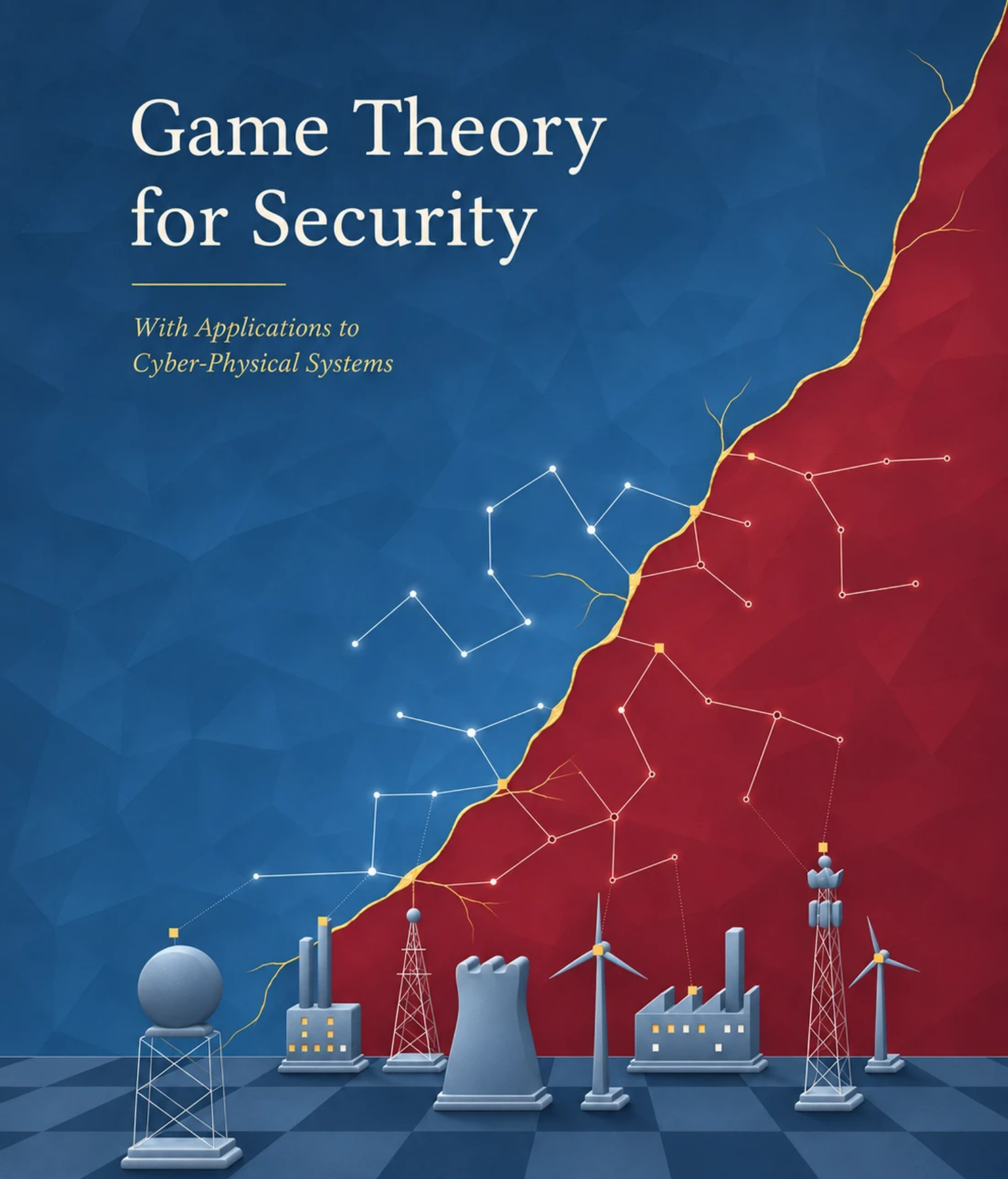
Game Theory for Security

With Applications to Cyber-Physical Systems



Quanyan Zhu • Juntao Chen • Junaid Farooq

Game Theory for Security

With Applications to Cyber-Physical Systems

Quanyan Zhu
Juntao Chen
Junaid Farooq

Cambridge University Press
2026

About this sample

This is a sample of *Game Theory for Security: With Applications to Cyber-Physical Systems* by Quanyan Zhu, Juntao Chen, Junaid Farooq. It contains the preface and the following chapters, reproduced exactly as they appear in the complete book, so page numbers, cross-references, and the table of contents refer to the full volume.

Preface		xxi–xxii
Chapter 1	Introduction	1–21
Chapter 3	Noncooperative Games *(Part I)*	41–85
Chapter 10	Composition of Games *(Part II)*	370–399
Chapter 13	Cyber Deception *(Part III)*	489–524
Bibliography		764–799

The complete book has sixteen chapters in three parts and six appendices.

Sample prepared 15 November 2026.

Contents

Preface **xxi**

1 Introduction **1**
1.1 Games and Security . . . 3
1.2 Strategic Themes in Security Games . . . 12
1.3 Recent Advances and Challenges . . . 12
1.4 Structure of the Book . . . 16
1.5 Discussion, Summary, and Further Reading . . . 19
1.6 Exercises . . . 20

2 Cyber-Physical System Security: An Introduction **22**
2.1 Foundations of CPS Security . . . 24
2.2 Threat Models and Attack Surfaces . . . 26
2.3 Risk Assessment, Goals, and Defenses . . . 29
2.4 Standards and Regulations . . . 30
2.5 Mapping CPS Security to Game Models . . . 32
2.6 Incidents and Future Directions . . . 34
2.7 Discussion, Summary, and Further Reading . . . 37
2.8 Exercises . . . 38

I Fundamentals of Security Games **40**

3 Noncooperative Games **41**
3.1 Normal-Form Games . . . 43
3.2 Extensive-Form Games . . . 58
3.3 Applications in Security . . . 73
3.4 Discussion, Summary, and Further Reading . . . 82
3.5 Exercises . . . 83

4 Games of Incomplete Information **86**
4.1 Bayesian Games . . . 88
4.2 Hypergames . . . 98
4.3 Applications: Strategic Trust in Cloud-Enabled CPS . . . 103
4.4 Applications: Hypergames for Security . . . 110
4.5 Discussion, Summary, and Further Reading . . . 117
4.6 Exercises . . . 119

5 Dynamic and Stochastic Games **122**
5.1 Markov Games with Complete Information . . . 124
5.2 Markov Games with Incomplete Information . . . 143

5.3 Security Applications of Markov Games . . . 151
5.4 Discussion, Summary, and Further Reading . . . 168
5.5 Exercises . . . 170

6 Differential Games 173
6.1 Linear-Quadratic Differential Games . . . 175
6.2 Differential Games with Switching Modes . . . 193
6.3 Differential Games with Algebraic Constraints . . . 200
6.4 Application: Resilient Control of Power Systems . . . 206
6.5 Application: Secure Control of Satellites . . . 212
6.6 Discussion, Summary, and Further Reading . . . 217
6.7 Exercises . . . 219

II Advanced Security Games 223

7 Learning in Games and Algorithms 224
7.1 Learning in Repeated Games . . . 226
7.2 Reinforcement Learning for Markov Games . . . 242
7.3 Evolutionary Games and Dynamics . . . 250
7.4 Security Applications . . . 254
7.5 Discussion, Summary, and Further Reading . . . 268
7.6 Exercises . . . 270

8 Network Games 274
8.1 Games on Networks . . . 276
8.2 Network Formation Games . . . 284
8.3 Network Design . . . 292
8.4 Network Interdiction Games . . . 299
8.5 Application: Multi-Layer Resilient Network Formation . . . 307
8.6 Discussion, Summary, and Further Reading . . . 317
8.7 Exercises . . . 319

9 Mechanism Design Theory 323
9.1 Introduction to Mechanism Design . . . 325
9.2 Principal–Agent Models . . . 333
9.3 Contract Theory . . . 340
9.4 Application: Security-Aware Cyber Insurance . . . 351
9.5 Application: Security as a Service for Cloud-Enabled Control . . . 360
9.6 Discussion, Summary, and Further Reading . . . 366
9.7 Exercises . . . 367

10 Composition of Games 370
10.1 Introduction . . . 372
10.2 Equilibrium Foundations for Composed Games . . . 373
10.3 Games-in-Games for Cross-Layer Resilient Control . . . 378
10.4 Cyber-Physical Signaling Games for Multi-Agent Autonomous Systems . . . 383
10.5 iSTRICT: Interdependent Strategic Trust in the Cloud-Enabled IoCT . . . 387
10.6 Multi-Resolution Dynamic Games for Cyber Campaigns . . . 391
10.7 Discussion, Summary, and Further Reading . . . 396
10.8 Exercises . . . 397

11 Bounded Rational Games 400

11.1 Bounded Rationality . . . 402
11.2 Bounded Perception and Attention . . . 404
11.3 Subjective Modeling . . . 412
11.4 Bounded Computation and Reasoning . . . 419
11.5 Application: Cognitive Arbitrage . . . 428
11.6 Application: Security Investment Games . . . 434
11.7 Bilateral Cognitive Security Games . . . 438
11.8 Discussion, Summary, and Further Reading . . . 444
11.9 Exercises . . . 445

III Applications of Game Theory in Cyber-Physical Systems 449

12 Industrial Control System Security 450
12.1 Modeling Views and Attack Models for Industrial Control Systems . . . 452
12.2 Dynamic Games for Layered APT Defense-in-Depth . . . 463
12.3 Moving-Horizon Hybrid Stochastic Games for Secure Control . . . 471
12.4 Cross-Layer Cyber Defense and Switching Control for Wind Energy Systems . . . 478
12.5 Discussion, Summary, and Further Reading . . . 483
12.6 Exercises . . . 485

13 Cyber Deception 489
13.1 Introduction to Cyber Deception . . . 491
13.2 Game-Theoretic Analysis of Cyber Deception . . . 495
13.3 Mechanism Design Theory of Cyber Deception . . . 504
13.4 Dynamic Cyber Deception . . . 509
13.5 Moving-Target Defense . . . 516
13.6 Discussion, Summary, and Further Reading . . . 521
13.7 Exercises . . . 522

14 IoT Security 525
14.1 IoT Security at Three Levels . . . 527
14.2 Device-Level Security Games . . . 531
14.3 Dynamic Botnet Formation in Wireless IoT Networks . . . 538
14.4 Malware Versus Patching as a Zero-Sum Differential Game . . . 545
14.5 Decentralized Protection as a Population Game . . . 551
14.6 Cross-Layer Botnet Games Against the Power Grid . . . 558
14.7 Discussion, Summary, and Further Reading . . . 567
14.8 Exercises . . . 569

15 Supply Chain Security 572
15.1 Supply Chains as Inherited Attack Surfaces . . . 574
15.2 Graph-Based Supply Chain Risk and Vendor Selection . . . 577
15.3 Security Investment Games in Supply Chain Networks . . . 585
15.4 Verification, Contracts, and Counterfeits . . . 591
15.5 Blockchain Ledgers and the Incentive Security of Nakamoto Consensus . . . 599
15.6 Discussion, Summary, and Further Reading . . . 608
15.7 Exercises . . . 610

16 Next-Generation Wireless Network Security 614
16.1 Architectures and Attack Surfaces of Next-Generation Wireless Networks . . . 616
16.2 Jamming and Eavesdropping Games . . . 619
16.3 Stochastic Anti-Jamming Games in Dynamic Spectrum Access . . . 628

16.4 Incentive-Compatible Spectrum Reservation . . . 635
16.5 Resilient Multi-Layer Network Design Under Compromise . . . 641
16.6 Securing AI-Native Network Control . . . 649
16.7 Discussion, Summary, and Further Reading . . . 658
16.8 Exercises . . . 661

Appendices **665**

A Linear and Convex Optimization **665**
A.1 Optimization Formulations . . . 667
A.2 Convexity and Duality . . . 669
A.3 Optimality Conditions . . . 673
A.4 Optimization Algorithms . . . 676
A.5 Fixed-Point Theorems . . . 679
A.6 Exercises . . . 682
A.7 Discussion, Summary, and Further Reading . . . 684

B Probability and Stochastic Processes **686**
B.1 Probability Spaces and Random Variables . . . 688
B.2 Conditional Probability and Information . . . 690
B.3 Markov Chains and Stochastic Processes . . . 692
B.4 Martingales and Stochastic Approximation . . . 696
B.5 Exercises . . . 699
B.6 Discussion, Summary, and Further Reading . . . 700

C Linear Systems and Optimal Control **703**
C.1 State-Space Models . . . 705
C.2 Stability, Controllability, and Observability . . . 706
C.3 Optimal Control and Linear-Quadratic Regulators . . . 710
C.4 State Estimation, Kalman Filtering, and Separation . . . 713
C.5 Dynamic Programming and Hamilton–Jacobi–Bellman Equations . . . 717
C.6 Exercises . . . 720
C.7 Discussion, Summary, and Further Reading . . . 722

D Foundations of Network Security **725**
D.1 Security Objectives . . . 727
D.2 Networked Systems and Attack Surfaces . . . 728
D.3 Threats and Vulnerabilities . . . 731
D.4 Cyber Attack Lifecycle . . . 732
D.5 Security Models and Representations . . . 734
D.6 Security Metrics . . . 737
D.7 Security, Risk, and Resilience . . . 739
D.8 Exercises . . . 741
D.9 Discussion, Summary, and Further Reading . . . 741

E Foundations of Cyber Risk Analysis **744**
E.1 What Is Cyber Risk? . . . 746
E.2 Components of Cyber Risk . . . 748
E.3 Uncertainty in Cybersecurity . . . 749
E.4 Risk Assessment Frameworks . . . 750
E.5 Cyber Risk in Networked Systems . . . 752
E.6 Adversarial Risk . . . 754

E.7 Risk, Security, and Resilience . . . 756
E.8 From Risk Analysis to Security Games . . . 757
E.9 Exercises . . . 759
E.10 Discussion, Summary, and Further Reading . . . 760

F Notation **762**
F.1 General Mathematical Notation . . . 762
F.2 Players, Strategies, and Payoffs . . . 762
F.3 Dynamics and Information . . . 763
F.4 Physical Dynamics and Control . . . 763
F.5 Networks . . . 763
F.6 Abbreviations . . . 763

Bibliography **764**

Index **800**

Preface

Cyber-physical systems are now woven into nearly every part of daily life. Power grids, water utilities, pipelines, factories, hospitals, transportation networks, and homes are operated by networked sensors, controllers, and software that tie computation to the physical world. Industrial controllers that were designed decades ago for isolated plants are now connected to enterprise networks, cloud services, and remote vendors; medical devices, vehicles, and household appliances communicate continuously; and wireless networks carry the measurements and commands on which physical processes depend. This integration has brought remarkable gains in efficiency and capability, but it has also changed what a security failure means. A compromise in the cyber layer no longer stays there: it can open a valve, trip a breaker, disable a safety system, or stop a production line.

The security and resilience of these systems have long been a concern, and the geopolitical conflicts of recent years have made them an urgent one. Stuxnet showed that malicious code could physically damage industrial equipment. The attacks on the Ukrainian power grid in 2015 and 2016 cut electricity to hundreds of thousands of customers, the second with Industroyer, malware built to speak the grid's own control protocols. Triton targeted the safety instrumented systems whose sole purpose is to prevent industrial disasters. A ransomware attack in 2021 led to the shutdown of the Colonial Pipeline, and government advisories have since warned that state-sponsored actors such as Volt Typhoon are positioning themselves inside critical infrastructure in advance of any future conflict. In an era of war and rivalry between states, the infrastructure on which society depends has become part of the front line.

Securing cyber-physical systems is not simply information security applied to a new setting. Physical processes evolve in real time and cannot be paused while an incident is investigated. An attack is measured not only by the data it exposes but by the damage it does to equipment, the environment, and human safety. Controllers and field devices remain in service for decades, long after their software has stopped receiving updates. Safety mechanisms and security mechanisms interact, sometimes in conflict. And the systems are owned and operated by many parties (utilities, manufacturers, integrators, vendors, regulators, and users) whose interests and information differ. Each of these features shapes what an attacker can do and what a defender can afford.

Yet such problems are still largely addressed case by case. Each incident is analyzed after the fact, each vulnerability is patched when it is found, and security is too often the last consideration in system design, added once the architecture is fixed and the system is already deployed. This reactive practice cannot keep pace with adversaries who learn, adapt, and choose their targets deliberately. What is needed is a systematic viewpoint, and a systematic viewpoint needs a foundation.

We believe that game theory provides that foundation. Security differs from reliability in one essential respect: the failures it guards against are caused not by nature but by intelligent agents who anticipate the defender's choices and respond to them. A defense that is optimal against random faults can be exactly the one that a strategic attacker learns to evade. Game theory is the system-theoretic discipline built to model this kind of interdependent decision making. It captures conflict and adversarial behavior in cyber and physical environments, and it gives precise meaning to ideas that security practice uses informally: deterrence, commitment, deception, information asymmetry, incentives, and adaptation. It also connects naturally with the other tools of systems engineering, including control, optimization, learning, and risk analysis, so that strategic reasoning can be built into the design of a system rather than bolted on afterward.

A second motivation for this book is the need for a system view of cyber-physical security. Research in this area has grown up in separate silos. Computer scientists study attacks, intrusion detection, and secure protocols; control theorists study secure estimation and resilient control; economists study incentives, externalities, and insurance; and communications researchers study jamming, spectrum access, and network security. Each community has made

important advances, but the literature remains fragmented, with different vocabularies and different models of what are often the same underlying problems. Since 2010 the Conference on Decision and Game Theory for Security (GameSec) has done much to bring these researchers together. There remains, however, a need for a common reference that can educate the next generation of researchers and practitioners in this area and give them a shared language. This book is our attempt to meet that need.

The book is organized in three parts, preceded by two introductory chapters. Chapter 1 introduces security as a problem of strategic interdependence, and Chapter 2 supplies the background on cyber-physical systems, their threats, and their defenses. Part I develops the fundamentals of security games: normal- and extensive-form games, games of incomplete information, Markov games, and differential games. Part II turns to advanced topics: learning in games, network games, mechanism design, the composition of games across layers and resolutions, and bounded rationality. Part III applies these models to industrial control systems, cyber deception, the Internet of Things, supply-chain security, and next-generation wireless networks. Five appendices collect the background in optimization, probability, control, network security, and cyber risk analysis, so that the book can be read without other references at hand. Each chapter opens with an overview and learning objectives, develops its models through worked examples drawn from security practice, and closes with exercises and annotated further reading.

We have also tried to make the results in the book reproducible. The book is accompanied by Python code, organized by chapter, with more than a hundred short scripts that reproduce the numerical results and figures of the worked examples. The code is freely available from the book's web page, which will also collect errata and updates:

`https://quanyan8.github.io/game-theory-for-security/`

A marker beneath an example names the script that computes it. Each script prints its computation step by step, so that it can be read alongside the derivation in the text, and ends with checks confirming that it recovers the numbers reported in the book. The scripts require only Python with NumPy, SciPy, and Matplotlib. They are meant to be modified as well as run: changing a parameter and running a script again is often the quickest way to see how an equilibrium depends on the assumptions behind it, and the scripts provide a starting point for the exercises and for readers' own models.

The book is written for graduate students in electrical and computer engineering, computer science, operations research, and systems engineering, for whom it can serve as the text of a course on game theory for security. It is also written for researchers in control, networking, machine learning, and economics who wish to enter the security literature, for security researchers who wish to learn game theory, and for practitioners and policy analysts who design or evaluate security measures. We assume only linear algebra, calculus, and elementary probability. Chapter 1 suggests several paths through the material for readers with different goals.

The book draws on more than a decade of research by the authors and by many colleagues in the community, whose work we have tried to represent faithfully and whose contributions are acknowledged throughout the text.

Acknowledgments

We are grateful for the financial support we have received from several funding agencies, including the National Science Foundation, the Defense Advanced Research Projects Agency, and the Department of Energy. This work would not have been possible without the many communities that have shaped it, among them the control systems community, the GameSec community, and the IEEE Communications Society, whose meetings, workshops, and discussions have influenced our thinking over many years. The first author is especially grateful to the students with whom he has worked over the past decades; their ideas, questions, and effort have made this book possible.

Quanyan Zhu
New York, New York

Juntao Chen
New York, New York

Junaid Farooq
Dearborn, Michigan

Chapter 1

Introduction

Security is strategic because attackers, defenders, users, platforms, and institutions adapt to one another. Inspection added at one entry point can move attacks to another; a patrol schedule, once observed, tells the attacker where the defender will not be; publishing a patch protects updated systems while pointing attackers toward unpatched ones; and an organization may underinvest because much of the benefit of its security spills over to others. A vulnerability scan or a risk model with fixed attack likelihoods can list weaknesses, but it cannot say how an intelligent adversary will respond to a defense; and because perfect prevention is unavailable, the defender must also plan to respond and recover. This book asks how to reason about security when the outcome depends on the choices of several decision makers with different objectives and information.

This chapter introduces the book's game-theoretic lens: incentives, information, commitment, deception, resilience, risk, uncertainty, and solution concepts. It first frames security as a strategic problem, reads the history of computer security as a sequence of generations, and traces security games from minimax and Colonel Blotto to deployed and cyber-physical security games. It then models a multi-stage campaign as a kill chain paired with a defense chain and matches solution concepts, from minimax and Nash to Stackelberg, Bayesian, and dynamic equilibria, to the kinds of security reasoning they formalize. Eight strategic themes that recur throughout the book follow. A survey of recent advances and open challenges covers cyber resilience, learning, generative AI and AI agents, deception, mechanism design, scalability, and behavioral realism. It closes with the organization of the book, its prerequisites, and suggested reading paths.

The chapter stays at the level of strategic reasoning and book organization; it uses almost no formal notation and contains no proofs. Chapter 2 then provides the cybersecurity and cyber-physical-system background needed for the technical applications, and Chapter 3 introduces the formal models and equilibrium concepts on which every later chapter builds.

Chapter Overview

1. Section 1.1: Games and security
   - security as strategic interdependence among attackers, defenders, users, platforms, and institutions;
   - generations of security, from physical custody and access control to risk management, cyber resilience, AI-enabled security, and AI agents;
   - a brief history of security games, from minimax and Colonel Blotto through deterrence, inspection, and network security games to deployed and cyber-physical security games;
   - threat landscapes, cyber kill chains, risk quantification, and solution concepts as bridges from security practice to game-theoretic models.
2. Section 1.2: Strategic themes in security games
   - resource allocation, adaptation, information, commitment, deception, resilience, incentives, and risk quantification;

- recurring modeling questions that later chapters treat with different classes of games.

3. Section 1.3: Recent advances and challenges
   - cyber resilience, learning in games, adversarial machine learning, generative AI, AI agents, and AI-enabled attack and defense;
   - cyber deception, mechanism design, scalability, behavioral realism, empirical validation, privacy, and safety.
4. Section 1.4: Structure of the book
   - Part I develops the fundamentals of security games: noncooperative, incomplete-information, dynamic, and differential games;
   - Part II develops advanced security games: learning, networks, mechanism design, composition, and bounded rationality;
   - Part III applies the models to industrial control systems, cyber deception, IoT security, supply chains, and next-generation wireless networks;
   - audience, prerequisites, and suggested reading paths through the book.
5. Section 1.5 and exercises
   - how later chapters specialize the strategic lens, and annotated further reading;
   - problems on modeling a security problem, adaptation and displacement, solution concepts, delegation to an AI agent, and incentives beyond direct attack.

Chapter Summary

Security games model security as a coupled decision problem. The defender's action changes the attacker's incentives, the attacker's action changes the defender's risk, and both sides act under constraints and uncertainty. A defense may deter an attack, displace it to a different target, reveal information, change user behavior, or create incentives for new forms of adaptation. Game theory gives these feedback effects a precise language, and its solution concepts turn quantified risk into decisions: minimax reasoning gives worst-case guarantees, Stackelberg reasoning captures deterrence by commitment, Bayesian reasoning captures inference about hidden attacker types, and dynamic and learning-based concepts capture adaptation over time.

A second lesson is that modern security cannot be reduced to prevention. Some attacks will evade protection, exploit unknown vulnerabilities, or arrive through dependencies outside the defender's direct control. Security games therefore explain not only how to deter or block attacks, but also how to prepare, respond, and recover when prevention is incomplete. Equilibrium serves throughout as a benchmark for consistent incentives, not as a claim that attackers and defenders behave perfectly rationally; bounded rationality, learning, and empirical validation relax it where it is too strong.

Finally, modern security games increasingly combine equilibrium reasoning with learning, control, optimization, behavioral modeling, deception, network science, mechanism design, and risk assessment, and AI agents change the actions and information available to both sides. These developments motivate the structure of the book, which begins with the fundamentals of security games, moves to advanced models, and ends with cyber-physical applications.

Notation and Scope

This introductory chapter uses notation sparingly. The book's formal notation is introduced in Chapter 3, the notation common to all chapters is listed in Appendix F, and each later chapter describes its own symbols at its start. In words, its core objects are the set of players (attackers, defenders, and other stakeholders); the actions available to each player and the strategies, possibly randomized, that choose among them; what each player knows, including private types and the beliefs held about them; the state of the system and how it evolves; and the utility a player maximizes or the cost it minimizes. Solution concepts such as Nash and Stackelberg equilibrium then single out the strategy profiles consistent with these incentives. The only symbols in this chapter are the stage games $G_1, \ldots, G_7$ of the multi-stage campaign in Figure 1.4, with G_k the attacker–defender contest at the kth stage.

The scope is conceptual: the chapter explains why security is strategic and why game theory provides the organizing language for the rest of the book, and it contains no proofs. No background in game theory or security is assumed;

Chapter 2 supplies the cybersecurity and cyber-physical-system background, and the appendices supply the mathematical background.

Learning Objectives

After reading this chapter, the reader should be able to:

(1) explain why security problems are strategic rather than purely technical, and identify attackers, defenders, users, platforms, and institutions as strategic decision makers;

(2) distinguish prevention, detection, response, recovery, risk, and resilience as security concepts;

(3) describe how incentives, information, commitment, deception, and adaptation enter security models;

(4) match common security reasoning (robust protection, deterrence by commitment, inference under hidden types, adaptation over time) to the corresponding solution concepts;

(5) explain why game theory complements risk analysis, control, learning, and mechanism design;

(6) trace how game-theoretic thinking about security developed, from military and arms-control games to network, deployed, and cyber-physical security games;

(7) describe the organization of the book and the role played by the later game-theoretic chapters.

1.1 Games and Security

1.1.1 Security as a Strategic Problem

Security is often described as a technical problem: a system has vulnerabilities, an attacker exploits them, and the defender deploys controls to reduce risk. This description is useful, but incomplete. Many of the most important security failures arise because the relevant actors adapt. Attackers observe defenses and change targets. Defenders allocate scarce resources and decide which risks to tolerate. Users respond to warnings, incentives, and inconvenience. Vendors decide how quickly to patch. Firms decide how much information to disclose after an incident. Regulators, insurers, and platforms shape the environment in which these choices are made. The security outcome is therefore not determined by technology alone. It is determined by interdependent decisions.

Game theory is the mathematical study of such interdependent decisions. A game specifies who the decision makers are, what actions they can take, what they know, when they move, and how they evaluate outcomes [410, 449]. This framework fits security because security is rarely a one-sided optimization problem. A defender cannot simply choose the technically strongest control without considering cost, usability, observability, and adversarial response. An attacker cannot simply choose the most damaging action without considering stealth, effort, uncertainty, and the probability of detection. A user cannot simply choose the safest behavior without considering time, trust, convenience, and incentives. Game theory makes these tradeoffs explicit.

The case for game theory becomes stronger when security is viewed through the lens of resilience. Prevention is necessary, but it cannot be the only organizing principle because perfect prevention is not available in realistic systems. Unknown vulnerabilities, supply-chain dependencies, insider behavior, delayed patching, misconfigurations, and coordinated campaigns mean that defenders must also prepare for attacks that succeed partially. Cyber resilience asks how a system can anticipate, withstand, respond to, and recover from such events while preserving mission objectives. This perspective naturally brings together game theory, control theory, and learning theory: game theory models strategic conflict and incentives, control theory models system dynamics and recovery, and learning theory models adaptation under incomplete knowledge [637, 364, 278]. In this book, game theory provides the strategic spine of that broader resilience view.

1.1.2 Generations of Security

The history of computer security can be read as a sequence of generations in which the strategic dimension becomes more explicit. In the 1950s and early 1960s, security was largely tied to physical custody, administrative control, batch

processing, and the isolation of scarce computing resources. The central question was how to prevent unauthorized people from reaching machines, media, programs, and classified information. By the late 1960s, time-sharing and resource-sharing systems made isolation insufficient. Ware's work and the Defense Science Board report showed that multi-access computing required systematic controls for users, programs, storage, terminals, communications, certification, and management [395, 594]. The security problem had already become a problem of many users sharing a system under different privileges and incentives.

The next generation, developed through the 1970s, made security more formal. Multilevel-security models such as Bell–LaPadula represented confidentiality policies as state-transition rules, while Saltzer and Schroeder's design principles articulated enduring ideas such as least privilege, complete mediation, fail-safe defaults, work factor, and psychological acceptability [51, 515]. This generation clarified that security is not only a perimeter, but a set of rules governing permitted actions. It also introduced a theme that remains central to game-theoretic security: a control is meaningful only relative to what a capable user or adversary can do around it.

The 1980s and 1990s brought a monitoring and network generation. As systems became connected, the defender could no longer rely on a static access-control policy alone. Audit trails, intrusion monitoring, anomaly detection, and expert systems became important because attacks could unfold over time and because evidence had to be interpreted under uncertainty [21, 147]. Security began to look less like a locked room and more like an ongoing inference problem: which events are benign, which are probes, which are compromises, and which should trigger response.

The 2000s and 2010s made security an enterprise and campaign problem. Risk management frameworks emphasized likelihood, impact, uncertainty, controls, residual risk, and organizational decision making [422]. Kill-chain and ATT&CK models framed intrusions as staged campaigns rather than isolated events [285, 396]. The NIST Cybersecurity Framework organized security around governance, identification, protection, detection, response, and recovery, while zero trust moved the field away from static network perimeters toward resource-centered decisions about users, assets, devices, and workflows [430, 505]. These shifts made clear that security decisions are distributed across organizations, technologies, and time.

The 2010s also carried security into the physical world. Industrial control systems, power grids, vehicles, medical devices, and building automation became connected to enterprise networks and the Internet, so an intrusion could alter a physical process rather than only steal or corrupt data [98]. Stuxnet showed that malicious code could damage centrifuges while replaying normal readings to operators, and the 2015 attack on the Ukrainian power grid showed that a coordinated campaign could cut electricity to hundreds of thousands of customers [347, 133]. In this cyber-physical generation, the consequences of an attack are governed by physics, control loops, and safety constraints. The defender's goal shifts from keeping the adversary out to keeping the process safe and restoring it when protection fails. Security thus becomes inseparable from control and resilience: decisions at the cyber layer and the physical layer interact, and both unfold over time [660, 276].

The 2020s add an AI-enabled and agentic generation on top of the cyber-physical one. AI systems can help defenders process telemetry, summarize threat intelligence, prioritize vulnerabilities, identify anomalies, triage incidents, and adapt response. The same technologies also help attackers scale reconnaissance, social engineering, vulnerability discovery, malware generation, and campaign planning. The rise of AI agents deepens this change because AI systems are no longer limited to producing recommendations. They may plan over multiple steps, call tools, retrieve information, write or execute code, open tickets, modify configurations, and act under delegated identities. AI risk management, generative-AI risk management, adversarial machine learning, and agentic-AI threat modeling therefore become part of the security problem rather than external tools applied after the fact [426, 428, 583, 452]. This generation is strategic at its core: intelligent systems learn from adversarial data, attackers adapt to automated defenses, and defenders must decide when to trust, override, randomize, constrain, or revoke AI-generated actions.

Across these generations, the object of security expands. It begins with protecting machines and data, moves to protecting users and access paths, then to protecting networks and enterprises, then to the physical processes those networks control, and now to AI-enabled systems and agents that act with delegated authority. Each expansion adds more decision makers, more uncertainty, more adaptation, and more consequences. This is the basic reason that game theory becomes more important over time. It provides a way to model not only the controls, but the strategic environment created by those controls. Figure 1.1 summarizes the progression.

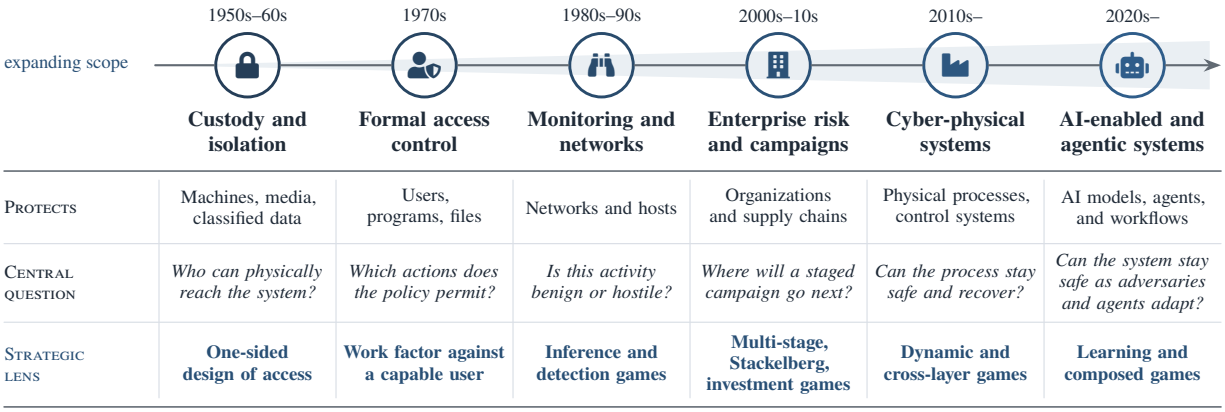


Figure 1.1: Generations of security and their strategic counterparts. Each generation widens what is protected, poses a new central question, and calls for a richer strategic lens. The widening band behind the timeline reflects the growth in decision makers, uncertainty, adaptation, and consequence.

1.1.3 A Brief History of Security Games

The generations above trace how security practice evolved. The game-theoretic view of security has a parallel history, and it is older than the computer. Figure 1.2 summarizes it in six eras.

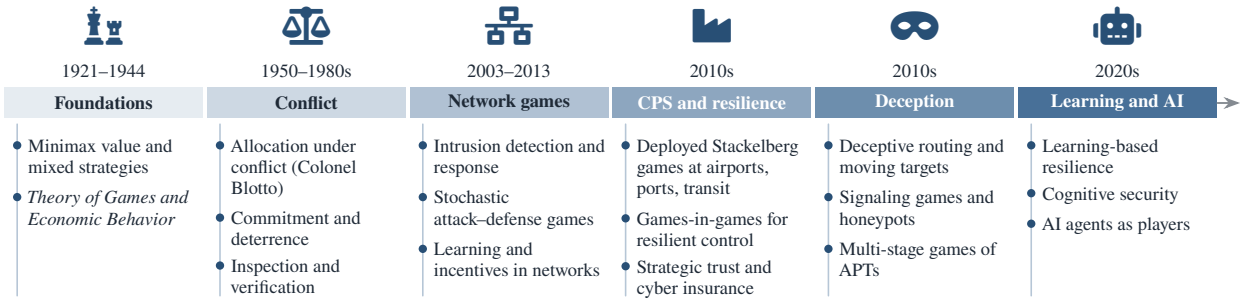


Figure 1.2: A brief history of security games. Each era adds a new kind of strategic question: minimax play, commitment and deterrence, network attack and defense, field deployment and cyber-physical resilience, deception, and finally learning, cognition, and AI agents. Eras are ordered in time but not drawn to scale.

Foundations. Game theory was born partly from questions of conflict. Borel's notes of 1921 studied games of strategy and the idea of mixed play, von Neumann's minimax theorem of 1928 established that every finite two-person zero-sum game has a value, and von Neumann and Morgenstern's *Theory of Games and Economic Behavior* of 1944 gave the field its foundations [77, 585, 586]. Zero-sum games remain the natural model when one side's gain is exactly the other side's loss, and the minimax value is still the benchmark for a defender who assumes the worst.

Conflict and deterrence. After the Second World War, much of game theory's early development took place in defense research. The Colonel Blotto game, in which two commanders divide limited forces across several battlefields, became the prototype of resource allocation under conflict [227]. Games of search and hide-and-seek, of pursuit and evasion, and of tactical allocation filled the RAND memoranda of the 1950s and were collected in texts such as Dresher's *Games of Strategy* [156]. Schelling's *The Strategy of Conflict* moved the emphasis from zero-sum fighting to bargaining, threats, and commitment, showing that the ability to bind oneself to a course of action can be a source of power [520]. Deterrence, credibility, and commitment reappear today in Stackelberg security games and in the design of visible defensive policies. Inspection games, developed for arms control and nuclear safeguards, asked how an inspector with limited resources should verify that an inspectee complies with an agreement [33]. They are the direct ancestors of modern models of auditing, monitoring, and intrusion detection.

Network security games. Game theory reached computer security in the early 2000s. Alpcan and Başar modeled intrusion detection as a game between an attacker and a detection system, and Lye and Wing modeled the interaction between an attacker and a network administrator as a stochastic game over the states of a network [11, 376]. Dynamic and stochastic games were soon used to configure intrusion-detection systems as threats evolve and to choose network security configurations when attacker and defender have different objectives [655, 678]. Because neither side knows the full game in practice, learning entered early: heterogeneous and distributed learning algorithms let defenders and attackers adapt their strategies in zero-sum stochastic games with incomplete information [677, 679]. Incentives entered as well. In collaborative intrusion-detection networks, peers share alerts only if the sharing mechanism rewards honest participation and penalizes free riding, and incentive-compatible mechanisms were designed to sustain that cooperation [666]. Alpcan and Başar's monograph *Network Security: A Decision and Game-Theoretic Approach* and the survey *Game Theory Meets Network Security and Privacy* consolidated this first decade [12, 380, 510]. FlipIt, a game of stealthy takeover in which each player periodically seizes control of a resource without knowing when the other last moved, captured the persistent and covert character of advanced attacks [577].

Deployed security games. A parallel line of work brought game theory into operational use. Beginning with the ARMOR system at Los Angeles International Airport in 2007, Stackelberg security games were used to randomize checkpoints and patrols so that an adversary who observes the defense over time cannot exploit a predictable pattern [483, 560]. Algorithms for massive games made the approach scale to problems such as scheduling federal air marshals, and later deployments covered port and ferry protection, transit systems, and wildlife conservation [321, 543]. These deployments mattered beyond their own domains. They showed that game-theoretic security could be computed, implemented, and evaluated in the field, and they made commitment to a randomized policy a standard idea in security practice.

Cyber-physical systems and resilience. From the 2010s, the field turned to systems in which cyber actions have physical consequences. Robust and resilient control designs for power systems treated cyber attacks and physical disturbances within one game framework [656], and the games-in-games principle coupled a cyber-layer stochastic game with a physical-layer control game so that security and resilient control could be designed together [657, 660, 680]. The same principle was extended to multi-agent autonomous systems, interdependent infrastructure networks, and dynamic games for secure control-system design [603, 262, 276]. As control moved to the cloud, strategic trust mechanisms decided when a device should accept commands from a possibly compromised cloud [478, 109, 474], and cyber insurance was analyzed as a game among insurers, users, and attackers [621, 623].

Deception. Deception became a design variable rather than a trick. Deceptive routing games, feedback-driven moving-target defense, and dynamic games of belief manipulation showed how a defender can shape what an attacker learns [665, 658, 256]. Signaling games with evidence captured deception that is imperfect and can leak, reinforcement learning optimized honeypot engagement, and a taxonomy organized the growing family of deception games [476, 264, 475, 479]. Advanced persistent threats were modeled as multi-phase and multi-stage games, with defense-in-depth strategies and Bayesian belief updates across the kill chain [496, 673, 268].

Learning, cognition, and AI. The most recent era combines games with learning and with models of human and machine reasoning. Reinforcement learning became the basis of a feedback view of cyber resilience, in which the defender learns to respond and recover while the attacker adapts [269, 278], and conjectural online learning lets a defender adjust its response as its model of the attacker is revised [231]. Cognitive security treats human attention and bias as an attack surface and brings experiments with human attackers into the loop [272, 53]. Zero-trust authentication, meta-games for penetration testing, and a unified theory of cyber resilience built from game, control, and learning theory extend the toolkit [213, 214, 637]. Large language models now enter as players themselves: LLM–Nash and LLM–Stackelberg games study equilibria among agents that reason in natural language, with applications to spear-phishing and agentic deception [641, 640, 359]. Section 1.3 returns to these developments. Across all six eras, the underlying question has stayed the same: how should a defender act when the threat it faces responds to the defense?

1.1.4 Threat Landscape and Strategic Heterogeneity

Figure 1.3 gives a qualitative picture of this contemporary threat landscape. The horizontal axis represents attacker sophistication: technical skill, operational maturity, resources, and domain knowledge. The vertical axis represents

stealthiness: the ability to avoid detection, hide within legitimate behavior, or exploit trusted channels. Marker size represents relative impact, including operational, financial, safety, and mission consequences. The purpose of the figure is not to provide a fixed empirical ranking. It is to show why security decisions cannot be reduced to a single list of vulnerabilities. Different threats occupy different strategic regions because they require different attacker capabilities, expose different defender blind spots, and create different losses.

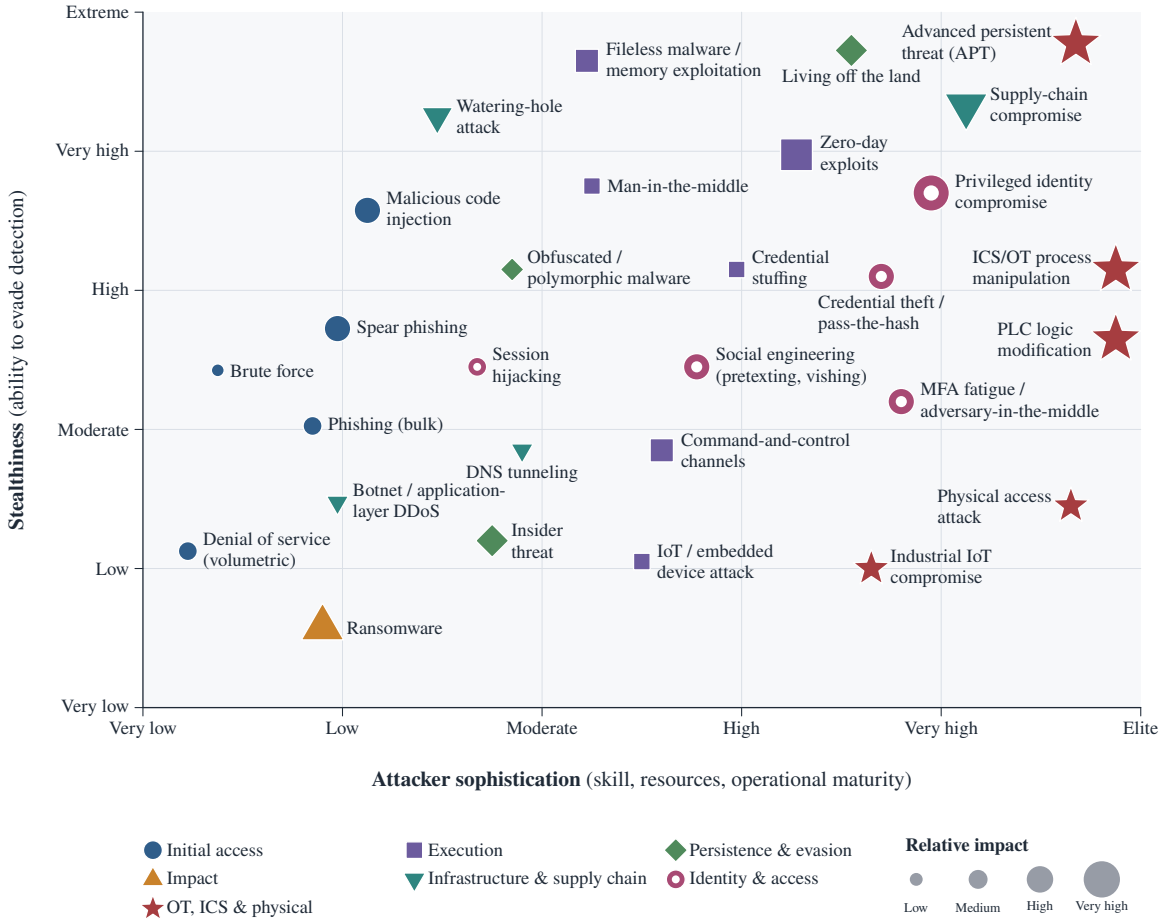


Figure 1.3: A qualitative cyber threat landscape. Each threat is placed by the attacker sophistication it requires and the stealthiness it affords; marker size indicates relative impact and marker shape indicates the threat family. Positions are illustrative rather than an empirical ranking. The map shows why game-theoretic security models must distinguish attacker capability, defender observability, and consequence severity.

