
# Malware Analysis and Detection Engineering

A Comprehensive Approach to Detect and Analyze Modern Malware

Abhijit Mohanta, Anoop Saldanha

*Foreword by Pedram Amini*

Apress®

# Malware Analysis and Detection Engineering

## A Comprehensive Approach to Detect and Analyze Modern Malware

**Abhijit Mohanta, Anoop Saldanha**

***Foreword by Pedram Amini***

Apress®

# *Malware Analysis and Detection Engineering: A Comprehensive Approach to Detect and Analyze Modern Malware*

Abhijit Mohanta
Independent Cybersecurity Consultant,
Bhubaneswar, Odisha, India

Anoop Saldanha
Independent Cybersecurity Consultant,
Mangalore, Karnataka, India

ISBN-13 (pbk): 978-1-4842-6192-7
ISBN-13 (electronic): 978-1-4842-6193-4
https://doi.org/10.1007/978-1-4842-6193-4



Managing Director, Apress Media LLC: Welmoed Spahr
Acquisitions Editor: Celestin Suresh John
Development Editor: Laura Berendson
Coordinating Editor: Divya Modi

Cover designed by eStudioCalamar

Cover image designed by Freepik (www.freepik.com)

Distributed to the book trade worldwide by Springer Science+Business Media New York, 1 New York Plaza, New York, NY 10004. Phone 1-800-SPRINGER, fax (201) 348-4505, e-mail orders-ny@springer-sbm.com, or visit www.springeronline.com. Apress Media, LLC is a California LLC and the sole member (owner) is Springer Science + Business Media Finance Inc (SSBM Finance Inc). SSBM Finance Inc is a **Delaware** corporation.

For information on translations, please e-mail booktranslations@springernature.com; for reprint, paperback, or audio rights, please e-mail bookpermissions@springernature.com.

Apress titles may be purchased in bulk for academic, corporate, or promotional use. eBook versions and licenses are also available for most titles. For more information, reference our Print and eBook Bulk Sales web page at http://www.apress.com/bulk-sales.

Any source code or other supplementary material referenced by the author in this book is available to readers on GitHub via the book's product page, located at www.apress.com/9781484261927. For more detailed information, please visit http://www.apress.com/source-code.

Printed on acid-free paper

# Table of Contents

# About the Authors

**Abhijit Mohanta** is an independent cybersecurity consultant and corporate trainer who has worked extensively in malware reverse engineering, vulnerability research, antivirus engine development, anti-malware signature writing, and sandbox development. He has worked with Symantec, McAfee, and Juniper Networks anti-malware labs. He holds several patents. He blogs regularly and has been a speaker at security conferences and workshops. His articles have been republished and quoted in several blogs and whitepapers, including *eForensics* magazine. He is also the author of the book *Preventing Ransomware: Understand, Prevent, and Remediate Ransomware Attacks* (Packt Publishing, 2018).

**Anoop Saldanha** is one of the core authors of the Suricata Intrusion Detection and Prevention System, funded by the US Department of Homeland Security (DHS). He works as an independent security consultant and as a corporate security trainer. He designs and develops various detection technologies to secure both the host and the network, ranging from network security tools such as IDS/IPS to malware sandboxes, malware analysis tools, firewalls, endpoints, and IoT security tools. He holds multiple patents in the field of security and speaks at security conferences and workshops. He has previously worked in threat research labs and detection engineering teams at RSA Security, Juniper Networks, Cyphort Cybersecurity, and various other cybersecurity startups.

# About the Technical Reviewer

**Ravikant Tiwari** is a cybersecurity professional with in-depth knowledge of malware analysis and reverse engineering. He has more than nine years of experience in the antivirus industry. He has worked for cybersecurity firms such as Comodo Security Solutions, Norman ASA, McAfee, FireEye, and Acronis. He is a certified ethical hacker. His area of expertise includes malware analysis, reverse engineering, signature creation, and security research for designing and developing new features and solutions to enhance the detection capabilities of cybersecurity products. He has designed machine learning models for use in malware and exploit detection. He has been a member of the architect council at McAfee Labs, brainstorming and producing new solutions for McAfee. Currently leading the threat research lab, he is responsible for a multitude of tasks, including automation, producing malware detection rules, and developing a prototype for the Acronis Cyber Protect solution. He has written many blogs, articles, and threat reports. He is a speaker at RSA and Total Security conferences.

Occasionally, he provides expert comments and insights on security breaches and major hacks for media houses and newsrooms.

# About the Foreword Author

**Pedram Amini** has spent much of his time in the shoes of a reverse engineer—developing automation tools and processes. In conjunction with his passion, he launched OpenRCE.org, a community website dedicated to the art and science of reverse engineering. He has presented at Black Hat, DEF CON, REcon, Ekoparty, BlueHat, ShmooCon, ToorCon, and Virus Bulletin, and taught numerous sold-out reverse engineering courses. He holds a computer science degree from Tulane University and is co-author of the book *Fuzzing: Brute Force Vulnerability Discovery*.

