.

- Routinely kept access logs
- /var/log/messages

1.3.7 Digital Evidence

"Digital evidence" is any documentation that satisfies the requirements of "evidence" in a proceeding, but that exists in electronic digital form. Digital evidence may rest in microscopic spots on spinning platters, magnetized to greater or lesser degrees in a somewhat nonvolatile scheme, but regardless, unintelligible except through multiple layers of abstraction and filesystem protocols. In other cases, digital evidence may be charges held in volatile storage, which dissipate within seconds of a loss of power to the system. Digital evidence may be no more tangible, nor permanent, than pulses of photons, radio frequency waves, or differential levels of voltage on copper wires.

Naturally, digital evidence poses challenges for investigators seeking to preserve it and attorneys seeking to admit it in court. In order for evidence to be admissible in United States federal courts, digital evidence must adhere to the same standards as other types of evidence: it must be deemed relevant to the case and authentic. "The standard for authenticating computer records is the same as for authenticating other records . . . ," wrote the U.S. Department of Justice (DoJ) in 2009. "Importantly, courts have rejected arguments that electronic evidence is inherently unreliable because of its potential for manipulation. As with paper documents, the mere possibility of alteration is not sufficient to exclude electronic evidence. Absent specific evidence of alteration, such possibilities go only to the evidence's weight, not admissibility."[13]

Examples of "digital evidence" include:

- Emails and IM sessions
- Invoices and records of payment received
- Routinely kept access logs
- /var/log/messages

1.3.8 Network-Based Digital Evidence

"Network-based digital evidence" is digital evidence that is produced as a result of communications over a network. The primary and secondary storage media of computers (e.g., the RAM and hard drives) tend to be fruitful fodder for forensic analysis. Due to data remanence, persistent storage can retain forensically recoverable and relevant evidence for hours, days, even years beyond file deletion and storage reuse. In contrast, network-based digital evidence can be extremely volatile. Packets flit across the wire in milliseconds, vanish from switches in the blink of an eye. Web sites change depending on from where they're viewed and when.

The requirements for admissibility of network-based digital evidence are murky. Often, the source that generated the evidence is not obtainable or cannot be identified. When the

---

13. H. Marshall Jarrett, Director, EOUSA, "Searching and Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations" (Office of Legal Education Executive Office for United States Attorneys, 2009), 198–202, http://www.justice.gov/criminal/cybercrime/ssmanual/ssmanual2009.pdf.

evidence is a recording of a chat log, blog posting, or email, the identity of the parties in the conversation (and therefore the authors of the statements) may be difficult to prove. When the evidence is a web site, the litigant may need to provide supporting evidence to demonstrate that the image presented in court is what actually existed at the time and location that it was supposedly viewed. For example, "[s]everal cases have considered what foundation is necessary to authenticate the contents and appearance of a website at a particular time. Print-outs of web pages, even those bearing the URL and date stamp, are not self-authenticating. . . . Thus, courts typically require the testimony of a person with knowledge of the website's appearance to authenticate images of that website."[14]

There is little case precedent on the admissibility of network packet captures. Depending on the method of capture and the details of the case, packet captures of network traffic may be treated as *recordings* of events, similar to a taped conversation.

Examples of "network-based digital evidence" can include:

- Emails and IM sessions
- Browser activity, including web-based email
- Routinely kept packet logs
- /var/log/messages

1.4 Challenges Relating to Network Evidence

Network-based evidence poses special challenges in several areas, including acquisition, content, storage, privacy, seizure, and admissibility. We will discuss some common challenges below.

- **Acquisition** It can be difficult to locate specific evidence in a network environment. Networks contain so many possible sources of evidence—from wireless access points to web proxies to central log servers—that sometimes pinpointing the correct location of the evidence is tricky. Even when you do know where a specific piece of evidence resides, you may have difficulty gaining access to it for political or technical reasons.
- **Content** Unlike filesystems, which are designed to contain all the contents of files and their metadata, network devices may or may not store evidence with the level of granularity desired. Network devices often have very limited storage capacity. Usually, only selected metadata about the transaction or data transfer is kept instead of complete records of the data that traversed the network.
- **Storage** Network devices commonly do not employ secondary or persistent storage. As a consequence, the data they contain may be so volatile as to not survive a reset of the device.
- **Privacy** Depending on jurisdiction, there may be legal issues involving personal privacy that are unique to network-based acquisition techniques.

14. H. Marshall Jarrett, Director, EOUSA, "Searching and Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations," 204.