Threats in the lower-left region, such as volumetric denial-of-service, bulk phishing, and brute-force attacks, are often easier to detect and require less specialized capability, although they may still be common and costly at scale. In game-theoretic terms, these threats often motivate models of resource allocation, filtering, deterrence, and investment under high-volume pressure. Threats in the middle of the figure, such as credential stuffing, session hijacking, DNS tunneling, command-and-control channels, and social engineering, are more strategic because they exploit trust relationships, legitimate credentials, communication channels, or human behavior. They force the defender to reason about monitoring, false alarms, user incentives, and incomplete information.

The upper part of the figure contains stealthier and more sophisticated threats, including fileless malware, memory exploitation, watering-hole attacks, man-in-the-middle attacks, zero-day exploits, supply-chain compromises, privileged identity compromise, and advanced persistent threats. These threats are strategically significant because the attacker may remain hidden while learning about the defender, and the defender may have to act before knowing whether a compromise has occurred. This is precisely the setting in which Bayesian games, signaling games, hypergames, and dynamic games become useful. The defender is not merely asking which control blocks a known exploit; the defender is

asking what an attacker with unknown capability will do, what evidence should change beliefs, and when response or recovery should be triggered.

The right side of the figure also highlights identity, supply-chain, and operational-technology threats. These threats show that modern security is increasingly about trust, dependency, and physical consequence. A supply-chain compromise exploits a trusted relationship rather than only a local software flaw. Privileged identity compromise turns an authorized pathway into an attack path. Industrial control and operational-technology attacks may require specialized knowledge but can produce high-impact physical disruption. These cases motivate the broader scope of the book: complete-information games for resource allocation, incomplete-information games for hidden capabilities and beliefs, Markov and stochastic games for evolving attacks, learning games for adaptive defense, and mechanism-design models for incentives across organizations.

1.1.5 Strategic Adversaries and Multi-Stage Campaigns

The central contribution of game theory to security is that it endogenizes the adversary. In a nonstrategic risk model, an attack distribution is often treated as fixed. In a strategic model, the attack distribution may change when the defender changes policy. Increasing inspection at one entry point can move attacks elsewhere. Publishing a patch can reduce the vulnerability of updated systems while revealing information to attackers who target unpatched systems. Deploying a honeypot can collect intelligence, but it can also change an attacker's belief about the network. Hardening a high-value target can lower its direct risk, yet increase the relative attractiveness of connected assets. These are not peripheral effects. They are often the heart of the security problem.

The same point appears in the operational language of cyber kill chains and adversary tactics, techniques, and procedures. A serious intrusion is rarely a single isolated event. It unfolds as a campaign: reconnaissance, preparation, initial access, execution, persistence, privilege escalation, lateral movement, command and control, collection, exfiltration, and impact. The intrusion kill-chain perspective emphasizes that defenders can disrupt an attack at different stages, while MITRE ATT&CK provides a widely used knowledge base of adversary tactics and techniques grounded in real observations [285, 396]. Recent campaigns against critical infrastructure, in which state-sponsored actors pre-positioned themselves in operational networks for months without acting, show how long and deliberate these campaigns can be [138]. These frameworks show why security is sequential and strategic. An attacker chooses whether to spend more effort on reconnaissance, whether to remain stealthy or move quickly, whether to exploit a foothold immediately or preserve access, and whether to adapt when a control blocks one path. A defender chooses where to detect, deny, deceive, degrade, contain, and recover. Game theory is the natural framework for this setting because each stage changes the options, beliefs, and incentives at later stages. For game-theoretic modeling, the kill chain is best read as the temporal skeleton of the interaction: actions at early stages shape information, feasible actions, and payoffs at later stages. This reading has been made precise in multi-stage and multi-phase game models of advanced persistent threats, in which each phase is a game whose outcome sets the initial conditions of the next, and in Bayesian dynamic games in which the defender updates its belief about the attacker's type as the campaign unfolds [673, 496, 265, 268].

Figure 1.4 illustrates this idea. The left column shows the attacker kill chain, moving from reconnaissance to actions on objectives. The right column shows the corresponding defender chain, moving from situational awareness to recovery and resilience. The center column is the game-theoretic layer: each stage is represented by a strategic game, labeled G_1 through G_7, that captures the local attacker–defender contest at that point in the campaign.

The first stage, reconnaissance, is naturally modeled as a deception game: the attacker tries to gather information about assets and vulnerabilities, while the defender tries to shape what can be observed through concealment, misinformation, monitoring, and situational awareness. Weaponization becomes a configuration game, because attacker preparation depends on the exposed vulnerability landscape while the defender hardens systems and manages configurations. Delivery becomes an evasion–detection game, exploitation becomes an exploit–defense game, installation becomes a resource-allocation game between persistence and containment, command and control becomes a communication-channel game, and actions on objectives become an impact-mitigation game.

The point is not that these seven labels are the only possible models. The point is that a cyber campaign is not one undifferentiated attack. At each stage, the players have different actions, observations, costs, and objectives. A defender who understands this staged structure can ask more precise questions: where should information be hidden or revealed, where should detection be strongest, where is prevention more valuable than recovery, and where can a small change in incentives shift the attacker's continuation strategy? Game theory turns the kill chain from a descriptive sequence

into an analytical framework for reasoning about defender choices, attacker adaptation, and resilience across the entire intrusion lifecycle.

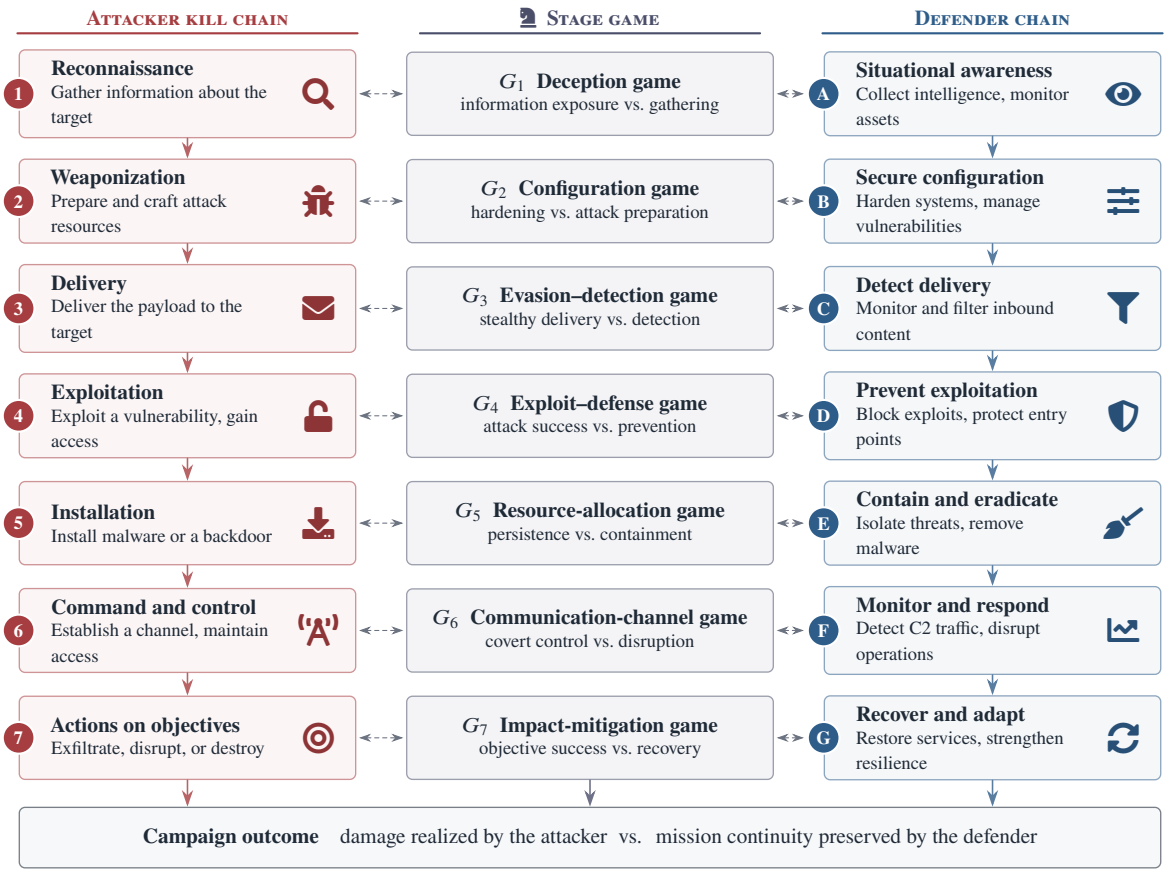


Figure 1.4: A multi-stage game-theoretic view of a cyber campaign. Each stage of the attacker kill chain (left) is paired with the corresponding defender capability (right) through a stage game G_k (center) that names the local strategic tension. Outcomes of earlier stage games shape the information, feasible actions, and payoffs of later ones, and together they determine the campaign outcome.

1.1.6 Risk and Solution Concepts

Security games as strategic risk models. Security games provide a disciplined abstraction for reasoning about defense under conflict and uncertainty [12, 380, 151]. The defender may choose monitoring levels, patching schedules, patrol routes, access-control policies, deception signals, backup capacity, or investment levels. The attacker may choose a target, attack timing, exploit path, reconnaissance effort, level of stealth, or whether to attack at all. Payoffs encode damage, cost, disruption, privacy loss, safety impact, reputation, effort, and risk of detection. Information sets encode what each side can observe and what must be inferred. Timing determines whether the interaction is simultaneous, sequential, repeated, or dynamic.

Risk quantification is the point at which this strategic abstraction becomes operational. A common engineering view represents cyber risk through likelihood, impact, and uncertainty [422]. Game theory keeps these quantities, but makes them action-dependent. The likelihood of compromise depends on what the defender protects, what the attacker observes, which attack path the attacker chooses, and how the defender is expected to respond. The impact of compromise depends not only on the asset, but also on timing, propagation, containment, recovery, and the attacker's continuation strategy. Game theory therefore does not replace risk assessment; it makes risk assessment responsive to adversarial adaptation.

In this book, a security game refers broadly to a strategic model in which defenders choose protective, monitoring, response, or incentive actions while adversaries choose attack, evasion, reconnaissance, or manipulation actions. The purpose is not to force every security problem into the same template, but to make the players, actions, timing, information, payoffs, and solution concepts explicit.

Why game theory is the right framework. This abstraction is powerful because it separates the structure of a security problem from the details of a particular technology. The same strategic pattern can appear in airport screening, network intrusion detection, cloud access control, vulnerability disclosure, supply-chain assurance, and industrial control. In each case, a defender has limited resources, an attacker can observe or infer something about the defense, and both sides make choices under uncertainty. The technical details matter, but the strategic form helps identify what is common across domains: scarce protection, adaptive targeting, imperfect monitoring, asymmetric information, and competing objectives.

This is why game theory is the right framework when security depends on adaptive response. Optimization is useful when the environment is fixed, but security environments are not fixed because adversaries respond. Statistics is useful for estimating risk, but security risks are not only sampled from nature because attackers choose when and how to generate them. Control theory is essential for maintaining system behavior, but an adversary may intentionally manipulate the state, observations, and control channels. Game theory connects these perspectives by treating the defender's action and the attacker's response as part of the same model. It asks what the attacker will do after observing a defense, what the defender should commit to before the attacker moves, what each side believes when information is incomplete, and how present actions shape future choices. These are precisely the questions that distinguish security from ordinary reliability engineering.

Resilience and strategic randomization. The resilience viewpoint also changes the meaning of a good defense. A defense is not good only because it reduces the probability of compromise. It may also be good because it slows the attacker, preserves observability, limits physical damage, maintains a degraded but safe operating mode, or creates time for recovery. These effects are strategic. If a defender can credibly recover quickly, the attacker may shift to stealthier or more persistent behavior. If a defender can isolate a compromised subsystem, the attacker may invest more in lateral movement. If a defender can deceive the attacker about recovery capacity, the attacker may misallocate effort. Game theory is valuable because it represents these preparation, response, and recovery choices as part of one adversarial interaction.

Consider a simple resource-allocation problem. A defender cannot protect every target at full strength, so it must decide how to distribute monitoring or hardening across assets. If the defender uses a deterministic pattern, an attacker may learn that pattern and attack the least protected target. Randomization can improve security because it makes the attacker's payoff uncertain. This is not randomness for its own sake; it is strategic randomization. A mixed defense policy can make several attack options similarly unattractive, reducing the value of reconnaissance and preventing the attacker from exploiting a predictable weakness. Stackelberg security games formalize this idea when the defender commits to a randomized policy and the attacker best-responds [560, 543]. Chapter 3 develops these models and shows how the optimal randomized coverage is computed.

Beyond allocation. Allocation is only the simplest reason that security is strategic. Defenders and attackers rarely know each other's capabilities, so the contest is over beliefs as much as actions; a defender who can credibly commit to a policy can shape the attacker's choices before any attack occurs, though a commitment that reveals too much can be exploited; and decoys, honeypots, and misleading signals can turn uncertainty into a defensive resource. Many security failures, moreover, involve no attacker decision at all: organizations underinvest because the benefits of their security spill over to others, users ignore advice whose cost is immediate and whose benefit is uncertain, and vendors, insurers, and platforms pursue objectives that only partly align. Security is therefore not always zero-sum, and improving it can require changing incentives through contracts, audits, insurance, liability, or reporting rules rather than finding a better response to an adversary. Section 1.2 organizes these strands into the recurring themes of the book.

Solution concepts as security reasoning. Solution concepts are the bridge between quantified risk and security policy [410, 449, 46]. Each one encodes a different kind of security reasoning, as summarized in Table 1.1. Minimax reasoning gives conservative guarantees when the defender wants protection against worst-case behavior. A Nash equilibrium, in which no player wants to deviate unilaterally, is appropriate when attackers and defenders choose

simultaneously and each side's risk depends on the other's action. A Stackelberg equilibrium captures commitment: a defender commits to an inspection, monitoring, deception, or patrol policy, and the attacker observes or learns it before responding. Bayesian equilibrium captures private information about attacker capability, defender monitoring, or asset value. Repeated-game reasoning captures reputation, deterrence, and the long-run consequences of retaliation or cooperation. Markov and stochastic-game solution concepts apply when present actions change future states, as in persistent intrusion, moving-target defense, resilient control, and recovery. A solution concept is therefore not merely a mathematical ending to a model; it states what kind of security reasoning is being performed. These concepts do not eliminate engineering judgment, but they organize it.

Table 1.1: Solution concepts and the security reasoning they formalize

Solution concept	Security reasoning	Developed in
Minimax and saddle point	Robust protection against worst-case attacks	Chapter 3
Nash equilibrium	Mutual best response under simultaneous choice	Chapter 3
Stackelberg equilibrium	Deterrence and redirection by commitment	Chapter 3
Bayesian and perfect Bayesian equilibrium	Inference and signaling under hidden types	Chapter 4
Markov perfect and stochastic-game equilibrium	Adaptation as the system state evolves	Chapters 5–6
Learning dynamics and bounded-rational equilibria	Behavior when players learn or reason imperfectly	Chapters 7, 11

The equilibrium perspective is sometimes misunderstood as a claim that attackers and defenders always behave perfectly rationally. That is not the right interpretation. Equilibrium is a modeling device for studying consistent incentives. It asks what patterns of behavior could persist when each player understands the relevant incentives and anticipates the others. When this assumption is too strong, the model can be relaxed through bounded rationality, quantal response, empirical game-theoretic analysis, or learning. The point is not to worship equilibrium, but to use it as a benchmark. Once the benchmark is clear, deviations from it become meaningful rather than vague.

Strategic side effects and modeling limits. Game-theoretic models also help explain why purely local security improvements can fail. A control that reduces risk at one node may increase risk elsewhere by changing attack paths. A detection rule that works against yesterday's malware may induce attackers to change behavior. A strict authentication policy may improve access control while increasing unsafe workarounds by users. A disclosure policy may accelerate patching while increasing short-term exploitation. Security design therefore requires reasoning about second-order effects. Game theory gives the analyst a vocabulary for those effects: best responses, equilibrium shifts, signaling, externalities, incentives, learning, and strategic substitutes or complements.

Remark 1.1.1. A security control changes the strategic environment. It can reduce risk directly, but it can also change beliefs, incentives, timing, and the future actions of other decision makers. Game-theoretic modeling is valuable because it asks what the defender's action causes other actors to do next.

The value of game theory is not that it makes security perfectly predictable. Real attackers may be boundedly rational, poorly informed, opportunistic, ideological, or organizationally constrained. Real defenders may have incomplete data, limited budgets, legacy infrastructure, and conflicting priorities. A useful model does not pretend these complications disappear. Instead, it states assumptions clearly, reveals the logic of strategic dependence, and provides a baseline against which more realistic behavioral, empirical, and learning-based models can be developed.

Role in this book. This book uses game theory as a language for security reasoning. The goal is to connect mathematical models to concrete security questions: how to allocate scarce defenses, how to reason under incomplete information, how deception changes beliefs, how attacks and defenses evolve over time, how learning agents adapt, how incentives can be designed, and how cyber-physical systems can remain safe and resilient under strategic pressure.

1.2 Strategic Themes in Security Games

This section identifies the strategic themes that organize the book. The goal is not to describe the technical details of cyber-physical systems; Chapter 2 takes up that domain background. The goal is to explain why many security problems have the structure of a game. Eight themes recur, and each pairs a practical security question with a role for game theory.

Allocation and adaptation. The first theme is *resource allocation*: where should limited inspection, monitoring, hardening, deception, backup, and recovery resources be placed? Game models answer by representing the constrained choices of both sides, including the tradeoffs among targets, costs, and coverage. The second theme, *adaptation*, asks how an attacker will change its behavior after the defender changes policy. Here game theory treats attack likelihoods and attack paths as responses to the defense rather than as fixed inputs, which is the main difference between a strategic and a nonstrategic risk model.

Information, commitment, and deception. The third theme is *information*: what should each side infer from alerts, silence, probes, disclosures, and observed behavior? Models of asymmetric information, belief updating, signaling, hidden types, and mismatched perceptions give this question a precise form. The fourth theme, *commitment*, asks when a visible or credible defensive policy can deter, redirect, or reshape an attack; Stackelberg and sequential models capture both the value of committing and the risk of being exploited by an attacker who observes the commitment. The fifth theme, *deception*, turns information around. It asks when controlled uncertainty improves security rather than merely adding complexity, and it studies how decoys, honeypots, moving-target defense, and selective disclosure alter the attacker's beliefs.

Resilience, incentives, and risk. The sixth theme is *resilience*: how should a system prepare, continue operating, respond, recover, and learn after partial compromise? Game models link the strategic interaction to dynamic response, recovery capacity, and future incentives. The seventh theme, *incentives and institutions*, asks why users, firms, vendors, insurers, and regulators underinvest or fail to coordinate; mechanism design, contracts, and network games align private incentives with collective security. The eighth theme, *risk quantification*, asks how likelihood, impact, uncertainty, and residual risk should change when adversaries adapt. Game theory connects risk assessment to actions, beliefs, payoffs, and solution concepts, so that a risk estimate becomes a function of the defense rather than a constant.

These themes are not separate topics that appear only once; they are lenses that return throughout the book. Chapter 3 introduces the basic strategic forms for resource allocation, adaptation, and commitment. Chapter 4 studies asymmetric information, signaling, and perception. Chapters 5 and 6 add dynamics and state evolution for multi-stage attacks, resilient control, and recovery. Chapter 7 studies learning and adaptation. Chapters 8 through 11 carry the same themes into networked, institutional, compositional, and bounded-rational settings, and Chapters 12 through 16 show how they combine in concrete cyber-physical domains.

1.3 Recent Advances and Challenges

Game-theoretic security has moved far beyond small matrix games between an attacker and a defender. Surveys of the field now cover intrusion detection and response, privacy, network security, moving-target defense, and dynamic games for cyber-physical systems [151, 323, 173, 276]. The models have grown to include large systems, uncertain opponents, learning agents, deceptive environments, networked dependencies, incentive mechanisms, and boundedly rational behavior. This growth follows the reality of modern security. Defenders operate where the model is incomplete, the opponent adapts, the infrastructure is interconnected, and the consequences of error may be physical, economic, or social. This section reviews the main directions and the challenges that remain open in each.

1.3.1 Cyber Resilience and Learning

Cyber resilience has become an organizing theme for this broader view. Resilience is not a stronger form of protection. It is a different property: the ability of a system to plan for, absorb, recover from, and adapt to adverse events, including attacks that partially succeed [364, 507]. The question is no longer only how to prevent an intrusion, but how to continue essential functions under stress, respond before damage propagates, and recover afterward. This makes security explicitly dynamic. Proactive defenses shape the attacker's initial choices. Responsive defenses shape the attacker's

continuation strategy once an incident begins. Retrospective learning changes future incentives by improving detection, patching, recovery, and institutional memory. Because the three stages are coupled, a resilience theory benefits from the confluence of game, control, and learning perspectives [637, 276].

Learning is the second major advance. In many security problems, neither side knows the full payoff structure, transition model, or opponent strategy. Reinforcement learning and multi-agent learning let a defender adapt its policy from data while keeping a strategic interpretation of the interaction [269, 278, 571]. Self-play between learned attackers and defenders in simulated networks has begun to produce defense strategies that can be evaluated against adaptive opponents rather than fixed attack scripts [232]. These methods matter most for Markov games, honeypot engagement, adaptive network defense, and cyber-physical systems whose state spaces are too large for exact dynamic programming. Learning does not remove the need for game theory. It raises new game-theoretic questions about convergence, exploration, regret, and robustness against opponents who are also learning.

Learning also creates a new vulnerability. A learning defender can be manipulated through poisoned observations, misleading probes, or attacks chosen to shape its future policy. An adversary who falsifies the cost signals seen by a reinforcement learner can steer it toward a policy of the adversary's choosing without ever touching the learner's code [280]. A classifier trained on data that an adversary can contaminate is, in effect, playing a game against that adversary, and its training procedure should be designed with that game in mind [622]. Security learning therefore differs from ordinary prediction. The data are not passive; they may be generated by an opponent who knows that the defender is learning. Game-theoretic learning is concerned not only with performance, but with strategic data generation, safe exploration, and robustness against manipulation.

1.3.2 AI-Enabled and Agentic Security

Recent advances in artificial intelligence make the security problem harder in two related ways. First, AI can strengthen the defender through automated triage, anomaly detection, malware analysis, vulnerability prioritization, threat intelligence, and adaptive response. Second, the same capabilities can strengthen the attacker by lowering the cost of reconnaissance, improving phishing and social engineering, accelerating exploit development, and enabling more adaptive campaign planning [92, 617]. The result is not simply a faster version of the old problem. AI changes the strategic balance because it reduces the cost of experimentation, expands the set of feasible attack paths, personalizes attacks, and lets both sides update their policies more quickly.

Defenders increasingly need AI, but AI alone is not enough. Modern infrastructures generate more telemetry, alerts, vulnerability reports, and incident evidence than human analysts can process. AI can help estimate the state of the system, infer attacker movement, forecast likely impact, rank response options, and update risk estimates as evidence arrives. Yet AI does not by itself specify what the defender should optimize, which uncertainty matters, what the attacker will do after the AI-assisted response is deployed, or when a recommendation is too risky to follow. Game theory supplies that specification: players, actions, information, payoffs, dynamics, and a solution concept. AI estimates and adapts; game theory explains how those estimates should enter strategic decisions under adversarial response [426, 583, 637].

Remark 1.3.1. AI-enabled security should be modeled as a change in the strategic environment, not only as a new tool. If AI lowers the cost of reconnaissance, improves deception, automates exploit generation, or accelerates response, then it changes the feasible action sets, the information available to the players, and the speed at which strategies adapt. These are game-theoretic primitives.

Adversarial machine learning made the connection between games and AI security visible a decade before the current wave of generative models [70, 465, 141]. A classifier, detector, or decision system plays the role of a defender, while an adversary chooses examples, perturbations, poisoning strategies, or evasion paths. Improving a detector changes the distribution of attacks it faces, so the interaction is strategic by construction. Game-theoretic reasoning clarifies when robustness is a minimax problem, when the attacker has incomplete information, when training data can be manipulated, and when defense should be randomized. Taxonomies such as the NIST adversarial machine learning report and the MITRE ATLAS knowledge base now organize these attacks in the same way that ATT&CK organizes conventional intrusions [583, 398].

Generative AI and AI agents sharpen the need for this reasoning. A chatbot mainly changes the information environment; an agent changes the action environment. An agent built on a large language model can plan, remember, retrieve context, call tools, write and execute code, and act under a non-human identity [590]. The security question is

then no longer whether a model can be fooled by a single input, but how an adversary can influence an AI-enabled workflow over time. Indirect prompt injection is the clearest example. Instructions hidden in a web page, document, or email that the agent retrieves can redirect the agent's behavior without the user ever seeing them [226, 371]. Memory and retrieval poisoning, tool misuse, confused-deputy behavior, excessive agency, and compromised agent identities extend the same pattern [146, 453, 452, 428]. These risks are most serious when the agent can modify files, trigger transactions, change access policies, or act on cyber-physical processes.

The same capabilities cut both ways. A defensive agent can compress time: it can correlate alerts across telemetry streams, summarize incident evidence, generate detection logic, prioritize vulnerabilities by exposure and mission impact, emulate attacker paths, and recommend containment. In cyber-physical settings it may also fuse sensor evidence, diagnose abnormal process behavior, and check whether a control action is safe. An offensive agent can automate reconnaissance, generate and test exploit variants, tailor social engineering to a target, and decide when to stay stealthy rather than act. In both roles, the value of the agent depends on how much authority it has, what tools it can call, what it can observe, and when a human must approve. These are not only software-design choices. They are strategic choices about delegation, commitment, observability, and risk.

Game theory gives the right abstraction because agentic security is a contest over policies, not isolated prompts. The defender chooses an architecture for delegation, monitoring, tool access, approval gates, logging, deception, and recovery. The attacker chooses whether to attack the model, the memory, the retrieval channel, the tool interface, the human overseer, or a downstream system that trusts the agent. A tightly constrained agent is safer but slower; a more autonomous agent responds faster but magnifies the impact of manipulation. Human approval reduces catastrophic actions but creates bottlenecks that an attacker can exploit. These tradeoffs lead naturally to Stackelberg, partially observable, stochastic, and learning-in-games models. Benchmarks that pit prompt-injection attacks against defended tool-using agents are a first step toward evaluating such policies empirically [144].

The relationship also runs in the opposite direction. Language-model agents can themselves be treated as players: they can be prompted or trained to compute equilibria, to reason about commitment, or to plan multi-step deception operations whose consistency is checked by a game model [641, 640, 359]. Game theory then plays a double role. It describes the strategic environment in which AI agents operate, and it supplies a discipline that keeps an agent's reasoning consistent with the incentives of the other players. Zhu [638] examines this two-way relationship between game theory and language-model agents in cyber defense, and Li and Zhu [358] argue for a shift from prevention-centric security toward agentic cyber resilience, analyzing attacker and defender workflows as coupled adaptive processes with game-theoretic formulations as a unifying design language.

1.3.3 Deception and Incentive Design

Deception has become a major theme in its own right. Traditional security often treats uncertainty as a weakness to be reduced. Game-theoretic deception shows that uncertainty can also be a defensive resource when it is deliberately shaped. Honeypots, decoys, moving-target defense, traffic morphing, and selective disclosure alter an attacker's beliefs and induce costly mistakes. A taxonomy of defensive deception distinguishes perturbation, moving-target defense, obfuscation, mixing, honey-x, and attacker engagement, and pairs each with a class of game models: signaling games, Bayesian persuasion, hypergames, and partially observable stochastic games [475, 635, 479]. Moving-target defense in particular has its own body of surveys organized around what to move, when to move it, and how to measure the benefit [118, 526].

Deception must be designed as part of a system, not as an isolated trick. A decoy that is too obvious will be ignored. A decoy that is too realistic may consume resources or create operational risk. A moving-target defense increases the attacker's uncertainty but also increases complexity for the defender. A misleading signal may buy time but erode trust if legitimate users or operators are affected. Attackers deceive too, and a defender that ignores the possibility of being deceived will respond to the wrong threat [270]. Game theory forces the designer to represent these tradeoffs explicitly and to ask whether the other side has an incentive to believe, ignore, test, or exploit the signal. Chapter 13 develops these models in detail.

Mechanism design and incentive alignment widen the scope of security games beyond direct attacker–defender conflict. Many failures occur because secure behavior is socially valuable but privately costly. Device owners postpone updates, firms underreport incidents, vendors underinvest in long-term maintenance, and organizations hesitate to share threat intelligence. Mechanism design asks how contracts, insurance, penalties, subsidies, audits, and information

policies can make secure behavior incentive compatible. Collaborative intrusion detection networks are an early example: peers share alerts only if the sharing mechanism rewards honest participation and penalizes free riding [666]. Cyber insurance is a more recent one. Designed as a contract between insurer and insured, it becomes a lever for resilience rather than only a transfer of loss [369]. The same perspective applies to supply-chain security, IoT patching, platform governance, and coordinated defense across organizations, and it is the subject of Chapter 9.

1.3.4 Scalability, Composition, and Behavioral Realism

Scalability remains a central challenge. Real cyber-physical infrastructures may contain thousands or millions of components, and exact equilibrium computation quickly becomes infeasible. Recent work uses decomposition, graph structure, mean-field approximations, sampling, factored representations, and hierarchical models to make security games computationally useful. The challenge is not only computational. A scalable model must also preserve the dependencies that matter: cascading failures, shared vulnerabilities, correlated attacks, network topology, and the interaction between local and global incentives.

Composition is one response. Instead of treating a large infrastructure as a single flat game, the analyst models subsystems as games and studies how those games interact. The games-in-games principle makes this idea concrete. A cyber-layer stochastic game between attacker and defender determines the operating mode of the system, and that mode in turn parameterizes a physical-layer control game; the two are solved jointly so that cyber defense and physical control are designed together [657, 680]. The same structure appears in the design of resilient infrastructure networks and in cross-layer defense of wireless and energy systems [103, 113]. A local industrial-control game may determine the resilience of one plant, a network game how failures propagate between plants, and a mechanism-design problem whether operators share information. Composition is difficult because equilibria do not always combine cleanly, but it is essential for cloud-enabled IoT, critical infrastructure, supply chains, and multi-layer networks, and it is the subject of Chapter 10.

Behavioral realism is another challenge. Classical game theory often assumes fully rational players with common knowledge of the game. Security practice suggests a more complicated picture. Attackers may be opportunistic, strategic, ideological, automated, or organizationally constrained. Defenders may be overloaded, risk averse, institutionally fragmented, or limited by legacy systems. Users misread warnings and trade security for convenience. Human attention and cognitive bias are themselves an attack surface, exploited by phishing, alert fatigue, and social engineering, and they can be modeled with the same system-scientific and game-theoretic tools as technical vulnerabilities [272]. Bounded rationality, bounded perception, bounded reasoning, and empirical game-theoretic analysis are therefore needed for models that aim to predict and improve real behavior; Chapter 11 takes up these questions.

Empirical validation is the companion to behavioral realism. A model that produces a clean equilibrium but fails to predict behavior is still useful as a benchmark, but it should not be mistaken for a deployable policy. Stackelberg security games earned their influence in physical security partly because they were deployed and evaluated in the field [560, 543]. Cyber security games need comparable grounding from incident data, laboratory experiments, cyber ranges, red-team exercises, operator studies, simulation environments, and benchmarks [232, 144]. Validation is hard because adversarial behavior changes when defenses change, and because many important attacks are rare or poorly observed. It is nonetheless necessary if game-theoretic security is to influence practice rather than remain a conceptual framework.

1.3.5 Engineering Constraints and Strategic Tradeoffs

The hardest open problems sit at the boundary between game theory and engineering. A CPS defense must be strategic, but it must also satisfy safety constraints, timing constraints, and physical laws. A deception strategy must influence beliefs, but it must also be implementable in a real network. A learning defense must adapt, but it must not explore unsafe actions. A mechanism must align incentives, but it must also be enforceable and acceptable to institutions.

Privacy and safety complicate the agenda further. A security policy that increases monitoring may reduce attack risk while increasing privacy risk. A control policy that prioritizes safety may reduce performance or availability. A disclosure policy that helps defenders may also inform attackers. These tensions are not merely engineering constraints. They are strategic tradeoffs among stakeholders with different preferences, so future security games must handle multiple objectives and multiple kinds of harm rather than reducing every outcome to a single loss number.

1.3.6 Directions and Open Challenges

These developments strengthen the case for game theory. As attacks become more automated, multi-stage, and AI-assisted, the defender's problem looks less like static hardening and more like strategic control under uncertainty. The defender must anticipate adaptation, manage beliefs, allocate scarce attention, decide when to reveal or conceal information, and choose policies that remain effective when the opponent learns. Game theory places all of these elements in one model: players, information, timing, actions, dynamics, incentives, and solution concepts. Without such a framework, security analysis risks becoming a catalog of controls without a theory of adversarial response. The directions reviewed above can be read together as a research agenda. Each contributes something the classical attacker–defender game lacks, and each leaves a challenge that the later chapters take up. A complementary agenda for game theory in distributed systems security, organized into analytical, systems, and integration challenges, is laid out in [3].

Adaptation and dynamics. Learning in games supplies adaptive policies for adversaries whose behavior is unknown or changing. The open challenge is to make such learning safe and trustworthy: exploration must not drive a physical process into an unsafe state, learning dynamics must converge or at least remain bounded, and the learner must stay robust when the opponent deliberately shapes the data it learns from. Cyber resilience extends the same concern from a single policy to the whole life of an incident. It contributes a view of security as preparation, response, recovery, and continuity of mission under attack, but it still needs a theory that integrates strategic behavior with control dynamics, recovery constraints, and learning across the three stages. Risk quantification ties both strands to practice. Game models already make likelihood, impact, and residual risk depend on the defender's action; the remaining difficulty is to estimate those quantities when the attacker's response, rather than a fixed threat distribution, determines what is likely.

Artificial intelligence. Adversarial machine learning contributes strategic models of evasion, poisoning, and robust prediction, yet scalable defenses against adaptive and only partially informed attackers remain elusive. Generative and agentic AI widen the problem from a single model to a workflow in which an AI system plans, remembers, calls tools, and acts. Game models of AI-assisted attack and defense, of delegation to agents, and of tool use and workflow control are only beginning to appear. The open questions are concrete ones: how to bound the effect of prompt injection, how much agency to grant, how to protect memory and retrieval channels from poisoning, where human oversight adds safety rather than a new bottleneck, and how to reason about two sides that both automate their strategies.

Beliefs and incentives. Cyber deception gives a formal account of belief manipulation and misdirection. To be useful in practice it must be credible to the attacker, observable at the right moments, affordable to operate, and integrated with detection and response rather than deployed as a separate trick. Mechanism design addresses a different gap: it aligns incentives for secure behavior across organizations and users who do not share a single objective. Its open challenges are institutional as much as mathematical, namely practical enforcement, truthful reporting of incidents and vulnerabilities, and adoption by the firms, insurers, and regulators who must run the mechanism.

Scale, people, and evidence. Network and compositional games, including games-in-games models, capture interdependent systems and cascading effects; the challenge is to keep the analysis tractable without discarding the dependencies that cause cascades in the first place. Bounded rationality brings models of attackers, users, and operators closer to how they actually decide, but those models must be calibrated empirically and must predict behavior under cognitive limits rather than merely describe it. Empirical validation, finally, connects every other direction to practice through field deployments, cyber ranges, simulators, and benchmarks. Its distinctive difficulty is that the adversary adapts to the evaluation itself, so a defense that scores well on yesterday's benchmark may not hold against tomorrow's attacker.

The purpose of this book is to build a path through these ideas. It begins with foundational models because they provide the grammar of strategic reasoning. It then develops richer models for uncertainty, dynamics, learning, networks, incentives, composition, and bounded rationality. Finally, it applies these ideas to cyber-physical domains where the strategic and physical dimensions of security cannot be separated.

1.4 Structure of the Book

Chapter 1 introduces the central thesis of the book: security is a strategic problem, and game theory provides a principled language for modeling adversarial adaptation, incentives, information, and resilience [637]. Section 1.1 develops the

connection between game theory and security in general. Section 1.2 identifies the recurring strategic themes—resource allocation, adaptation, information, commitment, deception, resilience, incentives, and risk quantification—that later chapters model with different classes of games. Section 1.3 describes recent advances and challenges and explains why the rest of the book is organized around increasingly rich classes of games.

Chapter 2, *Cyber-Physical System Security: An Introduction*, provides the security background needed for the later modeling chapters. It explains why CPS security differs from conventional cybersecurity, introduces a minimal model of a CPS under attack, and surveys threat models and attack surfaces such as false-data injection, replay, actuator compromise, and supply-chain compromise, together with kill chains, risk assessment, security goals, defense mechanisms, standards, and the incidents that motivate game-theoretic models.

The book then follows the architecture shown in Figure 1.5. The first two chapters orient the reader: Chapter 1 gives the strategic lens, and Chapter 2 gives the security and CPS background. Part I develops the core mathematical grammar of security games. Part II adds learning, network structure, incentives, composition, and behavioral realism. Part III returns to concrete cyber-physical domains, showing how the earlier models combine in industrial control, deception, IoT, supply chains, and wireless networks. Appendices A–E collect the background in optimization, probability, control, network security, and cyber risk analysis on which the main chapters draw.

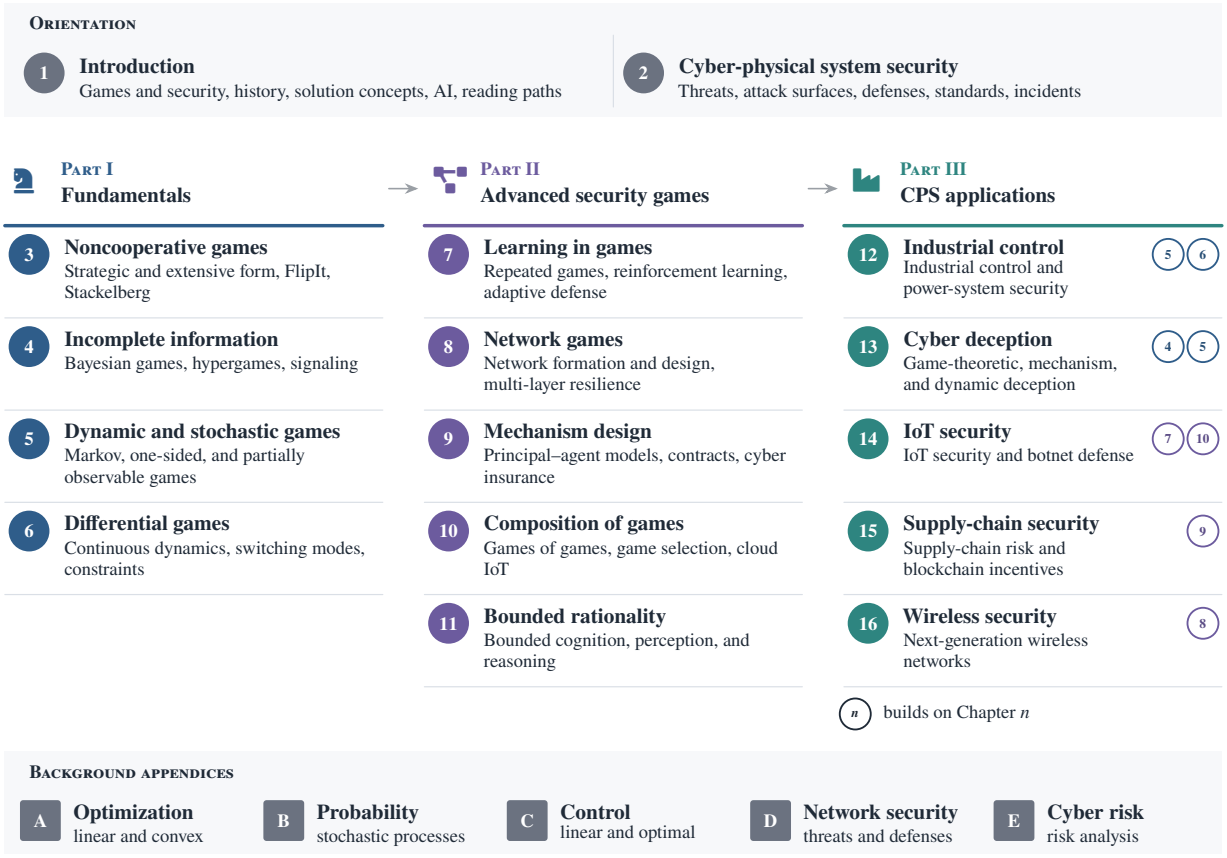


Figure 1.5: Organization of the book. Chapters 1 and 2 orient the reader; Part I develops the foundational game models, Part II the advanced security games, and Part III the cyber-physical applications. The small circled numbers beside each application chapter mark the theory chapters it builds on, and the appendices supply mathematical and security background.

Part I develops the fundamentals of security games. Chapter 3, *Noncooperative Games*, introduces strategic-form games and extensive-form games as the basic models of rational interaction [410, 449]. It then connects these models to FlipIt games, Cut the Rope, and Stackelberg security games, showing how timing, commitment, and resource allocation shape security outcomes [543]. Chapter 4, *Games of Incomplete Information*, develops Bayesian games and hypergames, then applies them to strategic trust in cloud-enabled CPS, modeled by coupled signaling and FlipIt games, and to

hypergame models of misperceived security interactions. Chapter 5, *Dynamic and Stochastic Games*, studies Markov games, one-sided information games, and partially observable games, with applications to factored games for critical infrastructure protection and partially observable games for cyber deception. Chapter 6, *Differential Games*, develops linear quadratic differential games, differential games with switching modes, and differential games with algebraic constraints, with applications to resilient control of power systems and secure control of satellites.