Pedram focuses the majority of his time on InQuest, whose product provides deep file inspection (DFI) for real-time threat detection and "retrohunting," a novel approach that leverages the power of hindsight to apply today's threat intelligence to yesterday's data. Built by SOC analysts for SOC analysts, InQuest is designed to save enterprises their most limited resource, human cognition. He was formerly a director of software development at Avast after the acquisition of his startup, Jumpshot, a fully automated solution for the removal of deeply entrenched Windows malware infections. He is the founder of the Zero Day Initiative at TippingPoint (acquired by 3Com/HP). He has managed the world's largest group of independent researchers, and he served as the assistant director and one of the founding members of iDEFENSE Labs (acquired by Verisign).

# Acknowledgments

Thanks to our technical reviewer, Ravikant Tiwari, for his time and expertise reviewing more than 800 pages of our book and the various exercises and examples, making sure we have our content accurate to the dot. We'd like to thank Brad from Malware Traffic Analysis (www.malwaretrafficanalysis.com) for permitting us to use his samples. We'd like to thank Hex-rays for providing us licensed versions of the famous IDA Pro tool which we have covered in this book. Special thanks to the authors of various cybersecurity-related tools, without which writing this book would be impossible.

We'd also like to thank everybody at Apress, including the copyediting staff working hard behind the scenes, for their effort in helping with this book and making sure it meets the highest standards. Special thanks to Divya Modi, Matthew Moodie, Laura Berendson, Celestin Suresh John, and Nikhil Karkal in helping us through the various stages of this book development.

**Abhijit Mohanta:** To my dear father, thanks for your encouragement, without which I would never have been confident enough to write down my ideas into a book.

To the love of my life—my dear wife, Shreeti, thank you for being patient with me all these months while I spent hours writing the book.

**Anoop Saldanha:** I'd like to thank my wife, Sonia Carrol, without whose love and patience it would have been impossible for me to write this book. Thanks again to my wife and my daughters, Sadhana and Suvidha, for putting up with my insane schedule while writing this book!

I'd also like to thank my dad, William, and my mom, Nayana, for easing my load and helping me manage my various everyday tasks to free my time up to write this book. Not to mention the often underrated guidance and wisdom they have given me throughout the years.

# Foreword

*This book is a beast! If you're looking to master the ever-widening field of malware analysis, look no further. This is the definitive guide for you.*

Reverse engineering (or reversing) is a fascinating subject and one that I've always had a love affair with. Puzzle lovers and tinkerers alike will find appeal in the art of reversing. Talented practitioners can discover and exploit software vulnerabilities, dissect the intent behind a novel malware sample, and hack a toy like a Big Mouth Billy Bass to operate as an Amazon Echo.

When approached by newcomers looking for advice on how to get started with reversing, I generally recommend that they start with malware analysis. The software targets are smaller than enterprise software and, therefore, more digestible. While code volume is lower, malware can, and will, employ any number of tricks that add hurdles for the analyst. Overcoming these challenges will quickly improve your skillset, and there are fresh malware samples for one to play with daily.

Malware analysts are needed now more than ever. The volume of unique malware, similar to the general volume of Internet-transmitted data, is growing rapidly every year. When I first got into the industry almost 20 years ago, there were hundreds to thousands of samples daily. Today, it's well into the millions. This increase in volume is of some benefit to defenders. Large volumes of data are a requisite for data science. There's tremendous value in machine learning, but it's no silver bullet. Manual analysis is still mandatory and will be for some time to come.

The stakes have never been higher. In 2010, Stuxnet was first discovered, and, to date, it's the most technically impressive piece of software I've ever seen. It is a modular and air-gap jumping worm, armed with four zero-day exploits and targeted toward Iranian nuclear enrichment centrifuges (reportedly ruining almost 20% of them). It is a clear sign of the military-industrial complex engaging on a new frontier. With today's large budgets and a shifting focus to digital, we can certainly expect some similarly sensational headlines in the future.

Authors Abhijit and Anoop have done an incredible job putting together a truly all-encompassing work. I mean, wow, Chapter 16 is a book unto itself! I admire these two seasoned practitioners for making an effort to create such an incredible guide through such a wide field.

Another piece of advice I'm quick to share with folks looking to delve into reversing and malware analysis: you must truly be passionate and be willing to put in the time. To the reader: master the materials in this book, and you'll be ready to join the global resistance against malware.

—Pedram Amini
InQuest CTO and Founder of OpenRCE.org and Zero Day Initiative

# Introduction

As cybersecurity specialists and corporate trainers, we are often contacted by people who say that their organization has been infected by malware, and they want to know what they should do to contain the infection, or they ask how they should secure their systems and network to prevent such attacks. The stories that we hear often follow the same storyline: *There was a malware infection, which our anti-malware product caught, we quarantined the system, cleaned it up, updated our IDS signatures, but now the infection is back, affecting our other systems and our staff.*

When we cross-question, some important questions are often left unanswered.

