TECH TODAY

THE REIGN OF BOTNETS

Defending Against Abuses, Bots and Fraud on the Internet

DAVID SÉNÉCAL

WILEY

The Reign of Botnets

The Reign of Botnets

Defending Against Abuses, Bots and Fraud on the Internet

David Sénécal

WILEY

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada and the United Kingdom.

ISBNs: 9781394262410 (Paperback), 9781394262427 (ePDF), 9781394262434 (ePub)

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services, please contact our Customer Care Department within the United States at (800) 762- 2974, outside the United States at (317) 572- 3993. For product technical support, you can find answers to frequently asked questions or reach us via live chat at `https://support.wiley.com`.

If you believe you've found a mistake in this book, please bring it to our attention by emailing our reader support team at `wileysupport@wiley.com` with the subject line "Possible Book Errata Submission."

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic formats. For more information about Wiley products, visit our web site at `www.wiley.com`.

Library of Congress Control Number: 2024936163

Cover image: © CSA-Printstock/Getty Images
Cover design: Wiley

For Dana, Daphne, Dawson, and, of course, Mr. Dean!

About the Author

David Sénécal grew up in France, lived in Germany and England, and immigrated to the United States in 2005. He lives with his family in the San Francisco Bay Area, California, and works for Akamai Technologies as a principal product architect. He brings 25+ years of experience working with web performance, security, and enterprise networking technologies through various roles (support, integration, consulting, development, product management, architecture, and research). He started working on bot detection concepts for Akamai in 2010, which became the very successful Bot Manager product, consistently recognized as a market leader by Forrester. He helped define the concept of *bot management* in the early 2010s, which combines bot detection, classification, visibility, and response strategy. The competition and the industry later adopted this term. In his current role, David leads a team of researchers, developers, and architects to keep up with the evolution of attacks and define the next generation of bot and fraud detection products.

Follow David on LinkedIn at `www.linkedin.com/in/davidsenecal`.

About the Technical Editor

Tyson Thomas has been a researcher in application security for eight years, focusing on bot detection algorithms leveraging behavioral biometric, device telemetry, and network data from web and mobile clients. Originally part of Akamai's acquisition of Cyberfend in 2017, he now leads the data science team at Akamai for the Bot Manager Premier security product. Prior to entering cybersecurity, he worked on developing pattern recognition and anomaly detection algorithms for drug discovery, manufacturing, automotive, retail, and military hyperspectral imaging applications. Before entering the private sector, Tyson worked at the NASA Jet Propulsion Laboratory researching neural network and evolvable hardware while completing a PhD in electrical engineering at the University of Southern California. He has a bachelor's degree in physics and economics.

Acknowledgments

It takes a village to raise a child, and this is also true when writing a book. Many talented researchers, data scientists, developers, and business leaders have indirectly contributed to this book while working with me on building the Bot Manager product I have been responsible for the last 10 years. Their input and feedback have been invaluable in shaping the direction of the product and furthering my understanding of the bot and fraud problem.

I want to acknowledge my mentors throughout the years who helped me in my journey as a professional: Patrice Boffa, who allowed me to build the very first prototype of Bot Manager more than 10 years ago; John Dilley, who recruited me as a product architect and trusted me to build Akamai's Bot Manager product; and finally Sreenath Kurupati, who helped me appreciate and understand the world of machine learning and artificial intelligence. I'm grateful for their trust in my instincts, for giving me guidance and support to solve difficult problems, and for allowing me to experiment and develop a fantastic product that protects thousands of websites around the world.

From the research and development team, I'd like to call out in no particular order key people who helped me throughout the years execute and deliver my vision: Spandan Brahmbhatt, Luke Stork, Chunliang Wu, Pujan Motiwala, Yossef Daya, Ory Segal, Nils Rehm, Nikolai Tschacher, Idan Pinto, Michael Bergmann, Harish Somaraddi, Prajaka Bhurke, Tu Vuong, Sai Modalavalasa, and more.

From the product management team, I'd like to acknowledge my partners in crime for many years, Pawan Bajaj and Maik Maurer.

A special thank you to my technical editor, Tyson Thomas, a cybersecurity veteran and lead data scientist who peer-reviewed this work and provided valuable feedback to improve the

quality of this book. Finally, a shout-out to a rising artist, my niece Julie Sénécal, who designed the robot illustrations and icons, adding character to the book.

All traffic graphs come from the Akamai Control Center with authorization from Akamai Technologies.

I'll always be grateful to be surrounded by such talented people.

—David Sénécal

Contents

Introduction

I've been interested in technology since a very young age with a particular attraction to computers, even if in the late 1980s and 1990s their capabilities were limited compared to what we have today. When I finished high school, the Internet existed but was not widely available. When it came time for me to choose a major for my college application, I looked for something that would allow me to learn and work with this emerging technology. I graduated from the Paul Sabatier University in Toulouse in the South of France in 1998 with a major in electrical engineering with a specialty in computer networking and telecommunications. Armed with this unusual high-tech degree and my knowledge of network protocols and computer programming, I started my career as a network administrator for a major insurance company (Les Mutuelles du Mans Assurances – MMA) in France, overseeing and enhancing the headquarters' network, supporting more than 5,000 users. After a few years, with my solid understanding of networks and telecommunication, I felt I needed an extra challenge. I moved to England to work as a multilingual technical support engineer for Azlan, a company later acquired by Tech Data, specializing in distributing networking equipment. Remotely helping customers configure and install their switches, routers, and firewalls was occasionally challenging. Doing so in French, English, and German and dealing with multiple regional accents made things even more interesting. Not only did I have to learn several products, but I also sometimes helped customers configure them in unexpected ways.