Part II develops advanced security games. Chapter 7, *Learning in Games and Algorithms*, studies learning in repeated games, reinforcement learning for Markov games, and evolutionary game dynamics, with applications to adaptive honeypot engagement and conjectural online learning for intrusion response. Chapter 8, *Network Games*, studies games played on networks, network formation games, network design, and network interdiction games, then applies them to multi-layer resilient network formation. Chapter 9, *Mechanism Design Theory*, introduces mechanism design, principal-agent models, and contract theory, with applications to security-aware cyber insurance and to security as a service for cloud-enabled control. Chapter 10, *Composition of Games*, develops equilibrium foundations for composed games and the games-in-games principle, with applications to cross-layer resilient control, signaling games for multi-agent autonomous systems, interdependent strategic trust in the cloud-enabled Internet of Controlled Things, and multi-resolution games for cyber campaigns. Chapter 11, *Bounded Rational Games*, studies games with bounded cognition, bounded perception, and bounded reasoning, with applications to surveillance games and security investment games.

Part III focuses on applications of game theory in cyber-physical systems. Chapter 12, *Industrial Control System Security*, presents modeling views and attack models for industrial control systems, dynamic games for layered defense-in-depth against advanced persistent threats, moving-horizon hybrid stochastic games for secure control, and cross-layer defense of wind energy systems [149, 383]. Chapter 13, *Cyber Deception*, develops game-theoretic analysis of cyber deception, mechanism-design approaches to deception, and dynamic cyber deception. Chapter 14, *IoT Security*, introduces IoT security and game-theoretic defense against IoT botnets. Chapter 15, *Supply Chain Security*, studies supply-chain security and includes a discussion of blockchain and Bitcoin-related security incentives. Chapter 16, *Next-Generation Wireless Network Security*, extends the book's game-theoretic perspective to emerging wireless systems whose security depends on spectrum sharing, distributed control, infrastructure resilience, and strategic adversarial behavior.

1.4.1 How to Use This Book

Audience. The book is written for three groups of readers. Graduate students in electrical and computer engineering, computer science, operations research, and systems engineering can use it as a text for a course on game theory for security. Researchers who work in control, networking, machine learning, or economics can use it as an entry point to the security literature, and security researchers can use it as an entry point to game theory. Practitioners and policy analysts who design or evaluate security programs can use the first two chapters and the application chapters of Part III to see what a strategic model adds to risk assessment and defensive design.

Prerequisites. The reader is assumed to be comfortable with linear algebra, calculus, and elementary probability. No prior course in game theory is required; Chapter 3 starts from first principles. The appendices collect the remaining background so that it can be consulted when needed rather than read in advance. Appendix A covers linear and convex optimization, which underlies equilibrium computation and mechanism design. Appendix B covers probability and stochastic processes for the Bayesian, Markov, and learning models of Chapters 4, 5, and 7. Appendix C covers linear systems and optimal control for the differential games of Chapter 6 and the control applications of Chapter 12. Appendices D and E give compact introductions to network security and cyber risk analysis for readers who come from mathematics or engineering rather than from security.

Chapter features. Each chapter opens with an overview, a chapter summary, and a list of learning objectives, and closes with exercises, a discussion that connects it to the rest of the book, and annotated further reading. Worked examples are set apart from the main text, and each ends with an observation that states the modeling lesson.

Exercises. The exercises at the end of each chapter are part of the development rather than an afterthought. Some ask the reader to verify a result or to repeat a worked example under different parameters; others ask the reader to formulate a game for a new security setting, to identify its players, information, and payoffs, and to interpret the equilibrium that

results. Several extend the chapter's models in directions that the text only mentions, and they are a natural source of course projects.

Code and reproducibility. The book is accompanied by Python code, organized by chapter, that reproduces the numerical results and figures of the worked examples. The code can be downloaded from the book's web page, `https://quanyan8.github.io/game-theory-for-security/`. A marker beneath an example, such as this one from the Stackelberg security game of Chapter 3,

▶ Code: `code/stackelberg_security_game.py`

names the script that computes it; the path is relative to the chapter's code folder. Each script is self-contained, requires only Python with NumPy, SciPy, and Matplotlib, prints its computation step by step so that it can be read alongside the text, and ends with checks that confirm it recovers the numbers reported in the book. Readers are encouraged to change the parameters, rerun the scripts, and use them as starting points for the exercises and for their own models; seeing how an equilibrium moves when a cost, a probability, or an information structure changes is often the fastest way to understand it.

Reading paths. Chapters 1 and 3 are the common core: every later chapter assumes the strategic vocabulary of Chapter 1 and the equilibrium concepts of Chapter 3. Beyond that core, the dependencies follow the bridges in Figure 1.5, and several paths are possible.

- *A one-semester course on security games* can cover Chapters 1 and 3–7, adding one application chapter from Part III as a capstone.
- *A cyber-physical security track* can follow Chapters 1–3, 5, and 6 into Chapter 12 on industrial control systems, with Appendix C as background.
- *A deception and AI track* can follow Chapters 1, 3, 4, 7, and 11 into Chapter 13 on cyber deception, drawing on the discussion of AI agents in Section 1.3.
- *An incentives and policy track* can follow Chapters 1–3, 8, and 9 into Chapter 15 on supply-chain security, with Appendix E as background.

Readers who already know game theory can skim Part I and move directly to Part II; readers who already know security can treat Chapter 2 and Appendices D and E as reference.

1.5 Discussion, Summary, and Further Reading

This chapter introduced the book's organizing claim: security is strategic because attackers, defenders, users, platforms, and institutions adapt to one another. The rest of the book develops this claim through increasingly structured models. Chapter 2 supplies the cyber and cyber-physical security background, while Appendices D and E provide compact references on network security and cyber risk analysis. Chapter 3 begins the formal game-theoretic development.

How later chapters build on this one. Later chapters do not repeat the general motivation for game theory; they specialize it. Chapter 3 gives the static and extensive-form notation, Chapter 4 adds incomplete information and subjective game models, Chapters 5 and 6 add state evolution in discrete and continuous time, and Chapter 7 adds learning. Chapters 8 through 11 study network structure, mechanism design, composition of games, and bounded rationality. Chapters 12 through 16 return to the themes of this chapter in specific domains, with Chapter 13 taking up cyber deception in depth.

Further reading. For standard game-theoretic foundations, see Osborne [449] and Myerson [410]; for dynamic games, see Başar and Olsder [46]. For network security and decision making, see Alpcan and Başar [12] and the survey by Manshaei et al. [380]. For Stackelberg security games and deployed systems, see Tambe [560] and Sinha et al. [543]. Broad surveys of game theory for security are given by Do et al. [151] and, for intrusion detection and response, Kiennert et al. [323]; dynamic games for cyber-physical security are surveyed in [173, 276]. For the resilience perspective that combines game, control, and learning theories, see [637, 278]; for resilience as a concept distinct from risk, see Linkov and Kott [364] and NIST SP 800-160 Vol. 2 [507]. Game-theoretic models of multi-stage

campaigns and advanced persistent threats appear in [673, 265], and the games-in-games approach to cross-layer CPS design is developed in [680]. For deception, see the taxonomy of Pawlick et al. [475] and the monograph [479]; for human and cognitive vulnerabilities, see [272]. For attack lifecycles and operational cyber-defense language, see the intrusion kill-chain perspective [285], MITRE ATT&CK [396], and NIST risk and cybersecurity guidance [422, 430]. For adversarial machine learning, see [70, 583] and the edited volume [311]; for generative and agentic AI security, see [226, 146, 426, 452], and for the interplay of game theory, language models, and agentic AI in cyber defense, see the survey of Li et al. [357] together with [638, 358]. Game-theoretic control of socio-technical networks, in which human behavior and incentives are coupled with cyber-physical infrastructure, is reviewed by Zhu and Başar [663], with emphasis on Stackelberg games, mechanism design, and dynamic games for misinformation management and resilience, and by Başar et al. [44], who contrast cooperative and noncooperative paradigms and connect equilibrium analysis with feedback learning and incentive design. Recent edited volumes collect work on the foundations of cyber deception [672] and on autonomous cyber resilience [312], and Rass et al. [495] give a German-language monograph on a game-theoretic approach to cybersecurity in critical infrastructures.

Summary. The main modeling lesson is that a security decision is rarely a one-sided optimization. A defense changes attacker incentives; an attack changes defender information and future options; and both sides may reason with partial information, limited resources, and imperfect models. Solution concepts make explicit which kind of reasoning is being used: robust protection, mutual best response, commitment, inference, or adaptation. This is the thread that ties the chapters together.

1.6 Exercises

The exercises below apply the concepts of this chapter to concrete security settings. Each exercise states its setting and the results to be derived.

Exercise 1.1. **Modeling a strategic security problem.** A university network administrator must decide whether to deploy an expensive endpoint-detection system before a semester begins. An attacker observes public information about the university's defenses and then decides whether to attack the student-information system, attack a less valuable public website, or wait. Write a compact game description for this setting. Your answer should identify the players, give at least two feasible actions for each player, state what information each player has when choosing an action, and propose one payoff term for the defender and one payoff term for the attacker. Conclude with two or three sentences explaining why the defender's best action depends on the attacker's expected response.

Exercise 1.2. **Adaptation and displacement.** A company hardens its virtual private network after a wave of credential-stuffing attacks. Six months later, phishing attempts against its help desk have doubled.

(a) Explain why a nonstrategic risk model, in which attack likelihoods are fixed inputs, would misjudge the benefit of the VPN hardening.

(b) Using the kill-chain view of Figure 1.4, identify the stage at which the attacker adapted and the defender capability that should be strengthened next.

(c) Give one reason why the overall risk might still be lower after the hardening, even though one attack type increased.

Exercise 1.3. **Matching security reasoning to solution concepts.** For each of the following statements, name the solution concept from Table 1.1 that best captures the reasoning, and justify your choice in one or two sentences.

(a) "We must assume the worst-case attacker and minimize our maximum loss."

(b) "Our patrol schedule will eventually be observed, so we should choose it knowing that the attacker will respond to it."

(c) "We do not know whether this intruder is a criminal or a state actor, so our response must work against both."

(d) "Neither side can gain by changing its strategy alone, given what the other side is doing."

Exercise 1.4. Delegating defense to an AI agent. A security operations team considers giving an AI agent authority to isolate endpoints and revoke credentials without human approval when it detects an intrusion.

(a) List the players in this setting. Is the agent itself a player, a tool of the defender, or both? Explain.

(b) Describe one way an attacker could exploit the agent's authority, for example through manipulated alerts or injected instructions in the data the agent reads.

(c) Explain the tradeoff the defender faces in choosing how much authority to delegate, and describe one design, such as approval thresholds or randomized human review, that changes the attacker's incentives.

Exercise 1.5. Incentives beyond direct attack. Several hospitals share a regional health-information exchange. A breach at one hospital can expose patient records held by the others, but each hospital pays the full cost of its own security upgrades.

(a) Explain why each hospital may invest less in security than would be best for the region as a whole.

(b) Is this a zero-sum interaction? Explain why or why not.

(c) Propose one mechanism, such as a shared insurance pool, a minimum standard, or a cost-sharing rule, and explain how it changes each hospital's incentives.

Part I

Fundamentals of Security Games

Chapter 3

Noncooperative Games

A defender who must decide where to place a few patrols, how often to rotate a cryptographic key, or whether to isolate a host after an alert faces an opponent who decides as well. The value of each defensive choice depends on the attack, and the best attack depends on the defense. Two questions follow: what can each side guarantee against the worst behavior of the other, and which choices are stable once each side anticipates the other? The answers depend on whether the players move simultaneously, whether one commits first, and what each observes before acting. This chapter introduces the language of noncooperative game theory in which these questions are posed, and fixes the notation of the book.

We begin with normal-form games, in which players choose without observing one another's choices. Zero-sum matrix games come first: security levels, saddle points, and mixed strategies lead to the minimax theorem, computed graphically or by linear programming. General-sum games follow, with best responses, Nash equilibrium, fictitious play, and correlated equilibrium as the outcome of regret matching. Extensive-form games then add the order of moves and what each player knows: information sets and behavioral strategies, perfect recall and Kuhn's theorem, subgame perfection, Stackelberg games, level-k thinking, and signaling games with perfect Bayesian Nash equilibrium. The applications section collects the security examples: the cut-the-rope game of stealthy intrusion, the FlipIt game of stealthy takeover, and Stackelberg security games solved by target enumeration [12, 380, 560].

Later chapters build on these representations: Chapter 4 generalizes signaling games to incomplete information, Chapters 5 and 6 add an evolving state, Chapter 7 studies learning dynamics, and Chapter 11 extends level-k reasoning. Elementary results, such as the properties of security levels, backward induction, and target induction, are proved; the minimax theorem, Nash's existence theorem, and Kuhn's theorem are stated with references.

Chapter Overview

1. Normal-form games
   - two-player zero-sum matrix games: security levels, saddle points, mixed strategies, the minimax theorem, and graphical and linear-programming computation;
   - N-player general-sum games: best responses, Nash equilibrium, fictitious play, Braess' paradox, correlated equilibrium, and regret matching.
2. Extensive-form games
   - game trees, information sets, mixed and behavioral strategies, perfect recall, and Kuhn's theorem;
   - subgames, backward induction, and subgame perfect Nash equilibrium;
   - Stackelberg games with ambiguous follower responses and inverse Stackelberg games;
   - level-k thinking in the centipede game, and signaling games with perfect Bayesian Nash equilibrium.
3. Applications in security
   - normal-form models: the cut-the-rope game of stealthy intrusion and the FlipIt game of stealthy takeover;
   - extensive-form models: Stackelberg security games, strong Stackelberg equilibrium, and the target-enumeration algorithm.

4. Discussion, summary, further reading, and exercises (Section 3.4)
   - sequential monitoring, signaling, bounded reasoning, algorithmic methods, and connections to later chapters;
   - practice problems on matrix games, equilibria, game trees, and the security applications.

Chapter Summary

Normal-form and extensive-form games provide two complementary ways to model strategic interaction. In a zero-sum matrix game, each player's security level bounds what it can guarantee. Pure security levels may leave a gap between the two players, but mixed strategies close it: by the minimax theorem every finite matrix game has a value, which a linear program computes. In general-sum games no single number summarizes the interaction. Nash equilibria exist in mixed strategies but may be inefficient, as Braess' paradox shows, and correlated equilibria, the long-run outcome of regret matching, enlarge the set of stable outcomes. The extensive form records who moves when and who knows what. Under perfect recall, behavioral strategies lose nothing relative to mixed strategies; subgame perfection removes threats that would not be carried out; a Stackelberg leader's guarantee depends on how the follower breaks ties; and in signaling games, beliefs updated by Bayes' rule determine whether messages reveal or pool private types. The security applications show that the choice of representation, and in particular the specification of who observes what, determines which defensive policies are credible: cut-the-rope and FlipIt turn stealthy intrusion into tradeoffs of allocation and timing, and a Stackelberg security game reduces the defender's commitment to one linear program per target.

Notation and Scope

Book conventions. Throughout the book, $\mathcal{N}$ is the set of players, $\mathcal{A}_i$ the action set of player i, $a = (a_i, a_{-i})$ an action profile, and $\Delta(X)$ the set of probability distributions on a set X; π_i is a mixed or behavioral strategy, u_i a utility that is maximized, and J_i a cost that is minimized. In dynamic models, $s \in \mathcal{S}$ is the state and $T(\cdot \mid s, a)$ the transition kernel; in models with private information, $\theta_i \in \Theta_i$ is a type and μ a belief. Single-letter labels of actions and messages are set upright (e.g. L, R), the attacker and the defender are indexed by A and D (as in u_A, J_D), equilibrium concepts are abbreviated upright (NE, SE), and $\triangleq$ marks a definition. The common notation is also listed in Appendix F.
Key symbols. In normal-form games, P_1 and P_2 are the row and column players, $A = (A_{ij})$ the payoff matrix, p and q their mixed strategies, $\overline{V}$, $\underline{V}$ and $V_m(A)$ the pure security levels and the mixed value, and $\varphi \in \Delta(\mathcal{A})$ a correlated equilibrium. Extensive-form games have a tree $\mathcal{G}$ with root x^0, information sets η_i, pure strategies $\gamma_i \in \Gamma_i$, behavioral strategies β_i, and level-k strategies π_i^k; a signaling game has prior ρ, messages $m \in \mathcal{M}$, posterior $\mu(\theta \mid m)$, and strategies σ_S, σ_R. In the applications, $G = (\mathcal{V}, \mathcal{E})$ is an attack graph with target τ, FlipIt has move rates α_i, move costs κ_i, payoff rates ν_i, and controller $C(t) \in \{D, A\}$, and c is a coverage vector.
Notation specific to this chapter: the players of matrix and Stackelberg games are P_1 and P_2, and in a zero-sum matrix game the row player P_1 minimizes the entry of A (Remark 3.1.2); mixed strategies are written p and q in matrix games and σ_S, σ_R in signaling games; signaling and Stackelberg security games write utilities with a capital U; and FlipIt writes payoff rates ν_i and uses T for a time horizon, not a transition kernel.

The chapter is restricted to finite normal-form and extensive-form games and to the security applications that they can express; it assumes basic probability and linear algebra, with the linear-programming duality of Appendix A. Elementary results are proved, while the minimax theorem, Nash's existence theorem, Kuhn's theorem, and the convergence of regret matching are stated with references. Evolving states, continuous time, and general incomplete information are deferred to Chapters 4–6.

Learning Objectives

After reading this chapter, the reader should be able to:

(1) compute pure and mixed security levels and saddle-point equilibria of matrix games, graphically or by linear programming;

(2) compute best responses and Nash equilibria of small general-sum games, and explain how fictitious play and regret matching relate to Nash and correlated equilibrium;

(3) represent a sequential interaction by a game tree with information sets, and convert between its extensive and strategic forms and between mixed and behavioral strategies;

(4) apply backward induction to find subgame perfect equilibria and to identify threats that are not credible;

(5) compute Stackelberg solutions under pessimistic tie-breaking, and explain how level-k reasoning departs from equilibrium play;

(6) derive posterior beliefs and perfect Bayesian Nash equilibria of simple signaling games;

(7) formulate the cut-the-rope, FlipIt, and Stackelberg security games, and compute their equilibrium rates, distributions, or commitments.

3.1 Normal-Form Games

Game theory is a mathematical framework for the study of strategic interaction among decision-makers [586, 450]. The normal form is the most compact representation of such an interaction. It suppresses the temporal order of play and records only the alternatives available to each player and the outcome that each player obtains once all choices have been made. The representation is appropriate when actions are chosen simultaneously, when the modeler deliberately abstracts from timing, or, more generally, when each player must select an action without observing the current choices of the other players.

A normal-form game is specified by three primitives: a finite set of *players*, denoted by $\mathcal{N}$; for each player $i \in \mathcal{N}$, a set of available actions, called the *action set* and denoted by $\mathcal{A}_i$; and, for each player $i \in \mathcal{N}$, a *payoff* (or *cost*) function that assigns a real number to every combination of actions. If each player $i \in \mathcal{N}$ chooses an action $a_i \in \mathcal{A}_i$, the resulting vector $a = (a_i)_{i \in \mathcal{N}}$ is called an *action profile*. For a given player i, we write $a_{-i} = (a_j)_{j \in \mathcal{N} \setminus \{i\}}$ for the actions of all players other than i, so that the profile may also be written as $a = (a_i, a_{-i})$. This notation is convenient in the definition of best responses, which describe the optimal choice of player i when the actions a_{-i} of the other players are held fixed.

The interpretation of the numerical value must be fixed in advance. In the two-player zero-sum matrix games studied below, the matrix entry is a cost to the row player P_1 and, equivalently, a payoff to the column player P_2; accordingly, P_1 minimizes and P_2 maximizes. In general-sum games, each player has its own payoff function, which it maximizes. In security applications the same mathematical object may be expressed either as a payoff or as a cost, and the direction of optimization is therefore always stated explicitly.

Normal-form analysis addresses three basic questions. The first concerns the outcome that each player can guarantee irrespective of the behavior of its opponents; this is the minimax, or security-level, question [585]. The second concerns the action profiles that are stable when every player optimizes against the choices of the others; this is the question of Nash equilibrium [418]. The third concerns the outcomes that become attainable when players may condition their actions on recommendations or on shared random signals; this leads to the notion of correlated equilibrium [29].

3.1.1 Two-Player Zero-Sum Games

3.1.1.1 Matrix Games

The most elementary class of games is that of two-player zero-sum games, in which the interests of the two players are diametrically opposed. If u_1 and u_2 denote the utilities of the two players, then $u_1(a_1, a_2) + u_2(a_1, a_2) = 0$ for every action profile (a_1, a_2). A gain for one player is a loss of equal size for the other, and the game is therefore a model of pure competition [586, 450].

Definition 3.1.1 (Two-player zero-sum normal-form game)**.** A *finite two-player zero-sum normal-form game* is a tuple $\langle \mathcal{N}, \{\mathcal{A}_i\}_{i \in \mathcal{N}}, \{u_i\}_{i \in \mathcal{N}} \rangle$ with the following components:

(1) $\mathcal{N} = \{1, 2\}$ is the set of players;

(2) for each $i \in \mathcal{N}$, $\mathcal{A}_i$ is the finite action set of player i;

(3) for each $i \in \mathcal{N}$, $u_i : \mathcal{A}_1 \times \mathcal{A}_2 \to \mathbb{R}$ is the payoff function of player i;

(4) the payoff functions satisfy the zero-sum condition $u_1(a_1, a_2) + u_2(a_1, a_2) = 0$ for all $(a_1, a_2) \in \mathcal{A}_1 \times \mathcal{A}_2$.

A finite two-player zero-sum game admits a representation by a single matrix. Let P_1 denote the row player and P_2 the column player, and suppose that P_1 has n_1 actions and P_2 has n_2 actions. The game is then represented by an $n_1 \times n_2$ matrix $A = (A_{ij})$, whose entry A_{ij} is the amount paid by P_1 to P_2 when P_1 chooses row i and P_2 chooses column j. If $A_{ij} > 0$, then P_1 pays A_{ij} to P_2; if $A_{ij} < 0$, then P_2 pays $|A_{ij}|$ to P_1. Under this convention, P_1 seeks to minimize the entry A_{ij} and P_2 seeks to maximize it.

Remark 3.1.2 (Cost versus reward conventions). This chapter follows the cost convention of [46] for two-player zero-sum games: the row player P_1 minimizes and the column player P_2 maximizes the matrix entry. Later chapters that build on the reward-based dynamic-programming tradition reverse these roles; the zero-sum Markov games of Chapter 5, for instance, take P_1 as the maximizer of a reward and P_2 as the minimizer. The two conventions are interchangeable: negating the payoff matrix exchanges the roles of minimizer and maximizer without altering the strategic content, so that every result stated in one convention translates directly into the other.

EXAMPLE 3.1 **Matching pennies as a zero-sum matrix**

Two players, P_1 and P_2, simultaneously place a penny on the table, each choosing whether it shows heads or tails. If the two pennies show the same side, P_1 wins one unit from P_2; if they show different sides, P_2 wins one unit from P_1. In accordance with the convention above, each entry of the following matrix is the payment from P_1 to P_2.

	P_2	
P_1	**Heads**	**Tails**
Heads	-1	1
Tails	1	-1

Rows correspond to the choices of P_1 and columns to those of P_2. The entry -1 in position (Heads, Heads) indicates that the payment from P_1 to P_2 is negative, that is, P_1 receives one unit from P_2. The entry 1 in position (Heads, Tails) indicates that P_1 pays one unit to P_2.

Observation. Matching pennies is a zero-sum game because every outcome transfers one unit from one player to the other. A single matrix suffices to describe the game because, under the payment convention, the payoff of P_2 is the negative of the payoff of P_1.

3.1.1.2 Pure Strategies

The behavior of a player is described by a *strategy*. In a one-shot normal-form game, a *pure strategy* of player i is a single action $a_i \in \mathcal{A}_i$, chosen with probability 1.

Consider the matrix game $A = (A_{ij})$. The row player P_1 seeks an upper bound on its loss that holds irrespective of the choice of P_2. If P_1 chooses row i, the largest loss it can incur is the largest entry of that row, which P_2 attains by choosing a column that maximizes the payment. Player P_1 therefore chooses a row i^* whose row maximum is as small as possible, so that

$$\overline{V} \triangleq \max_j A_{i^*,j} = \min_i \max_j A_{ij} \le \max_j A_{ij}, \qquad i = 1, 2, \dots, n_1. \tag{3.1.1}$$

The quantity $\overline{V}$ is called the *loss ceiling* of P_1, the *upper value* of the game, or the *security level* of P_1, and the row i^* is called a *security strategy* for P_1.

Symmetrically, the column player P_2 seeks a lower bound on its gain that holds irrespective of the choice of P_1. If P_2 chooses column j, the smallest gain it can obtain is the smallest entry of that column, which P_1 attains by choosing a row that minimizes the payment. Player P_2 therefore chooses a column j^* whose column minimum is as large as possible, so that

$$\underline{V} \triangleq \min_i A_{i,j^*} = \max_j \min_i A_{ij} \ge \min_i A_{ij}, \qquad j = 1, 2, \dots, n_2. \tag{3.1.2}$$

The quantity $\underline{V}$ is called the *gain floor* of P_2, the *lower value* of the game, or the *security level* of P_2, and the column j^* is called a *security strategy* for P_2. These lower and upper values are the pure-strategy counterparts of the minimax quantities that underlie the theory of zero-sum games [585, 586]. When the dependence on the matrix must be made explicit, we write $\overline{V}(A)$ and $\underline{V}(A)$.

THEOREM 3.1.3
In every matrix game $A = (A_{ij})$, the following statements hold:

(i) the security level of each player is unique;

(ii) each player has at least one security strategy;

(iii) the security level of P_2 (the maximizer) does not exceed the security level of P_1 (the minimizer), that is,

$$\min_i A_{i,j^*} = \underline{V} \le \overline{V} = \max_j A_{i^*,j},$$

where i^* and j^* denote security strategies for P_1 and P_2, respectively.

Proof. Parts (i) and (ii) hold because the game has finitely many rows and columns, so that the minima and maxima in (3.1.1) and (3.1.2) are well defined and attained. For (iii), observe that every row k and every column l satisfy

$$\min_i A_{il} \le A_{kl} \le \max_j A_{kj}.$$

Taking $k = i^*$ and $l = j^*$ yields

$$\underline{V} = \min_i A_{ij^*} \le A_{i^*j^*} \le \max_j A_{i^*j} = \overline{V}.$$

In particular, the outcome obtained when both players use security strategies lies between the lower value and the upper value. □

The upper and lower values admit an interpretation in terms of the order of moves. Suppose that P_1 must commit to a row before P_2 chooses a column. It is then optimal for P_1 to choose a security row i^*, and P_2 responds with a best-response column j^o satisfying

$$A_{i^*,j^o} = \max_j A_{i^*,j} = \overline{V} = \min_i \max_j A_{ij}. \tag{3.1.3}$$

This is the *minimax* problem, in which P_1 minimizes the maximum loss that P_2 can impose.

If the order is reversed, so that P_2 must commit to a column before P_1 chooses a row, it is optimal for P_2 to choose a security column j^*, and P_1 responds with a best-response row i^o satisfying

$$A_{i^o,j^*} = \min_i A_{i,j^*} = \underline{V} = \max_j \min_i A_{ij}. \tag{3.1.4}$$

This is the *maximin* problem, in which P_2 maximizes the minimum gain that P_1 can leave to it.

Definition 3.1.4 (Saddle-point equilibrium (SPE))**.** A pair (i^*, j^*) of a row and a column is a *saddle-point equilibrium* (SPE) of the matrix game $A = (A_{ij})$ if

$$A_{i^*,j} \le A_{i^*,j^*} \le A_{i,j^*}$$

for all $i = 1, 2, \ldots, n_1$ and all $j = 1, 2, \ldots, n_2$. The entry A_{i^*,j^*} is then called the *saddle-point value* of the game and is denoted by $V(A)$.

THEOREM 3.1.5
Let $A = (A_{ij})$ be an $n_1 \times n_2$ matrix game with $\underline{V}(A) = \overline{V}(A)$. Then the following statements hold:

(i) A has a saddle point in pure strategies;

(ii) an ordered pair (i, j) is a saddle point of A if and only if i is a security strategy for P_1 and j is a security strategy for P_2;

(iii) the saddle-point value is unique and is given by $V(A) = \underline{V}(A) = \overline{V}(A)$.

Proof. Let i^* be a security strategy for P_1 and j^* a security strategy for P_2; both exist by Theorem 3.1.3(ii). Since $\underline{V}(A) = \overline{V}(A)$,

$$A_{i^*,j} \le \max_{j'} A_{i^*,j'} = \overline{V} = \underline{V} = \min_{i'} A_{i',j^*} \le A_{i,j^*}$$

for all $i = 1, 2, \dots, n_1$ and $j = 1, 2, \dots, n_2$. Setting $i = i^*$ and $j = j^*$ shows that $A_{i^*,j^*} = \overline{V} = \underline{V}$, and consequently

$$A_{i^*,j} \le A_{i^*,j^*} \le A_{i,j^*}$$

for all i and j. This proves (i) and the sufficiency part of (ii). Conversely, let (i^*, j^*) be a saddle point. Then

$$\max_j A_{i^*,j} = A_{i^*,j^*} \le A_{i,j^*} \le \max_j A_{i,j}, \qquad i = 1, 2, \dots, n_1,$$

so that row i^* is a security strategy for P_1. An analogous argument shows that column j^* is a security strategy for P_2, which completes the proof of (ii). Part (iii) follows from (ii) and the equality $A_{i^*,j^*} = \overline{V} = \underline{V}$ established above. □

In a matrix game, a saddle point is therefore an entry that is simultaneously the maximum of its row and the minimum of its column. By choosing the saddle-point row i^*, P_1 ensures that its loss does not exceed $A_{i^*j^*}$; by choosing the saddle-point column j^*, P_2 ensures a gain of at least $A_{i^*j^*}$.

Saddle points in pure strategies need not exist. Indeed, if (i^*, j^*) is a saddle point of an arbitrary matrix game A, then the saddle-point inequalities give

$$\overline{V}(A) = \min_i \max_j A_{ij} \le \max_j A_{i^*,j} = A_{i^*,j^*} = \min_i A_{i,j^*} \le \max_j \min_i A_{ij} = \underline{V}(A),$$

so that $\overline{V}(A) \le \underline{V}(A)$. Together with Theorem 3.1.3(iii), this shows that the existence of a pure saddle point forces $\underline{V}(A) = \overline{V}(A)$, and part (i) of the preceding theorem gives the converse. Hence no pair of pure strategies satisfies the saddle-point inequalities when $\underline{V}(A) < \overline{V}(A)$. Matching pennies (Example 3.1), for which $\underline{V} = -1 < 1 = \overline{V}$, is the canonical instance.

3.1.1.3 Mixed Strategies

Definition 3.1.6 (Mixed strategy)**.** A *mixed strategy* of a player is a probability distribution over the pure strategies of that player.

We denote the mixed strategies of P_1 and P_2 by $p \in \Delta(\mathcal{A}_1)$ and $q \in \Delta(\mathcal{A}_2)$, respectively. When $|\mathcal{A}_1| = n_1$ and $|\mathcal{A}_2| = n_2$, the strategy sets are the simplices

$$\Delta(\mathcal{A}_1) = \left\{ p = (p_1, \dots, p_{n_1})^\top : p_i \ge 0,\ i = 1, \dots, n_1,\ \sum_{i=1}^{n_1} p_i = 1 \right\},$$

$$\Delta(\mathcal{A}_2) = \left\{ q = (q_1, \dots, q_{n_2})^\top : q_j \ge 0,\ j = 1, \dots, n_2,\ \sum_{j=1}^{n_2} q_j = 1 \right\}.$$

A pure strategy is the special case of a mixed strategy that assigns probability 1 to a single action and probability 0 to all others. When the players use the mixed strategies p and q independently, the expected payment from P_1 to P_2 is $J(p, q) = p^\top A q$. Player P_1 chooses p to minimize $J(p, q)$, and player P_2 chooses q to maximize it.

Definition 3.1.7 (Mixed security strategy)**.** A probability vector $p^* \in \Delta(\mathcal{A}_1)$ is a *mixed security strategy* for P_1 in the matrix game A if

$$\overline{V}_m(A) \triangleq \max_{q \in \Delta(\mathcal{A}_2)} p^{*\top} A q \le \max_{q \in \Delta(\mathcal{A}_2)} p^{\top} A q \qquad \text{for all } p \in \Delta(\mathcal{A}_1).$$

The quantity $\overline{V}_m(A)$ is the *mixed upper value* of the game. Analogously, a probability vector $q^* \in \Delta(\mathcal{A}_2)$ is a *mixed security strategy* for P_2 if

$$\underline{V}_m(A) \triangleq \min_{p \in \Delta(\mathcal{A}_1)} p^{\top} A q^* \ge \min_{p \in \Delta(\mathcal{A}_1)} p^{\top} A q \qquad \text{for all } q \in \Delta(\mathcal{A}_2),$$

and the quantity $\underline{V}_m(A)$ is the *mixed lower value* of the game.

Definition 3.1.8 (Mixed-strategy saddle-point equilibrium)**.** A pair $(p^*, q^*) \in \Delta(\mathcal{A}_1) \times \Delta(\mathcal{A}_2)$ is a *saddle point in mixed strategies* if

$$p^{*\top} A q \le p^{*\top} A q^* \le p^{\top} A q^* \qquad \text{for all } p \in \Delta(\mathcal{A}_1) \text{ and all } q \in \Delta(\mathcal{A}_2).$$

The number $V_m(A) = p^{*\top} A q^*$ is called the *saddle-point value*, or the *value*, of the game.

THEOREM 3.1.9 **Mixed security levels [46]**
In every matrix game A in which mixed strategies are admitted, the following statements hold:

(i) the mixed security level of each player is unique;

(ii) each player has at least one mixed security strategy;

(iii) the pure and mixed security levels satisfy

$$\underline{V}(A) \le \underline{V}_m(A) \le \overline{V}_m(A) \le \overline{V}(A).$$

The middle inequality in (iii) asserts that the value the maximizer can guarantee does not exceed the bound on loss that the minimizer can guarantee; it is proved by the argument used for Theorem 3.1.3(iii), with $J(p, q)$ in place of A_{ij}. The two outer inequalities hold because the pure strategies are contained in the mixed-strategy sets.

COROLLARY 3.1.10
In every matrix game A in which mixed strategies are admitted, the mixed upper and lower values are given by

$$\overline{V}_m(A) = \min_{p \in \Delta(\mathcal{A}_1)} \max_{q \in \Delta(\mathcal{A}_2)} p^{\top} A q \qquad \text{and} \qquad \underline{V}_m(A) = \max_{q \in \Delta(\mathcal{A}_2)} \min_{p \in \Delta(\mathcal{A}_1)} p^{\top} A q.$$

THEOREM 3.1.11 **Minimax theorem**
In every finite matrix game A, the mixed security levels of the two players coincide:

$$\overline{V}_m(A) = \min_{p \in \Delta(\mathcal{A}_1)} \max_{q \in \Delta(\mathcal{A}_2)} p^{\top} A q = \max_{q \in \Delta(\mathcal{A}_2)} \min_{p \in \Delta(\mathcal{A}_1)} p^{\top} A q = \underline{V}_m(A).$$

The minimax theorem, due to von Neumann [585], asserts that every finite zero-sum matrix game has a value in mixed strategies and that both players possess minimax strategies. Consequently, every such game has a saddle point in mixed strategies, and its value satisfies $V_m(A) = \overline{V}_m(A) = \underline{V}_m(A)$; the saddle-point strategies, however, need not be unique.

LEMMA 3.1.12 **Invariance of minimax strategies [46]**
Let $A = (A_{ij})$ and $A' = (A'_{ij})$ be $n_1 \times n_2$ matrices with $A'_{ij} = \zeta A_{ij} + \xi$ for all i and j, where $\zeta > 0$ and $\xi \in \mathbb{R}$ are

constants. Then the matrix games A and A' have the same minimax strategies for P_1 and for P_2, and their values satisfy $V_m(A') = \zeta V_m(A) + \xi$.

3.1.1.4 Computation of Saddle-Point Equilibria

Graphical methods. The mixed saddle point of a small matrix game can often be found by plotting the expected payment as a function of a single mixing probability. The minimizer selects the lowest point of the upper envelope of its possible losses, and the maximizer selects the highest point of the lower envelope of its possible gains.

EXAMPLE 3.2 **Graphical solution of a 2×2 matrix game**

▸ Code: `code/graphical_matrix_game.py`

Consider the 2×2 matrix game below, in which, as before, the row player P_1 minimizes the payment and the column player P_2 maximizes it.

	P_2	
P_1	**1**	**2**
1	3	0
2	−1	1

The row maxima are 3 and 1, so the pure upper value is $\overline{V} = \min\{3, 1\} = 1$, attained by row 2. The column minima are -1 and 0, so the pure lower value is $\underline{V} = \max\{-1, 0\} = 0$, attained by column 2. Since $\underline{V} < \overline{V}$, the game has no saddle point in pure strategies, and mixed strategies must be considered. Write the mixed strategy of P_1 over the rows $(1, 2)$ as $p = (p_1, p_2)^\top = (p_1, 1 - p_1)^\top$ and that of P_2 over the columns $(1, 2)$ as $q = (q_1, q_2)^\top = (q_1, 1 - q_1)^\top$, where $p_1, q_1 \in [0, 1]$.

Consider first the security problem of P_1. Against column 1, the expected payment is $J = 3p_1 - (1 - p_1) = 4p_1 - 1$; against column 2, it is $J = 1 - p_1$. These two lines are plotted in Figure 3.1. Their upper envelope, $\max\{4p_1 - 1,\ 1 - p_1\}$, is the largest expected loss that P_2 can impose for a given p_1. Because the first line is increasing and the second decreasing in p_1, the envelope attains its minimum at their intersection, where $4p_1 - 1 = 1 - p_1$, that is, at $p_1^* = 2/5$, with $p_2^* = 3/5$. By playing $p^* = (2/5, 3/5)^\top$, player P_1 therefore limits its worst expected loss to $1 - 2/5 = 3/5$.

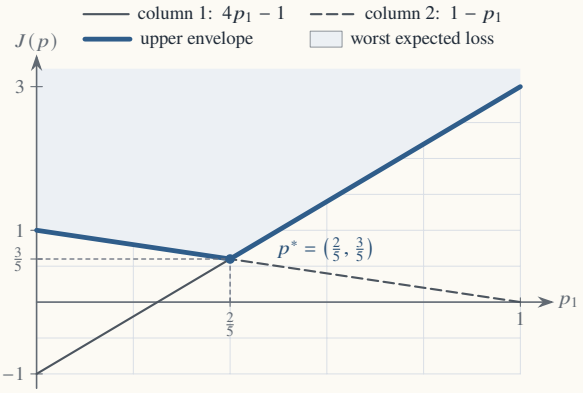


Figure 3.1: Graphical solution for P_1's mixed security strategy. The upper envelope of the two column payoffs is minimized at $p^* = (2/5, 3/5)^\top$.

Consider next the security problem of P_2. Against row 1, the expected payment is $J = 3q_1$; against row 2, it is $J = -q_1 + (1 - q_1) = 1 - 2q_1$. These two lines are plotted in Figure 3.2. Their lower envelope, $\min\{3q_1,\ 1 - 2q_1\}$, is the expected gain that P_2 can guarantee for a given q_1, and it attains its maximum at the intersection, where

$3q_1 = 1 - 2q_1$, that is, at $q_1^* = 1/5$, with $q_2^* = 4/5$. By playing $q^* = (1/5, 4/5)^\top$, player P_2 guarantees an expected gain of at least $3 \cdot (1/5) = 3/5$.
Since the two guaranteed levels coincide, the game has the mixed saddle-point equilibrium

$$(p^*, q^*) = \big((2/5,\ 3/5)^\top,\ (1/5,\ 4/5)^\top\big),$$

and its value is $V_m(A) = 3/5$.
This procedure is known as the *graphical solution* of a matrix game. It is practical for 2×2 games and extends to $2 \times n_2$ and $n_1 \times 2$ games, but it becomes cumbersome when both dimensions are large.

Observation. The minimizer equalizes the opponent's active columns, and the maximizer equalizes the opponent's active rows. The intersection of these binding payoff lines gives the mixed security strategies.

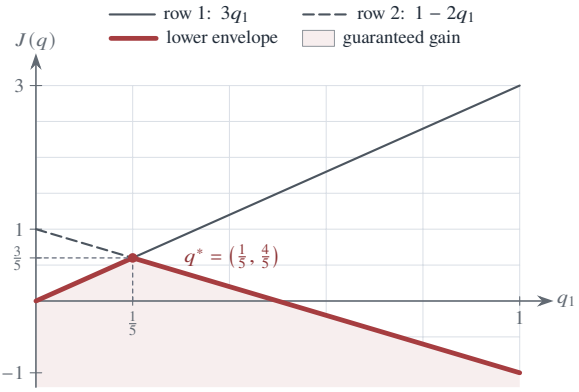


Figure 3.2: Graphical solution for P_2's mixed security strategy. The lower envelope of the two row payoffs is maximized at $q^* = (1/5, 4/5)^\top$.

Linear programming. For larger matrices, a more scalable approach is to formulate the matrix game as a linear program (LP). This formulation rests on the duality between the security problem of the minimizer and that of the maximizer [586, 46]. Assume first that all entries of $A = (A_{ij})$ are strictly positive. This entails no loss of generality: if some entries of the original matrix are not positive, the invariance lemma permits the addition of a sufficiently large constant ξ to every entry, which leaves the minimax strategies unchanged and increases the value by ξ. By the minimax theorem, the mixed value of the positive game is

$$V_m(A) = \min_{p \in \Delta(\mathcal{A}_1)} \max_{q \in \Delta(\mathcal{A}_2)} p^\top A q = \max_{q \in \Delta(\mathcal{A}_2)} \min_{p \in \Delta(\mathcal{A}_1)} p^\top A q.$$

For a fixed $p \in \Delta(\mathcal{A}_1)$, the worst expected loss of P_1 is $V_1(p) = \max_{q \in \Delta(\mathcal{A}_2)} p^\top A q$. Because q may place all its mass on any single column, $V_1(p) \geq (A^\top p)_j$ for every column j, and $V_1(p)$ is the smallest number with this property; equivalently, $V_1(p)$ is the smallest V satisfying $A^\top p \leq V \mathbf{1}_{n_2}$, where $\mathbf{1}_k$ denotes the k-dimensional vector of ones and the inequality holds componentwise. The security problem of P_1 is therefore the LP

$$\begin{aligned} \min_{p, V} \quad & V \\ \text{s.t.} \quad & A^\top p \leq V \mathbf{1}_{n_2}, \\ & \mathbf{1}_{n_1}^\top p = 1, \qquad p \geq 0. \end{aligned}$$

Since all entries of A are positive, every feasible point satisfies $V > 0$, and the change of variables $\widetilde{p} = p/V$ is well

defined. Minimizing V is equivalent to maximizing $1/V = \mathbf{1}_{n_1}^\top \widetilde{p}$, which yields the equivalent LP

$$\begin{aligned} \max_{\widetilde{p}} \quad & \mathbf{1}_{n_1}^\top \widetilde{p} \\ \text{s.t.} \quad & A^\top \widetilde{p} \le \mathbf{1}_{n_2}, \\ & \widetilde{p} \ge 0. \end{aligned}$$

If $\widetilde{p}^*$ solves this LP, then

$$p^* = \frac{\widetilde{p}^*}{\mathbf{1}_{n_1}^\top \widetilde{p}^*} \qquad \text{and} \qquad V_m(A) = \frac{1}{\mathbf{1}_{n_1}^\top \widetilde{p}^*}.$$

The dual LP yields the mixed security strategy of P_2. For a fixed $q \in \Delta(\mathcal{A}_2)$, the gain that P_2 can guarantee is $V_2(q) = \min_{p \in \Delta(\mathcal{A}_1)} p^\top A q$, which is the largest V satisfying $Aq \ge V\mathbf{1}_{n_1}$. Since $V_2(q) > 0$ for a positive matrix, the change of variables $\widetilde{q} = q/V_2(q)$ is well defined, and maximizing $V_2(q)$ is equivalent to the LP

$$\begin{aligned} \min_{\widetilde{q}} \quad & \mathbf{1}_{n_2}^\top \widetilde{q} \\ \text{s.t.} \quad & A\widetilde{q} \ge \mathbf{1}_{n_1}, \\ & \widetilde{q} \ge 0. \end{aligned}$$

If $\widetilde{q}^*$ solves this LP, then

$$q^* = \frac{\widetilde{q}^*}{\mathbf{1}_{n_2}^\top \widetilde{q}^*} \qquad \text{and} \qquad V_m(A) = \frac{1}{\mathbf{1}_{n_2}^\top \widetilde{q}^*}.$$

Every positive matrix game can therefore be solved by a pair of mutually dual LPs. This observation is the basis for the use of the simplex method, interior-point methods, and other LP algorithms in the computation of minimax strategies for high-dimensional security games.