- Did you figure out the entry point of attack?
- Did you check for any infection spread (a.k.a. lateral movement/ spread across systems from the malware infection point)?
- Did you make sure you were able to figure out all the artifacts from the malware infection and cleaned all of them up?
- Were you able to understand the intention of the cyberattack and the threat actor behind the cyberattack?
- Did you inform your management and give them a full report of the true damage caused by the malware infection?

In most cases, the answers to these questions are not ascertained, leaving holes in the SOC, IR, and forensic stages, which can lead to the infection remaining present in your network. Not knowing the intentions behind the attack and the attacker means that IT and SOC teams are not fully aware of the true impact of the infection, leaving management in a plight to build a plan to prepare for the potential damage to their business and brand because of this infection.

This is exactly where *Malware Analysis and Detection Engineering* comes in. It does not only help you learn how to detect, analyze, and reverse engineer malware, but it also teaches you the importance of effective and efficient workflows.

This book was described by Pedram Amini, the founder of Zero Day Initiative and OpenRCE, as a "beast!". And a beast it is indeed, with more than 900 comprehensive pages of content and exercises. With this book at your fingertips, you should be able to take on any malware that comes your way.

## Malware Analysis and Reverse Engineering

Pretty much any cyberattack involves malware, and the number of such attacks is increasing every day, and attackers are getting bolder as well. Millions of pieces of malware are seen every day, but there aren't enough analysts out there to deal with it all. Malware analysis is an esoteric field that is mastered by only a few. It involves dissecting all types of malware efficiently and masterfully, with minimum expenditure of time and effort, high accuracy, and absolute inference of the malware's intentions. Today, there are various analysts out there, but not all of them have the requisite skill to dissect a piece of malware.

This book incorporates our combined multiyear experiences in the field of cybersecurity. It translates myriad questions and cases and converts them into efficient and understandable material, which should help any analyst learn how to analyze malware systematically by using various unspoken tricks used by industry researchers. The samples in this book largely focus on Windows executables, but we also cover how to analyze and reverse other types of malware, including Microsoft Office macro malware, PowerShell and JavaScript malware, and other scripting malware.

We also introduce in this book a new, open source tool—APIMiner, which we developed while writing this book. It should be a gamechanger for malware analysts and reverse engineers around the world, which should greatly increase the speed and accuracy with which you can analyze malware.

But malware analysis may not be enough for most cases, and we understand this well based on our experience. And this is why we dedicate a section of this on the esoteric topic of reverse engineering. In Chapter 16, which deserves to be a book on its own, we introduce you to the world of x86 assembly and debuggers. We walk you through various tricks to quickly reverse and debug malware. We don't treat reversing as a standalone topic, but instead, teach you how to combine various tools and tricks from malware analysis to make reverse engineering easier.

## Detection Engineering: The Lonely Stepchild

The first thing we discussed when we devised the content for this book was, why hasn't anyone covered how to detect malware? The first part of dealing with any malware infection is to detect the malware infection itself. Then comes analyzing and reverse engineering samples. In our experience with various cybersecurity companies, we have seen that there is a huge gap between detection engineers and malware researchers, which in the end translates to poor detection products. But if you combine the knowledge from these two areas, you will have the skill set to apply the tricks from malware analysis to detect malware samples. At the same time, detection engineering uses various automation and development tools, which, if used effectively, can help malware analysts speedily analyze and reverse malware samples.

To that end, we dedicate Part 6 of this book to detection engineering, taking you through the internals of the most important cybersecurity tools used in this industry: antiviruses, malware sandboxes, network intrusion detection and prevention systems, and binary instrumentation. By covering various detection tools and frameworks, which range from host-based anti-malware tools like antiviruses and binary instrumentation frameworks, to network security tools like IDS/IPS and Suricata, we teach you how to apply the intricate workings of these detection tools to automate your everyday analysis and reversing workflow.

## Hands-on

Ever seen kids take homework home and come back to school the next day with their work completed? That's the exact story of labs at the end of each chapter. And this is precisely why we don't use labs at the end of the chapter and instead incorporate the lab exercises as hands-on exercises in the chapters. You run and inspect examples under our supervision to make sure that you understand every aspect of what you might encounter.

The trouble with real-world malware exercises is that they rush you and place you in a state of panic when you are learning how to analyze them—because malware waits for no one. Our exercises are samples that were developed in-house and exhibit malware behavior under controlled conditions to let you analyze them at your own pace. At the same time, to prepare you for the real world, we have a ton of hands-on, real-world malware exercises throughout the book, allowing you to test the tricks you learned from the simulated samples against real-world samples.