

CONTENTS IN DETAIL

FOREWORD	xix
-----------------	------------

ACKNOWLEDGMENTS	xxi
------------------------	------------

INTRODUCTION	xxiii
---------------------	--------------

Who Should Read This Book	xxiv
What You'll Find in This Book	xxiv

PART I A PRIMER ON ANDROID MALWARE

1 THE BASICS OF ANDROID SECURITY	3
---	----------

The Android Security Model	3
Application Isolation	4
Attack Surface Reduction	5
Exploit Mitigation	6
Device Integrity	6
Permissions	7
Security Updates	8
Add-on Security and Safety Services	8
Collaboration Across Google	9
Sideloaded and Preloaded Malware Protection	9
The Android Package	10
Categories of Android Malware	10
Denial of Service	11
Backdoors	12
Rooting	13
Trojans	14
Spyware	16
Stalkerware	16
Phishing	17
Hostile Downloaders	18
Privilege Escalation	19
Ransomware	21
SMS Fraud	22

Toll Fraud	23
Call Fraud.....	24
Spam	24
Ad Fraud.....	24
Non-Android Threats	25
Up Next.....	26

2 ANDROID MALWARE IN THE WILD 27

The Early Years: 2008 to 2012.....	28
DroidSMS.....	28
DroidDream	29
The Wallpaper Family	29
The Camera Family	30
Cricketland	31
Dougaleaker	32
BeeKeeper	33
Dogowar.....	35
Other Early Android Malware.....	35
The Professionalization of Malware: 2013 and 2014.....	35
Ghost Push	36
BadNews, RuFraud, and RuPlay	36
WallySMS	38
Mono WAP	39
Cryptocurrency Malware	40
Taicliphot	41
The First Preinstalled Malware	42
The Rise of Large Malware Networks: 2015 and 2016	42
Turkish Clicker	42
Gaiaphish.....	44
Judy	45
DressCode	46
Joker	47
Triada	49
Chamois	51
Gooligan and Snowfox	52
Hummingbad.....	53
YouTube Downloader	54
The Consolidation of Abuse: 2017 and Onward	54
OneAudience	56
Android.Click.312.origin	57
Cheetah Mobile	58
Anti-Fraud SDKs	59
Loapi/Podec	60
HDC Bookmark	61
EagerFonts	62

GMobi	63
Adups	64
Redstone	65
Digitime	65
Up Next	67

PART II MANUAL ANALYSIS

3		71
STATIC ANALYSIS		
What Is Static Code Analysis?	72	
Guided vs. Unguided Analysis	72	
Knowing When You're Done	73	
Loading the Malware Sample into jadx	73	
Malicious Code in the Permissions	74	
Viewing the Permissions	75	
Finding the APIs Gated by Permissions	76	
Analyzing the READ_CONTACTS Permission	76	
Analyzing the BIND_NOTIFICATION_LISTENER_SERVICE Permission	77	
Malicious Code in App Entry Points	78	
Exported Activities	78	
Broadcast Receivers	79	
Services	80	
Application Subclasses	80	
Hiding Malicious Code	81	
Anti-Analysis Techniques	81	
Reflection	82	
Non-Java Code	84	
Encryption and Encoding	85	
The Malware's First Stage	86	
Understanding the Malicious Class	87	
Reverse Engineering the String Decryption Method	87	
Decrypting All Strings in the Class	89	
The Malware's Second Stage	89	
Entry Points	90	
The yin.Chao.yin Method	91	
The com.* Package	93	
The Malware's Third Stage	94	
jadx Decompilation Issues	94	
Entry Points	95	
Name Mangling	98	
Command-and-Control Server Communication	100	
Examining the Encryption Algorithm	101	
Probing the Server from the Command Line	102	

Registering with the Server	102
Processing the Registration Response	104
Downloading Commands	105
Processing the Command-and-Control Server's Response	106
Secretly Signing Up for the Premium Service	107
Setting Up the JavaScript Bridge	107
Interacting with the Java Bridge Object	109
Completing the Sign-up Process	110
The Mysterious Fourth Stage	111
Up Next	112

