Autonomous Cyber Deception
Autonomous Cyber Deception
Reasoning, Adaptive Planning, and Evaluation of HoneyThings
Preface

Why Cyber Deception? Cyberattacks have evolved to be highly evasive against traditional prevention and detection techniques, such as antivirus, perimeter firewalls, and intrusion detection systems. At least 360,000 new malicious files were detected every day, and one ransomware attack was reported every 40 s in 2017 (Chap. 10). An estimated 69% of breaches go undetected by victims but are spotted by an external party, and 66% of breaches remained undiscovered for more than 5 months (Chap. 10). Asymmetries between attacker and defender information and resources are often identified as root causes behind many of these alarming statistics. Cybercriminals frequently reconnoiter and probe victim defenses for days or years prior to mounting attacks, whereas defenders may only have minutes or seconds to respond to each newly emerging threat. Defenders seek to protect infrastructures consisting of thousands or millions of assets, whereas attackers can often leak sensitive information or conduct sabotage by penetrating just one critical asset. Finding ways to level these ubiquitous asymmetries has therefore become one of the central challenges of the digital age.

What Is Cyber Deception? Cyber deception has emerged as an effective and complementary defense technique to overcome asymmetry challenges faced by traditional detection and prevention strategies. Approaches in this domain deliberately introduce misinformation or misleading functionality into cyberspace in order to trick adversaries in ways that render attacks ineffective or infeasible. These reciprocal asymmetries pose scalability problems for attackers similar to the ones traditionally faced by defenders, thereby leveling the battlefield.

Cyber Deception Models Cyber deception can be accomplished in two major ways: (1) mutation, to frequently change the ground truth (i.e., the real value) of cyber parameters such as cyber configuration, IP addresses, file names, and URLs, and (2) misrepresentation, to change or corrupt only the value returned of cyber parameters to the attacker without changing the ground truth such as false fingerprinting, files, and decoy services. We therefore call the cyber parameters used
for deceiving the attackers HoneyThings. Using the concept of HoneyThings in both approaches expands the cyber exploration space for adversaries to launch effective attacks.

**Cyber Deception 4D Goals** Effective cyber deception aims to (1) *deflect* adversaries away from their goals by disrupting their progress through the kill chain; (2) *distort* adversaries’ perception of their environment by introducing doubt into the efficacy of their attacks; (3) *deplete* their financial, computing, and cognitive resources to induce biased and error-prone decisions that defenders can influence; and (4) *discover* unknown vulnerabilities and new TTPs (tactics, techniques, and procedures) of adversaries while predicting the tactical and strategical intents of adversaries.

**Book Overview** In light of this vision, this book brings together recent research results pursuant to these goals, in four major parts:

Part I addresses in developing Cyber Deception Reasoning Frameworks and consists of three chapters:

- Chapter 1 presents a framework that uses deep learning and differential privacy techniques to generate deceptive data that is hard to differentiate from real data.
- Chapter 2 presents a framework and a research prototype for intelligent cyber deception agents that can make autonomous decisions on how to counter ongoing attacks and that integrate with active defense tools.
- Chapter 3 studies how honeypot deception can be made more effective when applied with variety and discusses the range of deception tactics that can be considered, such as random error messages, honey files with some convincing real data, and out-of-date vehicle positions.

Part II is about Dynamic Decision-Making for Cyber Deception, and it consists of two chapters:

- Chapter 4 models cyber deception as a hypergame in which attackers and defenders can have different perceptions toward a given situation and carries out case studies to examine how players’ perception (or misperception) affects their decision-making in choosing a best strategy based on hypergame theory.
- Chapter 5 applies a series of game theory models to capture the strategic interactions between attackers and defenders, the multistage persistence, as well as the adversarial and defensive cyber deceptions.

Part III examines new approaches for network-based deception, spanning four chapters:

- Chapter 6 presents a new cyber deception framework that composes mutation, anonymity, and diversity to maximize key deception objectives (i.e., concealability, detectability, and deterrence) while constraining the overall deployment cost.
- Chapter 7 presents a highly dynamic network obfuscation and deception solution that overcomes limitations of existing solutions. Specifically, it mutates and
randomizes multiple aspects of network configurations simultaneously, leveraging network and host-level SDN, state-of-the-art virtualization techniques, and DNS deception.

- Chapter 8 examines how deceptive web service responses can be realized as software security patches that double as feature extraction engines for a network-level intrusion detection system, which can increase detection accuracy and adaptability due to the fast, automatic, and accurate labeling of live web data streams enabled by this approach.