Several years later, I felt like introducing a change in my life again, and I moved to the United States, where I started working for Akamai Technologies. There, I became more familiar with the intricacies of the Internet. My focus was initially on helping companies accelerate their websites. I worked with the top brands on the Internet from various industries, including e-commerce, travel and hospitality, media, social media, healthcare, and banking. It quickly evolved to help secure their websites as well. What became rapidly apparent to me was that most of the traffic on any website came from bots, causing stability issues. The tools available at the time to defend against such activity (mainly web application firewalls) were only partially effective. New tools needed to be developed to deal with the problem more effectively. So, once more, I decided to get out of my comfort zone and started building a product focusing mostly on bot detection. After all, how hard could it be? This started a new phase of my career as a product architect. At the time, I thought I'd work on solving this problem for a couple of years and then move on to the next challenge. I certainly managed to solve the original threat, but I did not anticipate how it would evolve then. More than 10 years later, I am still working on bot management.

Bot management products evolved rapidly and grew in complexity while becoming a must-have product for protecting life online. However, existing knowledge on bot and fraud detection is fragmented, surrounded by many misconceptions fueled by marketing pitches, myths, and sometimes outdated best practices. This makes the subject much more confusing and frustrating for web security professionals and website owners to understand. The lack of understanding of the problem prevents them from dealing with it effectively, ultimately benefiting fraudsters.

While building bot management products, educating security professionals became a big part of my mission. My peers, the sales force, the product support staff, and, more importantly, customers looking to use my products to protect their online business needed to be trained. Good content that goes to the heart of the problem in simple terms is hard to find and mostly nonexistent. So, I thought: maybe I should write a book! Because, after all, how hard could it be? It turns out it's not easy but somewhat easier and less time-consuming than building a bot management product! I persevered and wrote this book to cover the knowledge gap on the threat landscape and defense strategies. I want to unveil the mystery, clear up some misconceptions, clarify best practices, and make bots and fraud detection more accessible. This book focuses on the bot management concepts and applies to any product, whether from a vendor or homegrown.

This book aims to provide a comprehensive overview of the threat landscape and defense strategies. It provides some insight into the evolution of attacks and defense strategies over time, the motivation of attackers, how detection methods work, and how to analyze the traffic to assess accuracy and decide on the most appropriate response strategy. The knowledge acquired from this book will help security teams regain their advantage over attackers.

Who Should Read This Book

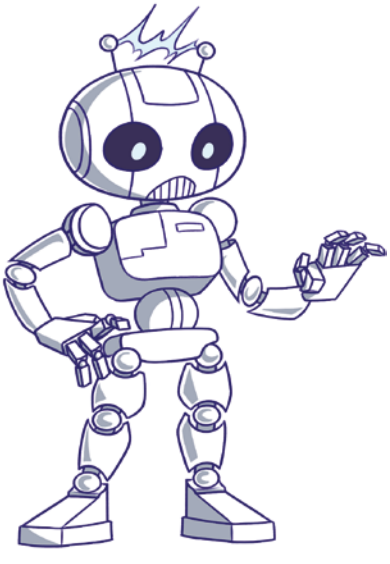
The target audience for this book includes web security professionals, website administrators, and anyone interested in or wanting to learn more about web security and, more specifically, bot management and automated fraud detection. No specific prior knowledge or experience is required to understand the content of this book.

Beginners will learn the basics of the Internet and web security while progressively diving deeper into bots, fraud, and abuse detection and mitigation. *Web security practitioners* with intermediate or advanced knowledge will better understand the threat evolution and the methods and best practices to mitigate attacks consistently and successfully. *Executives and decision-makers* reading this book will better appreciate the topic without the common vendor buzzwords or marketing bias, which will help them ask the right questions and make informed buy or build decisions. *Technology managers* (product managers) and *implementers* (security architects, developers, solution architects) will better understand the context of the bot problem and the best practices to integrate and use bot management technology to drive the most optimal outcome. *Data scientists, data analysts, and security operation support staff* monitoring and evaluating the activity detected will be able to interpret the data with a full understanding of the problem and help make data- and context-driven decisions to support the needs of their organization. *Students* in the field of computer science who are attracted to the cybersecurity space will gain a general understanding of the most critical security issues that affect online businesses today.

Any online business that generates significant revenue is at risk of fraudsters attacking their website using botnets to steal information, take over their users' identity, and make off with any assets included in the accounts. E-commerce sites (Amazon, Nike, Macy's), social media and dating sites (Facebook, LinkedIn, `Match.com`), fintech/banking sites (Bank of America, U.S. Bank, Wells Fargo), digital media (Netflix, Hulu, NBC), and gaming websites (Roblox, Electronic Arts, Epic Games) are all targets of bot and fraud attacks abusing the resources available on the website.