4 DYNAMIC ANALYSIS 115

What Is Dynamic Code Analysis?	115
Dynamic vs. Static Analysis	116
The Android Studio Emulator	117
Creating a System Image	117
Starting the Emulator	117
Resetting the Emulator	118
Interacting with the Emulator	118
Dynamic Analysis Tools	119
tcpdump	119
Wireshark	120
Frida	120
The Malware Sample	120
Detecting Malicious Functionality	121
Observing Filesystem Changes	122
Downloading Files for Inspection	123
Capturing Network Traffic	124
Analyzing Network Traffic	125
Analyzing Logs with Logcat	126
Analysis with Frida	127
Running frida-server	128
Using frida-trace to Find Interesting APIs	128
Finding Entry Points into the Malware with Frida Scripting	129
Executing the Frida Script	130
Decrypting the Command-and-Control Communications	131
With CyberChef	132
With Frida	135
Command-and-Control Server Messages	138
The /ping URL	139
The /metrics URL	139
The Rotating Encryption Keys	141
Other Malware Functionality	141
com.sniff with frida-trace	142
Accessibility Abuse	144

Adding Static Analysis	145
Other Command-and-Control Servers	145
Other Server Commands	146
More Accessibility Abuse	150
Automatically Granting Permissions	152
Injecting Phishing Windows	154
Stealing Credentials	155
Up Next	157

PART III MACHINE LEARNING DETECTION

5 MACHINE LEARNING FUNDAMENTALS 161

How Machine Learning for Malware Analysis Works	162
Identifying App Features	163
Creating Training Sets	164
Using Classification Algorithms	166
Classification Algorithms	167
Decision Trees	167
Bagging and Random Forest	169
Support Vector Machines	170
k-Nearest Neighbors	172
Naive Bayes	172
Evaluating Machine Learning Models	174
Struggles of Machine Learning Classifiers	177
Identical Feature Vectors	177
Balance vs. Imbalance	177
Interpretability	178
Cross-Validation vs. Rolling Window Prediction	178
Up Next	179

6 MACHINE LEARNING FEATURES 181

Static Features	182
Dynamic Features	184
Method Call Features (A Weak Tactic)	186
Triadic Suspicion Graph Features	187
Suspicion Scores	189
The Suspicion Rank	191
TSG Features	192
Landmark-Based Features	195
Selecting Landmarks	195
Computing Landmark-Based Features	198

Feature Clustering	199
Generating Feature Clusters	199
Choosing Clustering and Feature Aggregation Algorithms	200
Correlation Graph–Based Feature Transformation	201
Further Reading	202
Up Next	203

7

ROOTING MALWARE **205**

Rooting Malware Families	206
Testing Classifier Performance	206
Rooting Malware vs. Goodware	208
Permission-Related Features	210
Network-Based Features	212
Rooting Malware vs. Other Malware	214
Permission-Related Features	214
Other Features	216
DroidDream: A Case Study	216
Up Next	218

8

SPYWARE **219**

Spyware Families	219
Spyware vs. Goodware	220
Permission-Related Features	222
Prediction Efficacy	223
Spyware vs. Other Malware	224
Permission-Related Features	224
Prediction Efficacy	226
Qibla Compass Ramadan: A Case Study	227
Predictions for Spyware Apps	232
Up Next	233

9

BANKING TROJANS **235**

Banking Trojan Families	236
Banking Trojans vs. Goodware	238
SMS Permission Features	238
Other Permission Features	240
Prediction Efficacy	241
Banking Trojans vs. Other Malware	242
Permission-Related Features	242
Prediction Efficacy	245

Marcher: A Case Study	246
Up Next	249

10 RANSOMWARE 251

How Ransomware Attacks Work	252
Android Ransomware Families	252
Ransomware vs. Goodware	255
Permission-Related Features	255
Other Features	257
Prediction Efficacy	257
Ransomware vs. Other Malware	258
Permission-Related Features	258
Prediction Efficacy	260
Simpllocker: A Case Study	261
Predictions for Important Ransomware Samples	264
Up Next	265

11 SMS FRAUD 267

SMS Fraud vs. Goodware	268
Non-SMS Permissions	268
The Absence of SMS Permissions	270
Prediction Efficacy	274
SMS Fraud vs. Other Malware	275
Permission-Related Features	275
Prediction Efficacy	277
BeeKeeper: A Case Study	277
Predictions for SMS Fraud Samples	280
Up Next	280

12 THE FUTURE OF ANDROID MALWARE 283

Windows vs. Android	284
Windows	284
Android	285
Hiding Malicious Behavior with Anti-Analysis Techniques	286
Native ARM Code	286
Downloaded Modules	287
Less Popular Languages	288
SDK-less Techniques	288
Distribution	289
Preloaded Malware and Supply Chain Compromises	289
Smarter Sideloaded	290

Malware Economics 291
Machine Learning Trends for Attackers and Defenders..... 291
Next Steps..... 293

INDEX

295