- Chapter 9 presents a technique to contain the risks of compromising buggy IoT devices by creating a protection layer on top of the local network and providing fine-grained control over the communications of individual IoT devices in the network. It uses software-defined networking (SDN) technologies to realize device- and device-group-specific views of the network that reduce the attack surface against vulnerable devices in the network, contain effects of device infections in case of device compromise, and enforce effective measures for blocking unwanted release of contextual data.

Finally, Part IV discussed automated techniques for deceiving malware, consisting of two chapters:

- Chapter 10 presents a new analytics framework and tool that can analyze the malware binary and automatically extract deception parameters in order to enable the automated creation of cyber deception plans.

- Chapter 11 proposes a system for automatically extracting the system resource constraints from malware code and generating HoneyResource (e.g., malware vaccines) based on the system resource conditions.

In each book chapter, theoretical and experimental exercises for researchers and students are presented to deepen the understanding of the deception concepts and techniques presented in this book.

The investigations, discoveries, and experiences reported in these recent results open many potential avenues for future work. We recommend the following research directions:

- **Autonomy and Resiliency**: The high speed and complexity of modern cyberattacks demands cyber deceptions that are highly *autonomous* to be self-adaptive, and *resilient* to survive failures that might reveal deception assets or plans. Methods are needed that can clearly specify both the defender’s mission and the attacker’s mission and that can identify the theoretical foundations, objectives, levels, and risks of automation.

- **Modularity and Interdisciplinarity**: The innovation and deployment of cyber deception requires scientific and engineering foundations that make HoneyThings into plug-and-play commodities that are easy to instantiate, integrate, deploy, adapt, and maintain. Efforts are needed to close the presently large gap between engineering constraints and psychology theory. Mechanisms are also needed for sharing cyber deception data that can be used to study the psychological and cognitive influence of cyber deception on the adversary.
Quantitative Evaluation: To make measurable scientific progress toward effective, deployable, deception-powered cyber defenses, the scientific community must establish accepted metrics and methodologies for evaluating proposed cyber deception techniques, frameworks, and systems. Such evaluations must go beyond mere anecdotal observations of effectiveness to obtain experimental results that are systematic, comprehensive, reproducible, and statistically valid and that afford “apples-to-apples” comparisons of competing research ideas.

Charlotte, NC, USA  Ehab Al-Shaer
Charlotte, NC, USA  Jinpeng Wei
Richardson, TX, USA  Kevin W. Hamlen
Durham, NC, USA  Cliff Wang
August 2018
Acknowledgments

We would like to thank the United States Army Research Office (ARO) for sponsoring the 2018 Workshop on HoneyThings: Autonomous and Resilient Cyber Deception, which created an opportunity for many interesting discussions about hot topics on autonomous cyber deception. We thank all authors of the chapters in this book for sharing their latest and greatest research results.

Finally, the editorial team thanks UNC Charlotte PhD students Md Sajidul Islam Sajid and Mohiuddin Ahmed for their help in preparing the final version of this book.
Contents

Part I Cyber Deception Reasoning Frameworks

1 Using Deep Learning to Generate Relational HoneyData .......... 3
  1.1 Introduction .......................................................... 3
  1.2 Related Work .......................................................... 4
  1.3 Background ............................................................ 6
    1.3.1 Deep Learning .................................................. 6
    1.3.2 Differential Privacy ............................................ 7
    1.3.3 Differentially Private Composition Theorem ............... 8
  1.4 Methodology .......................................................... 8
    1.4.1 Differentially Private Synthetic Data Generation
          Model ............................................................. 8
  1.5 Experiments .......................................................... 11
    1.5.1 Task 1: Cyber Deception for Attacker with No
          Honeydata Knowledge ........................................ 12
    1.5.2 Task 2: Cyber Deception for Attacker with
          Honeydata Knowledge ........................................ 13
  1.6 Conclusions .......................................................... 17
References ........................................................................ 18

2 Towards Intelligent Cyber Deception Systems ....................... 21
  2.1 Introduction ............................................................ 21
  2.2 Preliminaries .......................................................... 22
    2.2.1 Intelligent Cyber Defense Agents .......................... 22
    2.2.2 Active Defense .................................................. 23
  2.3 Towards Intelligent Cyber Deception Systems .................... 23
    2.3.1 Usage Scenario .................................................. 25
    2.3.2 The Architecture of the AHEAD System .................. 26
  2.4 Evolving the Pot: ADARCH ........................................ 27
    2.4.1 ADARCH Design ............................................... 28
    2.4.2 Python Embedding and Extension ........................... 29
    2.4.3 Advantages of the ADARCH Framework ................. 30
3 Honeypot Deception Tactics ........................................... 35
  3.1 Introduction .................................................................. 35
  3.2 Honeypot Deception Options ...................................... 36
  3.3 Some Example Tactics .............................................. 38
  3.4 Deception as a Game ................................................ 40
  3.5 Honeypot Experiments ............................................... 42
  3.6 Exercises ............................................................... 43
References......................................................................... 43