3.1.2 N-Player General-Sum Games

3.1.2.1 Normal-Form Game

Minimax reasoning is natural in zero-sum games because the interests of the players are exactly opposed. Many security and economic interactions, however, are not purely zero sum: the players may compete in one dimension, share risk in another, and benefit from coordination in a third. Such interactions are modeled by *general-sum games*, in which each player has its own payoff function.

Definition 3.1.13 (N-player general-sum normal-form game)**.** A *finite N-player general-sum normal-form game* is a tuple $\langle \mathcal{N}, \{\mathcal{A}_i\}_{i\in\mathcal{N}}, \{u_i\}_{i\in\mathcal{N}} \rangle$ with the following components:

(1) $\mathcal{N} = \{1, 2, \dots, N\}$ is the finite set of players;

(2) for each $i \in \mathcal{N}$, $\mathcal{A}_i$ is the finite action set of player i, and $\mathcal{A} = \prod_{i\in\mathcal{N}} \mathcal{A}_i$ is the set of action profiles $a = (a_i, a_{-i})$, where a_{-i} denotes the actions of all players other than i;

(3) for each $i \in \mathcal{N}$, $u_i : \mathcal{A} \to \mathbb{R}$ is the payoff function of player i, which player i seeks to maximize.

A two-player general-sum game ($N = 2$) is commonly displayed as a bimatrix: P_1 chooses a row, P_2 chooses a column, and each entry lists the payoff pair (u_1, u_2). Matching pennies, for instance, takes the payoff-pair form

	P_2	
P_1	**Heads**	**Tails**
Heads	$(1, -1)$	$(-1, 1)$
Tails	$(-1, 1)$	$(1, -1)$

The first coordinate of each entry is the payoff of P_1 and the second is the payoff of P_2. The interaction is the same zero-sum game as in Example 3.1, now written in terms of the utility of each player rather than the single payment matrix A.

3.1.2.2 Solution Concept: Nash Equilibrium

Nash equilibrium (NE) is the central solution concept for general-sum games [418, 205]. It describes a strategy profile in which the strategy of every player is optimal given the strategies of the other players.

For each $i \in \mathcal{N}$, a mixed strategy of player i is a probability distribution $\pi_i \in \Delta(\mathcal{A}_i)$, and $\pi_i(a_i)$ denotes the probability that it assigns to the action $a_i \in \mathcal{A}_i$. Players randomize independently, so the expected payoff of player i under the mixed-strategy profile $\pi = (\pi_j)_{j\in\mathcal{N}}$ is

$$\bar{U}_i(\pi) = \sum_{a\in\mathcal{A}} u_i(a) \prod_{j\in\mathcal{N}} \pi_j(a_j).$$

As for action profiles, we write $\pi_{-i} = (\pi_j)_{j\in\mathcal{N}\setminus\{i\}}$ and $\pi = (\pi_i, \pi_{-i})$.

Definition 3.1.14 (Nash equilibrium (NE))**.** A mixed-strategy profile $\pi^* = (\pi_i^*)_{i\in\mathcal{N}}$ is a *Nash equilibrium* if

$$\bar{U}_i(\pi_i^*, \pi_{-i}^*) \geq \bar{U}_i(\pi_i, \pi_{-i}^*) \tag{3.1.5}$$

for every $i \in \mathcal{N}$ and every $\pi_i \in \Delta(\mathcal{A}_i)$.

Condition (3.1.5) states that no player can raise its expected payoff by a unilateral deviation: for every $i \in \mathcal{N}$, the strategy π_i^* is optimal for player i when the other players use π_{-i}^*. A Nash equilibrium is thus a self-consistent prediction, in which the strategy of each player is a best response to the strategies of the others. This characterization is made precise by the best-response correspondence.

Definition 3.1.15 (Best response (BR))**.** Let $i \in \mathcal{N}$ and $\pi_{-i} \in \prod_{j\neq i} \Delta(\mathcal{A}_j)$. The *best response* of player i to π_{-i} is the set

$$BR_i(\pi_{-i}) \triangleq \arg\max_{\pi_i\in\Delta(\mathcal{A}_i)} \bar{U}_i(\pi_i, \pi_{-i}).$$

Since $BR_i(\pi_{-i})$ may contain more than one strategy, the best response is in general a correspondence rather than a single-valued function. In this notation, π^* is a Nash equilibrium if and only if

$$\pi_i^* \in BR_i(\pi_{-i}^*) \qquad \text{for every } i \in \mathcal{N}. \tag{3.1.6}$$

THEOREM 3.1.16
Every finite normal-form game has a Nash equilibrium in mixed strategies [418].

3.1.2.3 Computation of Nash Equilibria

Best-response analysis.

EXAMPLE 3.3 **Battle of the Sexes and best responses**

▸ Code: `code/battle_of_the_sexes.py`

Two players wish to coordinate on a common action but disagree on which coordinated outcome is preferable. Each player chooses between the actions B and S. If both choose S, player P_1 receives 2 and P_2 receives 1; if both choose B, the payoffs are reversed. If their actions differ, both players receive 0. Both players maximize their payoffs.

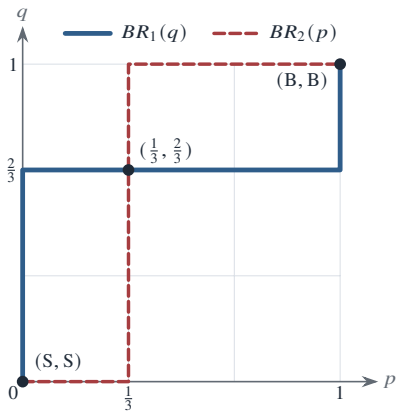


Figure 3.3: Best-response correspondences in the Battle of the Sexes, where p and q are the probabilities with which P_1 and P_2 choose B. The graph of BR_1 (solid) has a horizontal segment at $q = 2/3$, where P_1 is indifferent, and the graph of BR_2 (dashed) has a vertical segment at $p = 1/3$. The three intersections are the Nash equilibria (S, S), (B, B), and $(1/3, 2/3)$.

	P_2	
P_1	B	S
B	$(1, 2)$	$(0, 0)$
S	$(0, 0)$	$(2, 1)$

The pure action profiles (B, B) and (S, S) are Nash equilibria: at either profile, a unilateral deviation reduces the payoff of the deviating player from a positive value to 0. To determine the mixed equilibria, let $p = \mathbb{P}(P_1 \text{ chooses B})$ and $q = \mathbb{P}(P_2 \text{ chooses B})$, and identify each mixed strategy with the corresponding probability in $[0, 1]$. The expected payoff of P_1 is

$$\bar{U}_1(p, q) = pq + 2(1 - p)(1 - q) = p(3q - 2) + (2 - 2q),$$

which is increasing in p if $q > 2/3$, decreasing in p if $q < 2/3$, and constant in p if $q = 2/3$. The best response of P_1 is therefore

$$BR_1(q) = \begin{cases} 0, & q < \frac{2}{3}, \\ 1, & q > \frac{2}{3}, \\ [0, 1], & q = \frac{2}{3}. \end{cases}$$

The best response is set-valued at $q = 2/3$ because P_1 is then indifferent between B and S; for $q < 2/3$ it strictly prefers S, and for $q > 2/3$ it strictly prefers B.
Similarly, the expected payoff of P_2 is

$$\bar{U}_2(p, q) = 2pq + (1 - p)(1 - q) = q(3p - 1) + (1 - p),$$

so that

$$BR_2(p) = \begin{cases} 0, & p < \frac{1}{3}, \\ 1, & p > \frac{1}{3}, \\ [0, 1], & p = \frac{1}{3}. \end{cases}$$

A profile (p, q) is a Nash equilibrium if and only if $p \in BR_1(q)$ and $q \in BR_2(p)$. The two best-response graphs

intersect at exactly three points: the pure equilibria $(p, q) = (1, 1)$ and $(p, q) = (0, 0)$, and the mixed equilibrium $(p^*, q^*) = (1/3, 2/3)$, as shown in Figure 3.3.

Observation. A mixed Nash equilibrium makes each player indifferent among the actions used with positive probability. In this game, P_1 mixes only when P_2 chooses B with probability $2/3$, and P_2 mixes only when P_1 chooses B with probability $1/3$.

Learning: fictitious play. Fictitious play is an adaptive learning rule introduced by Brown [91]. In each period, every player records the empirical frequencies of the past actions of its opponent and plays a best response to the resulting empirical distribution.

Consider a two-player game played repeatedly at times $t = 1, 2, \dots$, and let $a_i^t \in \mathcal{A}_i$ denote the action of player i at time t. For each $i \in \{1, 2\}$, let $w_i^t : \mathcal{A}_{-i} \to \mathbb{R}_{\geq 0}$ denote the weighted count of the opponent's actions held by player i after t rounds. The initial weights w_i^0 are given, and w_i^t is obtained from w_i^{t-1} by adding 1 to the weight of the action a_{-i}^t observed at time t. Player i forms the empirical belief

$$\mu_i^t(a_{-i}) = \frac{w_i^t(a_{-i})}{\sum_{\hat{a}_{-i} \in \mathcal{A}_{-i}} w_i^t(\hat{a}_{-i})}, \qquad a_{-i} \in \mathcal{A}_{-i},$$

and at time $t + 1$ chooses a best response to this belief,

$$a_i^{t+1} \in \arg\max_{a_i \in \mathcal{A}_i} \mathbb{E}_{a_{-i} \sim \mu_i^t} \left[u_i(a_i, a_{-i}) \right].$$

EXAMPLE 3.4 **Two rounds of fictitious play**

▸ Code: `code/fictitious_play.py`

Consider the two-player game below, with action sets $\mathcal{A}_1 = \{\mathrm{U}, \mathrm{D}\}$ and $\mathcal{A}_2 = \{\mathrm{L}, \mathrm{R}\}$, in which each entry is the payoff pair (u_1, u_2).

	P_2	
P_1	L	R
U	$(3, 3)$	$(0, 0)$
D	$(4, 0)$	$(1, 1)$

Suppose that the initial weights are $w_1^0 = (3, 0)$ and $w_2^0 = (1, 2.5)$, where w_1^0 is defined over (L, R) and w_2^0 over (U, D). The initial beliefs are then $\mu_1^0 = (1, 0)$ and $\mu_2^0 = (1/3.5,\ 2.5/3.5)$.
In round 1, player P_1 expects payoff 3 from U and 4 from D, and therefore chooses D; indeed, D strictly dominates U, so P_1 chooses D under every belief. Player P_2 expects payoff $3/3.5$ from L and $2.5/3.5$ from R, and therefore chooses L. After the profile (D, L) has been observed, the weights become $w_1^1 = (4, 0)$ and $w_2^1 = (1, 3.5)$, and the beliefs become $\mu_1^1 = (1, 0)$ and $\mu_2^1 = (1/4.5,\ 3.5/4.5)$. In round 2, player P_1 again chooses D, while P_2 now expects payoff $3/4.5$ from L and $3.5/4.5$ from R, and therefore chooses R. Subsequent rounds update the beliefs in the same manner. Since P_2 prefers L if and only if $3\mu_2^t(\mathrm{U}) > 1 - \mu_2^t(\mathrm{U})$, that is, $\mu_2^t(\mathrm{U}) > 1/4$, and since $\mu_2^t(\mathrm{U}) = 1/(3.5 + t) < 1/4$ for every $t \geq 1$, the profile (D, R) is played in every round after the first; it is the unique Nash equilibrium of the game. Meanwhile $\mu_1^t(\mathrm{L}) = 4/(3 + t)$ for $t \geq 1$ decreases to zero, so the belief of P_1 converges to the observed play of P_2 (Figure 3.4).
In those classes of games for which fictitious play converges, the empirical distributions of play approach a Nash equilibrium.

Observation. Fictitious play separates learning from optimization. Each player first estimates the opponent's empirical distribution and then solves a one-period best-response problem against that estimate.

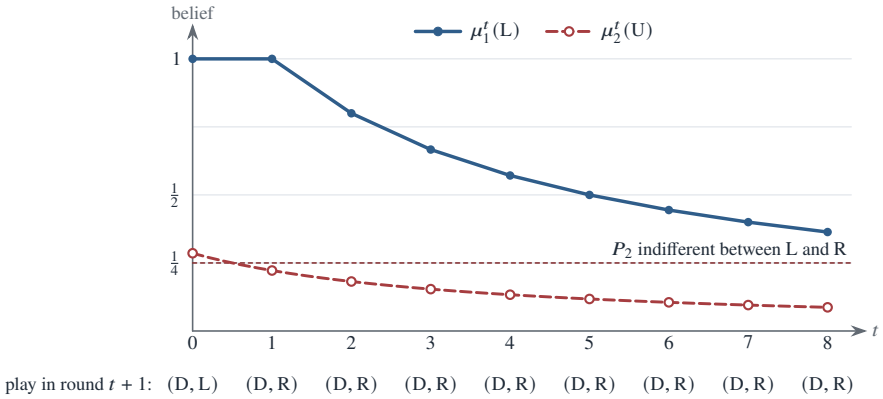


Figure 3.4: Fictitious play in Example 3.4. The belief $\mu_1^t(\mathrm{L})$ of P_1 that P_2 plays L (solid) and the belief $\mu_2^t(\mathrm{U})$ of P_2 that P_1 plays U (dashed) after t rounds. Once $\mu_2^t(\mathrm{U})$ falls below the indifference level $1/4$, player P_2 switches from L to R, and play settles at the Nash equilibrium (D, R); the bottom row lists the profile played in round $t+1$.

3.1.2.4 Solution Concept: Correlated Equilibrium

Before introducing correlated equilibrium, it is instructive to observe that Nash-style reasoning can produce stable but inefficient outcomes. In traffic networks, the analogue of Nash equilibrium is Wardrop equilibrium [593]: no infinitesimal traveler can reduce its travel time by switching routes. Braess' paradox shows that the outcome of such selfish routing can become worse when the network is expanded [86].

EXAMPLE 3.5 **Braess' paradox**

▸ Code: `code/braess_paradox.py`

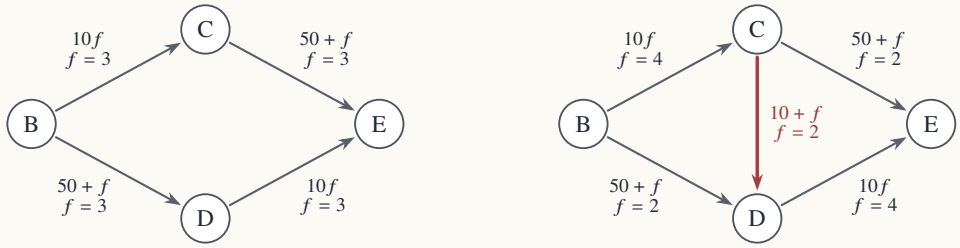


Figure 3.5: Braess network with six units of traffic from B to E. Edge labels give latency as a function of edge flow f; the second label on each edge gives its equilibrium flow. Adding the link C → D raises the equilibrium route time from 83 to 92.

Consider the network of one-way roads in Figure 3.5(a). Travelers want to go from B to E. There are two routes, BCE and BDE. The travel time on each edge depends on the amount of traffic on that edge.

Suppose that six units of traffic travel from B to E, and let f_1 and f_2 denote the flows on routes BCE and BDE, with $f_1 + f_2 = 6$. Each edge of route BCE carries flow f_1, so the route travel times are

$$\ell_{\mathrm{BCE}}(f) = 10 f_1 + (50 + f_1) = 50 + 11 f_1, \qquad \ell_{\mathrm{BDE}}(f) = (50 + f_2) + 10 f_2 = 50 + 11 f_2.$$

A route flow f is a *Wardrop equilibrium* if every route that carries positive flow has the minimum travel time among all routes. Both routes are used in equilibrium: if all six units took one route, its travel time would be 116, while the empty route would take 50. The equilibrium is therefore characterized by

$$50 + 11 f_1 = 50 + 11 f_2, \qquad f_1 + f_2 = 6,$$

whose unique solution is $f_1 = f_2 = 3$. Every traveler then spends $\ell_{\mathrm{BCE}} = \ell_{\mathrm{BDE}} = 83$, and the total travel time is $\ell_{\mathrm{tot}} = \sum_r f_r \ell_r = 6 \times 83 = 498$.
Now add the link from C to D, as in Figure 3.5(b). The routes are BCE, BDE, and BCDE, with flows f_1, f_2, and f_3 and $f_1 + f_2 + f_3 = 6$. The edge flows are $f_{\mathrm{BC}} = f_1 + f_3$, $f_{\mathrm{CE}} = f_1$, $f_{\mathrm{BD}} = f_2$, $f_{\mathrm{DE}} = f_2 + f_3$, and $f_{\mathrm{CD}} = f_3$, so

$$\begin{aligned}
\ell_{\mathrm{BCE}}(f) &= 10(f_1 + f_3) + (50 + f_1) = 50 + 11f_1 + 10f_3,\\
\ell_{\mathrm{BDE}}(f) &= (50 + f_2) + 10(f_2 + f_3) = 50 + 11f_2 + 10f_3,\\
\ell_{\mathrm{BCDE}}(f) &= 10(f_1 + f_3) + (10 + f_3) + 10(f_2 + f_3) = 10 + 10f_1 + 10f_2 + 21f_3.
\end{aligned}$$

If all three routes are used, their travel times must be equal. Equating ℓ_{BCE} and ℓ_{BDE} gives $f_1 = f_2$; equating ℓ_{BCE} and ℓ_{BCDE} then gives $40 = 9f_1 + 11f_3$, and together with $2f_1 + f_3 = 6$ this yields

$$f_1 = f_2 = f_3 = 2, \qquad \ell_{\mathrm{BCE}} = \ell_{\mathrm{BDE}} = \ell_{\mathrm{BCDE}} = 92, \qquad \ell_{\mathrm{tot}} = 6 \times 92 = 552.$$

Because every edge latency is strictly increasing in its flow, the equilibrium edge flows are unique, so this is the only Wardrop equilibrium of the augmented network.
Adding the extra road increases the equilibrium route time from 83 to 92 and the aggregate person-time from 498 to 552. The paradox is not that an additional road is physically harmful; rather, the new road changes incentives, and selfish route choice moves the system to a worse equilibrium.

Observation. Equilibrium is determined by incentives, not only by physical capacity. Adding an option can change best responses in a way that makes every selfish user worse off.

Correlated equilibrium addresses a different, though related, limitation of Nash equilibrium [29]. In a mixed-strategy Nash equilibrium the randomizations of the players are independent. A correlated equilibrium instead admits a trusted recommendation device, which draws an action profile from a joint distribution over $\mathcal{A}$ and privately recommends to each player the corresponding component of that profile. The recommendations are self-enforcing if no player, having observed its own recommended action, can increase its conditional expected payoff by deviating.

EXAMPLE 3.6 Game of chicken and correlated recommendations

▸ Code: `code/chicken_correlated_equilibrium.py`

The game of chicken admits the interpretation of two drivers approaching an intersection. Each player either goes (G) or yields (Y), and both players maximize their payoffs, which are given below.

	P_2	
P_1	G	Y
G	$(-10, -10)$	$(5, 0)$
Y	$(0, 5)$	$(-1, -1)$

The game has two pure-strategy Nash equilibria, (G, Y) and (Y, G), and in addition a mixed-strategy Nash equilibrium. Let p denote the probability with which P_1 chooses G and q the probability with which P_2 chooses G. Against q, the action G yields P_1 the expected payoff $-10q + 5(1 - q) = 5 - 15q$, and the action Y yields $-(1 - q) = q - 1$. These are equal if and only if $q = 3/8$, and G is strictly better if and only if $q < 3/8$. The same computation applies to P_2 by symmetry, so the best responses are

$$BR_1(q) = \begin{cases} [0,1], & q = \frac{3}{8},\\ 0, & q > \frac{3}{8},\\ 1, & q < \frac{3}{8}, \end{cases} \qquad BR_2(p) = \begin{cases} [0,1], & p = \frac{3}{8},\\ 0, & p > \frac{3}{8},\\ 1, & p < \frac{3}{8}. \end{cases}$$

The unique completely mixed equilibrium is therefore $p^* = q^* = 3/8$: each player chooses G with probability $3/8$ and Y with probability $5/8$. The induced distribution over action profiles is

	P_2	
P_1	G	Y
G	$9/64$	$15/64$
Y	$15/64$	$25/64$

Under this independent randomization the collision (G, G) occurs with probability $9/64$. By contrast, the following distribution is a fair recommendation scheme under which no collision occurs:

	P_2	
P_1	G	Y
G	0	$1/2$
Y	$1/2$	0

This distribution cannot be generated by independent mixed strategies, since it places positive probability on (G, Y) and on (Y, G) but zero probability on (G, G) and (Y, Y); under independence, the first property would force the last two probabilities to be positive.

The recommendation scheme is a correlated equilibrium: a player told to go expects the other to yield and obtains $5 > 0$, and a player told to yield expects the other to go and obtains $0 > -10$. Each player's expected payoff under the scheme is $\frac{1}{2} \cdot 5 + \frac{1}{2} \cdot 0 = 5/2$, whereas the mixed equilibrium yields each player $q^* - 1 = -5/8$. Figure 3.6 compares the three solutions. Each pure equilibrium avoids collisions but favors one player, and the mixed equilibrium treats the players symmetrically at the price of a collision probability $9/64$ and a negative expected payoff; the correlated equilibrium is both symmetric and collision-free.

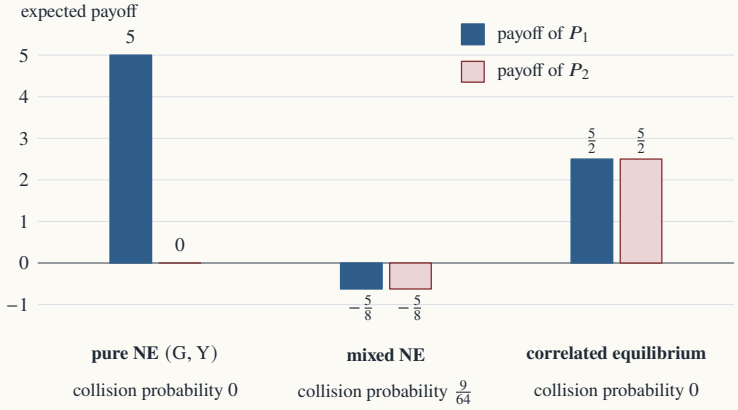


Figure 3.6: Expected payoffs in the game of chicken of Example 3.6 under three solutions: the pure Nash equilibrium (G, Y), the mixed Nash equilibrium in which each player goes with probability $3/8$, and the correlated equilibrium that recommends (G, Y) and (Y, G) with probability $1/2$ each. Dark bars give the payoff of P_1, light bars that of P_2, and the probability of the collision (G, G) is listed below each solution. The pure equilibrium (Y, G) is the mirror image of (G, Y).

Observation. Because Nash mixing randomizes independently, it leaves a positive collision probability. A correlation device can coordinate the players by telling exactly one player to go and the other to yield.

Definition 3.1.17 (Correlated equilibrium (CE))**.** Let $\varphi \in \Delta(\mathcal{A})$ be a probability distribution over the set $\mathcal{A} = \prod_{i \in \mathcal{N}} \mathcal{A}_i$ of action profiles, and write $\mathcal{A}_{-i} = \prod_{j \neq i} \mathcal{A}_j$. The distribution φ is a *correlated equilibrium* if, for every player $i \in \mathcal{N}$

and every pair of actions $a_i, a_i' \in \mathcal{A}_i$,

$$\sum_{a_{-i} \in \mathcal{A}_{-i}} \varphi(a_i, a_{-i}) \left[u_i(a_i, a_{-i}) - u_i(a_i', a_{-i}) \right] \geq 0. \tag{3.1.7}$$

For the game of chicken in Example 3.6, assign probabilities to action profiles as follows:

	P_2	
P_1	G	Y
G	φ_{11}	φ_{12}
Y	φ_{21}	φ_{22}

Consider P_1. If P_1 is told to go, obeying is optimal when deviating to Y does not raise the expected payoff, that is, when

$$\varphi_{11}\left[u_1(\mathrm{G},\mathrm{G}) - u_1(\mathrm{Y},\mathrm{G})\right] + \varphi_{12}\left[u_1(\mathrm{G},\mathrm{Y}) - u_1(\mathrm{Y},\mathrm{Y})\right] = -10\varphi_{11} + 6\varphi_{12} \geq 0.$$

If P_1 is told to yield, obeying is optimal when

$$\varphi_{21}\left[u_1(\mathrm{Y},\mathrm{G}) - u_1(\mathrm{G},\mathrm{G})\right] + \varphi_{22}\left[u_1(\mathrm{Y},\mathrm{Y}) - u_1(\mathrm{G},\mathrm{Y})\right] = 10\varphi_{21} - 6\varphi_{22} \geq 0.$$

The same calculation for P_2, conditioning on the column recommendation, gives $-10\varphi_{11}+6\varphi_{21} \geq 0$ and $10\varphi_{12}-6\varphi_{22} \geq 0$. Hence φ is a correlated equilibrium of the game of chicken if and only if

$$\begin{aligned} &-10\varphi_{11} + 6\varphi_{12} \geq 0, \qquad 10\varphi_{21} - 6\varphi_{22} \geq 0, \\ &-10\varphi_{11} + 6\varphi_{21} \geq 0, \qquad 10\varphi_{12} - 6\varphi_{22} \geq 0, \end{aligned} \qquad \sum_{i,j} \varphi_{ij} = 1, \qquad \varphi_{ij} \geq 0.$$

The fair no-collision distribution $\varphi_{12} = \varphi_{21} = 1/2$ and $\varphi_{11} = \varphi_{22} = 0$ satisfies these constraints and is therefore a correlated equilibrium: a player told to go knows that the other player has been told to yield, and a player told to yield knows that the other player has been told to go.

Every Nash equilibrium induces a correlated equilibrium, so NE $\subseteq$ CE; the inclusion can be strict, because not every correlated equilibrium can be generated by independent mixed strategies. The CE constraints are linear in φ, so finding a correlated equilibrium is an LP feasibility problem [29, 450].

3.1.2.5 Correlated Equilibrium Computation

No-regret learning. Correlated equilibria also arise as the long-run outcome of no-regret learning [242]; learning in repeated and Markov games, in which no-regret dynamics play a central role, is developed in Chapter 7. Consider the normal-form game $\langle \mathcal{N}, \{\mathcal{A}_i\}_{i\in\mathcal{N}}, \{u_i\}_{i\in\mathcal{N}} \rangle$ played repeatedly at times $t = 1, 2, \ldots$. Before time $t+1$, all players observe the history of play $h_t = (a^{t'})_{t'=1}^{t}$, where $a^{t'} = (a_1^{t'}, \ldots, a_N^{t'})$ is the action profile realized at time t'. Each player $i \in \mathcal{N}$ then chooses an action $a_i^{t+1} \in \mathcal{A}_i$ according to a probability distribution that is updated on the basis of its past regret, defined as follows.

Fix a player $i \in \mathcal{N}$ and two actions $j, k \in \mathcal{A}_i$. The *counterfactual payoff* of player i at time t' is the payoff it would have obtained had it replaced action j by action k, the actions of the other players being unchanged:

$$W_i^{t'}(j,k) \triangleq \begin{cases} u_i(k, a_{-i}^{t'}) & \text{if } a_i^{t'} = j, \\ u_i(a^{t'}) & \text{otherwise.} \end{cases}$$

The *average regret* of player i for not having followed the rule "play k whenever j is played" during the first t periods is the average gain from this replacement,

$$D_i^t(j,k) \triangleq \frac{1}{t}\sum_{t'=1}^{t} W_i^{t'}(j,k) - \frac{1}{t}\sum_{t'=1}^{t} u_i(a^{t'}).$$

A positive value of $D_i^t(j,k)$ indicates that the replacement rule would have increased the average payoff of player i, and a nonpositive value indicates that it would not. Only the positive part enters the update, and the *regret* is accordingly defined as $R_i^t(j,k) \triangleq \max\{D_i^t(j,k), 0\}$. Player i shifts probability toward alternative actions in proportion to these regrets: if $j = a_i^t$ is the action chosen at time t, the probabilities of play at time $t+1$ are

$$\pi_i^{t+1}(k) = \frac{1}{Z_i} R_i^t(j,k) \quad \text{for all } k \in \mathcal{A}_i \setminus \{j\}, \qquad \pi_i^{t+1}(j) = 1 - \sum_{k \in \mathcal{A}_i \setminus \{j\}} \pi_i^{t+1}(k),$$

where the normalizing constant $Z_i > 0$ is chosen large enough that $\pi_i^{t+1}(j) \geq 0$, so that π_i^{t+1} is a probability distribution. This adaptive procedure is known as *regret matching*.

The connection with correlated equilibrium is made through the empirical distribution $z_t \in \Delta(\mathcal{A})$ of the action profiles played up to time t, defined by $z_t(a) \triangleq |\{t' \leq t : a^{t'} = a\}|/t$ for every $a \in \mathcal{A}$, where $|Q|$ denotes the cardinality of a finite set Q. Thus $z_t(a)$ is the relative frequency with which the profile a has been played in the first t periods.

THEOREM 3.1.18
If every player follows a no-internal-regret adaptive procedure such as regret matching, then the empirical distributions of play z_t converge almost surely as $t \to \infty$ to the set of correlated equilibrium distributions of the game [242].

In this learning process, correlation originates in the shared history of play. Although the players randomize independently at each stage, their updates depend on the same sequence of past action profiles, and the empirical distribution of joint actions may therefore approach a correlated equilibrium that is not a product distribution.

3.2 Extensive-Form Games

Normal-form games suppress the timing of decisions. Extensive-form games restore the dynamic structure of strategic interaction: the order in which players move, the information available to each player when moving, the points at which chance intervenes, and the payoffs attached to each complete course of play [337]. The central object is a game tree. A path from the root to a node is a *history* of play, and a path from the root to a terminal node, at which a payoff vector is assigned, is a *terminal history*. At each nonterminal node, either a player or chance selects the next branch.

This additional structure changes the meaning of a strategy. In a normal-form game, a pure strategy coincides with an action. In an extensive-form game, a pure strategy is a complete contingent plan: it prescribes an action at every information set of the player, including information sets that are not reached in equilibrium [337, 450]. The distinction is essential in security applications, where many decisions are conditional. A defender may first select a monitoring policy, then observe an alarm, and then decide whether to isolate a subsystem; an attacker may observe the defensive posture before deciding whether to probe, escalate, or withdraw.

The section develops, in turn, the formal description of extensive-form games and information sets, mixed and behavioral strategies, perfect recall, subgame perfection, Stackelberg games, level-k thinking, and signaling games.

3.2.1 Games with Imperfect Information

Matching pennies provides the simplest illustration of imperfect information. Each player P_i, $i = 1, 2$, chooses an action from $\{\mathrm{H}, \mathrm{T}\}$. In Table 3.1, each entry is the payoff pair (u_1, u_2) of P_1 and P_2, respectively; P_1 receives 1 if the two actions match and -1 otherwise. The game is zero-sum, since $u_1 + u_2 = 0$ at every outcome.

	P_2	
P_1	H	T
H	$(1,-1)$	$(-1,1)$
T	$(-1,1)$	$(1,-1)$

Table 3.1: Matching pennies in normal form.

The same strategic situation admits an extensive-form representation. Figures 3.7 and 3.8 show two equivalent game trees, in which P_1 and P_2, respectively, move first. In each tree the second mover does not observe the first mover's action, and the second mover's two decision nodes are therefore joined into a single information set.

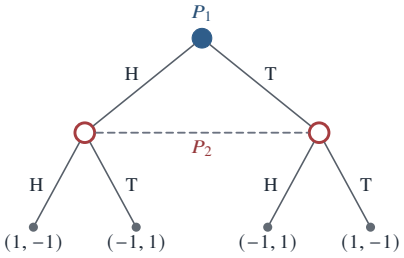


Figure 3.7: P_1 moves first; P_2 does not observe P_1's action.

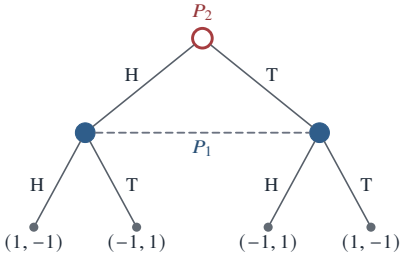


Figure 3.8: P_2 moves first; P_1 does not observe P_2's action.

Although both trees order the moves in time, each is strategically equivalent to the simultaneous-move game of Table 3.1, because the second mover cannot condition the choice of action on the first mover's action. The information set records precisely this restriction. A normal-form matrix specifies the feasible action profiles and their payoffs; an extensive-form tree specifies, in addition, the order of moves and the information available at each decision node. The following definition formalizes the tree structure.

Definition 3.2.1 (Extensive-form game tree with chance moves)**.** An extensive-form game tree with chance moves is a tuple $\mathcal{G} = (\mathcal{N}, \mathcal{V}, \mathcal{E}, x^0, (\mathcal{V}_i)_{i\in\mathcal{N}\cup\{0\}}, (\pi_0(\cdot \mid x))_{x\in\mathcal{V}_0}, O, u)$, where

(1) $\mathcal{N}$ is the finite set of players;

(2) $(\mathcal{V}, \mathcal{E}, x^0)$ is a finite rooted tree with vertex set $\mathcal{V}$, edge set $\mathcal{E}$, and root (initial node) x^0;

(3) $O \subseteq \mathcal{V}$ is the set of terminal nodes (outcomes);

(4) $(\mathcal{V}_i)_{i\in\mathcal{N}\cup\{0\}}$ is a partition of the set $\mathcal{V} \setminus O$ of nonterminal nodes, where $\mathcal{V}_i$, $i \in \mathcal{N}$, is the set of decision nodes of player i and $\mathcal{V}_0$ is the set of chance nodes;

(5) for every chance node $x \in \mathcal{V}_0$, $\pi_0(\cdot \mid x)$ is a probability distribution over the edges leaving x;

(6) $u : O \to \mathbb{R}^{|\mathcal{N}|}$ assigns to each terminal node a payoff vector, with one component for each player.

For every nonterminal node x, the set of edges leaving x is denoted by $\mathcal{A}(x)$; when $x \in \mathcal{V}_i$ for some $i \in \mathcal{N}$, the elements of $\mathcal{A}(x)$ are the actions available to player i at x.

3.2.1.1 Information Sets

Definition 3.2.2 (Information set)**.** Let $\mathcal{G}$ be an extensive-form game and let $i \in \mathcal{N}$. An *information set* of player i is a pair $(\eta_i, \mathcal{A}(\eta_i))$ with the following two properties.

(1) The set $\eta_i = \left\{x_i^1, x_i^2, \ldots, x_i^{|\eta_i|}\right\}$ is a nonempty subset of $\mathcal{V}_i$ at whose nodes player i has the same number of available actions; that is, there is a positive integer l_i, which may depend on η_i, such that $|\mathcal{A}(x_i^j)| = l_i$ for $j = 1, 2, \ldots, |\eta_i|$.

(2) The set $\mathcal{A}(\eta_i)$, of cardinality l_i, is the common action set at the nodes of η_i; each action in $\mathcal{A}(\eta_i)$ identifies exactly one edge in $\mathcal{A}(x_i^j)$ for every $j = 1, 2, \ldots, |\eta_i|$.

An information set thus collects decision nodes of a single player that this player cannot distinguish from one another when play reaches the set. The information sets of player i form a partition of $\mathcal{V}_i$, denoted by I_i. In Figure 3.7, the unique information set of P_2 contains the two decision nodes that follow P_1's actions H and T. We write $\eta_2 = \{\mathrm{H}, \mathrm{T}\}$ as shorthand for these two decision nodes; the notation does not refer to terminal outcomes.

3.2.1.2 Behavioral Strategies

Consider now the modification of Figure 3.7 in which P_2 observes P_1's action before moving. In the resulting tree, shown in Figure 3.9, each of the two decision nodes of P_2 constitutes a singleton information set.

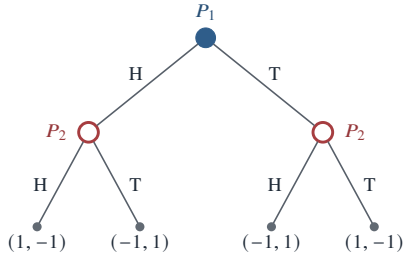


Figure 3.9: P_2 observes P_1's action: each of P_2's decision nodes is a singleton information set.

Since P_2 can now distinguish between the two actions of P_1, a pure strategy of P_2 must specify an action after $a_1 = \mathrm{H}$ and an action after $a_1 = \mathrm{T}$, where a_1 denotes the action of P_1. Player P_2 therefore has the four pure strategies

$$\gamma_2^1 = \begin{cases} \mathrm{H} & \text{if } a_1 = \mathrm{H} \\ \mathrm{H} & \text{if } a_1 = \mathrm{T} \end{cases} \qquad \gamma_2^2 = \begin{cases} \mathrm{T} & \text{if } a_1 = \mathrm{H} \\ \mathrm{T} & \text{if } a_1 = \mathrm{T} \end{cases} \qquad \gamma_2^3 = \begin{cases} \mathrm{H} & \text{if } a_1 = \mathrm{H} \\ \mathrm{T} & \text{if } a_1 = \mathrm{T} \end{cases} \qquad \gamma_2^4 = \begin{cases} \mathrm{T} & \text{if } a_1 = \mathrm{H} \\ \mathrm{H} & \text{if } a_1 = \mathrm{T} \end{cases}.$$

The strategic-form representation accordingly expands to Table 3.2, in which each column of P_2 is labeled by the pair of actions taken after H and after T, in that order. For example, the label HT denotes the strategy γ_2^3, which plays H after H and T after T. With this additional information, the strategy $\mathrm{TH} = \gamma_2^4$, which plays T after H and H after T, is

	P_2			
P_1	HH	HT	TH	TT
H	$(1,-1)$	$(1,-1)$	$(-1,1)$	$(-1,1)$
T	$(-1,1)$	$(1,-1)$	$(-1,1)$	$(1,-1)$

Table 3.2: Matching pennies when P_2 observes P_1.

weakly dominant for P_2: it yields P_2 the payoff 1 against both actions of P_1, which is the largest payoff available to P_2 in the game. Changing what a player observes therefore changes the strategy space and, with it, the predicted outcome of the game.

In an extensive-form game, a pure strategy of player i is a complete contingent plan, namely a function that assigns to each information set $\eta_i \in I_i$ an action in $\mathcal{A}(\eta_i)$. A mixed strategy randomizes over these complete contingent plans.

Definition 3.2.3 (Mixed strategy in an extensive-form game)**.** Let Γ_i be the finite set of pure strategies of player i in an extensive-form game. A *mixed strategy* of player i is a probability distribution over Γ_i, that is, an element of the set $\Delta(\Gamma_i) \triangleq \{\pi_i : \Gamma_i \to [0,1] : \sum_{\gamma_i \in \Gamma_i} \pi_i(\gamma_i) = 1\}$.

A behavioral strategy, by contrast, randomizes locally at each information set rather than globally over complete plans [337].

Definition 3.2.4 (Behavioral strategy)**.** Let I_i be the collection of information sets of player i. A *behavioral strategy* of player i is a function β_i on I_i such that $\beta_i(\eta_i) \in \Delta(\mathcal{A}(\eta_i))$ for every $\eta_i \in I_i$; that is, β_i assigns to each information set of player i a probability distribution over the actions available there.

A mixed strategy thus randomizes once, before play begins, over complete deterministic plans, whereas a behavioral strategy randomizes separately at each information set as play unfolds. In games with perfect recall, the two forms of randomization are outcome-equivalent (Theorem 3.2.6); the distinction nevertheless matters for modeling and for computation. Chapter 5 carries these ideas into stochastic games in which information evolves with a state, including

partially observable games, and Chapter 12 uses types, beliefs, and behavioral strategies in multi-stage games of layered APT defense.

EXAMPLE 3.7 **Market-entry game**

Firm P_2 (the entrant) decides whether to enter a market currently served by firm P_1 (the incumbent). If P_2 enters, P_1 decides whether to fight entry or to accommodate it. If P_2 stays out, P_1 does not move. The payoffs, which both firms maximize, are shown in Figure 3.10.

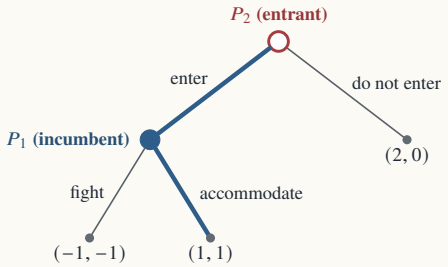


Figure 3.10: Market-entry game. Payoffs are (P_1, P_2); thick edges mark the subgame perfect choices.

The entrant's strategies are E (enter) and N (do not enter); the incumbent's strategies are F (fight) and A (accommodate), each understood as the action taken if entry occurs. Table 3.3 gives the strategic-form representation.

	P_2	
P_1	E	N
F	$(-1,-1)$	$(2,0)^\star$
A	$(1,1)^\star$	$(2,0)$

Table 3.3: Strategic form of the market-entry game. A star ($\star$) marks a Nash equilibrium of the strategic-form game.

The strategic-form game has exactly two pure-strategy Nash equilibria, (A,E) and (F,N). At (A,E), accommodation is the incumbent's unique best response to entry, since $1 > -1$, and entry is the entrant's unique best response to accommodation, since $1 > 0$. At (F,N), the incumbent is indifferent between F and A because entry does not occur, and staying out is the entrant's best response to fighting, since $0 > -1$. The equilibrium (F,N), however, rests on a noncredible threat: if P_2 entered, P_1 would obtain 1 by accommodating and -1 by fighting, and would therefore accommodate. The extensive form exposes this off-path incentive and selects the subgame perfect outcome (A,E) [524, 205].

Observation. The strategic-form table lists all complete contingent plans, but the tree reveals which threats are credible after a history is reached. This is exactly why subgame perfection refines Nash equilibrium.

3.2.2 Perfect Recall and Kuhn's Theorem

The distinction between mixed and behavioral strategies is not merely notational. A mixed strategy selects one complete contingent plan before play begins, whereas a behavioral strategy specifies a probability distribution locally at each information set. In many extensive-form games the two descriptions generate the same set of probability distributions over terminal histories, but this equivalence depends on what the players remember.