1 A Short History of the Internet

Our journey begins with a description of the evolution of the Internet and the emergence of a new type of fraud and abuse that leverages botnets.



From ARPANET to the Metaverse

The Internet is so ingrained in our day-to-day life that it seems as though it's always been around. However, the Internet is a relatively new invention—and it keeps evolving. The precursor of the Internet, called the Advanced Research Projects Agency Network (ARPANET), was invented in the 1960s, in the middle of the Cold War, to ensure continuity of availability of the network and computing resources even after a portion of it is removed or destroyed. Government researchers could also share information quickly without requiring them to travel to another location. ARPANET was a closed system using proprietary protocols, and only explicitly authorized people could access it. The idea of a network where one could share information and computing resources sparked the interest of academics, and the need for standardized communication protocols arose. Various communication protocols, including Transmission Control Protocol/Internet Protocol (TCP/IP), Hypertext Transfer Protocol (HTTP), and Domain Name System (DNS), were developed in the 1980s, marking the birth of the Internet as we know it today. TCP/IP defines how information is exchanged between two machines on the Internet. DNS, the equivalent of the phone book, transforms a hostname into the IP address where the service can be found. HTTP defines how web content is to be requested and shared between the browser running on the client and the web server. These protocols enable communication between systems from different vendors and connect them. Secure Sockets Layer (SSL) and, later, the Transport Layer Security (TLS) protocols add a layer of security and safety to the HTTP protocol. Newer languages like HyperText Markup Language (HTML) and JavaScript were invented to help develop websites and make content available in a structured and dynamic way.

Initially, the Internet was reserved for the technical elite who knew the protocols, had the right equipment, understood how to connect to the network, and knew how to query it to retrieve information. The development of web browser software in the 1990s, like Netscape and Internet Explorer, compatible with all of the aforementioned protocols and languages, made the Internet accessible to all. Web search engines such as AltaVista, Yahoo! Search, and Google Search also made it easier to query and find information online. When I was a college student in the 1990s, the Internet was in its infancy. All you could do was visit various websites to find information. Most news outlets would have a website with the latest sports results or events of the world. Major retailers started to create websites to showcase their products, and airlines advertised their flights. But e-commerce wasn't quite a thing just yet, and we still had to go to a brick-and-mortar shop to buy products or to a travel agency to book a flight.

Rapid technological advancement, including faster modems and expansion of the network infrastructure, supported the growth of the Internet. As the Internet grew more popular, investors started pouring money into a multitude of Internet companies with the hope of turning a profit one day. These companies' valuations, which were purely based on speculative future earnings and profits, surged in the late 1990s with record-breaking initial public offerings (IPOs) that saw their stock triple within a day. These events fueled an irrational investment strategy from venture capital firms to companies that sometimes did not have a strong business plan or viable products for fear of missing out. In March 2000, large stock sell orders from leading high-tech companies like Cisco or Dell caused a panic sale and marked the beginning of the decline of the "Internet bubble." Investors became more rational, and capital became harder to find for startups that were not profitable. Many of these cash-strapped startups disappeared rapidly. Companies that reorganized and refocused their effort on developing valuable services and products survived, and some, like Akamai Technologies, Google, Amazon, and Apple, became very successful and key players in the development of the Internet.

When the bubble burst, it felt like a setback, but eventually, the Internet not only survived but started to thrive. As the quality of the Internet network improved, so did the content. The classic dial-up modem connection that had a maximum speed of 56Kbps was soon replaced by a more advanced and reliable network and telecom infrastructure. Integrated Services Digital Network (ISDN) offered speeds of up to 128Kbps, more than double what a dial-up modem could achieve. At the turn of the century, digital subscriber lines (DSLs), which offered high-speed Internet, became more widely available through conventional telephone networks, cable, and fiber optics. Today, Internet service providers offer connections as fast as 10Gbps, which is 178,571 times faster than the fastest dial-up modem. Advancements in mobile telecommunication and the emergence of smartphones meant that consumers could access the Internet from anywhere at any time for the first time. Mobile network expansion also helped expand the reach of the Internet to rural areas. Today, one can even browse the Internet while on a plane or cruising on the ocean, thanks to satellite networks.

As more and more people were drawn to the Internet, the distribution of rich content became a real issue. The networks that carried the Internet traffic did not always have the adequate capacity to handle the demand. Telecom operators would do their best to route the traffic, but frequent congestion and often long distances between the client and the server led to slow page load or stream buffering for Internet users, especially during popular events. Content Delivery Network (CDN) companies like Akamai Technologies, Fastly, and Cloudflare, to name a few, became the backbone of the Internet. CDNs helped fix the problem by avoiding transporting the content long distances and bringing it closer to the user. CDNs helped make the Internet faster and more reliable. I've worked on and off for the biggest CDN company in the world, Akamai Technologies, since 2006 and saw the Internet evolve from a front-row seat.

Let's look at different types of websites and services that became available on the Internet and how they managed to turn their online presence into a revenue stream.