Part II Dynamic Decision-Making for Cyber Deception

4 Modeling and Analysis of Deception Games Based on Hypergame Theory ......................................................... 49
  4.1 Introduction .................................................................. 49
  4.2 Related Work .......................................................... 51
  4.3 Hypergame Theory ................................................... 52
  4.4 Case Study .................................................................. 56
    4.4.1 Attacker’s Game ................................................ 57
    4.4.2 Defender’s Game ............................................... 60
    4.4.3 Estimation of HEUs ............................................. 62
  4.5 Example Analytical Model Using Stochastic Petri Nets ...... 64
  4.6 Experiments and Discussion ......................................... 70
  4.7 Conclusion .................................................................. 71
  4.8 Exercise Problems .................................................... 72
References......................................................................... 73

5 Dynamic Bayesian Games for Adversarial and Defensive Cyber Deception ......................................................... 75
  5.1 Introduction .................................................................. 76
    5.1.1 Literature ........................................................... 77
    5.1.2 Notation ............................................................. 78
  5.2 Static Game with Complete Information for Cybersecurity ...... 78
  5.3 Static Games with Incomplete Information for Cyber Deception ................................................................. 80
  5.4 Dynamic Bayesian Game for Deception and Counter-Deception ................................................................. 82
    5.4.1 Signaling Game for Cyber Deception ....................... 82
    5.4.2 Multi-stage with Two-Sided Incomplete Information 85
  5.5 Conclusion and Future Works ......................................... 94
5.6 Exercise ............................................................. 95
  5.6.1 Question 1: Equilibrium Computation ............... 95
  5.6.2 Question 2: The Negative Information Gain in Game Theory ......................................... 95
References ................................................................ 96

Part III  Network-Based Deception

6  CONCEAL: A Strategy Composition for Resilient Cyber Deception: Framework, Metrics, and Deployment ........................................... 101
  6.1 Introduction and Motivation .................................... 102
  6.2 Threat Model and Objectives .................................... 103
    6.2.1 Threat Model ........................................ 103
    6.2.2 Defense Objectives .................................. 104
  6.3 Technical Approach ............................................. 104
    6.3.1 CONCEAL Framework Key Components ............ 104
    6.3.2 CONCEAL Effectiveness Metrics ...................... 106
    6.3.3 CONCEAL Architecture and Planner ................. 110
    6.3.4 Formalization of Constraints ......................... 111
  6.4 Evaluation ........................................................ 113
  6.5 Implementation .................................................. 117
    6.5.1 Web Interface ...................................... 118
    6.5.2 SMT Solver ....................................... 119
    6.5.3 ActiveSDN ......................................... 119
  6.6 Related Works .................................................... 121
  6.7 Conclusion ....................................................... 122
  6.8 Exercise Problems ................................................ 122
References .................................................................... 123

7  NetShifter: A Comprehensive Multi-Dimensional Network Obfuscation and Deception Solution .............................................. 125
  7.1 Introduction .......................................................... 125
  7.2 Summary of Existing Network Obfuscation and Deception Solutions ............................................. 127
    7.2.1 SDN-Based Solutions .................................. 127
    7.2.2 Non-SDN-Based Solutions ......................... 128
  7.3 NetShifter System Architecture .................................. 129
  7.4 NetShifter Network Obfuscation and Deception Techniques ............................................ 131
    7.4.1 IP Address Mutation Across Enclaves ............. 131
    7.4.2 Flow Migration ...................................... 133
    7.4.3 Topology Mutation .................................. 134
    7.4.4 DNS/IP Binding Mutation ......................... 135
    7.4.5 Service Randomization .............................. 135
References .................................................................... 135
8 Deception-Enhanced Threat Sensing for Resilient Intrusion Detection