Definition 3.2.5 (Perfect recall)**.** Player $i \in \mathcal{N}$ has *perfect recall* if, for any two decision nodes $x, x' \in \mathcal{V}_i$ that lie in the same information set, player i has the same personal history at x and at x'; that is, along the path from the root x^0 to x and along the path from x^0 to x', player i passes through the same sequence of own information sets and takes the same sequence of own actions at them [337]. An extensive-form game has perfect recall if every player has perfect recall.

Under perfect recall, a player may be uncertain about the actions of the other players or about the outcomes of chance moves, but never forgets what the player previously knew or chose. The assumption is natural in most security

models: a defender may not observe the attacker's type, yet remembers whether an alert was investigated, whether a patch was deployed, and which access-control rule was selected.

THEOREM 3.2.6 **Kuhn's theorem**

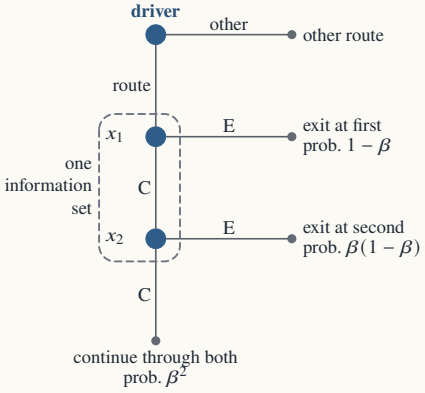
Let $\mathcal{G}$ be a finite extensive-form game with perfect recall. Then mixed strategies and behavioral strategies are outcome-equivalent in $\mathcal{G}$: for every player i and every mixed strategy $\pi_i \in \Delta(\Gamma_i)$, there exists a behavioral strategy β_i that induces the same probability distribution over terminal histories as π_i against every strategy profile of the other players; conversely, every behavioral strategy of player i has an outcome-equivalent mixed strategy.

Kuhn's theorem [337] explains why behavioral strategies are the standard representation of play in dynamic games with perfect recall. A behavioral strategy is read directly from the tree, since each probability is attached to a specific information set. It is also more economical computationally: the modeler optimizes over local probabilities at the information sets rather than over distributions on the set of all complete plans.

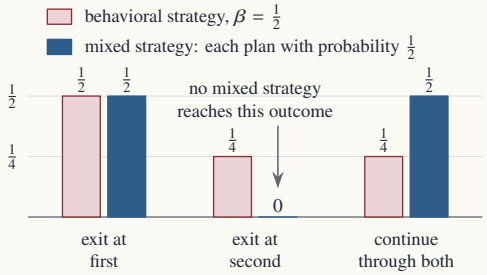
EXAMPLE 3.8 **Why recall matters**

Consider a one-player decision problem in which a driver encounters two identical intersections. At each intersection the driver either continues, denoted by C, or exits, denoted by E. The driver cannot remember whether the current intersection is the first or the second, so the two intersection nodes belong to a single information set, and a behavioral strategy must assign the same probability $\beta \in [0, 1]$ of continuing at both nodes. The game tree is shown in Figure 3.11(a).

The driver first decides whether to take the route that leads to the intersections. Conditional on taking this route, the behavioral strategy with continuation probability β leads the driver to exit at the first intersection with probability $1 - \beta$, to exit at the second intersection with probability $\beta(1 - \beta)$, and to continue through both intersections with probability β^2. These probabilities arise because the same local randomization is applied each time the information set is reached.



(a) Tree with imperfect recall



(b) Outcome probabilities

Figure 3.11: The decision problem of Example 3.8. (a) The two intersections x_1 and x_2 form a single information set (dashed), so a behavioral strategy continues with the same probability β at both; each terminal node is labeled with the probability of reaching it, conditional on the route. (b) Probabilities of the three outcomes that follow the route under the behavioral strategy with $\beta = 1/2$ and under the mixed strategy that selects each of the two complete plans, "always exit" and "always continue," with probability $1/2$. Both strategies exit at the first intersection with probability $1/2$, but only the behavioral strategy exits at the second; every mixed strategy assigns probability 0 to that outcome.

A mixed strategy, by contrast, selects a complete plan before play begins. Since a plan prescribes a single action at the information set, the only plans available conditional on taking the route are "exit whenever an intersection is reached" and "continue whenever an intersection is reached." Every mixture of these two plans therefore exits at the first intersection or continues through both, and assigns probability zero to exiting at the second

intersection. The behavioral strategy reaches that outcome with probability $\beta(1-\beta)$, which is positive for every $\beta \in (0,1)$; Figure 3.11(b) compares the two descriptions for $\beta = 1/2$. In this problem with imperfect recall, the outcome distributions generated by mixtures of complete plans thus differ from those generated by behavioral randomization, which shows that the perfect-recall hypothesis of Theorem 3.2.6 cannot be dropped.

Observation. Behavioral strategies are locally randomized, while mixed strategies randomize over complete plans. Perfect recall is the condition that makes these two descriptions outcome-equivalent.

In most games of this chapter, the players have perfect recall, and behavioral strategies can therefore be used without loss of the strategic possibilities represented by mixed strategies. When a model deliberately includes bounded memory, deception, or cognitive limitations, the equivalence must be verified rather than assumed.

3.2.3 Subgame Perfect Nash Equilibrium

Nash equilibrium requires that no player can increase its payoff by a unilateral deviation, evaluated at the beginning of the game and given the complete strategies of the other players. In dynamic games this requirement is too weak: a Nash equilibrium may be sustained by a threat that would not be optimal to carry out if the history at which it applies were actually reached. Subgame perfect Nash equilibrium (SPNE), introduced by Selten [524], strengthens Nash equilibrium by requiring optimality after every history that begins a subgame.

Definition 3.2.7 (Subgame)**.** Let x be a decision node of an extensive-form game $\mathcal{G}$ such that $\{x\}$ is an information set. The *subgame* $\mathcal{G}_x$ of $\mathcal{G}$ rooted at x consists of x, all successors of x, and the restriction to these nodes of the players, chance probabilities, information sets, and payoffs of $\mathcal{G}$, provided that every information set of $\mathcal{G}$ that contains a successor of x contains only successors of x. The game $\mathcal{G} = \mathcal{G}_{x^0}$ is itself a subgame.

The requirement that no information set be cut by the subtree is essential. If a player cannot distinguish two nodes, the analysis cannot ask what would happen after one of them while disregarding the other; a subgame must be a self-contained game in which the information of every player is well defined.

Definition 3.2.8 (Subgame perfect Nash equilibrium)**.** A strategy profile $\gamma^* = (\gamma_i^*)_{i \in N}$ is a *subgame perfect Nash equilibrium* if, for every subgame $\mathcal{G}_x$ of $\mathcal{G}$, the restriction of γ^* to $\mathcal{G}_x$ is a Nash equilibrium of $\mathcal{G}_x$. Equivalently, after every history that begins a subgame, the continuation strategy of each player is a best response to the continuation strategies of the other players.

In finite games with perfect information, subgame perfect equilibria can be computed by backward induction. The decision problems at the nodes whose successors are all terminal are solved first; each such node is then replaced by the payoff vector that its solution induces, and the procedure is repeated on the resulting smaller tree until the root is reached. The strategies so constructed are optimal in every subgame [524, 450].

THEOREM 3.2.9 **Backward induction**
Every finite extensive-form game with perfect information has a pure-strategy subgame perfect Nash equilibrium. One such equilibrium can be constructed by backward induction.

Proof sketch. At each decision node whose successors are all terminal, the acting player selects an action that maximizes its payoff; such an action exists because the set of actions is finite. For a chance node whose successors are all terminal, the expected payoff vector is computed with respect to the chance probabilities. Each such node thereby receives a continuation payoff vector, and treating these vectors as terminal payoffs yields a game with a strictly shorter tree. Since the tree is finite, repeating the argument selects an optimal action at every decision node after finitely many steps. The same construction applies within every subtree rooted at a decision node, so the restriction of the resulting pure-strategy profile to every subgame is a Nash equilibrium of that subgame. □

The market-entry game of Figure 3.10 illustrates the refinement. Its strategic form has the two Nash equilibria (A, E) and (F, N). The profile (F, N) prescribes that the incumbent fight entry and, in response, that the entrant stay out. If entry occurred, however, the incumbent would obtain 1 by accommodating and -1 by fighting, so the restriction of

(F, N) to the subgame that follows entry is not a Nash equilibrium of that subgame. The threat to fight is therefore not credible, and (F, N) is not subgame perfect. Backward induction selects accommodation after entry and, given accommodation, entry by the entrant; the unique SPNE is (A, E).

Subgame perfection is relevant to security games because many defensive decisions are sequential. A defender may first choose an investment level, then observe a signal, and then allocate response resources; an attacker may probe the system before launching an attack. SPNE requires every continuation plan to remain optimal once the corresponding stage is reached, and thereby excludes strategy profiles whose stability rests on actions that would not be taken when the time came.

3.2.4 Stackelberg Games

Nash equilibrium treats the players symmetrically: each player's strategy is a best response to the strategies of the others. Many decision problems are instead hierarchical. One player, the *leader*, commits to a strategy first, and the other player, the *follower*, observes this commitment before choosing a response. Stackelberg equilibrium models this leader–follower structure [587, 46].

3.2.4.1 Stackelberg Games with Ambiguity

The market-entry game of Example 3.7 is a simple Stackelberg game in which the entrant, moving first, is the leader, and the incumbent's credible response determines the outcome. A further issue arises when the follower has more than one optimal response to a strategy of the leader. The leader's evaluation of that strategy then depends on which optimal response the follower is assumed to select, and the analysis must specify how this ambiguity is resolved [46, 205].

EXAMPLE 3.9 **Pessimistic Stackelberg tie-breaking**

Consider the two-person game in Table 3.4, in which P_1 is the leader, P_2 is the follower, and each entry is the pair of costs (J_1, J_2), which the two players minimize. Each player chooses between L and R, and the follower has the additional option M.

	P_2		
P_1	L	M	R
L	$(0,0)^\star$	$(1,0)^\star$	$(3,1)$
R	$(2,-1)^\star$	$(2,0)$	$(-1,1)$

Table 3.4: Stackelberg game. A star ($\star$) marks the follower's optimal (cost-minimizing) response in each row of the leader.

Figure 3.12 displays the same game as a tree in which the follower observes the leader's choice.

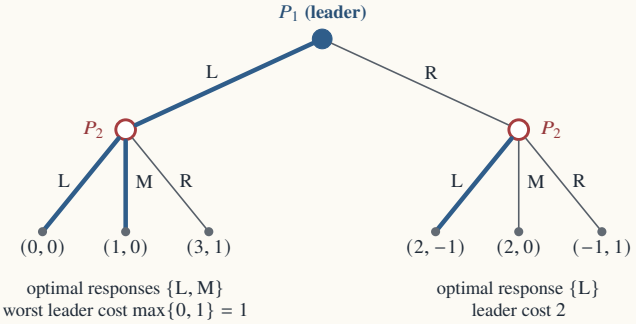


Figure 3.12: The Stackelberg game of Table 3.4 in extensive form, with costs (J_1, J_2) that both players minimize. Thick edges mark the follower's optimal responses. After L the follower is indifferent between L and M, so a pessimistic leader evaluates L by the larger leader cost 1; after R the follower's response L is unique and costs the leader 2. The pessimistic Stackelberg strategy is therefore L, with guaranteed cost 1.

A pessimistic leader evaluates each of its strategies by the worst cost among the follower's optimal responses. If P_1 chooses L, the follower's costs are 0, 0, and 1 for L, M, and R, respectively, so the follower's optimal responses are L and M, and the leader's worst-case cost is $\max\{0, 1\} = 1$. If P_1 chooses R, the follower's costs are -1, 0, and 1, so the follower's unique optimal response is L, and the leader's cost is 2. Since $1 < 2$, the pessimistic Stackelberg strategy of P_1 is L, with guaranteed cost 1. An optimistic leader, who expects the follower to select L after L, would instead anticipate the cost 0; the predicted cost thus depends on the tie-breaking convention, which a Stackelberg analysis must state explicitly.

Observation. When the follower has several optimal responses, the leader's predicted cost depends on the tie-breaking convention. Pessimistic Stackelberg analysis protects the leader against the worst follower response among the follower's best responses.

The formal definitions are stated for a two-person finite game without chance moves in which P_1 is the leader and P_2 is the follower. The players have finite, nonempty strategy sets Γ_1 and Γ_2 and cost functions $J_1, J_2 : \Gamma_1 \times \Gamma_2 \to \mathbb{R}$, which they minimize.

Definition 3.2.10 (Optimal response set)**.** The *optimal response set* of P_2 to a strategy $\gamma_1 \in \Gamma_1$ of P_1 is the set $BR_2(\gamma_1) = \{\gamma_2 \in \Gamma_2 : J_2(\gamma_1, \gamma_2) \le J_2(\gamma_1, \gamma_2') \text{ for all } \gamma_2' \in \Gamma_2\}$. Since Γ_2 is finite and nonempty, $BR_2(\gamma_1)$ is nonempty for every $\gamma_1 \in \Gamma_1$.

Definition 3.2.11 (Pessimistic Stackelberg strategy)**.** A strategy $\gamma_1^* \in \Gamma_1$ is a *(pessimistic) Stackelberg strategy* for the leader if

$$\max_{\gamma_2 \in BR_2(\gamma_1^*)} J_1(\gamma_1^*, \gamma_2) = \min_{\gamma_1 \in \Gamma_1} \max_{\gamma_2 \in BR_2(\gamma_1)} J_1(\gamma_1, \gamma_2) =: J_1^*.$$

The quantity J_1^* is the *(pessimistic) Stackelberg cost* of the leader.

Remark 3.2.12 (Uniqueness of the cost). A Stackelberg strategy of the leader need not be unique, but the pessimistic Stackelberg cost J_1^* is uniquely determined, being the value of the min–max problem in the preceding definition.

Remark 3.2.13 (Single-valued follower response). If $BR_2(\gamma_1)$ is a singleton for each $\gamma_1 \in \Gamma_1$, then the follower's response defines a function $b_2 : \Gamma_1 \to \Gamma_2$ by $BR_2(\gamma_1) = \{b_2(\gamma_1)\}$, and the condition defining a Stackelberg strategy γ_1^* reduces to

$$J_1(\gamma_1^*, b_2(\gamma_1^*)) = \min_{\gamma_1 \in \Gamma_1} J_1(\gamma_1, b_2(\gamma_1)) = J_1^*.$$

In this case J_1^* is the cost that the leader actually attains.

Definition 3.2.14 (Stackelberg outcome)**.** Let $\gamma_1^* \in \Gamma_1$ be a Stackelberg strategy of the leader P_1, and let $\gamma_2^* \in BR_2(\gamma_1^*)$ be an optimal response of the follower. The pair (γ_1^*, γ_2^*) is a Stackelberg solution, and the pair of costs $\big(J_1(\gamma_1^*, \gamma_2^*), J_2(\gamma_1^*, \gamma_2^*)\big)$ is the corresponding Stackelberg outcome.

Remark 3.2.15 (Outcome cost versus guaranteed cost). The leader's outcome cost satisfies $J_1(\gamma_1^*, \gamma_2^*) \le J_1^*$, and the inequality can be strict. In Example 3.9, the Stackelberg solution (L, L) gives P_1 the cost $0 < J_1^* = 1$, whereas the Stackelberg solution (L, M) gives P_1 the cost $1 = J_1^*$. If $BR_2(\gamma_1^*)$ is a singleton, the two costs coincide.

3.2.4.2 Inverse Stackelberg Games

In a standard Stackelberg game, the leader commits to a fixed action. In an inverse Stackelberg game, the leader commits instead to a *rule*, that is, a mapping from the follower's action to the leader's action, and thereby shapes the follower's incentives [46].

Let $\mathcal{A}_L$ and $\mathcal{A}_F$ denote the decision sets of the leader and the follower, and let $J_L, J_F : \mathcal{A}_L \times \mathcal{A}_F \to \mathbb{R}$ be their cost functions, both of which are minimized. The leader announces a function $\gamma_L : \mathcal{A}_F \to \mathcal{A}_L$, which maps each decision $a_F \in \mathcal{A}_F$ of the follower into the leader's decision $a_L = \gamma_L(a_F)$. Given the announced rule, the follower chooses a decision $a_F^*(\gamma_L) \in \arg\min_{a_F \in \mathcal{A}_F} J_F(\gamma_L(a_F), a_F)$, where, as in the preceding analysis, a tie-breaking convention is required whenever the minimizer is not unique. Anticipating this response, the leader chooses a rule γ_L^* satisfying

$$\gamma_L^* \in \arg\min_{\gamma_L} J_L\big(\gamma_L(a_F^*(\gamma_L)), a_F^*(\gamma_L)\big),$$

where the minimum is taken over the admissible rules $\gamma_L : \mathcal{A}_F \to \mathcal{A}_L$. The leader thus optimizes over a space of functions, and the objective is the composition of the leader's rule with the follower's induced response.

3.2.5 Level-k Thinking

Level-k thinking is a behavioral model of bounded strategic reasoning [414]. Rather than assuming that every player reasons all the way to an equilibrium, the model assigns to each player a finite depth of reasoning, defined recursively. A level-0 player is nonstrategic and follows a rule specified by the modeler, such as uniform randomization over actions. A level-1 player best responds to level-0 behavior, and, for each $k \geq 1$, a level-k player best responds to the belief that the other players reason at level $k-1$. If π_{-i}^{k-1} denotes player i's belief about the level-$(k-1)$ behavior of the other players, a level-k action of player i therefore satisfies

$$a_i^k \in \arg\max_{a_i \in \mathcal{A}_i} \mathbb{E}_{\pi_{-i}^{k-1}} \left[u_i(a_i, a_{-i}) \right].$$

In population versions of the model, weights ψ_i^k are assigned to the reasoning levels, and predicted behavior is a mixture of level-0, level-1, level-2, and higher-level choices. Such mixtures account for observed play that departs from equilibrium predictions, particularly in complex games and in games played only a few times. Chapter 11 returns to level-k games together with other models of bounded rationality, including quantal response, rational inattention, and prospect theory, and applies them to surveillance and security-investment problems.

EXAMPLE 3.10 **Centipede game**

In a centipede game, two players alternate between taking the larger share of a growing pot and passing the pot to the other player. A player who passes receives less than the immediate share if the opponent then takes, but more if the opponent also passes. Each player is therefore tempted to take at every move, although both would gain by waiting. The original version had 100 moves, which explains the name; any game with this structure is called a centipede game.

Backward induction yields a unique subgame perfect equilibrium, in which the first player takes at the first move, and every Nash equilibrium produces the same outcome. In experiments, however, relatively few subjects take immediately, and those who continue earn more than the equilibrium payoff [389]. The centipede game is therefore a standard illustration of backward induction and of iterated elimination of dominated strategies, as well as of the limits of these equilibrium predictions as descriptions of human play.

Figure 3.13 shows a four-move version.

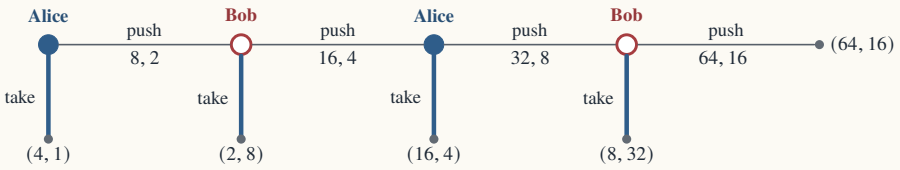


Figure 3.13: A four-move centipede game between Alice and Bob. Each push doubles both piles, and the numbers below each push give the piles the next mover faces; the mover who takes keeps the larger pile. Payoffs are (Alice, Bob). Thick edges mark the backward-induction choices: *take* at every node, so the game ends at the first move with payoffs $(4, 1)$.

The players are Alice and Bob, and Alice moves first. Initially there are two piles of coins, one of 4 coins and one of 1 coin. At each turn the active player either takes the larger pile, leaving the smaller pile to the other player, or pushes both piles to the other player, in which case both piles double. The game ends when a player takes or after a fixed number of moves. In Figure 3.13, if Bob pushes at the fourth node, the game ends with payoffs $(64, 16)$. At that node Bob prefers taking, which yields 32, to pushing, which yields 16. Given this choice, Alice prefers taking at the third node (16 rather than 8), and the same comparison applies at every earlier node, down to the first, where Alice takes (4 rather than 2). The coins added by each push come from outside the game; neither player pays for them.

Observation. Backward induction predicts immediate taking, whereas bounded-depth reasoning can predict delayed taking. The centipede game therefore distinguishes equilibrium reasoning from level-k behavioral reasoning.

3.2.5.1 Level-k Thinking Model

Backward induction in the centipede game presumes that both players reason forward to the final node and then backward to the first. Level-k thinking relaxes this presumption by replacing common knowledge of full rationality with a finite-depth model of strategic reasoning. A level-1 player may best respond to a level-0 rule that sometimes pushes, a level-2 player best responds to level-1 behavior, and so on. As k increases, the prediction may approach that of backward induction, while low levels account for the observation that human players often push for several rounds before taking.

Consider a finite normal-form game

$$\mathcal{G} = \langle \mathcal{N}, \{\mathcal{A}_i\}_{i\in\mathcal{N}}, \{u_i\}_{i\in\mathcal{N}} \rangle .$$

For each player $i \in \mathcal{N}$, fix a level-0 rule $\pi_i^0 \in \Delta(\mathcal{A}_i)$. This rule is not derived from optimization; it is a behavioral anchor supplied by the modeler, commonly uniform randomization, a salient default action, or a simple heuristic observed in data. For $k \geq 1$, player i believes that every other player $j \neq i$ uses a lower-level strategy, typically π_j^{k-1}. The level-k best-response set of player i is

$$BR_i^k = \arg\max_{\pi_i \in \Delta(\mathcal{A}_i)} \mathbb{E}_{\pi_i, \pi_{-i}^{k-1}} \left[u_i(a_i, a_{-i}) \right],$$

and a level-k strategy π_i^k is any element of BR_i^k. If the best response is unique and pure, the recursion selects a single action; otherwise the modeler must specify a tie-breaking rule or allow π_i^k to mix over the best-response set.

In population applications, the observed behavior of player i is modeled as a mixture of reasoning levels. Let $\bar{k}$ be the highest level considered, and let $\psi_i^k \geq 0$ be the fraction of the population in the role of player i that reasons at level $k \in \{0, 1, \ldots, \bar{k}\}$, with $\sum_{k=0}^{\bar{k}} \psi_i^k = 1$. The predicted action distribution of player i is

$$\widehat{\pi}_i = \sum_{k=0}^{\bar{k}} \psi_i^k \pi_i^k.$$

The weights ψ_i^k may be taken from experimental evidence, estimated from data, or varied in a sensitivity analysis. The model complements rather than replaces equilibrium analysis: it describes how predictions change when agents reason only a bounded number of steps.

The same construction applies to extensive-form games with behavioral strategies. A level-0 defender may, for instance, follow a default monitoring rule at every information set; a level-1 attacker computes a best response to that rule, and a level-2 defender best responds to the level-1 attacker. Let $\mathcal{B}_i$ denote the set of behavioral strategies of player i, and let β_{-i}^{k-1} denote player i's belief about the level-$(k-1)$ behavioral strategies of the other players. A level-k behavioral strategy of player i satisfies

$$\beta_i^k \in \arg\max_{\beta_i \in \mathcal{B}_i} \mathbb{E}_{\beta_i, \beta_{-i}^{k-1}} \left[u_i(h) \right],$$

where h is the terminal history induced by the behavioral-strategy profile $(\beta_i, \beta_{-i}^{k-1})$. The extensive-form version is particularly relevant to security because bounded reasoning may arise at specific information sets: an attacker may fail to anticipate a later forensic response, or a defender may respond myopically to an alert without accounting for the attacker's earlier incentives to deceive.

In the centipede game, a level-0 rule may assign positive probability to pushing at each decision node. A level-1 player who expects this behavior may rationally push for several rounds, since the opponent is believed likely to continue, whereas a level-2 player anticipates the level-1 response and may take earlier. Suppose, for example, that a level-0 player pushes with probability $1/2$ at every node of Figure 3.13. A level-1 Alice then pushes at both of her nodes. A level-1 Bob pushes at node 2 and takes at node 4, and a level-2 Alice, best responding to him, pushes at node 1 but takes at node 3. The model thus generates delayed taking without abandoning optimization: every level optimizes, but against a

bounded model of the other player rather than against a fully rational equilibrium strategy.

3.2.6 Signaling Games

A signaling game models strategic interaction under asymmetric information [548]. Nature first chooses a type for the sender. The sender observes the type and sends a message, and the receiver observes the message, forms a belief about the sender's type, and chooses an action. Chapter 4 develops the underlying theory of types and beliefs in Bayesian games, and Chapter 13 uses signaling games with evidence, together with Bayesian persuasion, to model cyber deception.

A signaling model distinguishes three objects that informal reasoning often conflates: the sender's private information, the observable message, and the receiver's posterior belief. A message may reveal, conceal, or strategically distort information. The receiver therefore cannot identify the message with the type; it must infer which types would have chosen to send that message in equilibrium.

Definition 3.2.16 (Finite signaling game)**.** A finite signaling game consists of:

- a finite type set $\Theta = \{\theta_1, \theta_2, \dots, \theta_{|\Theta|}\}$ selected by Nature;
- a prior distribution ρ over Θ, where $\rho(\theta_i) \geq 0$ and $\sum_{i=1}^{|\Theta|} \rho(\theta_i) = 1$;
- a finite message set $\mathcal{M} = \{m_1, \dots, m_{|\mathcal{M}|}\}$ for the sender;
- a finite action set $\mathcal{A}_R = \{a_1, \dots, a_{|\mathcal{A}_R|}\}$ for the receiver;
- payoff functions $U_S(\theta, m, a)$ and $U_R(\theta, m, a)$ for the sender and receiver.

A sender strategy is a conditional distribution $\sigma_S(m|\theta)$ over messages given type θ. A receiver strategy is a conditional distribution $\sigma_R(a|m)$ over actions given message m. A receiver belief is a conditional distribution $\mu(\theta|m)$ over types after observing message m.

The timing of the game is as follows.

(i) Nature draws $\theta \in \Theta$ according to the prior ρ.

(ii) The sender observes θ and chooses a message $m \in \mathcal{M}$ according to $\sigma_S(\cdot|\theta)$.

(iii) The receiver observes m, but not θ, and updates its belief to $\mu(\cdot|m)$.

(iv) The receiver chooses an action $a \in \mathcal{A}_R$ according to $\sigma_R(\cdot|m)$.

(v) Payoffs $U_S(\theta, m, a)$ and $U_R(\theta, m, a)$ are realized.

A signaling game is thus an extensive-form game with imperfect information, in which the receiver's information set after message m contains every history (θ, m), $\theta \in \Theta$, that could have generated that message.

Given a strategy profile (σ_S, σ_R), the probability of a terminal history (θ, m, a) is

$$\mathbb{P}(\theta, m, a) = \rho(\theta)\sigma_S(m|\theta)\sigma_R(a|m).$$

Whenever message m is sent with positive probability, the receiver's posterior belief is determined by Bayes' rule,

$$\mu(\theta|m) = \frac{\rho(\theta)\sigma_S(m|\theta)}{\sum_{\theta' \in \Theta} \rho(\theta')\sigma_S(m|\theta')}.$$

If the denominator is zero, m is off the equilibrium path. Bayes' rule then places no restriction on $\mu(\cdot|m)$, but the specified belief still determines whether the receiver's off-path action is sequentially rational.

Figures 3.14 and 3.15 show the timing and the game tree of a signaling game. The sender knows the type selected by Nature, whereas the receiver must infer the type from the sender's message.

Figure 3.16 shows a signaling game with two types, two messages, and two receiver actions.

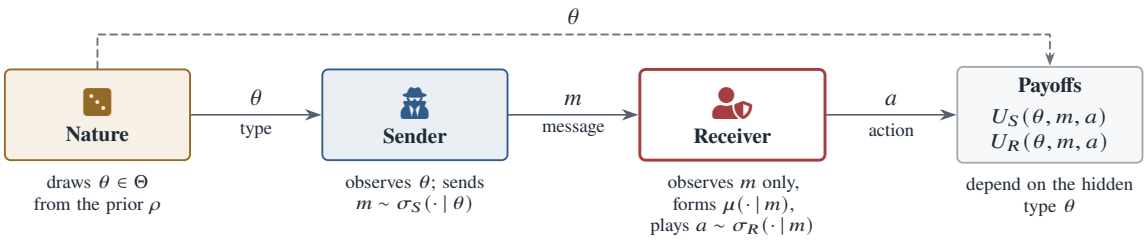


Figure 3.14: Timing of a signaling game. The sender privately observes the type drawn by Nature; the receiver sees only the message and must act on its belief about the type.

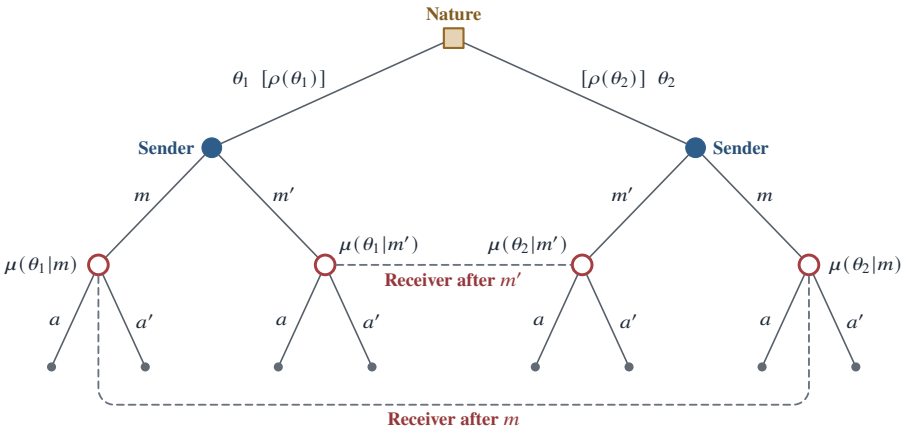


Figure 3.15: Game tree of a signaling game with two types and two messages. Dashed lines join the receiver's nodes that follow the same message; these are the receiver's information sets, and $\mu(\theta|m)$ is the belief held at each node.

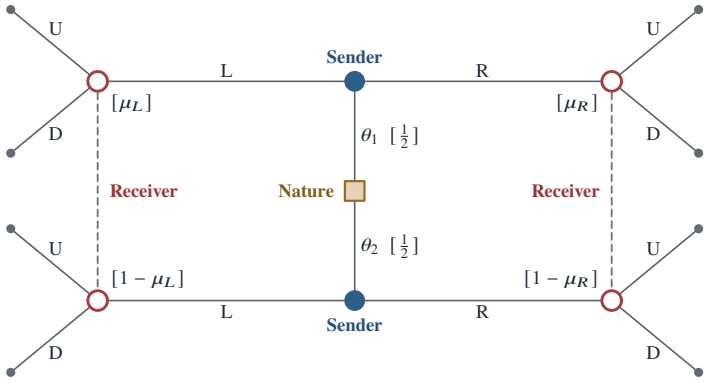


Figure 3.16: A signaling game with two types, two messages (L, R), and two receiver actions (U, D). Nature draws each type with probability $\frac{1}{2}$. The receiver's information sets are the vertical dashed lines, with beliefs $\mu_L = \mu(\theta_1|\mathrm{L})$ and $\mu_R = \mu(\theta_1|\mathrm{R})$.

In a two-type game, a pure sender strategy specifies a message for each type. The pair (L, R) denotes the strategy in which type θ_1 sends L and type θ_2 sends R; it is a complete contingent plan, since it specifies a message for every type. A pure receiver strategy is likewise a complete contingent plan; for example, (U, D) denotes the strategy in which the receiver chooses U after L and D after R. A sender strategy is *pooling* if both types send the same message, as in (L, L) and (R, R), and *separating* if the two types send different messages, as in (L, R) and (R, L). A *mixed sender strategy* is one in which at least one type randomizes over messages according to the probabilities $\sigma_S(m|\theta)$. A separating strategy reveals the type to the receiver, whereas under a pooling strategy the receiver must rely on the prior after the message that is sent and on off-path beliefs after the message that is not.

3.2.6.1 Perfect Bayesian Nash Equilibrium

Perfect Bayesian Nash equilibrium (PBNE) combines optimal strategies with consistent beliefs and is closely related to sequential equilibrium [336, 205]. It is the standard solution concept for signaling games: the receiver's action must be optimal given its beliefs, and those beliefs must be consistent with the sender's strategy whenever Bayes' rule applies.

Definition 3.2.17 (Perfect Bayesian Nash equilibrium)**.** An assessment $(\sigma_S^*, \sigma_R^*, \mu^*)$ is a perfect Bayesian Nash equilibrium if:

(i) for every message m, the receiver's strategy $\sigma_R^*(\cdot|m)$ assigns positive probability only to actions that maximize expected receiver payoff under $\mu^*(\cdot|m)$;

(ii) for every type θ, the sender's strategy $\sigma_S^*(\cdot|\theta)$ assigns positive probability only to messages that maximize the sender's expected payoff given the receiver strategy σ_R^*;

(iii) for every message reached with positive probability under σ_S^*, the belief $\mu^*(\cdot|m)$ is obtained from ρ and σ_S^* by Bayes' rule.

An equilibrium of a signaling game is called an *assessment* because it consists of strategies together with beliefs. Strategies alone do not suffice, since the receiver's optimal action after a message depends on its belief about the sender's type.

For pure strategies, write $m^*(\theta)$ for the message sent by type θ and $a^*(m)$ for the action chosen by the receiver after message m. The three conditions of the definition then reduce to the following. The receiver is sequentially rational if, for every $m \in \mathcal{M}$,

$$a^*(m) \in \arg\max_{a \in \mathcal{A}_R} \sum_{\theta \in \Theta} \mu(\theta|m) U_R(\theta, m, a);$$

the sender is sequentially rational if, for every $\theta \in \Theta$,

$$m^*(\theta) \in \arg\max_{m \in \mathcal{M}} U_S(\theta, m, a^*(m));$$

and beliefs are consistent if, for every message m_j for which the denominator is positive and every type θ_i,

$$\mu(\theta_i|m_j) = \frac{\rho(\theta_i)\sigma_S(m_j|\theta_i)}{\sum_{\theta' \in \Theta} \rho(\theta')\sigma_S(m_j|\theta')}.$$

Off the equilibrium path, Bayes' rule does not determine beliefs, and PBNE admits any belief, provided that the receiver's action is optimal given that belief.

In the game of Figure 3.16, the receiver's belief after message L is $(\mu_L, 1-\mu_L)$ and after message R is $(\mu_R, 1-\mu_R)$, where

$$\mu_L = \mu(\theta_1|\mathrm{L}), \qquad \mu_R = \mu(\theta_1|\mathrm{R}).$$

EXAMPLE 3.11 **Checking PBNE in a two-type signaling game**

Figure 3.17 specifies payoffs for the game of Figure 3.16. The sender is of type θ_1 or θ_2, each with prior probability $\rho(\theta_1) = \rho(\theta_2) = 1/2$, and sends L or R; the receiver observes the message but not the type and chooses U or D. The objective is to determine every assessment that satisfies the three PBNE conditions: receiver optimality, sender optimality, and Bayesian consistency of beliefs on the equilibrium path.

The payoffs (U_S, U_R) are as follows. For type θ_1, the outcomes (L, U), (L, D), (R, U), and (R, D) yield $(1, 3)$, $(4, 0)$, $(2, 1)$, and $(0, 0)$, respectively; for type θ_2, the same outcomes yield $(2, 4)$, $(0, 1)$, $(1, 0)$, and $(1, 2)$. The analysis proceeds in three stages. The receiver's sequentially rational responses are first derived for arbitrary beliefs μ_L and μ_R; the sender's optimal messages are then derived given those responses; finally, each candidate sender strategy is checked for consistency with Bayes' rule. Candidate equilibria frequently fail at the last stage even when the strategies are optimal given the beliefs.

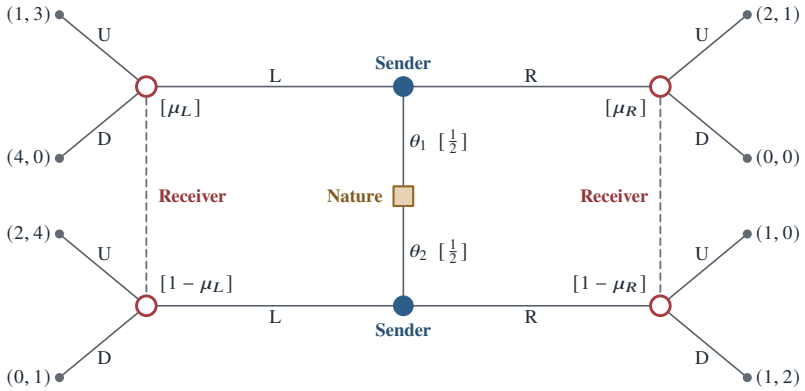


Figure 3.17: The signaling game of Figure 3.16 with payoffs (sender, receiver) at each terminal node.

Receiver's best responses. After message m, the receiver chooses $a \in \{\mathrm{U}, \mathrm{D}\}$ to maximize $\sum_{\theta \in \Theta} \mu(\theta|m) U_R(\theta, m, a)$. After L, the expected payoffs of the two actions are

$$U_R(\mathrm{U}|\mathrm{L}) = 3\mu_L + 4(1-\mu_L) = 4 - \mu_L \quad \text{and} \quad U_R(\mathrm{D}|\mathrm{L}) = 0 \cdot \mu_L + 1 \cdot (1-\mu_L) = 1 - \mu_L.$$

Since $4 - \mu_L > 1 - \mu_L$ for every $\mu_L \in [0,1]$, action U is the unique best response to L, whatever the receiver believes about the type that sent it. After R, the expected payoffs are

$$U_R(\mathrm{U}|\mathrm{R}) = 1 \cdot \mu_R + 0 \cdot (1-\mu_R) = \mu_R \quad \text{and} \quad U_R(\mathrm{D}|\mathrm{R}) = 0 \cdot \mu_R + 2(1-\mu_R) = 2 - 2\mu_R,$$

so that $U_R(\mathrm{U}|\mathrm{R}) \geq U_R(\mathrm{D}|\mathrm{R})$ if and only if $\mu_R \geq 2/3$. The receiver's sequentially rational responses are therefore

$$a^*(\mathrm{L}) = \mathrm{U}, \qquad a^*(\mathrm{R}) = \begin{cases} \mathrm{U}, & \mu_R > 2/3, \\ \mathrm{D}, & \mu_R < 2/3, \end{cases}$$

and at $\mu_R = 2/3$ the receiver is indifferent after R, so that any mixture $\sigma_R(\mathrm{U}|\mathrm{R}) = y \in [0,1]$ is sequentially rational. In contrast to the response to L, the response to R depends on the posterior: U is optimal when type θ_1 is sufficiently likely, and D when type θ_2 is.

Sender's best responses. Each type anticipates the receiver's response to each message and chooses a message that maximizes $U_S(\theta, m, a^*(m))$. Because $a^*(\mathrm{L}) = \mathrm{U}$ for every belief, message L yields $U_S(\theta_1, \mathrm{L}, \mathrm{U}) = 1$ to type θ_1 and $U_S(\theta_2, \mathrm{L}, \mathrm{U}) = 2$ to type θ_2. Let $y = \sigma_R(\mathrm{U}|\mathrm{R})$ denote the probability that the receiver chooses U after R. Message R then yields

$$U_S(\theta_1, \mathrm{R}) = 2y + 0 \cdot (1-y) = 2y \quad \text{and} \quad U_S(\theta_2, \mathrm{R}) = 1 \cdot y + 1 \cdot (1-y) = 1.$$

Type θ_2 therefore strictly prefers L, since $2 > 1$, regardless of the receiver's belief and behavior after R; that is, $m^*(\theta_2) = \mathrm{L}$ in every equilibrium. Type θ_1 strictly prefers R if $y > 1/2$, strictly prefers L if $y < 1/2$, and is indifferent if $y = 1/2$. In terms of pure receiver responses, type θ_1 sends R when R induces U (payoff $2 > 1$) and L when R induces D (payoff $1 > 0$). The belief μ_R thus governs the choice of type θ_1 through its effect on the receiver's action after R, whether or not R is on the equilibrium path.

Belief consistency. With the prior $\rho(\theta_1) = \rho(\theta_2) = 1/2$, Bayes' rule gives, for every message m_j sent with positive probability,

$$\mu(\theta_i|m_j) = \frac{\rho(\theta_i)\sigma_S(m_j|\theta_i)}{\sum_{\theta' \in \Theta} \rho(\theta')\sigma_S(m_j|\theta')}.$$

We examine the four pure sender strategies in turn.
Under pooling on R, that is, (R, R), Bayes' rule gives $\mu_R = 1/2$, while L is off path and μ_L is unrestricted. Since $a^*(\mathrm{L}) = \mathrm{U}$ for every μ_L, type θ_2 obtains 2 by deviating to L rather than 1 from R. No belief supports this strategy, and there is no PBNE with pooling on R.
Under the separating strategy (L, R), both messages are on path, and

$$\mu_L = \frac{1 \cdot \frac{1}{2}}{1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2}} = 1, \qquad \mu_R = \frac{0 \cdot \frac{1}{2}}{0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2}} = 0.$$

Since $\mu_R < 2/3$, the receiver chooses D after R, and type θ_2 obtains 1 from R but 2 from L. This strategy is therefore not part of any PBNE.
Under the separating strategy (R, L), both messages are again on path, and

$$\mu_L = \frac{0 \cdot \frac{1}{2}}{0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2}} = 0, \qquad \mu_R = \frac{1 \cdot \frac{1}{2}}{1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2}} = 1.$$

Since $\mu_R = 1 > 2/3$, the receiver chooses U after both messages. Type θ_1 obtains 2 from R and 1 from L, and type θ_2 obtains 2 from L and 1 from R, so neither type has a profitable deviation. In this equilibrium the message reveals the type: R identifies θ_1 and L identifies θ_2.
Under pooling on L, that is, (L, L), Bayes' rule gives

$$\mu_L = \frac{1 \cdot \frac{1}{2}}{1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2}} = \frac{1}{2},$$

so that the on-path belief equals the prior and L is uninformative. Message R is off path, and μ_R is not determined by Bayes' rule. Type θ_2 has no incentive to deviate, and type θ_1 has none if and only if $2y \le 1$, that is, $y \le 1/2$. Sequential rationality after R requires $y = 0$ if $\mu_R < 2/3$ and permits any $y \in [0, 1]$ if $\mu_R = 2/3$, whereas $\mu_R > 2/3$ forces $y = 1$, which violates $y \le 1/2$. Pooling on L is therefore supported exactly by the off-path beliefs $\mu_R \le 2/3$, with the receiver choosing D after R when $\mu_R < 2/3$ and choosing U after R with probability at most $1/2$ when $\mu_R = 2/3$. Although the off-path belief is free within PBNE, it is not irrelevant: it must make the receiver willing to choose D (with sufficient probability) after R, which in turn deters type θ_1 from deviating to R. It remains to consider sender strategies in which type θ_1 randomizes. Since type θ_2 always sends L, any positive probability $\sigma_S(\mathrm{R}|\theta_1) > 0$ implies $\mu_R = 1$ by Bayes' rule, hence $y = 1$; but then type θ_1 strictly prefers R, so randomization by θ_1 cannot be optimal. The results of the analysis are collected in the following proposition.