- **Seizure** Seizing a hard drive can inconvenience an individual or organization. Often, however, a clone of the original can be constructed and deployed such that critical operations can continue with limited disruption. Seizing a network device can be much more disruptive. In the most extreme cases, an entire network segment may be brought down indefinitely. Under most circumstances, however, investigators can minimize the impact on network operations.
- **Admissibility** Filesystem-based evidence is now routinely admitted in both criminal and civil proceedings. As long as the filesystem-based evidence is lawfully acquired, properly handled, and relevant to the case, there are clear precedents for authenticating the evidence and admitting it in court. In contrast, network forensics is a newer approach to digital investigations. There are sometimes conflicting or even nonexisting legal precedents for admission of various types of network-based digital evidence. Over time, network-based digital evidence will become more prevalent and case precedents will be set and standardized.

1.5 Network Forensics Investigative Methodology (OSCAR)

Like any other forensic task, recovering and analyzing digital evidence from network sources must be done in such a way that the results are both reproducible and accurate. In order to ensure a useful outcome, forensic investigators should perform our activities within a methodological framework. The overall step-by-step process recommended in this book is as follows:

- **O**btain information
- **S**trategize
- **C**ollect evidence
- **A**nalyze
- **R**eport

We refer to this methodology as "OSCAR," and walk through each of these steps in the following section.

1.5.1 Obtain Information

Whether you're law enforcement, internal security staff, or a forensic consultant, you will always need to do two things at the beginning of an investigation: obtain information about the incident itself, and obtain information about the environment.

1.5.1.1 The Incident

Usually you will want to know the following things about the incident:

- Description of what happened (as is currently known)
- Date, time, and method of incident discovery

- Persons involved
- Systems and data involved
- Actions taken since discovery
- Summary of internal discussions
- Incident manager and process
- Legal issues
- Time frame for investigation/recovery/resolution
- Goals

This list is simply a starting point, and you will need to customize it for each incident.

1.5.1.2 The Environment

The information you gather about the environment will depend on your level of familiarity with it. Remember that every environment is constantly changing, and complex social and political dynamics occur during an incident. Even if you are very familiar with an organization, you should always take the time to understand how the organization is responding to this particular incident, and clearly establish who needs to be kept in the loop. Usually you will want to know the following things about the environment:

- Business model
- Legal issues
- Network topology (request a network map, etc. if you do not have one)
- Available sources of network evidence
- Organizational structure (request an organizational chart if you do not have one)
- Incident response management process/procedures (forensic investigators are part of the response process and should be at least basically familiar with it)
- Communications systems (is there a central incident communication system/evidence repository?)
- Resources available (staff, equipment, funding, time)

1.5.2 Strategize

It is crucial that early on you take the time to accurately assess your resources and plan your investigation. While this is important for any investigation, it is especially important for network forensics because there are many potential sources of evidence, some of which are also very volatile. Investigators must work efficiently. You will want to regularly confer with others on the investigative/incident response team while planning and conducting the investigation to ensure that everyone is working in concordance and that important developments are communicated.

Source of Evidence	Likely Value	Effort	Volatility	Priority
Firewall logs	High	Medium	Low	2
Web proxy cache	High	Low	Medium	1
ARP tables	Low	Low	High	3

Figure 1–1. An example of evidence prioritization. In this example, we list potential sources of evidence, the likely value, the likely effort to obtain it, and the expected volatility. These values will be different for every investigation.

Here are some tips for developing an investigative strategy:

- Understand the goals and time frame of the investigation.
- List your resources, including personnel, time, and equipment.
- Identify likely sources of evidence.
- For each source of evidence, estimate the value and cost of obtaining it.
- Prioritize your evidence acquisition.
- Plan the initial acquisition/analysis.
- Decide upon method and times of regular communication/updates.
- Keep in mind that after conducting your initial analysis, you may decide to go back and acquire more evidence. Forensics is an iterative process.

Figure 1–1 shows an example of evidence prioritization. In this example, the organization collects firewall logs but stores them in a distributed manner on systems that are not easily accessed. The organization has a web proxy, which is centrally accessed by key security staff. ARP tables can be gathered from any system on the local LAN.

The table lists potential sources of evidence, the likely value for the investigation, the expected effort required to obtain the evidence, and the expected volatility. All of these values are unique to each investigation; every organization has different system configurations, data retention policies, and access procedures. Furthermore, the network equipment, investigative resources, and goals of each investigation vary widely.

Based on this information, we can create our evidence spreadsheet and prioritize accordingly. Next, we would develop a plan for evidence acquisition based on our available resources.

1.5.3 Collect Evidence

In the previous step, "Strategize," we prioritized our sources of evidence and came up with an acquisition plan. Based on this plan, we then collect evidence from each source. There are three components you must address every time you acquire evidence:

- **Document**—Make sure to keep a careful log of all systems accessed and all actions taken during evidence collection. Your notes must be stored securely and may be referenced in court. Even if the investigation does not go to court, your notes will still be very helpful during analysis. Be sure to record the date, time, source, method of acquisition, name of the investigator(s), and chain of custody.