8.1 Introduction ......................................................... 147
8.2 Deceptive Collection of Attack Data ............................... 148
8.3 Intrusion Detection Challenges ................................... 150
8.4 Mining Deception-Enhanced Threat Data ....................... 151
8.5 Use Case: Booby-Trapping Software for Intrusion Detection .... 152
  8.5.1 Architectural Overview ...................................... 152
  8.5.2 Detection Models .......................................... 154
  8.5.3 Attack Classification ...................................... 156
8.6 Evaluation Testbed .................................................. 157
  8.6.1 Realistic Web Traffic Generation .......................... 157
  8.6.2 Experimental Results ...................................... 159
  8.6.3 Discussion ................................................. 161
8.7 Conclusion ......................................................... 162
8.8 Exercises ............................................................ 162
  8.8.1 Software Engineering Exercises ........................... 162
  8.8.2 Machine Learning Exercises ............................... 163
References ................................................................ 164

9 HONEYSCOPE: IoT Device Protection with Deceptive Network Views .............................................................. 167

9.1 Introduction ........................................................... 167
  9.1.1 Principle of “Need to See” ................................... 168
  9.1.2 Deception Through Network Views ......................... 169
  9.1.3 HONEYSCOPE ............................................. 170
9.2 Design of HONEYSCOPE ............................................ 170
  9.2.1 HONEYSCOPE Implementation Approaches ............ 171
  9.2.2 HONEYSCOPE Network Structure ........................ 172
  9.2.3 Device Type Identification ................................. 175
9.3 HONEYSCOPE Components ....................................... 176
  9.3.1 HONEYSCOPE Controller ................................. 176
  9.3.2 HONEYSCOPE Security Gateway ........................ 176
  9.3.3 Communication Between Network Groups ................ 177
  9.3.4 Case Study .................................................. 179
9.4 Conclusion .................................................................................. 179
9.5 Hands-on Exercises ................................................................. 180
References ...................................................................................... 181

Part IV Malware Deception

10 gExtractor: Automated Extraction of Malware Deception

Parameters for Autonomous Cyber Deception ........................................ 185

10.1 Introduction .............................................................................. 186

10.2 Modeling Attack Behavior Using Binary Symbolic Execution .............. 188

10.2.1 Attack Behavior Model .......................................................... 188

10.2.2 Malware Symbolic Execution ................................................. 190

10.3 Deception Parameters Extraction ................................................ 191

10.3.1 Identifying Relevant Paths ..................................................... 192

10.3.2 Eliminating Don’t-Care Variables .......................................... 193

10.4 Evaluation .................................................................................. 194

10.4.1 Case Study I: Bitcoin Miner .................................................. 194

10.4.2 Case Study II: FTP Credential-Stealer ..................................... 199

10.4.3 Challenges and Future Work .................................................. 203

10.5 Related Work ............................................................................ 203

10.6 Conclusion .................................................................................. 204

10.7 Exercises ................................................................................... 204

References ...................................................................................... 205

11 Malware Deception with Automatic Analysis and Generation of HoneyResource .............................................................. 209

11.1 Introduction .............................................................................. 209

11.2 Problem Statement and Approach Overview ................................... 211

11.2.1 Malware HoneyResource Background ..................................... 211

11.2.2 Approach Overview ............................................................... 213

11.3 Phase-I: Candidate Selection ...................................................... 214

11.3.1 Taint Sources ........................................................................ 215

11.3.2 Taint Propagation ................................................................. 216

11.4 Phase-II: HoneyResource Generation .......................................... 217

11.4.1 Exclusiveness Analysis ........................................................... 217

11.4.2 Impact Analysis .................................................................... 217

11.4.3 Determinism Analysis ............................................................. 219

11.4.4 Malware Clinic Test ............................................................... 220

11.5 Phase-III: HoneyResource Delivery and Deployment ...................... 221

11.5.1 Direct Injection ................................................................. 221

11.5.2 HoneyResource Daemon ...................................................... 222
11.6 Evaluation ......................................................................... 222  
11.6.1 Experiment Dataset .................................................. 222  
11.6.2 Evaluation Result on Candidate Selection ................. 223  
11.6.3 Evaluation on HoneyResource Generation .............. 224  
11.6.4 Case Studies .......................................................... 226  
11.6.5 HoneyResource Effect Analysis ............................... 227  
11.6.6 Performance Overhead ........................................... 229  
11.7 Limitations and Future Work ......................................... 230  
11.7.1 Evasions from Malware ........................................... 230  
11.7.2 Limitation on Dynamic Analysis .............................. 231  
11.7.3 Potential False Positive ........................................... 231  
11.7.4 Deployment Issues .................................................. 231  
11.7.5 Deception Goals ..................................................... 232  
11.8 Related Work ............................................................. 232  
11.8.1 Immunization-Based Defense ................................. 232  
11.8.2 Dynamic Malware Analysis .................................... 233  
11.9 Conclusion .................................................................... 233  
11.10 Exercise ..................................................................... 233  
References ........................................................................ 234