PROPOSITION 3.2.18
In the signaling game of Figure 3.17, let $y = \sigma_R(\mathrm{U}|\mathrm{R})$. An assessment is a perfect Bayesian Nash equilibrium if and only if it belongs to one of the following two families.

(i) *Separating equilibrium.* The sender plays (R, L), the receiver plays (U, U), and the beliefs are $\mu_L = 0$ and $\mu_R = 1$.

(ii) *Pooling equilibria.* The sender plays (L, L), the receiver chooses U after L and U after R with probability y, the on-path belief is $\mu_L = 1/2$, and the off-path belief and receiver behavior after R satisfy either $\mu_R < 2/3$ and $y = 0$, or $\mu_R = 2/3$ and $y \in [0, 1/2]$.

In particular, the equilibria in which the receiver uses a pure strategy are the separating equilibrium in (i) and the pooling equilibria with receiver strategy (U, D) and off-path belief $\mu_R \le 2/3$.

Proof. Receiver optimality forces $a^*(\mathrm{L}) = \mathrm{U}$, and therefore type θ_2 sends L with probability one in every equilibrium, so the sender strategy is (L, L), (R, L), or a mixture in which only θ_1 randomizes. Mixtures are excluded because they imply $\mu_R = 1$, under which R is strictly optimal for θ_1. Under (R, L), Bayes' rule yields

$\mu_L = 0$ and $\mu_R = 1$, and the receiver's and sender's choices are optimal, which gives (i). Under (L, L), Bayes' rule yields $\mu_L = 1/2$, and the receiver's sequential rationality after R together with the no-deviation condition $2y \le 1$ for type θ_1 holds exactly for the pairs (μ_R, y) listed in (ii). □

The example shows why the solution of a signaling game must include beliefs. In a normal-form game, a strategy profile is verified by checking that each player best responds to the strategies of the others. In a signaling game, the receiver's best response is defined only once beliefs are specified after each message. On-path beliefs are restricted by Bayes' rule; off-path beliefs are not determined in this way, but they still constrain deviations, because they determine the receiver's response to an unexpected message.

Observation. PBNE requires both optimal actions and consistent beliefs. On-path beliefs are pinned down by Bayes' rule; off-path beliefs must be chosen so that the proposed continuation strategies remain sequentially rational and so that sender deviations are not profitable.

3.3 Applications in Security

This section collects the security applications of the chapter, now that the normal-form and extensive-form foundations are in place. Treating them together separates the core modeling ideas from the domain examples and shows how the same security problem takes different forms under different game representations [12, 380]. A normal-form model emphasizes simultaneous policy choice, payoff matrices, security levels, and equilibrium conditions; an extensive-form model emphasizes commitment, observation, information sets, belief updates, and sequential rationality.

The applications are organized around this distinction. The cut-the-rope and FlipIt games reduce stealthy intrusion, cleaning, and takeover to tradeoffs in strategic allocation and timing. Stackelberg security games then show how an explicit order of moves changes both the formulation of the problem and the solution concept; extensions to sequential monitoring, signaling, and bounded reasoning are taken up in the concluding discussion of the chapter.

3.3.1 Normal-Form Security Applications: Cut-the-Rope and FlipIt

A normal-form game is an appropriate model whenever a security problem can be reduced to a choice among policies. The defender chooses a configuration, an inspection level, or an allocation of resources; the attacker chooses a target, an attack mode, or a level of effort. Once both choices are fixed, the payoff records a quantity such as expected loss, detection probability, operational cost, or attacker gain. The abstraction is deliberately compact: it does not describe the full sequence of scanning, intrusion, detection, and response, but it provides a baseline model against which strategic incentives can be compared [12, 380].

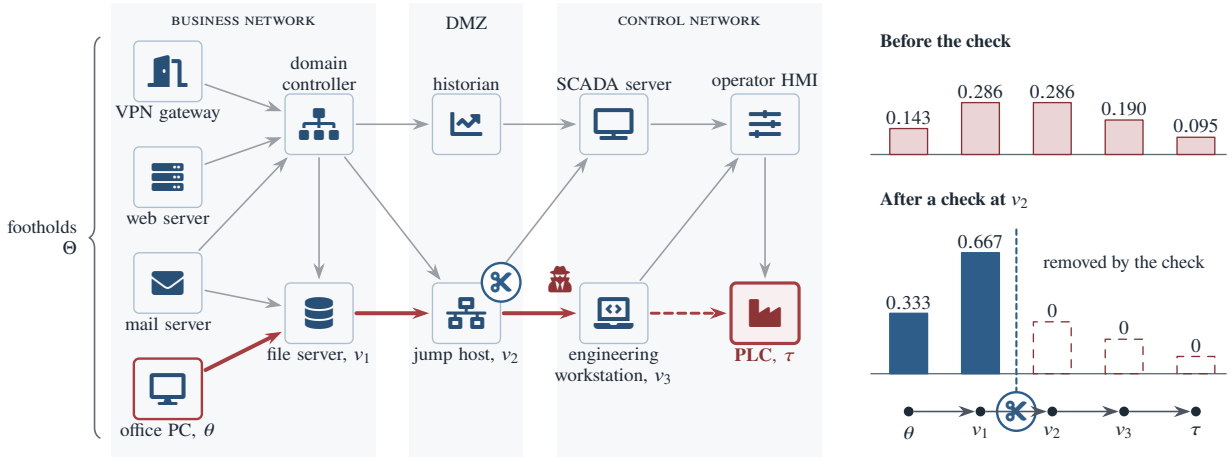
This subsection treats two security games of this kind. The cut-the-rope game models stealthy intrusion along an attack path, in which the attacker retains return paths and the defender attempts to remove them. FlipIt models stealthy takeover of a shared resource, in which either player may move at any time but neither immediately observes who controls the resource. Both games are dynamic in their full form, yet each admits a normal-form or stage-game interpretation that exposes the strategic tradeoffs before richer extensive-form or stochastic models are introduced.

3.3.1.1 The Cut-the-Rope Game

The cut-the-rope game was introduced to model advanced persistent threats that progress through a system gradually and stealthily [493]. The system is represented by an attack graph $G = (\mathcal{V}, \mathcal{E})$, whose vertices may represent hosts, privilege levels, vulnerabilities, or configuration states. The attacker starts from an unknown compromised location and seeks to reach a target vertex $\tau \in \mathcal{V}$; a directed path toward τ represents a sequence of successful penetrations, privilege escalations, or lateral movements. The "rope" is the attacker's hidden return path. If the defender cleans a vertex on this path, the attacker loses the portion of the route from the cleaned vertex onward and is pushed back to an earlier foothold. Richer multi-stage models of advanced persistent threats in industrial control systems are developed in Chapter 12.

A concrete setting. Consider the small industrial network in Figure 3.18(a), which is divided into three zones: a business network, a demilitarized zone (DMZ), and a control network. A water utility runs its treatment process from a

programmable logic controller (PLC) in the control network, which is the target τ. The control network also contains a SCADA server, an engineering workstation, and an operator human–machine interface (HMI); the DMZ contains a historian and a jump host that bridge the business network and the control network. An intruder has already gained a foothold on one of four internet-facing or user-facing machines in the business network (a VPN gateway, a web server, a mail server, or an office PC compromised by a phishing email), but the security team does not know which. Several routes lead from these footholds to the PLC, for example through the domain controller, the historian, the SCADA server, and the HMI, or through the file server, the jump host, and the engineering workstation. From its foothold the intruder moves laterally, one step at a time, harvesting credentials as it proceeds. The route it has already traveled is its rope: stolen sessions and implants that allow it to fall back and continue if a host further along is cleaned. The security team cannot monitor every host continuously. Instead, in each inspection period it performs a single spot check, such as reimaging a host or rotating the credentials stored on it. A check on a host through which the intruder has passed cuts the rope at that host, and the intruder loses that host and everything beyond it. The defender must therefore decide which host to check, and how often, without knowing where the intruder started or how far it has advanced.



(a) Attack graph (b) Attacker location on r

Figure 3.18: Cut-the-rope in an industrial network. (a) Attack graph spanning the business network, the DMZ, and the control network. The intruder holds an unknown foothold in Θ (here the office PC, θ) and has advanced along the path r (thick) to the engineering workstation; the solid part of r behind it is the rope. A spot check at the jump host v_2 cuts the rope. (b) Distribution of the intruder's location L on the vertices of r (horizontal axis) after one inspection interval, when it advances a Poisson number of steps with rate $\lambda = 2$. Top: without a check, the intruder reaches the PLC with probability 0.095. Bottom: a check at v_2 cuts r before v_2; the mass on v_2, v_3, and τ (dashed) is removed, and the remaining mass is renormalized onto θ and v_1, so the PLC is no longer reached.

Formal model. An instance of the game is a tuple $\mathcal{G}_{\mathrm{CTR}} = \big(G, \tau, \mathcal{K}, \{\mathcal{R}_\theta\}_{\theta\in\Theta}, \lambda, \rho\big)$, in which $\mathcal{K} \subseteq \mathcal{V} \setminus \{\tau\}$ is the set of the defender's spot checks (cleaning actions), $\Theta \subseteq \mathcal{V} \setminus \{\tau\}$ is the set of possible attacker starting locations, $\rho \in \Delta(\Theta)$ is the defender's belief over these locations, $\mathcal{R}_\theta$ is the set of attack paths from θ to τ, and $\lambda > 0$ is the attack rate. During one inspection interval of the defender, the attacker advances a random number $K \sim \mathrm{Pois}(\lambda)$ of steps, with $\mathrm{Pois}_\lambda(k) = \mathbb{P}\{K = k\} = \exp(-\lambda)\lambda^k/k!$ for $k = 0, 1, 2, \ldots$. The parameter λ is therefore not a payoff parameter; it governs how far the attacker is expected to move while the defender is inactive.

For a path $r = (\theta, v_1, \ldots, \tau) \in \mathcal{R}_\theta$, let $\mathcal{V}(r)$ denote its set of vertices and let $d_r(\theta, v)$ denote the number of edges from θ to v along r. Let L denote the attacker's location at the end of the inspection interval. Before the defender acts, the distribution of L is obtained by conditioning the Poisson step distribution on the finite path,

$$\mathbb{P}_\lambda\{L = v \mid r, \theta\} = \frac{\mathrm{Pois}_\lambda(d_r(\theta, v))}{\sum_{v' \in \mathcal{V}(r)} \mathrm{Pois}_\lambda(d_r(\theta, v'))}, \qquad v \in \mathcal{V}(r),$$

and $\mathbb{P}_\lambda\{L = v \mid r, \theta\} = 0$ for $v \notin \mathcal{V}(r)$. The normalization reflects the fact that the attacker cannot move beyond the end

of the chosen path, so that all probability mass is assigned to vertices of r.

Effect of a cleaning action. If the defender checks a vertex $e \in \mathcal{K}$ that lies on r, the route is truncated at the cut point. Define the truncated path

$$r|e = \begin{cases} (\theta) & \text{if } e = \theta, \\ (\theta, v_1, \ldots, v_{j-1}) & \text{if } e = v_j \text{ for some } j \geq 1, \\ r & \text{if } e \notin \mathcal{V}(r). \end{cases}$$

A check at the foothold itself removes all progress but, by convention, leaves the attacker at θ, from which it can re-enter. All probability mass at and beyond the cut is removed and the remaining mass is renormalized on $\mathcal{V}(r|e)$, so that the attacker's location after the check has the distribution $U_v(e, r, \theta, \lambda) = \mathbb{P}_\lambda\{L = v \mid r|e, \theta\}$, $v \in \mathcal{V}$.

In Figure 3.18, the path $r = (\theta, v_1, v_2, v_3, \tau)$ has five vertices, at distances $0, 1, 2, 3, 4$ from θ. With $\lambda = 2$, the truncated Poisson weights yield the location distribution (0.143, 0.286, 0.286, 0.190, 0.095) on these vertices, so that the PLC is compromised with probability 0.095 in one inspection period. A check at $e = v_2$ truncates the path to (θ, v_1); renormalizing $\mathrm{Pois}_\lambda(0)$ and $\mathrm{Pois}_\lambda(1)$ gives the distribution $(1/3,\ 2/3)$ on these two vertices and zero mass on the PLC.

For mixed strategies $p \in \Delta(\mathcal{K})$ of the defender and $q_\theta \in \Delta(\mathcal{R}_\theta)$ of an attacker starting at θ, the location distribution of this attacker type is

$$U_v(p, q_\theta, \theta, \lambda) = \sum_{e \in \mathcal{K}} \sum_{r \in \mathcal{R}_\theta} p_e \, q_{\theta,r} \, U_v(e, r, \theta, \lambda).$$

Since the defender does not observe θ, the relevant ex ante distribution is the Bayesian mixture

$$U'_v(p, q, \lambda) = \sum_{\theta \in \Theta} \rho_\theta \, U_v(p, q_\theta, \theta, \lambda),$$

where $q = (q_\theta)_{\theta \in \Theta}$.

From distributions to a game. ▸ Code: `code/cut_the_rope.py`

The payoff is not naturally a scalar: the defender is concerned primarily with the probability mass near the target τ rather than with the attacker's average penetration depth. One way to obtain a normal-form matrix is to choose nonnegative loss weights ω_v, $v \in \mathcal{V}$, that increase as v approaches τ, with ω_τ representing the loss from compromise of the target, and to define the expected security loss $J_D(e, r; \theta) = \sum_{v \in \mathcal{V}} \omega_v \, U_v(e, r, \theta, \lambda)$. For each type θ, this defines a matrix whose rows are indexed by the defender's checks $e \in \mathcal{K}$ and whose columns are indexed by the attacker's paths $r \in \mathcal{R}_\theta$. The Bayesian robust formulation is then

$$\min_{p \in \Delta(\mathcal{K})} \ \max_{\{q_\theta \in \Delta(\mathcal{R}_\theta)\}_{\theta \in \Theta}} \sum_{\theta \in \Theta} \rho_\theta \sum_{e \in \mathcal{K}} \sum_{r \in \mathcal{R}_\theta} p_e \, q_{\theta,r} \, J_D(e, r; \theta).$$

The original formulation instead retains the vector payoff U' and orders distributions lexicographically from the target backward, so that reducing the probability of reaching τ takes precedence over reducing the probability mass at vertices farther from it. This is a multicriteria security-game formulation of the same idea [493].

Equilibrium structure. Because the attacker is stealthy, the defender receives no signal that identifies the type θ. The relevant equilibrium is therefore a pooling perfect Bayesian Nash equilibrium (PBNE; Section 3.2.6), in which the defender acts on the prior ρ over possible attacker locations and each attacker type chooses a path that maximizes the tail risk it induces. In the multicriteria formulation of [493], this pooling equilibrium is shown to be equivalent to a multi-goal security strategy. For the defender, checks that cut many high-risk paths close to the target should receive priority, in particular checks at bottleneck vertices through which every successful intrusion must pass. When such a bottleneck is available, the defender's equilibrium strategy may be pure. In Figure 3.18, however, no single host lies on every path from every foothold. The engineering workstation and the operator HMI together form a cut, since every path to the PLC enters it from one of these two hosts, but a single check can cover only one of them. The defender therefore mixes over several checks, chiefly the engineering workstation, the operator HMI, and the jump host, so that no single attack path reliably avoids inspection.

Attack graphs of this kind also underlie automated penetration testing, which searches for exploitable attack paths and turns them into countermeasures; recent frameworks do so with game-theoretic and neuro-symbolic reasoning [348]

and with language-model agents that couple penetration testing to remediation [261].

3.3.1.2 The FlipIt Game

FlipIt was introduced as a game of stealthy takeover of a shared resource [577]. The resource may be a cryptographic key, a server, an account, a virtual machine, or any other component whose control can pass between an attacker and a defender. At every time $t \geq 0$, exactly one player controls the resource. A move by player i transfers control to player i, but the opponent does not automatically observe the move. This absence of immediate feedback distinguishes FlipIt from standard inspection games. In Chapters 9 and 10, FlipIt serves as the cyber-layer game within larger models of cloud-enabled control, where it is coupled with contracts and with signaling.

Formal model. Index the defender by D and the attacker by A, and let $C(t) \in \{D, A\}$ denote the player who controls the resource at time $t \geq 0$, with $C(0) = D$. Let $t_{i,1} < t_{i,2} < \cdots$ be the move times of player $i \in \{D, A\}$. The controller at time $t > 0$ is the player who made the most recent move before t; that is, $C(t) = i$ if some move time $t_{i,j} < t$ of player i is the latest move time of either player before t, and $C(t) = D$ if no move has been made before t. For a horizon $T > 0$, let $M_i(T) = \#\{j : t_{i,j} \leq T\}$ be the number of moves of player i and $X_i(T) = \int_0^T \mathbf{1}\{C(t) = i\}\, dt$ the total time during which player i controls the resource on $[0, T]$. If each move of player i costs $\kappa_i > 0$, the average benefit rate of player i is

$$v_i(T) = \frac{X_i(T) - \kappa_i M_i(T)}{T} = \chi_i(T) - \kappa_i \alpha_i(T), \qquad \text{where} \quad \chi_i(T) = \frac{X_i(T)}{T}, \quad \alpha_i(T) = \frac{M_i(T)}{T}$$

are the control fraction and the move rate. The asymptotic payoff is $v_i = \liminf_{T\to\infty} \mathbb{E}[v_i(T)]$. Each player thus seeks a large control fraction χ_i at a small move-rate cost $\kappa_i \alpha_i$ [577].

Figure 3.19 shows a sample path. Because neither player observes the other's moves, the defender sometimes moves while it already holds the resource and pays κ_D without effect (hollow markers), and an attacker move can leave the resource compromised for a long interval before the defender's next scheduled move.

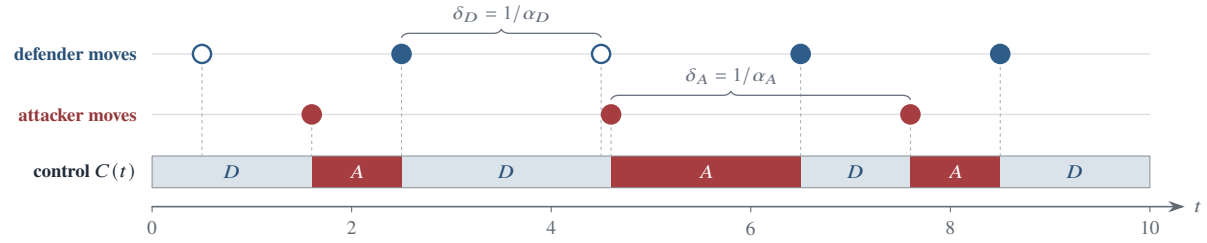


Figure 3.19: A FlipIt sample path with periodic play: the defender moves every $\delta_D = 2$ time units and the attacker every $\delta_A = 3$, each with a random phase. Filled markers are moves that take control; hollow markers are wasted moves made while the mover already controls the resource. The band shows the controller $C(t)$; here the attacker holds the resource for 3.7 of the 10 time units.

A pure strategy in FlipIt is not a single action but a moving policy, which maps the player's information to a waiting time before its next move. Common non-adaptive policies include periodic, renewal, and exponential play. A periodic policy with rate α_i moves every $\delta_i = 1/\alpha_i$ time units, and its first move is drawn uniformly from the first period so that the schedule cannot be predicted exactly.

Periodic strategies and payoff formulas. In the simplest normal-form version of the game, each player chooses the rate of a periodic policy. The payoff rates induced by a pair of periodic policies are as follows.

PROPOSITION 3.3.1 **Periodic FlipIt payoff rates**
Consider periodic policies with rates $\alpha_D, \alpha_A > 0$ and periods $\delta_D = 1/\alpha_D$ and $\delta_A = 1/\alpha_A$. If $\alpha_D \geq \alpha_A$, then

$$v_D(\alpha_D, \alpha_A) = 1 - \frac{\alpha_A}{2\alpha_D} - \kappa_D \alpha_D, \qquad v_A(\alpha_D, \alpha_A) = \frac{\alpha_A}{2\alpha_D} - \kappa_A \alpha_A.$$

If $\alpha_D \le \alpha_A$, then

$$\nu_D(\alpha_D, \alpha_A) = \frac{\alpha_D}{2\alpha_A} - \kappa_D \alpha_D, \qquad \nu_A(\alpha_D, \alpha_A) = 1 - \frac{\alpha_D}{2\alpha_A} - \kappa_A \alpha_A.$$

Proof. Suppose that $\alpha_D \ge \alpha_A$, so that $\delta_D \le \delta_A$. Because the attacker's phase is uniform, the probability that the attacker moves within a given defender interval of length δ_D is $\delta_D/\delta_A = \alpha_A/\alpha_D$. Conditional on such a move, its time is uniformly distributed over the interval, and the attacker controls the resource for an expected time $\delta_D/2$ until the next defender move. The expected control fractions are therefore $\chi_A = (\alpha_A/\alpha_D) \cdot \frac{1}{2} = \alpha_A/(2\alpha_D)$ and $\chi_D = 1 - \alpha_A/(2\alpha_D)$, and substituting them into $\nu_i = \chi_i - \kappa_i \alpha_i$ yields the first pair of expressions. The case $\alpha_D \le \alpha_A$ follows from the same argument with the roles of the two players interchanged. □

The proposition exhibits the central tradeoff of the game: moving faster increases the fraction of time in control, but the marginal gain in control eventually falls below the marginal cost of moving.

Nash equilibrium in the periodic game. ▸ Code: `code/flipit.py`

Solving the best-response conditions of the periodic FlipIt game yields the equilibrium rates [577]

$$(\alpha_D^*, \alpha_A^*) = \begin{cases} \left(\dfrac{1}{2\kappa_A}, \dfrac{\kappa_D}{2\kappa_A^2} \right) & \text{if } \kappa_D < \kappa_A, \\ \left(\dfrac{1}{2\kappa_D}, \dfrac{1}{2\kappa_D} \right) & \text{if } \kappa_D = \kappa_A, \\ \left(\dfrac{\kappa_A}{2\kappa_D^2}, \dfrac{1}{2\kappa_D} \right) & \text{if } \kappa_D > \kappa_A, \end{cases}$$

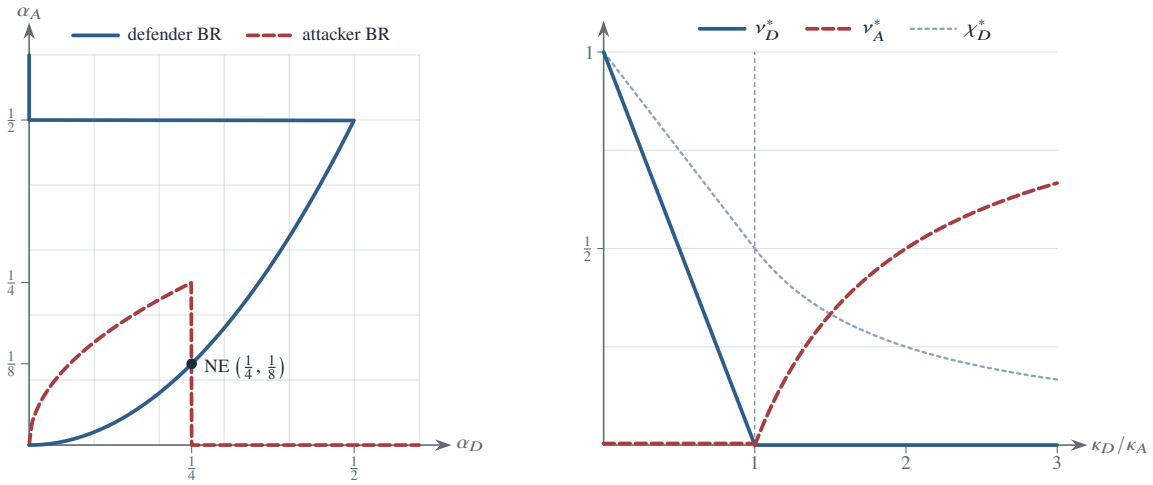
with corresponding equilibrium benefits

$$(\nu_D^*, \nu_A^*) = \begin{cases} \left(1 - \dfrac{\kappa_D}{\kappa_A}, 0 \right) & \text{if } \kappa_D < \kappa_A, \\ (0, 0) & \text{if } \kappa_D = \kappa_A, \\ \left(0,\ 1 - \dfrac{\kappa_A}{\kappa_D} \right) & \text{if } \kappa_D > \kappa_A. \end{cases}$$

The player with the lower move cost therefore obtains a positive surplus, and the equilibrium benefit of the other player is zero. If the defender's recovery actions are much cheaper than the attacker's compromise actions, the defender retains control for a large fraction of time, since $\chi_D^* = 1 - \kappa_D/(2\kappa_A)$ when $\kappa_D < \kappa_A$; if the attacker's moves are cheaper, the roles are reversed.

Figure 3.20(a) illustrates how the equilibrium arises. By the payoff formulas, the defender's best response to an attacker rate $\alpha_A < 1/(2\kappa_D)$ is $\alpha_D = \sqrt{\alpha_A/(2\kappa_D)}$, and the attacker's best response to a defender rate $\alpha_D < 1/(2\kappa_A)$ is $\alpha_A = \sqrt{\alpha_D/(2\kappa_A)}$; a player whose opponent moves faster than these thresholds does best not to move. For $\kappa_D = 1$ and $\kappa_A = 2$, the attacker is indifferent among all rates in $[0, 1/4]$ at $\alpha_D = 1/4$, and its best response is zero for $\alpha_D > 1/4$. The two best-response curves meet at $(\alpha_D^*, \alpha_A^*) = (1/4, 1/8)$, where the defender earns $\nu_D^* = 1/2$ and controls the resource three-quarters of the time. Figure 3.20(b) plots the equilibrium outcome against the cost ratio κ_D/κ_A.

Dominance and design implications. FlipIt also yields two qualitative results of direct relevance to security design [577]. First, among non-adaptive renewal strategies with a given rate of play, periodic play with a randomized phase strongly dominates general renewal play, so that the strategies surviving in the residual renewal game are periodic. Second, if the defender moves sufficiently fast, for example periodically with rate $\alpha_D > 1/(2\kappa_A)$, then every positive move rate yields a negative benefit to a non-adaptive attacker, and the attacker's best response is not to play. This second result must be interpreted with care, because moving too fast can also make the defender's own benefit negative when defensive moves are expensive.



(a) Best responses, $\kappa_D = 1$, $\kappa_A = 2$

(b) Equilibrium outcome

Figure 3.20: Solution of periodic FlipIt. (a) Best responses for $\kappa_D = 1$, $\kappa_A = 2$: the defender's (solid) and the attacker's (dashed) cross at the Nash equilibrium $(1/4, 1/8)$; for $\alpha_D > 1/(2\kappa_A)$ the attacker's best response is not to move. (b) Equilibrium benefits v_D^* (solid) and v_A^* (dashed) and the defender's control fraction χ_D^* (dotted) as functions of κ_D/κ_A. The dotted vertical line marks equal costs, $\kappa_D = \kappa_A$. The cheaper mover earns all the surplus.

Both the cut-the-rope game and FlipIt are driven by stealth. In cut-the-rope, stealth takes the form of hidden attacker progress and unobserved return paths; in FlipIt, it takes the form of delayed discovery of control changes. Both examples show that security games often require dynamic and informational structure beyond a single payoff matrix, while still relying on the normal-form concepts of strategies, payoffs, minimax reasoning, best responses, and equilibrium.

3.3.2 Extensive-Form Security Applications

Extensive-form games are the natural representation of security problems in which timing and observation matter. Consider the security operations center of a regional water utility. Each week it decides where to direct its single team of threat hunters; a ransomware group conducts reconnaissance for several weeks, through phishing probes and by observing how quickly anomalies are investigated, and then chooses which network segment to encrypt. The defender moves first, the attacker moves after observing the long-run frequencies of the defender's choices, and the attacker's action depends on what it has learned. A game tree with the defender at the root and the attacker's decision nodes below it makes this order of moves and this information structure explicit [337, 560]. The model developed below, the Stackelberg security game, is the standard formalization of such interactions; Example 3.12 solves the water-utility instance completely.

3.3.2.1 Stackelberg Security Games

Stackelberg security games model settings in which the defender, as leader, commits to a policy before the attacker, as follower, acts. The commitment may be a patrol distribution, an inspection schedule, a screening policy, or a monitoring rule. The attacker observes enough of this policy to form a response and then chooses where to attack. The model thus captures a central feature of many deployed security systems: a policy benefits the defender not only through the protection it provides directly but also through its effect on the attacker's incentives [560, 315, 543].

Formal model. Let $\mathcal{T} = \{1, \dots, n\}$ be the set of targets and let $\mathcal{S}_D \subseteq \{0,1\}^n$ be the set of the defender's feasible pure allocations, where $s_\tau = 1$ if allocation s covers target τ and $s_\tau = 0$ otherwise. A coverage vector $c = (c_\tau)_{\tau \in \mathcal{T}}$ records the marginal probability c_τ that target τ is covered. In the simplest resource-allocation setting, with n_{res} identical resources each of which covers one target, the feasible set C_D consists of the coverage vectors satisfying $0 \le c_\tau \le 1$ for all $\tau \in \mathcal{T}$ and $\sum_{\tau \in \mathcal{T}} c_\tau \le n_{\text{res}}$. In richer scheduling domains, such as the assignment of air marshals to feasible flight

tours, the feasible coverage vectors are instead those induced by mixtures of feasible pure allocations, so that

$$C_D = \Big\{ c \in [0,1]^n : c = \sum_{s \in \mathcal{S}_D} \pi_D(s)\, s \text{ for some } \pi_D \in \Delta(\mathcal{S}_D) \Big\}.$$

A mixed strategy π_D of the defender thus induces a coverage vector c, and the attacker responds by choosing a target $\tau \in \mathcal{T}$.

Each target τ carries four payoff parameters: if τ is attacked while covered, the defender and the attacker receive $U_D^c(\tau)$ and $U_A^c(\tau)$, respectively; if τ is attacked while uncovered, they receive $U_D^u(\tau)$ and $U_A^u(\tau)$. The standard monotonicity assumptions $U_D^c(\tau) > U_D^u(\tau)$ and $U_A^u(\tau) > U_A^c(\tau)$ for all $\tau \in \mathcal{T}$ state that coverage benefits the defender and harms the attacker. Since the payoffs depend only on whether the attacked target is covered, the expected utilities induced by a coverage vector c are

$$U_D(c,\tau) = c_\tau U_D^c(\tau) + (1 - c_\tau) U_D^u(\tau) \quad \text{and} \quad U_A(c,\tau) = c_\tau U_A^c(\tau) + (1 - c_\tau) U_A^u(\tau).$$

This marginal formulation is the principal reason that many large security games admit a compact representation [321, 543].

Strong Stackelberg equilibrium. The standard solution concept is the *strong Stackelberg equilibrium* (SSE). For a coverage vector c, the attacker's best-response set is $BR_A(c) = \arg\max_{\tau \in \mathcal{T}} U_A(c,\tau)$. The strong version of the equilibrium assumes that an attacker who is indifferent among several best responses selects the one most favorable to the defender, that is, a target $\tau^*(c) \in \arg\max_{\tau \in BR_A(c)} U_D(c,\tau)$, and the defender then commits to a coverage vector $c^* \in \arg\max_{c \in C_D} U_D\big(c, \tau^*(c)\big)$. This defender-favorable tie-breaking convention is widely adopted because the defender can often perturb its commitment slightly so that the desired attacker response becomes strictly optimal. It also guarantees the existence of an SSE in finite Stackelberg games [124, 543].

Target-induction formulation. The SSE can be computed by enumerating the targets that the defender might induce the attacker to choose. For each candidate target, the question is whether it can be made a best response of the attacker and at what utility to the defender.

PROPOSITION 3.3.2 **Target induction in compact security games**
For each target $a \in \mathcal{T}$, let $\hat{V}_a$ denote the optimal value of the problem

$$\begin{aligned} \hat{V}_a = \max_{c \in C_D} \quad & U_D(c,a) \\ \text{s.t.} \quad & U_A(c,a) \geq U_A(c,\tau), \qquad \forall \tau \in \mathcal{T}, \end{aligned} \tag{3.3.1}$$

with $\hat{V}_a = -\infty$ if the problem is infeasible. Under defender-favorable tie-breaking, if $a^* \in \arg\max_{a \in \mathcal{T}} \hat{V}_a$ and c_{a^*} solves (3.3.1) for $a = a^*$, then c_{a^*} is a strong Stackelberg commitment of the defender and $\hat{V}_{a^*}$ is the defender's SSE payoff.

Proof. Fix a target a. The constraints of (3.3.1) state precisely that $a \in BR_A(c)$, and the objective maximizes the defender's utility subject to this requirement. For every feasible c, defender-favorable tie-breaking gives $U_D(c, \tau^*(c)) \geq U_D(c,a)$, so that no candidate value $\hat{V}_a$ exceeds what the defender actually obtains from the corresponding commitment. Conversely, an SSE commitment c^* is feasible for the candidate problem with $a = \tau^*(c^*)$, where its objective value equals the SSE payoff. Hence $\max_{a \in \mathcal{T}} \hat{V}_a$ equals the SSE payoff, and a maximizing candidate solution attains it. □

The same principle underlies algorithmic approaches to larger Stackelberg security games: the defender's optimization problem contains constraints that force the selected target to be a best response of the attacker, while its objective maximizes the defender's induced utility.

Target-enumeration algorithm. ▸ Code: `code/stackelberg_security_game.py`

The preceding proposition yields a direct algorithm for compact Stackelberg security games: for each target a, solve (3.3.1) and retain the best feasible candidate. When C_D is described by linear constraints, (3.3.1) is a linear program, since $U_D(c,\tau)$ and $U_A(c,\tau)$ are affine in the coverage vector c. The procedure is summarized in Algorithm 1.

Algorithm 1 Target enumeration for a compact Stackelberg security game

Require: Target set $\mathcal{T}$, feasible coverage set C_D, utilities $U_D(c,\tau)$ and $U_A(c,\tau)$.
Ensure: Defender coverage c^* and induced attacked target a^*.
1: Set $V^* \leftarrow -\infty$, $c^* \leftarrow \emptyset$, and $a^* \leftarrow \emptyset$.
2: **for all** $a \in \mathcal{T}$ **do**
3: Solve the candidate problem (3.3.1).
4: **if** the candidate problem is feasible and $\hat{V}_a > V^*$ **then**
5: Set $V^* \leftarrow \hat{V}_a$, $c^* \leftarrow c_a$, and $a^* \leftarrow a$, where c_a solves (3.3.1).
6: **end if**
7: **end for**
8: **return** (c^*, a^*).

Algorithm 1 is deliberately elementary. It shows how the Stackelberg problem decomposes into a collection of defender optimization problems, each constrained by the attacker's best-response conditions. Large deployed systems rely on mixed-integer, column-generation, or marginal-coverage methods, but these methods preserve the same leader–follower structure.

EXAMPLE 3.12 **Allocating threat hunting across a water utility**

The water utility of the opening paragraph has $n_{\text{res}} = 1$ threat-hunting team and three targets: the SCADA server (τ_1), whose compromise halts treatment and pumping; the historian (τ_2), which stores process data and is a stepping stone to the control network; and the billing database (τ_3). A pure allocation assigns the team to one target, so $C_D = \{c \in [0,1]^3 : c_1 + c_2 + c_3 \le 1\}$, and c_τ is the fraction of weeks in which target τ is hunted. A hunted target detects the intrusion in time, so the defender loses nothing; otherwise the defender loses the value of the target. The payoff parameters, in units of \$100,000, are listed in Table 3.5.

Table 3.5: Payoff parameters of the water-utility security game.

Target τ	$U_D^c(\tau)$	$U_D^u(\tau)$	$U_A^c(\tau)$	$U_A^u(\tau)$
τ_1: SCADA server	0	−10	−2	10
τ_2: historian	0	−6	−2	6
τ_3: billing database	0	−2	−1	2

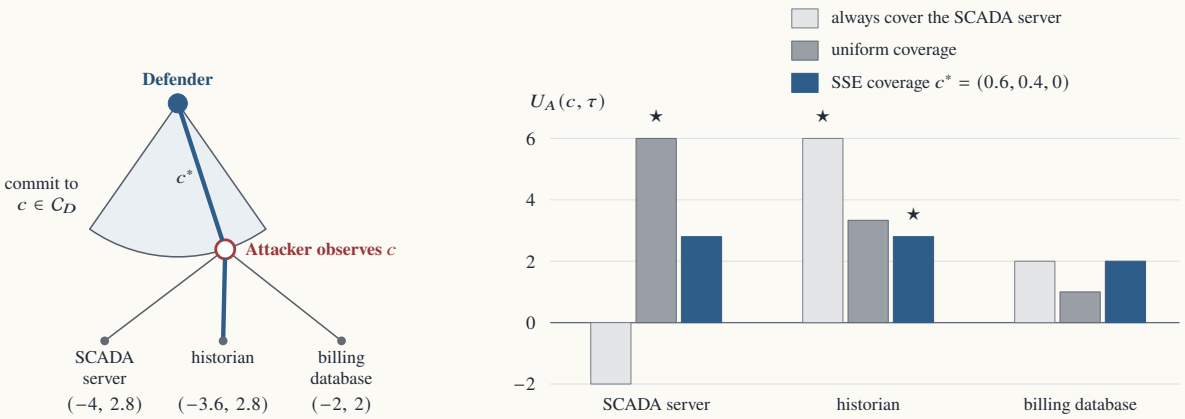
The induced utilities are affine in the coverage vector:

$$U_A(c,\tau_1) = 10 - 12c_1, \quad U_A(c,\tau_2) = 6 - 8c_2, \quad U_A(c,\tau_3) = 2 - 3c_3,$$
$$U_D(c,\tau_1) = -10 + 10c_1, \quad U_D(c,\tau_2) = -6 + 6c_2, \quad U_D(c,\tau_3) = -2 + 2c_3.$$

We solve the three candidate problems (3.3.1).
Inducing τ_2. The defender maximizes $-6 + 6c_2$. The constraint $U_A(c,\tau_2) \ge U_A(c,\tau_1)$ reads $c_1 \ge \frac{1}{3} + \frac{2}{3}c_2$. Since coverage of τ_3 only consumes budget, $c_3 = 0$ may be assumed; then $c_1 + c_2 \le 1$ gives $\frac{1}{3} + \frac{5}{3}c_2 \le 1$, that is, $c_2 \le \frac{2}{5}$. At $c = (0.6, 0.4, 0)$ the remaining constraint holds, because $U_A(c,\tau_2) = 2.8 \ge 2 = U_A(c,\tau_3)$. Hence $\hat{V}_2 = -6 + 6(0.4) = -3.6$.
Inducing τ_1. The defender maximizes $-10 + 10c_1$ subject to $c_1 \le \frac{1}{3} + \frac{2}{3}c_2$. With $c_2 = 1 - c_1$, this gives $c_1 \le \frac{3}{5}$, so the optimum is again $c = (0.6, 0.4, 0)$, with $\hat{V}_1 = -10 + 10(0.6) = -4$.
Inducing τ_3. The constraints $U_A(c,\tau_3) \ge U_A(c,\tau_1)$ and $U_A(c,\tau_3) \ge U_A(c,\tau_2)$ require $c_1 \ge \frac{2}{3} + \frac{1}{4}c_3$ and $c_2 \ge \frac{1}{2} + \frac{3}{8}c_3$, whose sum exceeds the budget. The problem is infeasible and $\hat{V}_3 = -\infty$.
By Proposition 3.3.2, the strong Stackelberg commitment is $c^* = (0.6, 0.4, 0)$: the team hunts on the SCADA server in three weeks out of five and on the historian in the remaining two, and never on the billing database. At c^* the attacker is indifferent between τ_1 and τ_2, each yielding 2.8, and defender-favorable tie-breaking selects τ_2. The defender's expected loss is 3.6, and the attacker's expected gain is 2.8. Shifting an arbitrarily small amount of coverage from τ_2 to τ_1 makes τ_2 the attacker's unique best response at a defender payoff arbitrarily close to −3.6,

which illustrates why the tie-breaking convention is innocuous.

Figure 3.21 compares this commitment with two intuitive policies. Always hunting on the most valuable asset, $c = (1, 0, 0)$, diverts the attacker to the historian, which it then compromises with certainty, and the defender loses 6. Uniform coverage, $c = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, leaves the SCADA server the most attractive target, and the defender loses $10 - \frac{10}{3} \approx 6.67$. The SSE commitment reduces the expected loss by 40% relative to the better of these two policies. It does so not by protecting each asset in proportion to its value, but by lowering the attacker's payoff on the two valuable targets until neither dominates; the billing database is left uncovered because its attacker payoff of 2 already lies below the equalized level 2.8.



(a) Commitment and response **(b) Attacker's payoff by target**

Figure 3.21: The water-utility Stackelberg security game of Example 3.12. (a) The defender commits to a coverage vector c from the continuum C_D (shaded fan); the attacker observes c and chooses a target. Leaf payoffs (U_D, U_A) are evaluated at the SSE commitment $c^* = (0.6, 0.4, 0)$, and the thick path is the SSE outcome. (b) The attacker's expected payoff $U_A(c, \tau)$ for each target under three commitments; a star marks the target attacked. Only the SSE coverage equalizes the two valuable targets and keeps the billing database below that level.

Algorithmic results. Deployed security problems rarely have small action spaces. A pure strategy of the defender may be an entire patrol route, checkpoint schedule, or flight tour, and the number of such strategies can be very large. DOBSS formulates Bayesian Stackelberg games as a mixed-integer linear program and was central to early deployed systems such as ARMOR at Los Angeles International Airport [468, 560]. Later systems exploit additional structure. Incremental strategy generation, as in ASPEN for Federal Air Marshal scheduling, solves a restricted master problem and repeatedly adds useful defender schedules rather than enumerating all of them. Marginal-coverage and polytope methods exploit the fact that utilities depend on coverage probabilities rather than on full schedules, provided that the optimal coverage vector can be decomposed into a mixture of feasible pure allocations [315, 543].

Robustness and behavioral response. The basic SSE assumes that the attacker observes the defender's mixed strategy and responds optimally. Surveys identify three departures from this assumption that matter in practice: uncertainty in payoffs, uncertainty in the defender's execution or the attacker's observation, and bounded rationality of the attacker [315, 543]. If the attacker follows a stochastic response model, such as a quantal-response or subjective-utility model [390], under which target τ is attacked with probability $q_\tau(c)$, the defender's problem becomes

$$\max_{c \in C_D} \sum_{\tau \in \mathcal{T}} q_\tau(c)\, U_D(c, \tau).$$

If the defender knows only that the attacker's response distribution $q = (q_\tau)_{\tau \in \mathcal{T}}$ belongs to an uncertainty set $Q(c) \subseteq \Delta(\mathcal{T})$, a robust formulation is

$$\max_{c \in C_D} \min_{q \in Q(c)} \sum_{\tau \in \mathcal{T}} q_\tau\, U_D(c, \tau).$$

These variants account for the combination of game-theoretic commitment with optimization, learning, and robust decision-making in modern applications of Stackelberg security games. The extensive-form interpretation is unchanged: the defender commits first, the attacker observes the commitment partially or fully, and the attacker's response is induced by the announced policy.

3.4 Discussion, Summary, and Further Reading

This chapter has concentrated on the core representations of noncooperative games and on worked security applications. Each of the directions outlined below extends the modeling language developed here with richer information structures, computational methods, or behavioral assumptions.

Sequential monitoring and response. In cyber operations, a defender frequently receives an alert before choosing a response. The alert may be noisy, and the attacker may know how alerts are generated. An extensive-form model can represent the alert either as a chance move or as a signal generated by the attacker's hidden type. The defender's strategy then specifies a response at each information set, such as ignoring the alert, investigating, isolating the affected host, deceiving the attacker, or escalating. Subgame perfection, and sequential rationality more generally, excludes response plans that would be irrational once a particular alert history has been reached.