- **Capture**—Capture the evidence itself. This may involve capturing packets and writing them to a hard drive, copying logs to hard drive or CD, or imaging hard drives of web proxies or logging servers.
- **Store/Transport**—Ensure that the evidence is stored securely and maintain the chain of custody. Keep an accurate, signed, verifiable log of the persons who have accessed or possessed the evidence.

Since the admissibility of evidence is dependent upon its relevance and reliability, investigators should carefully track the source, method of acquisition, and chain of custody. It's generally accepted that a bit-for-bit image of a hard drive is acceptable in court. For a lot of network-based evidence, the admissibility is not so clear-cut. When in doubt, take careful notes and consult legal counsel.

As with any evidence gathered in the course of an investigation, proper care must be taken to preserve evidence integrity and to document its use and disposition throughout its life cycle (from the initial acquisition to its return to its rightful owner). As we'll see, in some cases this may mean documenting and maintaining the physical chain of custody of a network device. However, in many cases the original incarnation of the evidence being acquired will never be taken into custody.

1.5.3.1 Tips for Evidence Collection

Best practices for evidence collection include:

- Acquire as soon as possible, and lawfully
- Make cryptographically verifiable copies
- Sequester the originals under restricted custody and access (or your earliest copy, when the originals are not available)
- Analyze only the copies
- Use tools that are reputable and reliable
- Document everything you do!

1.5.4 Analyze

Of course the analysis process is normally nonlinear, but certain elements should be considered essential:

- **Correlation** One of the hallmarks of network forensics is that it involves multiple sources of evidence. Much of this will be timestamped, and so the first consideration should be what data can be compiled, from which sources, and how it can be correlated. Correlation may be a manual process, or it may be possible to use tools to do it for you in an automated fashion. We'll look at such tools later on.

- **Timeline** Once the multiple data sources have been aggregated and correlated, it's time to build a timeline of activities. Understanding who did what, when, and how is the basis for any theory of the case. Recognize that you may have to adjust for time skew between sources!
- **Events of Interest** Certain events will stand out as potentially more relevant than others. You'll need to try to isolate the events that are of greatest interest, and seek to understand how they transpired.
- **Corroboration** Due to the relatively low fidelity of data that characterizes many sources of network logs, there is always the problem of "false positives." The best way to verify events in question is to attempt to corroborate them through multiple sources. This may mean seeking out data that had not previously been compiled, from sources not previously consulted.
- **Recovery of additional evidence** Often the efforts described above lead to a widening net of evidence acquisition and analysis. Be prepared for this, and be prepared to repeat the process until such time as the events of interest are well understood.
- **Interpretation** Throughout the analysis process, you may need to develop working theories of the case. These are educated assessments of the meaning of your evidence, designed to help you identify potential additional sources of evidence, and construct a theory of the events that likely transpired. It is of the utmost importance that you separate your interpretation of the evidence from fact. Your interpretation of the evidence is always a hypothesis, which may be proved or disproved.

1.5.5 Report

Nothing you'll have done to this point, from acquisition through analysis, will matter if you're unable to convey your results to others. From that perspective, reporting might be the most important aspect of the investigation. Most commercial forensic tools handle this aspect for the analyst, but usually not in a way that is maximally useful to a lay audience, which is generally necessary.

The report that you produce must be:

- Understandable by nontechnical laypeople, such as:
  - Legal teams
  - Managers
  - Human Resources personnel
  - Judges
  - Juries
- Defensible in detail
- Factual

In short, you need to be able to explain the results of your investigation in terms that will make sense for nontechnical people, while still maintaining scientific rigor. Executive summaries and high-level descriptions are key, but they must be backed by details that can easily be defended.

1.6 Conclusion

Network forensic investigations pose a myriad of challenges, from distributed evidence to internal politics to questions of evidence admissibility. To meet these challenges, investigators must carefully assess each investigation and develop a realistic strategy that takes into account both the investigative goals and the available resources.

We began this chapter with a series of case studies designed to illustrate how network forensic techniques are applied in real life. Subsequently, we reviewed the fundamental concepts in digital evidence, as employed in the United States common law system, and touched upon the challenges that relate specifically to network-based digital evidence. Finally, we provided you with a method for approaching network forensics investigations.

As Sun Tsu wrote 2,500 years ago: "A victorious army first wins and then seeks battle; a defeated army first battles and then seeks victory." Strategize first; then collect your evidence and conduct your analysis. By considering the challenges unique to your investigation up front, you will meet your investigative goals most efficiently and effectively.

This page intentionally left blank

Index