Incomplete information and signaling. Security decisions often involve hidden information. An attacker may be advanced or opportunistic, a defender may have strong or weak monitoring capability, and a message may be honest or deceptive. Signaling games capture such situations by endowing one player with a private type and an observable signal. Equilibrium analysis then determines whether the signal credibly reveals information, pools different types, or manipulates the receiver's belief.

Bounded reasoning. Full equilibrium reasoning can be too demanding when players have limited information, computation, or experience. Level-k models represent bounded strategic thinking in a structured way: a level-0 attacker follows a simple default rule, a level-1 defender best responds to that rule, a level-2 attacker best responds to the level-1 defender, and so on. Such models are appropriate when observed behavior deviates systematically from Nash or subgame perfect predictions.

From trees to algorithms. The extensive-form representation also provides the basis for computational methods. Backward induction solves finite games of perfect information. Sequence-form methods exploit the tree structure of larger games with perfect recall [332]. Dynamic programming and stochastic-game methods extend these ideas to settings in which a state evolves over time [528]. For large strategic systems, such as cyber campaigns in which the outcome of one interaction enables, synchronizes, or blocks later ones, game Petri nets attach local games to the places of a colored Petri net; for sequential nets, global equilibria are exactly collections of local equilibria augmented by continuation values, so that values can be computed place by place [647]. The concepts of this chapter thereby connect elementary game models with algorithmic security analysis.

Connections to later chapters. The representations of this chapter are the building blocks of the rest of the book. Chapter 4 develops games of incomplete information, of which the signaling games of Section 3.2.6 are a first instance. Chapters 5 and 6 replace the finite game tree with a state that evolves in discrete or continuous time, and Chapter 7 studies learning dynamics, including fictitious play and no-regret learning (Section 7.1.3), and their relation to the equilibrium concepts introduced here. Chapter 11 generalizes level-k reasoning to other models of bounded rationality, and Chapters 12 and 13 apply multi-stage, Stackelberg, and signaling models to industrial control and cyber deception.

Connections to appendices. Appendix A supplies the linear-programming duality behind the computation of saddle-point equilibria and the candidate problems of the target-enumeration algorithm. Appendix B supplies the probability background used for chance moves, Bayesian belief updating, and the Poisson model of the cut-the-rope game.

Further reading. Classical treatments of normal-form and extensive-form games are given by von Neumann and Morgenstern [586], Osborne and Rubinstein [450], and Fudenberg and Tirole [205]; Başar and Olsder [46] develop the zero-sum, Stackelberg, and dynamic theory in the notation adopted here. The foundational papers on the solution concepts are those of Nash [418], Aumann [29], Kuhn [337], Selten [524], and Kreps and Wilson [336]. For security applications, Alpcan and Başar [12] and Manshaei et al. [380] survey network-security games, Tambe [560] and

Sinha et al. [543] cover Stackelberg security games and their deployments, and the cut-the-rope and FlipIt games originate in [493] and [577], respectively. Lei, Ge, and Zhu [348] develop a game-theoretic and neuro-symbolic framework for distributed adaptive penetration testing of AI-enabled healthcare networks, and Huang and Zhu [261] pair a language-model penetration-testing stage that explores multiple attack paths with a remediation stage that recommends mitigations. Open analytical, systems, and integration challenges for game theory in distributed systems security are set out in [3].

Summary. A normal-form game describes a security interaction by the policies available to each side and the resulting payoffs; an extensive-form game additionally describes the order of moves and what each player knows when it acts. Saddle-point, Nash, and correlated equilibria characterize stable outcomes of simultaneous choice, whereas subgame perfection, Stackelberg equilibrium, and perfect Bayesian equilibrium characterize sequential choice under commitment and private information. The security applications of the chapter show that the choice of representation, and in particular the specification of who observes what, determines which defensive policies are credible and which are exploitable.

3.5 Exercises

The exercises below reinforce the modeling choices, notation, and solution concepts used throughout the chapter. Unless stated otherwise, matrix-game entries are payments from the row player P_1 to the column player P_2, so P_1 minimizes and P_2 maximizes. In general-sum games, payoff pairs are written as (u_1, u_2) and both players maximize. Most exercises can also serve as self-contained exam problems.

Exercise 3.1. Pure security levels in a matrix game. Consider the zero-sum matrix game

$$A = \begin{bmatrix} 2 & -1 & 4 \\ 0 & 1 & 3 \\ -2 & 2 & 1 \end{bmatrix}.$$

Compute the maximum of each row and the minimum of each column. Find $\overline{V} = \min_i \max_j A_{ij}$ and $\underline{V} = \max_j \min_i A_{ij}$. Does the game have a pure-strategy saddle point? Justify your answer using the saddle-point inequalities.

Exercise 3.2. Mixed security strategies by equalization. For the 2×2 game

$$A = \begin{bmatrix} 4 & 1 \\ 0 & 3 \end{bmatrix},$$

let $p = (p_1, 1 - p_1)^\top$ and $q = (q_1, 1 - q_1)^\top$. Derive the two expected-payment lines, as functions of p_1, that P_1 faces when P_2 chooses a pure column, and the two lines, as functions of q_1, that P_2 faces when P_1 chooses a pure row. Use equalization to compute p^*, q^*, and the value of the game. Then write the minimizer's linear program for this game, first shifting the entries so that they are strictly positive.

Exercise 3.3. Best responses in a coordination game. In the Battle of the Sexes example, let p be the probability that P_1 chooses B and q the probability that P_2 chooses B. Derive both best-response correspondences from the expected-payoff functions. Identify all pure Nash equilibria and the mixed Nash equilibrium. Explain why each player is indifferent between the two actions at the mixed equilibrium.

Exercise 3.4. Correlated recommendations in the game of chicken. Use the game of chicken in Example 3.6. Verify that the distribution

$$\varphi(\mathrm{G}, \mathrm{Y}) = \varphi(\mathrm{Y}, \mathrm{G}) = \frac{1}{2}, \qquad \varphi(\mathrm{G}, \mathrm{G}) = \varphi(\mathrm{Y}, \mathrm{Y}) = 0$$

satisfies the correlated-equilibrium inequalities. Compare each player's expected payoff under this correlated equilibrium with the expected payoff under the mixed Nash equilibrium.

Exercise 3.5. **A fictitious-play update.** Use the game in Example 3.4. Suppose that before a new round the empirical counts are $w_1 = (4, 2)$, kept by P_1 over P_2's actions (L, R), and $w_2 = (3, 3)$, kept by P_2 over P_1's actions (U, D). Compute each player's empirical belief and best response, and update the counts after both players play these best responses.

Exercise 3.6. **Strategic form from an extensive-form tree.** Consider the matching-pennies tree in which P_1 moves first and P_2 observes P_1's action. List all pure strategies of P_2 as complete contingent plans. Construct the strategic-form table and identify the best responses of P_2 to each pure action of P_1. Then explain why this game is not strategically equivalent to simultaneous matching pennies.

Exercise 3.7. **Credibility and subgame perfection.** Modify the market-entry game so that the incumbent's payoffs after entry are 0 from fighting and 2 from accommodating, while the entrant receives -1 if fought and 1 if accommodated. If the entrant does not enter, payoffs are $(2, 0)$, as in Example 3.7, with payoffs listed as (P_1, P_2). Write the strategic form, find all Nash equilibria, and then solve the game by backward induction. Which Nash equilibrium is ruled out by subgame perfection, and why?

Exercise 3.8. **Mixed versus behavioral strategies.** Construct a two-stage decision problem for a single player with imperfect recall, in which one information set contains two decision nodes reached through different earlier actions of that player. Describe a mixed strategy whose distribution over terminal histories cannot be induced by any behavioral strategy. Explain how this illustrates the role of perfect recall in Kuhn's theorem.

Exercise 3.9. **A two-target Stackelberg security game.** A defender has one divisible security resource and chooses coverage probabilities c_1, c_2 satisfying $0 \le c_\tau \le 1$ and $c_1 + c_2 \le 1$. The attacker chooses one target. Utilities are

	$\tau = 1$	$\tau = 2$
$U_D^c(\tau)$	4	3
$U_D^u(\tau)$	-8	-5
$U_A^c(\tau)$	-2	-1
$U_A^u(\tau)$	5	4

Write $U_D(c, \tau)$ and $U_A(c, \tau)$ for each target. For each candidate attacked target $a \in \{1, 2\}$, formulate the defender's optimization problem with the constraint that a is an attacker best response. Solve the two candidate problems and identify the strong Stackelberg equilibrium, including the attacked target and the defender's equilibrium payoff.

Exercise 3.10. **Cut-the-rope distribution after cleaning.** Consider one attacker type θ and one path

$$r = (\theta, v_1, v_2, \tau)$$

with attack rate $\lambda = 1$. First compute the normalized probability that the attacker is at each vertex of r before cleaning. Then compute the post-cleaning distributions if the defender checks v_1, checks v_2, or checks a vertex not on r. If the loss weights are $\omega_\theta = 0$, $\omega_{v_1} = 1$, $\omega_{v_2} = 3$, and $\omega_\tau = 10$, compute the expected security loss $J_D(e, r; \theta)$ for each cleaning action.

Exercise 3.11. **Periodic FlipIt rates.** In the periodic FlipIt model, suppose $\kappa_D = 1$ and $\kappa_A = 4$. Use the equilibrium formulas in the chapter to compute α_D^*, α_A^*, ν_D^*, ν_A^*, and the defender's control fraction χ_D^*, and interpret the result. Then repeat the computation for $\kappa_D = 3$ and $\kappa_A = 1$, and explain how a lower move cost shifts the equilibrium advantage.

Exercise 3.12. **Beliefs in a signaling game.** Consider a signaling game with two types, θ_1 and θ_2, prior $\rho(\theta_1) = \rho_1 \in (0, 1)$, and two messages, L and R. Suppose the sender strategy is $\sigma_S(\mathrm{L} \mid \theta_1) = \sigma_1$ and $\sigma_S(\mathrm{L} \mid \theta_2) = \sigma_2$. Derive $\mu(\theta_1|\mathrm{L})$ and $\mu(\theta_1|\mathrm{R})$ using Bayes' rule whenever the denominators are positive. For which

values of σ_1 and σ_2 is one of the messages off path? Explain why Bayes' rule does not determine the belief after an off-path message, and how PBNE nonetheless restricts that belief through the receiver's sequential rationality and the sender's incentive not to deviate.

Part II

Advanced Security Games

Chapter 10

Composition of Games

A security system is rarely a single game. Cyber defense, trust management, feedback control, penetration testing, and mission planning are separate strategic problems, yet the outcome of each enters the primitives of the others: a compromise probability becomes a prior, a trust decision becomes the information of a controller, and a physical loss becomes the value of a cyber asset. The same system may also be described at several resolutions, from individual tactics to campaign-level plans. A model that solves one of these games with the others held fixed can recommend a defense that the other games undermine: a vehicle that trusts a cloud service because compromise is rare makes compromise valuable to the attacker, who then raises its probability.

This chapter studies compositions of games, in which component games are solved under contexts generated by one another, and develops Gestalt equilibrium as the fixed-point condition that makes the components mutually consistent. Nash equilibrium is the special case of single-player components; existence follows from Kakutani's theorem, and uniqueness and the convergence of iterative computation from a contraction condition. The theory is applied to cross-layer resilient control, in which a cyber stochastic game and a physical differential game exchange transition rates and values; impact-aware signaling for autonomous agents, in which a signaling game is coupled with min–max model predictive control; strategic trust in cloud-enabled control, in which the iSTRICT architecture couples FlipIt, signaling, and control layers; and multi-resolution cyber campaigns, in which zoom operations link tactical extensive-form games with a strategic Markov game.

The components are solved with the methods of Chapters 3–6; the chapter adds the consistency condition that binds them. The general results are proved, and the linear-quadratic specialization and the iSTRICT results are sketched with references to the original papers. The compositions recur in Chapters 12–16, in industrial control, deception, the Internet of Things, supply chains, and wireless networks.

Chapter Overview

1. Composed games in security (Section 10.1)
   - interdependence through coupled primitives and through resolution;
   - the general architecture of component games, contexts, and coupling.
2. Foundations of composed games (Section 10.2)
   - component games, contexts, outcomes, solution correspondences, and coupling maps;
   - Gestalt equilibrium as a fixed point, its existence, uniqueness, and computation by iteration.
3. Games-in-games for cross-layer resilient control (Section 10.3)
   - the games-in-games principle: a cyber stochastic game coupled with a physical differential game;
   - the linear-quadratic specialization with coupled Riccati equations.
4. Cyber-physical signaling for autonomous agents (Section 10.4)
   - impact-aware signaling coupled with min–max model predictive control for autonomous agents;
   - danger zones and separating and pooling equilibria.

5. iSTRICT for cloud-enabled control (Section 10.5)
   - iSTRICT: FlipIt games, a signaling game, and a control layer for cloud-enabled control;
   - Gestalt Nash equilibrium, a worst-case compromise bound, and computation by iteration.
6. Multi-resolution games for cyber campaigns (Section 10.6)
   - zoom-in and zoom-out between tactical extensive-form games and a strategic Markov game.
7. Discussion, summary, further reading, and exercises (Sections 10.7 and 10.8)
   - connections to agentic and multi-agent defense, and to the other chapters of the book;
   - practice problems on cross-layer values, trust thresholds, fixed points, and zoom operations.

Chapter Summary

No component of a composed game can be solved correctly in isolation, because its parameters are produced by the others. A Gestalt equilibrium is a joint outcome that reproduces its own context. Its existence requires convex outcome spaces, so consistency may hold only in mixed outcomes even when every component has a pure solution, and alternating pure best responses can cycle; when the composed map is a contraction, the equilibrium is unique and the layer-by-layer iteration converges geometrically. In the games-in-games principle, cyber policies enter the physical game as transition rates and physical values enter the cyber game as stage costs; the equilibrium is characterized by coupled Hamilton–Jacobi–Isaacs and Shapley equations, which become coupled Riccati equations in the linear-quadratic case. In impact-aware signaling, a receiver rejects exactly the references inside a danger zone whose radius grows with its suspicion, so that in both the separating and the pooling equilibria no accepted reference brings an agent closer to a detected object. In iSTRICT, attacker-controlled services imitate defender-controlled ones, the selected equilibria pool on low-risk messages, and the device's trust threshold bounds the compromise probability whatever the cost of attacking. In multi-resolution games, zoom-in assigns strategic values to tactical outcomes and zoom-out turns tactical solutions into macro transition tendencies, so that refining a few components can correct campaign-level values. Each worked example compares the composed solution with the prediction of a model that ignores the coupling, and in each the difference is material.

Notation and Scope

Book conventions. The chapter uses the common notation of the book, introduced in Chapter 3 and listed in Appendix F. A component game $\Gamma_r(\xi_r)$ has context ξ_r, outcome $o_r \in O_r$, solution correspondence $\mathcal{S}_r$, and coupling map C_r; a Gestalt equilibrium is a fixed point $o^\star \in \mathcal{F}(o^\star)$, and ℓ is a defender action. In the games-in-games model, x, u, and w are the physical state, control, and disturbance, $i \in S$ is a cyber mode with rates λ_{ij}, and Z_i is the Riccati matrix of mode i, with scalar analogue $Z(\kappa, p)$. In the signaling and iSTRICT models, ρ is a reference, $d(\cdot,\cdot)$ a distance, and p_A^i the compromise probability of service $i \in \mathcal{I}$. In the multi-resolution game $\Gamma_{\mathcal{G}}$ on the components $\mathcal{V}$, Γ^s is a micro game with outcome distribution τ^s.
Notation specific to this chapter: $\mathcal{S}_r$ denotes a solution correspondence, and the set of cyber modes is written S; γ is the attenuation level of the physical game, and the cyber game is discounted by β; optimal and equilibrium quantities carry $\star$ rather than $*$; physical states are x; mixed and behavioral strategies are written σ, F, and G, and π is reserved for stationary Markov strategies; utilities are written U; the index i denotes a cyber mode in Section 10.3, an agent in Section 10.4, and a cloud service in Section 10.5; and the players of the multi-resolution game are P_A and P_D.

The scope is architectural. The chapter does not introduce new solution concepts for the component games; it uses the Nash, saddle-point, perfect Bayesian, and Markov-perfect concepts of Chapters 3–6 inside the components and studies the consistency condition that binds them. The general results are proved, and the linear-quadratic and iSTRICT results of the original papers are sketched. Appendix C reviews the linear-systems, Riccati, and model-predictive-control background used by the physical components.

Learning Objectives

After reading this chapter, the reader should be able to:

(1) identify, in a security architecture, the component games, their outcomes, and the variables that couple them;

(2) formulate a Gestalt equilibrium as a fixed point of coupled solution correspondences and verify existence through Kakutani's theorem or uniqueness through a contraction argument;

(3) explain why a composed game may have no Gestalt equilibrium in pure outcomes and how mixed outcomes restore existence;

(4) derive the coupled Hamilton–Jacobi–Isaacs and Shapley equations of the games-in-games principle and solve a scalar instance;

(5) characterize the impact-aware danger zone of a receiver and verify separating and pooling perfect Bayesian equilibria of a cyber-physical signaling game;

(6) compute the fixed point of an iSTRICT-type trust architecture and bound its compromise probability;

(7) apply zoom-in and zoom-out operations between tactical extensive-form games and a strategic Markov game, and quantify what the tactical detail changes.

10.1 Introduction

A modern cyber-physical system involves several interacting decision processes. A defender allocates monitoring resources, decides whether to trust data supplied by a cloud service, tunes a feedback controller, schedules penetration tests, and plans recovery at the mission level. An attacker chooses reconnaissance actions, exploit paths, lateral movements, and physical-impact objectives. Each of these processes can be modeled as a game with the methods of the preceding chapters, but the resulting games are not independent [205, 12]. The probability that a cloud service is compromised, which is the outcome of a cyber game, is the prior belief of a trust game; the trust decision determines the information available to a controller; and the physical performance of the controller determines how valuable the cloud service is to the attacker and to the defender. The parameters of each game are produced by the outcomes of the others.

A second source of interdependence is resolution. The same system can be described at several scales. A high-resolution model describes host-level privilege escalation or individual control actions; a medium-resolution model describes service availability or trust in cloud data; a low-resolution model describes campaign progress or mission success over a long horizon. These models are not competing descriptions of the system. They are strategic views of the same system at different resolutions, and decisions taken at one resolution change the quantities that the others take as given.

This chapter studies *compositions of games*, also called *meta-games*: architectures whose components are themselves games, control problems, or belief updates, and whose components exchange outcomes. Figure 10.1 shows the general structure. Each component game is solved under a *context*—a prior, a transition law, a value, or a cost—that is computed from the outcomes of all components. The solution concept for such an architecture is a consistency condition: every component returns an outcome that is admissible under its own solution concept, and the joint outcome regenerates the contexts under which the components were solved. Following Pawlick, Chen, and Zhu [474], this condition is called a *Gestalt equilibrium*.

The need for the composed viewpoint appears in each major application of the book. In an industrial control system, a host-level game describes privilege escalation, a network-level game describes lateral movement from the enterprise network into the operational network, a process-level game describes the manipulation of controller logic, and a system-level game describes restoration; a defensive action is valuable only through its effect on process safety, which is computed by a different game. In penetration testing, detailed games at individual hosts determine the probability that an exploit succeeds, while a campaign-level game decides which subnet or credential path to pursue; the campaign assigns values to local outcomes, and the local outcomes supply transition probabilities to the campaign [214]. In both cases, an equilibrium of one component computed under an assumed context can be inconsistent with the context that the other components actually produce.

The chapter is organized as follows. Section 10.2 defines component games, coupling maps, and Gestalt equilibrium, establishes existence and uniqueness, and solves a cloud-assisted vehicle example. Section 10.3 develops the games-in-games principle, in which a cyber stochastic game and a physical differential game exchange transition rates and values.

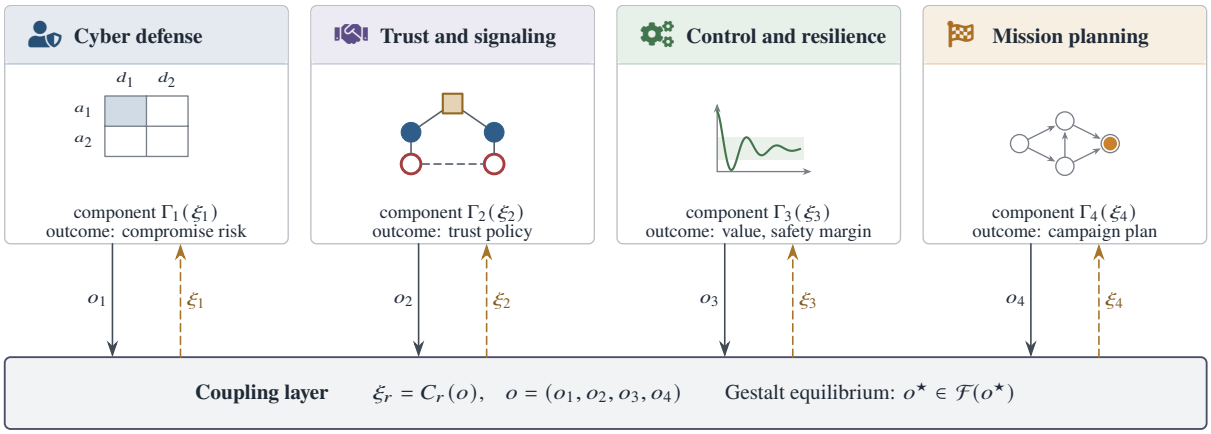


Figure 10.1: A composition of games for cyber-physical security. The upper row shows four component games, each drawn with a glyph of its model class: a cyber-defense matrix game, a trust and signaling game in extensive form, a control and resilience problem whose trajectory must stay in a safe band, and a mission-planning game on a campaign graph. Each component sends its outcome o_r down to the coupling layer (solid arrows) and receives its context $\xi_r = C_r(o)$, computed from the joint outcome o (dashed arrows). A Gestalt equilibrium is a joint outcome $o^\star$ that each component reproduces under the context generated by $o^\star$ itself.

Section 10.4 couples a signaling game with min–max model predictive control for autonomous agents. Section 10.5 presents the iSTRICT architecture for cloud-enabled control. Section 10.6 develops multi-resolution games that link tactical extensive-form games with a strategic Markov game. Section 10.7 discusses extensions and connections, and Section 10.8 contains the exercises.

10.2 Equilibrium Foundations for Composed Games

This section formalizes the architecture of Figure 10.1. A composed game is specified by its component games, the outcome that each component returns, the solution concept that each component uses, and the coupling maps that convert the joint outcome into the contexts of the components. Gestalt equilibrium is then a fixed point of the resulting correspondence, and its existence rests on the same fixed-point arguments that underlie the existence of Nash equilibrium [418, 307, 205].

10.2.1 Component Games, Contexts, and Coupling

Definition 10.2.1 (Context-dependent component game)**.** Let $K \geq 2$ and $r \in \{1, \dots, K\}$. A *context-dependent component game* is a family

$$\Gamma_r(\xi_r) \triangleq \Big(\mathcal{N}_r,\ \{\mathcal{A}_i^r\}_{i \in \mathcal{N}_r},\ \{J_i^r(\cdot\,; \xi_r)\}_{i \in \mathcal{N}_r} \Big), \qquad \xi_r \in \Xi_r,$$

where $\mathcal{N}_r$ is a finite set of players, $\mathcal{A}_i^r$ is the strategy set of player $i \in \mathcal{N}_r$, $J_i^r(\cdot\,; \xi_r)$ is the cost of player i, and ξ_r is a *context* taking values in a set Ξ_r. The context collects the primitives that are not determined inside the component: a prior belief, a transition law, a value of control, a compromise probability, or an information structure.

A component is solved by a solution concept chosen for that component. To allow different concepts in different components, the solution is recorded in an outcome space.

Definition 10.2.2 (Outcome and solution correspondence)**.** An *outcome space* of Γ_r is a set O_r whose elements describe solutions of Γ_r: a mixed-strategy profile, a saddle-point policy pair, a value function, an equilibrium payoff vector, or a statistic such as an equilibrium compromise probability. The *solution correspondence* $\mathcal{S}_r : \Xi_r \rightrightarrows O_r$ assigns to each context the set of outcomes that the chosen solution concept admits in $\Gamma_r(\xi_r)$.

Depending on the component, $\mathcal{S}_r$ may be a Nash-equilibrium correspondence, a saddle-point correspondence, a perfect-Bayesian-equilibrium correspondence, the solution set of a Riccati or Hamilton–Jacobi–Isaacs equation, or a prescribed estimator. A correspondence rather than a function is used because the components need not have unique solutions; a composed game may therefore have to select among several local outcomes while maintaining global consistency.

Definition 10.2.3 (Composed game)**.** A *composed game* (or *meta-game*) is a tuple $\big(\{\Gamma_r, O_r, \mathcal{S}_r, C_r\}_{r=1}^{K}\big)$, where each $C_r : O \to \Xi_r$ is a *coupling map* defined on the joint outcome space $O \triangleq O_1 \times \cdots \times O_K$. The *composed correspondence* $\mathcal{F} : O \rightrightarrows O$ is

$$\mathcal{F}(o) \triangleq \mathcal{S}_1\big(C_1(o)\big) \times \cdots \times \mathcal{S}_K\big(C_K(o)\big), \qquad o = (o_1, \ldots, o_K) \in O. \tag{10.2.1}$$

The correspondence (10.2.1) describes one round of the architecture: the joint outcome o generates the contexts $\xi_r = C_r(o)$, and each component is solved under its context. The coupling maps are the only channel through which the components interact.

Definition 10.2.4 (Gestalt equilibrium [474])**.** A joint outcome $o^\star \in O$ is a *Gestalt equilibrium* of the composed game if

$$o^\star \in \mathcal{F}(o^\star), \qquad \text{that is,} \qquad o_r^\star \in \mathcal{S}_r\big(C_r(o^\star)\big) \quad \text{for every } r \in \{1, \ldots, K\}. \tag{10.2.2}$$

Definition 10.2.4 requires more than the existence of a solution of each component. It requires that every component be solved under the context generated by the equilibrium outcomes themselves. For $K = 2$, Figure 10.2 shows the loop: the joint outcome (o_1, o_2) passes through C_1 and C_2, the two components are solved, and a Gestalt equilibrium is a pair that the loop returns unchanged.

A Gestalt equilibrium is not, in general, a Nash equilibrium of one larger strategic-form game, because the components may use different solution concepts and may share players. The converse inclusion does hold: Nash equilibrium is the special case in which every component is a single-player decision problem.

PROPOSITION 10.2.5 **Nash equilibrium as a Gestalt equilibrium**
Let $\Gamma = (\mathcal{N}, \{\mathcal{A}_i\}_{i\in\mathcal{N}}, \{J_i\}_{i\in\mathcal{N}})$ be a game with $\mathcal{N} = \{1, \ldots, K\}$. For each $r \in \mathcal{N}$, let $\Gamma_r(\xi_r)$ be the decision problem of player r with context $\xi_r = a_{-r} \in \Xi_r \triangleq \prod_{j\neq r} \mathcal{A}_j$, let $O_r = \mathcal{A}_r$, let $\mathcal{S}_r(a_{-r}) = \arg\min_{a_r \in \mathcal{A}_r} J_r(a_r, a_{-r})$, and let $C_r(a) = a_{-r}$. Then the Gestalt equilibria of this composed game are exactly the Nash equilibria of Γ.

Proof. By (10.2.2), $a^\star$ is a Gestalt equilibrium if and only if $a_r^\star \in \arg\min_{a_r \in \mathcal{A}_r} J_r(a_r, a_{-r}^\star)$ for every $r \in \mathcal{N}$, which is the definition of a Nash equilibrium of Γ. □

Cyber-physical interpretation. In the two-component case of greatest interest, Γ_1 is a cyber game Γ_C and Γ_2 is a physical control game Γ_P. The cyber-to-physical coupling $C_P(o)$ supplies the cyber mode, the transition law among modes, and the information available to the controller; the physical-to-cyber coupling $C_C(o)$ supplies the control value, the physical loss, and the recovery value of each mode. A cyber game solved without C_C may protect the wrong modes, and a controller designed without C_P may be robust against the wrong transition law. Section 10.3 develops this construction in full.

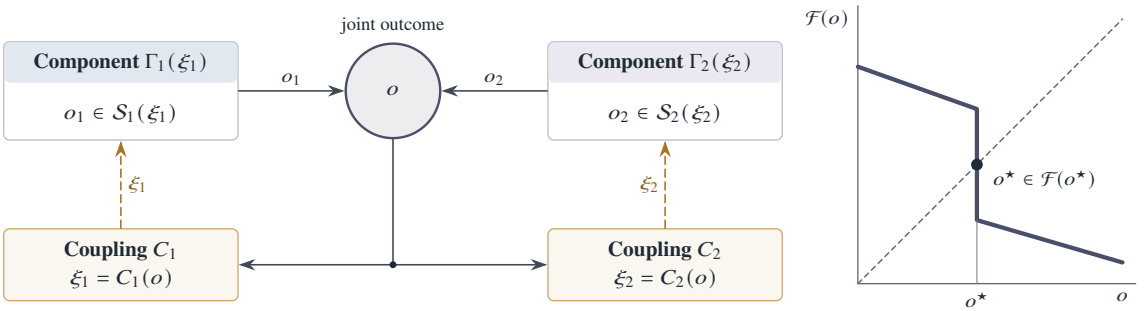
10.2.2 A Worked Composition: Cloud-Assisted Vehicle Trust

The following example contains the smallest nontrivial composition: two games coupled by a compromise probability in one direction and a value of compromise in the other.

EXAMPLE 10.1 **Cloud-assisted vehicle trust and control**

▸ Code: `code/vehicle_trust_metagame.py`

An autonomous shuttle on a logistics campus receives from a cloud traffic service a message that the lane ahead is clear. The message is produced by a cloud server that an attacker A may compromise and that an operator D defends. Figure 10.3 shows the three components of the architecture.



(a) Context–solution loop (b) Fixed point

Figure 10.2: The Gestalt equilibrium of a two-component composed game. (a) The context–solution loop: the components send their outcomes to the joint outcome $o = (o_1, o_2)$ (solid arrows), the coupling maps C_1 and C_2 convert o into the contexts ξ_1 and ξ_2 (dashed arrows), and each component is solved through its correspondence $\mathcal{S}_r$. (b) A composed correspondence $\mathcal{F}$ with a set-valued jump, plotted against the diagonal (dashed); the disc marks the fixed point $o^\star \in \mathcal{F}(o^\star)$, which lies on the vertical segment of $\mathcal{F}$. A Gestalt equilibrium is a joint outcome that the loop reproduces.

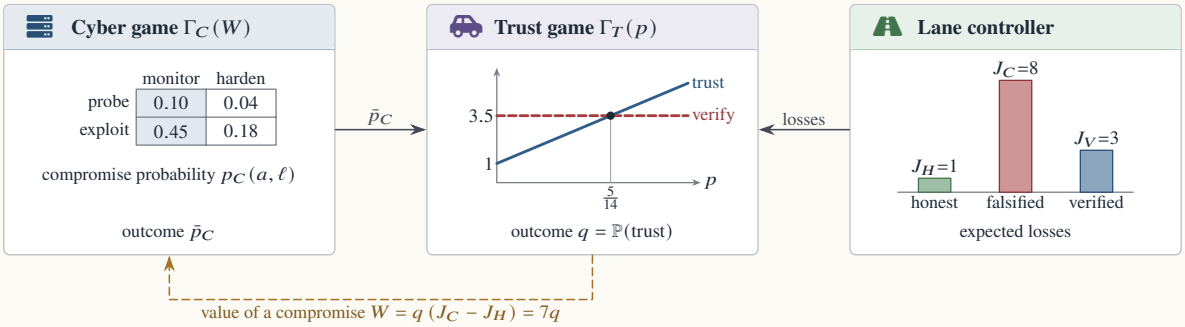


Figure 10.3: The composed game of Example 10.1. The cyber game Γ_C between attacker and operator, shown with its compromise probabilities $p_C(a, \ell)$, returns the equilibrium compromise probability $\bar{p}_C$ to the trust game Γ_T of the vehicle. The trust game compares the loss of trusting, $1 + 7p$, with the loss of verifying, 3.5; the two lines cross at $p = 5/14$. The lane controller supplies the losses $J_H = 1$, $J_C = 8$, and $J_V = 3$ of honest, falsified, and verified operation. The trust probability q returns to the cyber game as the value of a compromise $W = q(J_C - J_H) = 7q$ (dashed arrow).

Physical data. The lane controller of the shuttle incurs the expected loss $J_H = 1$ when it follows an honest message, $J_C = 8$ when it follows a falsified message, and $J_V = 3$ when it verifies the message with its on-board sensors and drives at reduced speed; verification costs an additional $c_V = 0.5$. These numbers are computed once from the lane-control model and do not depend on the strategic components.

Trust component $\Gamma_T(p)$. The context is the compromise probability $p \in [0, 1]$. The vehicle trusts the message with probability $q \in [0, 1]$ and otherwise verifies. Its expected loss is $q\,J_{\text{tr}}(p) + (1 - q)J_{\text{ve}}$, where $J_{\text{tr}}(p) \triangleq (1 - p)J_H + pJ_C = 1 + 7p$ and $J_{\text{ve}} \triangleq J_V + c_V = 3.5$. The two losses are equal at $p = 5/14$, so the solution correspondence is

$$\mathcal{S}_T(p) = \{1\} \text{ if } p < \tfrac{5}{14}, \qquad \mathcal{S}_T(p) = [0, 1] \text{ if } p = \tfrac{5}{14}, \qquad \mathcal{S}_T(p) = \{0\} \text{ if } p > \tfrac{5}{14}.$$

Cyber component $\Gamma_C(W)$. The context is the value $W \geq 0$ of a compromise, that is, the expected additional physical loss that a compromise causes. The attacker chooses $a \in \{\text{probe}, \text{exploit}\}$, the operator chooses $\ell \in \{\text{monitor}, \text{harden}\}$, and the compromise probability $p_C(a, \ell)$ is given in Table 10.1. The attacker's payoff is

$Wp_C(a,\ell) - c_x\mathbf{1}\{a = \text{exploit}\}$ and the operator's loss is $Wp_C(a,\ell) + c_h\mathbf{1}\{\ell = \text{harden}\}$, with $c_x = 1$ and $c_h = 2$. The outcome of $\Gamma_C(W)$ is the compromise probability $\bar{p}_C$ induced by a mixed Nash equilibrium, and $\mathcal{S}_C(W)$ is the set of such probabilities.

	Operator ℓ	
Attacker a	**monitor**	**harden**
probe	0.10	0.04
exploit	0.45	0.18

Table 10.1: Compromise probability $p_C(a,\ell)$ of the cloud server in Example 10.1.

Coupling. A compromise causes the additional loss $J_C - J_H = 7$ only when the falsified message is trusted. The coupling maps are therefore $C_T(\bar{p}_C, q) = \bar{p}_C$ and $C_C(\bar{p}_C, q) = W = 7q$, so that $W \in [0,7]$.
Solution of the cyber component. For $W \in [0,7]$, hardening increases the operator's loss by $2 - 0.06W > 0$ against probing and by $2 - 0.27W > 0$ against exploitation, so monitoring strictly dominates hardening. Against monitoring, exploitation yields the attacker $0.45W - 1$ and probing yields $0.10W$; exploitation is strictly better if and only if $W > 20/7$. Let g be the probability of exploitation. Then $\bar{p}_C = 0.10 + 0.35g$, and

$$\mathcal{S}_C(W) = \{0.10\} \text{ if } W < \tfrac{20}{7}, \qquad \mathcal{S}_C(W) = [0.10,\, 0.45] \text{ if } W = \tfrac{20}{7}, \qquad \mathcal{S}_C(W) = \{0.45\} \text{ if } W > \tfrac{20}{7}.$$

Gestalt equilibrium. A Gestalt equilibrium is a pair $(\bar{p}_C^\star, q^\star)$ with $q^\star \in \mathcal{S}_T(\bar{p}_C^\star)$ and $\bar{p}_C^\star \in \mathcal{S}_C(7q^\star)$. If $q^\star < 20/49$, then $W < 20/7$, so $\bar{p}_C^\star = 0.10 < 5/14$ and $q^\star = 1$, a contradiction. If $q^\star > 20/49$, then $\bar{p}_C^\star = 0.45 > 5/14$ and $q^\star = 0$, again a contradiction. Hence $q^\star = 20/49 \in (0,1)$, which requires $\bar{p}_C^\star = 5/14$, and the relation $0.10 + 0.35g = 5/14$ gives $g^\star = 36/49$. The unique Gestalt equilibrium is

$$q^\star = \tfrac{20}{49} \approx 0.408, \qquad \bar{p}_C^\star = \tfrac{5}{14} \approx 0.357, \qquad g^\star = \tfrac{36}{49} \approx 0.735, \qquad \ell^\star = \text{monitor}.$$

Figure 10.4 displays the two correspondences and their unique intersection. At the equilibrium the vehicle's expected loss is $J_{\text{tr}}(5/14) = J_{\text{ve}} = 3.5$.

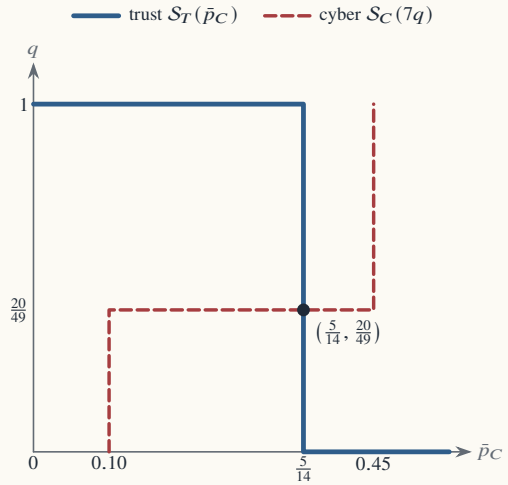


Figure 10.4: The two solution correspondences of Example 10.1 in the plane of compromise probability $\bar{p}_C$ and trust probability q. The solid line is the trust correspondence $\mathcal{S}_T(\bar{p}_C)$; the dashed line is the cyber correspondence $\mathcal{S}_C(7q)$, drawn as a function of q. The two correspondences intersect only at the Gestalt equilibrium $(5/14,\ 20/49)$, marked by the disc. Neither pure trust nor pure verification is consistent.

Observation. A siloed analysis solves the cyber game without the physical layer, which amounts to $W = 0$, and predicts $\bar{p}_C = 0.10$; the vehicle then trusts the cloud and expects the loss $1 + 7(0.10) = 1.7$. Once the vehicle

trusts, however, the value of a compromise becomes $W = 7$, the attacker exploits, $\bar{p}_C = 0.45$, and the realized loss is $1 + 7(0.45) = 4.15$. The composed analysis predicts the consistent loss 3.5. Consistency is achieved only in mixed outcomes: the vehicle verifies with probability $29/49$, and this randomization is precisely what holds the attacker at indifference.

10.2.3 Existence of Gestalt Equilibrium

The example shows that a Gestalt equilibrium may fail to exist in pure outcomes, because the trust correspondence jumps from 1 to 0 at the threshold. Convexity of the outcome spaces restores existence, as in Nash's theorem.

Assumption 10.2.6 (Regular composed game)**.** For every $r \in \{1, \dots, K\}$, the outcome space O_r is a nonempty, compact, convex subset of a finite-dimensional Euclidean space; the coupling map $C_r : O \to \Xi_r$ is continuous, where Ξ_r is a metric space; and the solution correspondence $\mathcal{S}_r$ has nonempty, convex values and a closed graph in $\Xi_r \times O_r$.

THEOREM 10.2.7 **Existence of Gestalt equilibrium**
Under Assumption 10.2.6, the composed game has at least one Gestalt equilibrium.

Proof. The set O is a nonempty, compact, convex subset of a Euclidean space, as a finite product of such sets. For each $o \in O$, $\mathcal{F}(o)$ is nonempty and convex, as a product of nonempty convex sets. To show that $\mathcal{F}$ has a closed graph, let $o^n \to o$ and $\tilde{o}^n \to \tilde{o}$ with $\tilde{o}^n \in \mathcal{F}(o^n)$. Then $\tilde{o}^n_r \in \mathcal{S}_r(C_r(o^n))$ for each r. By continuity, $C_r(o^n) \to C_r(o)$, and since the graph of $\mathcal{S}_r$ is closed, $\tilde{o}_r \in \mathcal{S}_r(C_r(o))$. Hence $\tilde{o} \in \mathcal{F}(o)$. The values of $\mathcal{F}$ are closed subsets of the compact set O, hence compact, and a correspondence with a closed graph into a compact set is upper hemicontinuous. Kakutani's fixed-point theorem [307] therefore yields $o^\star \in O$ with $o^\star \in \mathcal{F}(o^\star)$. □

Remark 10.2.8 (Role of the outcome representation). The convexity in Assumption 10.2.6 is a property of the chosen outcome representation, not of the system. It holds when O_r is a simplex of mixed strategies or beliefs, or a compact convex set of policy or value parameters; it fails for pure strategies, discrete trust decisions, and finite attack paths. In Example 10.1, the pure outcomes $q \in \{0, 1\}$ admit no Gestalt equilibrium, while the mixed outcomes $q \in [0, 1]$ admit exactly one. When the components have nonconvex outcome spaces, existence must be established by a problem-specific argument or assumed.

10.2.4 Uniqueness and Computation by Iteration

A Gestalt equilibrium is computed in practice by solving the components in turn under the current contexts. When the components have unique solutions, the composed correspondence reduces to a map, and Banach's fixed-point theorem gives a sufficient condition for uniqueness and for the convergence of this procedure.

PROPOSITION 10.2.9 **Contraction condition**
Suppose that each $\mathcal{S}_r$ is single-valued, $\mathcal{S}_r(\xi_r) = \{\hat{o}_r(\xi_r)\}$, and that the map $\Phi \triangleq (\hat{o}_1 \circ C_1, \dots, \hat{o}_K \circ C_K) : O \to O$ is a contraction on the complete metric space $(O, \|\cdot\|)$, that is, $\|\Phi(o) - \Phi(o')\| \le \chi \|o - o'\|$ for some $\chi \in [0, 1)$ and all $o, o' \in O$. Then the composed game has a unique Gestalt equilibrium $o^\star$, and the iteration $o^{n+1} = \Phi(o^n)$ converges to $o^\star$ from every $o^0 \in O$, with $\|o^n - o^\star\| \le \chi^n \|o^1 - o^0\|/(1 - \chi)$.

Proof. Under the hypotheses, $\mathcal{F}(o) = \{\Phi(o)\}$, so the Gestalt equilibria are exactly the fixed points of Φ. Banach's fixed-point theorem gives a unique fixed point $o^\star$ and the convergence of the iteration. For the bound, $\|o^{n+1} - o^n\| \le \chi^n \|o^1 - o^0\|$, and the triangle inequality gives $\|o^n - o^m\| \le \chi^n \|o^1 - o^0\|/(1 - \chi)$ for all $m > n$; letting $m \to \infty$ yields the stated estimate. □

The contraction condition is restrictive, but it is verifiable in low-dimensional architectures, as the iSTRICT example of Section 10.5 illustrates. Without it, the iteration may cycle; in Example 10.1, alternating pure best responses produce the cycle $q = 1 \Rightarrow \bar{p}_C = 0.45 \Rightarrow q = 0 \Rightarrow \bar{p}_C = 0.10 \Rightarrow q = 1$, and the mixed equilibrium is found only by intersecting the correspondences.

10.3 Games-in-Games for Cross-Layer Resilient Control

The games-in-games principle of Zhu and Başar [660] is the canonical two-component composition for cyber-physical resilience. A cyber-physical control system has a physical state that evolves continuously and a cyber condition that changes among normal, degraded, compromised, and recovered modes. A physical control game computes, for each cyber mode, a robust performance value, following the minimax-control tradition of dynamic games [294, 46, 43]. A cyber stochastic game computes defense and attack policies and a discounted game value [528, 198]. Neither game is well posed alone: the physical game requires the rates at which the cyber modes change, which are set by the cyber policies, and the cyber game requires the cost of each mode, which is set by the physical value. In the notation of Section 10.2, the cyber outcome consists of the cyber policies and game value, the physical outcome consists of the control policies and mode-dependent value, and the games-in-games equilibrium is the Gestalt equilibrium of the two components. Figure 10.5 shows the two directions of coupling.

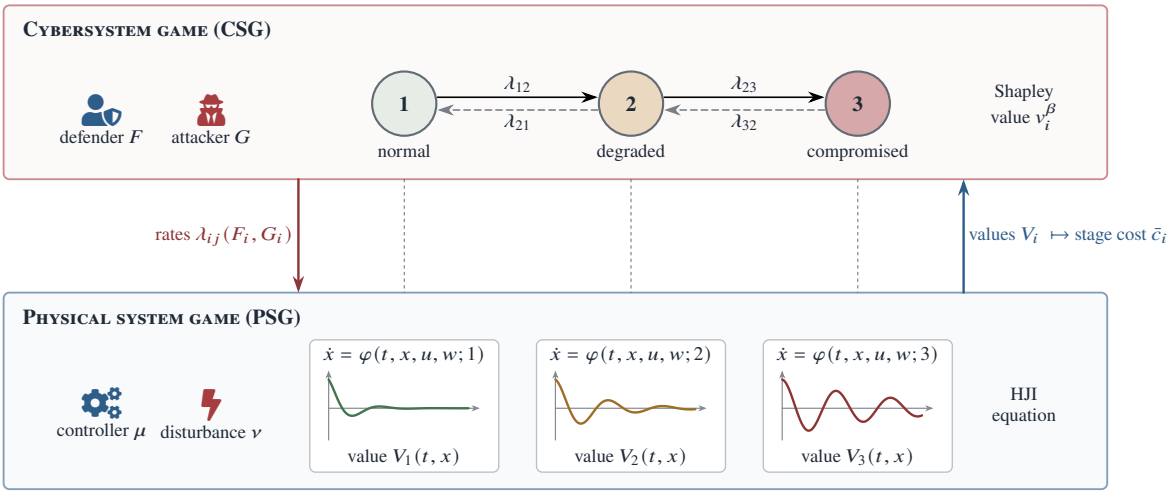


Figure 10.5: The games-in-games structure. The upper layer is the cybersystem game (CSG): the defender policy F and the attacker policy G determine the transition rates $\lambda_{ij}(F_i, G_i)$ among the cyber modes normal (1), degraded (2), and compromised (3), whose shading darkens with severity. Solid arrows between modes are attack-driven transitions, and dashed arrows are recovery transitions. The lower layer is the physical system game (PSG): in each mode i, the controller and the disturbance play a differential game whose value is V_i; the trajectories sketch the regulated state, which settles fast in the normal mode and oscillates longer as the mode degrades. Dotted lines connect each mode to its physical game. The downward arrow passes the transition rates to the PSG; the upward arrow passes the physical values, through the stage cost $\bar{c}_i$, to the CSG.

10.3.1 Hybrid Cyber-Physical Model

The system has a physical state $x(t) \in \mathbb{R}^n$ and a cyber state $i(t) \in S \triangleq \{1, \dots, N_S\}$. The physical state evolves on the fast time scale of the plant; the cyber state evolves on the slower time scale of attack, defense, failure, and recovery events. Conditional on $i(t) = i$, the physical dynamics are

$$\dot{x}(t) = \varphi\big(t, x(t), u(t), w(t); i\big), \qquad x(t_0) = x_0, \tag{10.3.1}$$

where $u(t) \in \mathbb{R}^m$ is the control input and $w(t) \in \mathbb{R}^{m_w}$ is a disturbance or adversarial physical input.

The defender and the attacker choose cyber actions from finite sets $\mathcal{D}_c$ and $\mathcal{A}_c$. For $\ell \in \mathcal{D}_c$ and $a \in \mathcal{A}_c$, the cyber state is a continuous-time Markov chain with transition rates $\lambda_{ij}(\ell, a) \geq 0$ for $j \neq i$, and the diagonal rates are $\lambda_{ii}(\ell, a) \triangleq -\sum_{j \neq i} \lambda_{ij}(\ell, a)$, so that every row of the generator sums to zero. A stationary mixed cyber policy assigns to

each mode a distribution $F_i \in \Delta(\mathcal{D}_c)$ for the defender and $G_i \in \Delta(\mathcal{A}_c)$ for the attacker, and the averaged rates are

$$\lambda_{ij}(F_i, G_i) \triangleq \sum_{\ell \in \mathcal{D}_c} \sum_{a \in \mathcal{A}_c} F_i(\ell)\, G_i(a)\, \lambda_{ij}(\ell, a). \tag{10.3.2}$$

These rates are the downward coupling: a stronger defense lowers the rates into compromised modes, and a stronger attack raises them.

10.3.2 Physical System Game

Definition 10.3.1 (Physical system game)**.** Fix stationary cyber policies $F = (F_i)_{i \in S}$ and $G = (G_i)_{i \in S}$. The *physical system game* (PSG) is the zero-sum differential game in which a closed-loop controller μ minimizes and a closed-loop disturbance policy ν maximizes

$$J_{F,G}(\mu, \nu) \triangleq \mathbb{E}\left[c_f\big(x(t_f); i(t_f)\big) + \int_{t_0}^{t_f} c\big(t, x(t), u(t), w(t); i(t)\big)\, dt \right],$$

with $u(t) = \mu(t, x(t), i(t))$, $w(t) = \nu(t, x(t), i(t))$, the dynamics (10.3.1), and the cyber state generated by the rates (10.3.2). The value of the PSG started at (t, x) in mode i is denoted by $V_i(t, x)$.

For fixed F and G, the Hamilton–Jacobi–Isaacs (HJI) equation of the PSG is

$$-\partial_t V_i(t,x) = \min_u \max_w \Big\{ \nabla_x V_i(t,x)^\top \varphi(t,x,u,w;i) + c(t,x,u,w;i) \Big\} + \sum_{j \in S} \lambda_{ij}(F_i, G_i)\, V_j(t,x), \qquad i \in S, \tag{10.3.3}$$

with terminal condition $V_i(t_f, x) = c_f(x; i)$. The first term is the Hamiltonian of a robust-control problem in mode i; the sum is the coupling term, through which cyber-induced jumps to other modes enter the physical value. Since $\sum_j \lambda_{ij} = 0$, the coupling term equals $\sum_{j \neq i} \lambda_{ij}(F_i, G_i)\big(V_j - V_i\big)$, the rate-weighted change in value caused by a jump.

THEOREM 10.3.2 **Physical-layer saddle point [660]**
Fix F and G. Suppose that (10.3.3) has a solution $(V_i)_{i \in S}$ that is continuously differentiable in (t, x), that the Isaacs condition holds, that is, $\min_u \max_w$ and $\max_w \min_u$ of the bracket in (10.3.3) coincide for all arguments, and that the feedback policies $\mu^\star(t, x, i)$ and $\nu^\star(t, x, i)$ attaining the saddle point of the bracket are admissible. Then $(\mu^\star, \nu^\star)$ is a closed-loop saddle point of the PSG, and $V_i(t, x)$ is its value.

Proof. Fix an admissible controller μ and let the disturbance use $\nu^\star$. By the generator of the jump process and Dynkin's formula, for the trajectory started at (t_0, x_0, i_0),

$$\mathbb{E}\big[V_{i(t_f)}(t_f, x(t_f))\big] - V_{i_0}(t_0, x_0) = \mathbb{E} \int_{t_0}^{t_f} \Big(\partial_t V_i + \nabla_x V_i^\top \varphi(\cdot, \mu, \nu^\star; i) + \textstyle\sum_j \lambda_{ij} V_j \Big)\Big|_{i = i(t)} dt.$$

Because $\nu^\star$ attains the maximum in (10.3.3) and the Isaacs condition holds, the integrand is at least $-c(t, x, \mu, \nu^\star; i)$, with equality when $\mu = \mu^\star$. Using the terminal condition, $J_{F,G}(\mu, \nu^\star) \geq V_{i_0}(t_0, x_0) = J_{F,G}(\mu^\star, \nu^\star)$. The symmetric argument with $\mu^\star$ fixed gives $J_{F,G}(\mu^\star, \nu) \leq V_{i_0}(t_0, x_0)$ for every admissible ν. Hence $(\mu^\star, \nu^\star)$ is a saddle point with value $V_{i_0}(t_0, x_0)$. □

The value $V_i(t, x)$ is conditional on the cyber policies: changing F or G changes the rates $\lambda_{ij}(F_i, G_i)$, hence (10.3.3), hence the robust-control value.

10.3.3 Cybersystem Game and Game Value

The cyber decisions are taken at epochs separated by a sampling interval $\Delta t > 0$ that is short relative to the mean holding times, $\Delta t \leq 1/\max_{i,\ell,a} |\lambda_{ii}(\ell, a)|$. Between epochs, the cyber state moves according to the transition probabilities

$$P_{ij}(\ell, a) \triangleq \delta_{ij} + \Delta t\, \lambda_{ij}(\ell, a), \tag{10.3.4}$$

where δ_{ij} is the Kronecker delta; the bound on Δt makes each row of $P(\ell, a)$ a probability distribution. The mixed extension $P_{ij}(F_i, G_i)$ is bilinear in (F_i, G_i), as in (10.3.2).

Definition 10.3.3 (Cybersystem game)**.** Fix the physical outcome (μ, ν, V). The *cybersystem game* (CSG) is the discounted zero-sum stochastic game on S in which, at mode i, the defender chooses $\ell \in \mathcal{D}_c$, the attacker chooses $a \in \mathcal{A}_c$, the system incurs the stage cost $\bar{c}_i(\ell, a; \mu, \nu, V)$, which the defender minimizes and the attacker maximizes, and the next mode is drawn from $P_{i\cdot}(\ell, a)$. Future costs are discounted by $\beta \in (0, 1)$. The stage cost combines direct cyber costs, such as the cost of hardening, with the physical value of the mode; it is the upward coupling from the PSG to the CSG.

The sign convention of Definition 10.3.3, in which the defender minimizes a cost, is the negative of the payoff convention of [660] and is equivalent to it. For a vector $v \in \mathbb{R}^{N_S}$, let $M_i(v)$ be the $|\mathcal{D}_c| \times |\mathcal{A}_c|$ matrix with entries

$$M_i(v)(\ell, a) \triangleq \bar{c}_i(\ell, a; \mu, \nu, V) + \beta \sum_{j \in S} P_{ij}(\ell, a)\, v_j .$$

The discounted game value v^β of the CSG satisfies the Shapley equation

$$v_i^\beta = \operatorname{val}\big[M_i(v^\beta)\big] \triangleq \min_{F_i \in \Delta(\mathcal{D}_c)} \max_{G_i \in \Delta(\mathcal{A}_c)} F_i^\top M_i(v^\beta)\, G_i, \qquad i \in S, \tag{10.3.5}$$

where, by the minimax theorem, min max may be replaced by max min [528, 198].

THEOREM 10.3.4 **Cyber-layer game value [528, 660]**
Fix the physical outcome (μ, ν, V) and suppose the stage costs $\bar{c}_i(\ell, a)$ are finite. Then (10.3.5) has a unique solution $v^{\beta\star} \in \mathbb{R}^{N_S}$. If, for every $i \in S$, $(F_i^\star, G_i^\star)$ is a pair of optimal strategies of the matrix game $M_i(v^{\beta\star})$, then the stationary policies $F^\star = (F_i^\star)_{i \in S}$ and $G^\star = (G_i^\star)_{i \in S}$ form a saddle point of the CSG, and $v_i^{\beta\star}$ is its value from mode i.

Proof. Define $(\mathcal{T}v)_i \triangleq \operatorname{val}[M_i(v)]$. The value of a matrix game is monotone and satisfies $|\operatorname{val}[M] - \operatorname{val}[M']| \le \max_{\ell,a} |M(\ell, a) - M'(\ell, a)|$. Since each row of $P(\ell, a)$ is a probability distribution, $|M_i(v)(\ell, a) - M_i(v')(\ell, a)| \le \beta \|v - v'\|_\infty$, and therefore $\|\mathcal{T}v - \mathcal{T}v'\|_\infty \le \beta \|v - v'\|_\infty$. As $\beta < 1$, $\mathcal{T}$ is a contraction on $(\mathbb{R}^{N_S}, \|\cdot\|_\infty)$, and Banach's theorem gives the unique fixed point $v^{\beta\star}$. If the defender uses $F^\star$, the attacker faces a discounted Markov decision process whose optimality equation is $v_i = \max_{G_i} F_i^{\star\top} M_i(v) G_i$; since $F_i^\star$ is optimal in $M_i(v^{\beta\star})$, the vector $v^{\beta\star}$ solves this equation, so no attacker policy obtains more than $v^{\beta\star}$. The symmetric argument with $G^\star$ fixed shows that no defender policy obtains less. Hence $(F^\star, G^\star)$ is a saddle point with value $v^{\beta\star}$. □

THEOREM 10.3.5 **Value-iteration convergence [660]**
Let $v^0 \in \mathbb{R}^{N_S}$ and define

$$v_i^{n+1} = \operatorname{val}\big[M_i(v^n)\big], \qquad i \in S,\ n \ge 0. \tag{10.3.6}$$

Then $\|v^n - v^{\beta\star}\|_\infty \le \beta^n \|v^0 - v^{\beta\star}\|_\infty$. Moreover, if (F_i^n, G_i^n) is a pair of optimal strategies of $M_i(v^n)$, every limit point of (F^n, G^n) is a stationary saddle point of the CSG.

Proof. The geometric bound follows from the contraction property established in the proof of Theorem 10.3.4, since $v^n - v^{\beta\star} = \mathcal{T}^n v^0 - \mathcal{T}^n v^{\beta\star}$. The matrices $M_i(v^n)$ converge to $M_i(v^{\beta\star})$, and the correspondence that maps a matrix to its set of optimal strategy pairs has a closed graph. Hence every limit point of (F_i^n, G_i^n) is a pair of optimal strategies of $M_i(v^{\beta\star})$, and Theorem 10.3.4 identifies it as a stationary saddle point. □

10.3.4 Games-in-Games Equilibrium

Equation (10.3.3) depends on (F, G) through the rates $\lambda_{ij}(F_i, G_i)$, and equation (10.3.5) depends on (μ, ν, V) through the stage costs $\bar{c}_i$. The cross-layer solution must satisfy both.

Definition 10.3.6 (Games-in-games equilibrium)**.** A *games-in-games equilibrium* is a tuple $(\mu^\star, \nu^\star, V^\star, F^\star, G^\star, v^{\beta\star})$ such that

(1) $V^\star$ solves (10.3.3) under the rates $\lambda_{ij}(F_i^\star, G_i^\star)$, and $(\mu^\star, \nu^\star)$ is the associated saddle point of Theorem 10.3.2;

(2) $v^{\beta\star}$ solves (10.3.5) under the stage costs $\bar{c}_i(\cdot\,; \mu^\star, \nu^\star, V^\star)$, and $(F^\star, G^\star)$ is a stationary saddle point of the CSG associated with $v^{\beta\star}$.

PROPOSITION 10.3.7 **Games-in-games as a Gestalt equilibrium**
Consider the two-component composed game with the PSG outcome $o_P = (\mu, \nu, V)$, the CSG outcome $o_C = (F, G, v^\beta)$, the coupling maps $C_P(o) = \big(\lambda_{ij}(F_i, G_i)\big)_{i,j\in S}$ and $C_C(o) = \big(\bar{c}_i(\cdot\,; \mu, \nu, V)\big)_{i\in S}$, and the solution correspondences $\mathcal{S}_P$ and $\mathcal{S}_C$ given by Theorems 10.3.2 and 10.3.4. A tuple is a games-in-games equilibrium if and only if (o_P, o_C) is a Gestalt equilibrium of this composed game.

Proof. Condition (1) of Definition 10.3.6 states $o_P^\star \in \mathcal{S}_P(C_P(o^\star))$, and condition (2) states $o_C^\star \in \mathcal{S}_C(C_C(o^\star))$. Together they are (10.2.2) for $K = 2$. □

Solving either game with the other held at an assumed outcome yields only a partial optimum: the PSG then guards against a transition law that the cyber layer does not produce, and the CSG protects modes according to costs that the physical layer does not incur. Stronger robust control lowers the cost of degraded modes and hence the value of protecting them, while stronger cyber defense lowers the rates into those modes and hence the robustness that the controller must provide; the games-in-games equilibrium balances robustness, security, and resilience.

10.3.5 Linear-Quadratic Specialization

Appendix C reviews the linear-quadratic and Riccati background used here. Consider the Markov jump linear system $\dot{x} = A_i x + B_i u + D_i w$ in mode i, with the infinite-horizon soft-constrained cost rate $c(x, u, w; i) = x^\top Q_i x + u^\top R_i u - \gamma^2 \|w\|^2$, where $Q_i \succeq 0$, $R_i \succ 0$, and $\gamma > 0$ is the attenuation level. The stationary HJI equation (10.3.3) with the quadratic value $V_i(x) = x^\top Z_i x$ reduces to the coupled game algebraic Riccati equations of minimax and H_∞ control [43]:

$$A_i^\top Z_i + Z_i A_i - Z_i B_i R_i^{-1} B_i^\top Z_i + \gamma^{-2} Z_i D_i D_i^\top Z_i + Q_i + \sum_{j\in S} \lambda_{ij}(F_i, G_i)\, Z_j = 0, \qquad i \in S. \tag{10.3.7}$$

THEOREM 10.3.8 **Linear-quadratic physical layer [660, 43]**
Suppose the cyber chain with rates $\lambda_{ij}(F_i, G_i)$ is irreducible, the pairs (A_i, B_i) are stochastically stabilizable, and the pairs $(A_i, Q_i^{1/2})$ are observable. Then there is a critical attenuation level $\gamma^\star \geq 0$ such that, for every $\gamma > \gamma^\star$, the equations (10.3.7) admit a minimal positive semidefinite solution $(Z_i)_{i\in S}$, the controller $u_i^\star(x) = -R_i^{-1} B_i^\top Z_i x$ and the disturbance $w_i^\star(x) = \gamma^{-2} D_i^\top Z_i x$ form a saddle point of the PSG with value $x^\top Z_i x$, and the closed-loop jump system under $u^\star$ is mean-square stable. For $\gamma < \gamma^\star$, the upper value of the PSG is infinite for some initial condition.

Proof sketch. Substituting $V_i(x) = x^\top Z_i x$ into (10.3.3), the bracket is the quadratic form $2x^\top Z_i(A_i x + B_i u + D_i w) + x^\top Q_i x + u^\top R_i u - \gamma^2\|w\|^2$, which is strictly convex in u and strictly concave in w. Its unique saddle point is $u_i^\star$ and $w_i^\star$, and substitution leaves the quadratic form $x^\top(\cdot)x$ whose vanishing for all x is (10.3.7). Theorem 10.3.2 then gives the saddle-point property. The existence of a minimal positive semidefinite stabilizing solution above $\gamma^\star$, and the loss of a finite upper value below it, are the standard conclusions of soft-constrained H_∞ theory for jump linear systems under the stated stabilizability and observability hypotheses [43, 660]. □

With the physical outcome given by Theorem 10.3.8, a natural stage cost of the CSG is $\bar{c}_i(\ell, a) = x_0^\top Z_i x_0 + c_i^{\mathrm{cy}}(\ell, a)$, the physical value from a reference state x_0 plus a direct cyber cost c_i^{cy}. The Shapley equation becomes

$$v_i^{\beta\star} = \mathrm{val}\Big[\, x_0^\top Z_i x_0 + c_i^{\mathrm{cy}}(\ell, a) + \beta \sum_{j\in S} P_{ij}(\ell, a)\, v_j^{\beta\star} \Big]_{\ell\in\mathcal{D}_c,\, a\in\mathcal{A}_c}, \qquad i \in S, \tag{10.3.8}$$

and the games-in-games equilibrium is a joint solution of (10.3.7) and (10.3.8), in which the matrices Z_i depend on the cyber policies of all modes through the coupling term of (10.3.7).

EXAMPLE 10.2 **Two cyber modes and a scalar plant**

▸ Code: `code/scalar_cross_layer.py`

A pumping station is controlled from a programmable logic controller that is either in the normal mode N or in a compromised mode C, in which manipulated sensor data amplify the plant response. Consider one cyber epoch starting in N. The operator chooses $\ell \in \{\text{harden}, \text{light}\}$, where light denotes routine protection, and the attacker chooses $a \in \{\text{quiet}, \text{attack}\}$. The probability $P(\ell, a)$ that the epoch ends in C is given in Table 10.2. If the operator hardens with probability q and the attacker attacks with probability g, the probability of entering C is

$$p_{\mathrm{NC}}(q, g) = (1 - g)\big(0.10 - 0.05q\big) + g\big(0.60 - 0.40q\big) = 0.10 - 0.05q + 0.50g - 0.35qg.$$

	Attacker a	
Operator ℓ	**quiet**	**attack**
harden	0.05	0.20
light	0.10	0.60

Table 10.2: Probability $P(\ell, a)$ of moving from the normal mode N to the compromised mode C in Example 10.2.

Physical component. If the epoch ends in mode j, the scalar plant is $x_{k+1} = A_j x_k + u_k$ with $A_{\mathrm{N}} = 0.8$ and $A_{\mathrm{C}} = 1.2$, and the stage cost is $x_{k+1}^2 + u_k^2$. The controller chooses the feedback $u_k = -\kappa x_k$ before the mode is revealed, so its expected cost is $Z(\kappa, p)x_k^2$ with

$$Z(\kappa, p) \triangleq (1 - p)(A_{\mathrm{N}} - \kappa)^2 + p\,(A_{\mathrm{C}} - \kappa)^2 + \kappa^2, \qquad p = p_{\mathrm{NC}}(q, g).$$

The context of this component is p. Setting $\partial Z/\partial\kappa = 0$ gives $\kappa^\star(p) = \bar{A}(p)/2$ and the physical value $Z^\star(p) = \bar{A}^{(2)}(p) - \bar{A}(p)^2/2$, where $\bar{A}(p) \triangleq (1-p)A_{\mathrm{N}} + pA_{\mathrm{C}} = 0.8 + 0.4p$ and $\bar{A}^{(2)}(p) \triangleq (1-p)A_{\mathrm{N}}^2 + pA_{\mathrm{C}}^2 = 0.64 + 0.80p$ are the first two moments of the post-epoch gain. For the four pure cyber profiles,

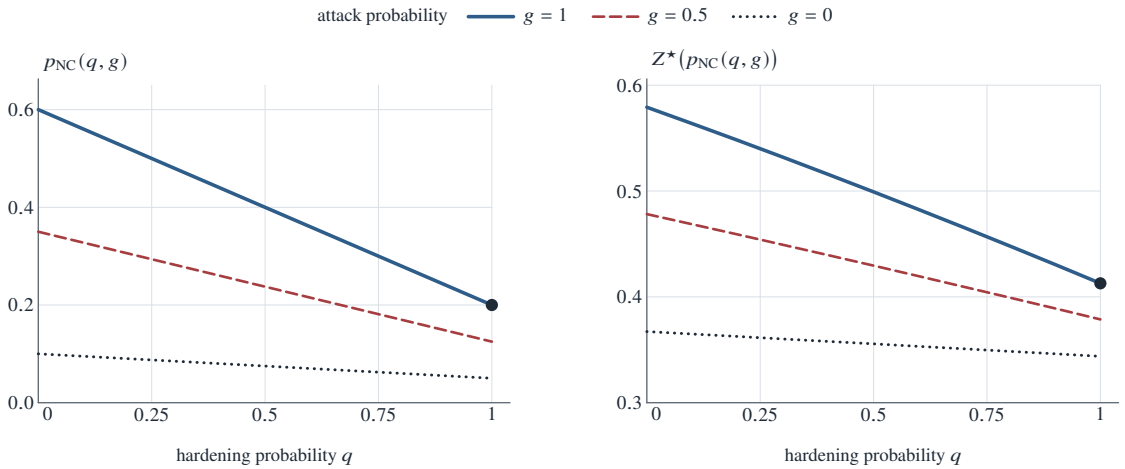
(ℓ, a)	p	$\bar{A}(p)$	$\kappa^\star(p)$	$Z^\star(p)$
(harden, quiet)	0.05	0.82	0.41	0.3438
(light, quiet)	0.10	0.84	0.42	0.3672
(harden, attack)	0.20	0.88	0.44	0.4128
(light, attack)	0.60	1.04	0.52	0.5792

Cyber component. The context of the CSG is the controller gain κ. Its stage cost is $\bar{c}(\ell, a; \kappa) = Z(\kappa, P(\ell, a)) + c_h \mathbf{1}\{\ell = \text{harden}\}$ with hardening cost $c_h = 0.06$; since only one epoch is considered, the continuation term of (10.3.5) is omitted. The cost $Z(\kappa, p)$ is affine in p with slope $(A_{\mathrm{C}} - \kappa)^2 - (A_{\mathrm{N}} - \kappa)^2 = 0.8(1 - \kappa)$, which is positive for $\kappa < 1$. Because $P(\ell, \text{attack}) > P(\ell, \text{quiet})$ for both ℓ, attacking strictly dominates for the attacker. Against an attack, hardening lowers the cost by $0.8(1-\kappa)(0.60-0.20) - c_h = 0.32(1-\kappa) - 0.06$, which is positive for every $\kappa < 0.8125$. Hence, for every gain in this range, the CSG has the unique saddle point (harden, attack).

Gestalt equilibrium. The saddle point gives $p = P(\text{harden}, \text{attack}) = 0.20$, the physical component returns $\kappa^\star(0.20) = 0.44 < 0.8125$, and the CSG under $\kappa = 0.44$ returns the same saddle point. Since every gain $\kappa^\star(p)$, $p \in [0, 0.6]$, lies in $[0.40, 0.52]$, this is the unique Gestalt equilibrium: $q^\star = 1$, $g^\star = 1$, $p^\star = 0.20$, $\kappa^\star = 0.44$, and the stage cost $Z^\star(0.20) + c_h = 0.4728$. Figure 10.6 plots the downward and upward maps over mixed cyber strategies.

Observation. A siloed design keeps routine protection and tunes the controller for the nominal transition probability 0.10, which gives $\kappa = 0.42$. The attacker then attacks, the transition probability is 0.60, and the stage cost is $Z(0.42, 0.60) = 0.5992$. The cross-layer equilibrium lowers the cost to 0.4728, including the hardening cost.

Most of the improvement comes from the cyber decision; retuning the controller to the equilibrium transition law contributes only $Z(0.42, 0.20) - Z(0.44, 0.20) = 0.0008$. The benefit of hardening, $0.32(1 - \kappa)$, decreases with the controller gain, so a more aggressive physical controller lowers the cyber incentive to harden; this is the upward coupling at work.



(a) Transition probability **(b) Physical value**

Figure 10.6: The maps of Example 10.2 over mixed cyber strategies. (a) The transition probability $p_{\mathrm{NC}}(q, g)$ as a function of the hardening probability q, for attack probabilities $g = 0$ (dotted), $g = 0.5$ (dashed), and $g = 1$ (solid). (b) The physical value $Z^\star(p_{\mathrm{NC}}(q, g))$ for the same three attack probabilities. The disc marks the Gestalt equilibrium $(q, g) = (1, 1)$, where $p = 0.20$ and $Z^\star = 0.4128$.

10.4 Cyber-Physical Signaling Games for Multi-Agent Autonomous Systems

This section develops the cyber-physical signaling framework of Xu and Zhu [603], in which the coupling between components runs through beliefs. A team of autonomous agents, such as unmanned aerial vehicles, receives reference positions from a control station over a vulnerable network. A man-in-the-middle attacker can replace a legitimate reference by a malicious one that drives an agent into a static obstacle, a *suicidal attack*, or into another agent, a *collision attack*. The cyber component is a signaling game [548, 336, 205] in which each agent decides whether to accept a received reference; the physical component is a min–max model predictive controller [388, 498] that tracks the accepted reference. The signaling game needs the physical consequence of accepting a reference, which only the controller can compute; the controller needs the accepted reference, which only the signaling game can decide. Figure 10.7 shows the architecture.

10.4.1 Physical Layer: Min–Max Model Predictive Control

Appendix C reviews the discrete-time linear-system and MPC background used here. There are n agents indexed by $i \in \mathcal{N} = \{1, \dots, n\}$. Agent i has the model

$$x^i_{k+1} = A_i x^i_k + B_i u^i_k + w^i_k, \qquad y^i_k = H_i x^i_k,$$

where x^i_k is the state, $u^i_k \in \mathcal{U}$ is the control input, $w^i_k \in \mathcal{W}$ is an unknown bounded disturbance, and $y^i_k \in \mathbb{R}^2$ is the position of the agent. For a reference $\rho \in \mathbb{R}^2$, agent i solves over the prediction horizon N_p the min–max problem

$$J^\star_c(x^i_k, \rho) \triangleq \min_{\hat{u} \in \mathcal{U}^{N_p}} \max_{\hat{w} \in \mathcal{W}^{N_p}} \sum_{t=0}^{N_p - 1} \Big(\| H_i \hat{x}^i_{k+t|k} - \rho \|^2 + c_u \| \hat{u}_{k+t|k} \|^2 \Big), \tag{10.4.1}$$

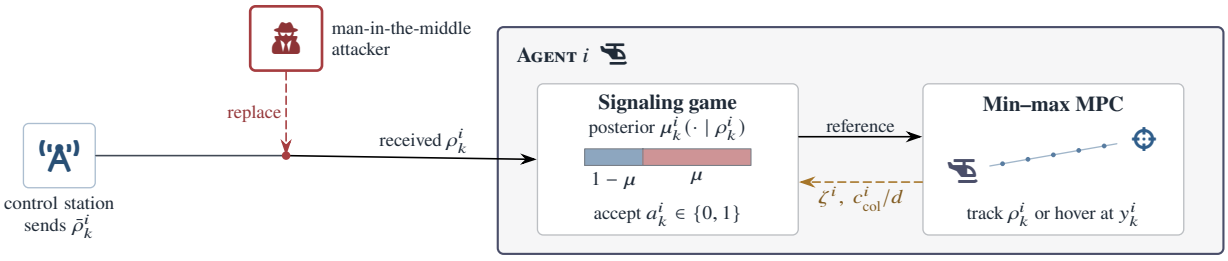


Figure 10.7: Cyber-physical signaling for autonomous agents. A man-in-the-middle attacker may replace the reference $\bar{\rho}_k^i$ sent by the control station with a malicious reference. The receiving agent i forms a posterior belief μ_k^i that the sender is malicious, drawn as the split bar of $1-\mu$ and μ, and accepts ($a_k^i = 1$) or rejects ($a_k^i = 0$) the received reference ρ_k^i. The min–max MPC layer plans a path to the accepted reference (dots mark the prediction horizon), or hovers at the current position y_k^i after a rejection, and returns the rejection cost ζ^i and the collision cost c_{col}^i/d that enter the receiver's payoff (dashed arrow).

where $\hat{x}_{k+t|k}^i$ is the state predicted from x_k^i under the inputs $\hat{u}$ and disturbances $\hat{w}$, and $c_u > 0$ weights the control effort. Only the first input of the minimizing sequence is applied, and the problem is solved again at the next step. If agent i rejects a reference, it hovers at its current position, that is, it tracks $\rho = y_k^i$.

10.4.2 Attack Model and Local Sensing

At step k, the control station intends to send the legitimate reference $\bar{\rho}_k^i$ to agent i. A malicious sender replaces it with a reference that moves agent i toward a target y_k^j, which is either a static obstacle or another agent. Each agent carries a local sensor with detection radius $\bar{d}_{\text{det}}^i$. With the Euclidean distance $d(\cdot,\cdot)$, the index set of objects detected by agent i is $\mathcal{M}_k^i \triangleq \{j : d(y_k^i, y_k^j) \le \bar{d}_{\text{det}}^i\}$, and the admissible references are restricted to the disk $\mathcal{R}_k^i \triangleq \{\rho : d(\rho, y_k^i) \le \bar{d}_{\text{ref}}^i\}$ with $\bar{d}_{\text{ref}}^i \le \bar{d}_{\text{det}}^i$, so that every object relevant to an admissible reference is sensed. The analysis below considers objects in $\mathcal{M}_k^i$ and writes $d_{ij} \triangleq d(y_k^i, y_k^j) > 0$ for the current distance to object j.

10.4.3 Cyber Layer: The Signaling Game

Signaling games and perfect Bayesian Nash equilibrium (PBNE) were introduced in Chapters 3 and 4. For receiver i at step k, the type of the sender is $\theta \in \Theta^i \triangleq \{\theta_0\} \cup \{\theta_1^j : j \in \mathcal{M}_k^i\}$, where θ_0 is the legitimate station and θ_1^j is a malicious sender that targets object j. The type is drawn from a prior with full support. The sender chooses a reference $\rho \in \mathcal{R}_k^i$; the receiver observes ρ, forms the posterior $\mu_k^i(\cdot \mid \rho) \in \Delta(\Theta^i)$, and chooses $a \in \{0,1\}$, where $a = 1$ means accept.

The receiver's costs encode the physical consequences. Let $\zeta^i \triangleq J_c^\star(x_k^i, y_k^i) + c_{\text{m}}^i > 0$ be the cost of rejecting, which consists of the hovering cost and the mission delay cost c_{m}^i, and let $c_{\text{col}}^i > 0$ scale the collision risk. Then

$$c_R^i(a,\rho,\theta_0) = (1-a)\,\zeta^i, \qquad c_R^i(a,\rho,\theta_1^j) = a\,\frac{c_{\text{col}}^i}{d(\rho, y_k^j)} + (1-a)\,\frac{c_{\text{col}}^i}{d_{ij}}.$$

Accepting a legitimate reference costs nothing; accepting a malicious reference costs a collision risk that grows without bound as the reference approaches the target; rejecting leaves the agent at its current distance from the target. The sender costs encode the opposing objectives,

$$c_S^i(a,\rho,\theta_0) = a\,d(\rho,\bar{\rho}_k^i) + (1-a)\,d(y_k^i,\bar{\rho}_k^i), \qquad c_S^i(a,\rho,\theta_1^j) = a\,d(\rho,y_k^j) + (1-a)\,d_{ij},$$

so that the legitimate station wants the agent to reach $\bar{\rho}_k^i$, and a malicious sender wants it to approach y_k^j. These costs are the upward coupling from the physical layer: the signaling game evaluates a reference by the physical outcome it produces, not by its apparent authenticity.

10.4.4 Impact-Aware Danger Zones

The receiver's best response has a geometric form. Consider a posterior that places probability μ on a malicious type targeting object j and $1-\mu$ on θ_0.

LEMMA 10.4.1 **Belief-dependent danger zone**
Under the posterior above, accepting ρ is a best response of receiver i if and only if $d(\rho, y_k^j) \geq \eta^{ij}(\mu)$, where

$$\eta^{ij}(\mu) \triangleq \frac{\mu\, c^i_{\mathrm{col}}\, d_{ij}}{\mu\, c^i_{\mathrm{col}} + (1-\mu)\, \zeta^i\, d_{ij}}.$$

The radius η^{ij} is increasing in μ, with $\eta^{ij}(0) = 0$ and $\eta^{ij}(1) = d_{ij}$.

Proof. The expected cost of accepting is $\mu c^i_{\mathrm{col}}/d(\rho, y_k^j)$, and that of rejecting is $\mu c^i_{\mathrm{col}}/d_{ij} + (1-\mu)\zeta^i$. Accepting is a best response if and only if $\mu c^i_{\mathrm{col}}/d(\rho, y_k^j) \leq \big(\mu c^i_{\mathrm{col}} + (1-\mu)\zeta^i d_{ij}\big)/d_{ij}$, which rearranges to $d(\rho, y_k^j) \geq \eta^{ij}(\mu)$. The derivative of η^{ij} with respect to μ equals $c^i_{\mathrm{col}}\zeta^i d_{ij}^2/\big(\mu c^i_{\mathrm{col}} + (1-\mu)\zeta^i d_{ij}\big)^2 > 0$, and the boundary values follow by substitution. □

The open disk $\Omega_k^{ij}(\mu) \triangleq \{\rho : d(\rho, y_k^j) < \eta^{ij}(\mu)\}$ is the *danger zone* of object j, and the receiver accepts exactly the references outside it. Under full suspicion, $\mu = 1$, the danger zone is $\hat{\Omega}_k^{ij} \triangleq \{\rho : d(\rho, y_k^j) < d_{ij}\}$, the set of references that would bring the agent closer to object j than it currently is. Over all detected objects, the full-suspicion danger zone and the associated receiver rule are

$$\Omega_k^i \triangleq \bigcup_{j \in \mathcal{M}_k^i} \hat{\Omega}_k^{ij}, \qquad \sigma_R^i(\rho) \triangleq \mathbf{1}\{\rho \notin \Omega_k^i\}. \tag{10.4.2}$$

The danger zone is where the two layers meet: the posterior sets its radius, and the physical geometry determines whether a reference lies inside it.

EXAMPLE 10.3 **Danger zone of a delivery drone**
A delivery drone i hovers at $y_k^i = (0, 0)$ beside a warehouse, and its sensor detects a radio mast at $y_k^j = (4, 0.5)$, so $d_{ij} = \sqrt{16.25} \approx 4.031$. The rejection cost is $\zeta^i = 2$, the collision scale is $c^i_{\mathrm{col}} = 2$, and the admissible references lie within $\bar{d}^i_{\mathrm{ref}} = 3.2$ of the drone. Lemma 10.4.1 gives the danger radii

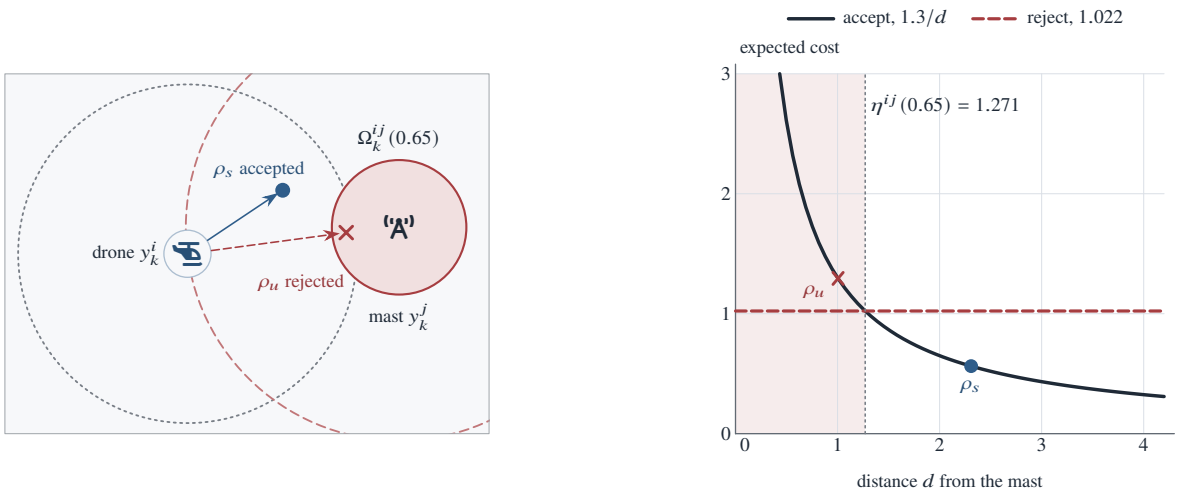
$$\eta^{ij}(0.2) = 0.235, \qquad \eta^{ij}(0.5) = 0.801, \qquad \eta^{ij}(0.65) = 1.271, \qquad \eta^{ij}(0.9) = 2.784, \qquad \eta^{ij}(1) = 4.031.$$

Suppose that, after the observed traffic, the drone assigns the posterior probability $\mu = 0.65$ to a malicious sender. The cost of rejecting is $0.65 \cdot 2/4.031 + 0.35 \cdot 2 = 1.022$. The reference $\rho_s = (1.8, 1.2)$ lies at distance 2.309 from the mast; its acceptance cost $1.3/2.309 = 0.563$ is below the rejection cost, so it is accepted. The reference $\rho_u = (3.0, 0.4)$ lies at distance 1.005; its acceptance cost $1.3/1.005 = 1.294$ exceeds the rejection cost, so it is rejected. Both references are admissible, since $d(\rho_s, y_k^i) = 2.163$ and $d(\rho_u, y_k^i) = 3.027$ are below $\bar{d}^i_{\mathrm{ref}}$. Figure 10.8 shows the scene and the two costs along the segment from the drone to the mast.

Observation. An authentication-only defense accepts every reference that passes its checks, including ρ_u, whose collision risk exceeds the cost of hovering by 0.27. The impact-aware rule rejects only references that move the drone into a region whose radius scales with the suspicion: 0.24 under a benign posterior and the full current distance 4.03 under certainty of attack. A higher rejection cost ζ^i, for example a time-critical delivery, shrinks the danger zone and makes the drone accept riskier references.

10.4.5 Equilibrium Structure

The equilibrium analysis distinguishes whether the legitimate reference is itself safe. The following theorem restates the separating equilibrium of Xu and Zhu in the type space Θ^i.



(a) Scene **(b) Expected costs**

Figure 10.8: The danger zone of Example 10.3. (a) The drone at the origin and the radio mast at $(4, 0.5)$. The dotted circle bounds the admissible references, the shaded disk is the danger zone $\Omega_k^{ij}(0.65)$ of radius 1.271, and the dashed circle is the full-suspicion boundary of radius $d_{ij} = 4.031$. The reference ρ_s is accepted and ρ_u is rejected. (b) The expected cost of accepting a reference at distance d from the mast (solid) and the cost of rejecting (dashed); the two curves cross at $d = \eta^{ij}(0.65) = 1.271$.

THEOREM 10.4.2 **Separating PBNE [603]**
Suppose $d(\bar{\rho}_k^i, y_k^j) > d_{ij}$ for every $j \in \mathcal{M}_k^i$. Then the following profile is a PBNE: the legitimate station sends $\bar{\rho}_k^i$; every malicious type sends the hover reference y_k^i; the receiver uses the rule (10.4.2); the posterior assigns probability one to θ_0 after $\bar{\rho}_k^i$, the prior conditioned on the malicious types after y_k^i, probability one to θ_1^j after an off-path reference in $\hat{\Omega}_k^{ij}$ (choosing any such j if several apply), and probability one to θ_0 after any other off-path reference. In this equilibrium, the receiver never accepts a reference inside Ω_k^i.

Proof. Receiver. After $\bar{\rho}_k^i$, the receiver believes θ_0; accepting costs $0 < \zeta^i$, and the rule accepts, since $\bar{\rho}_k^i \notin \Omega_k^i$ by hypothesis. After y_k^i, every malicious type θ_1^j gives accepting and rejecting the same cost c_{col}^i/d_{ij}, so the receiver is indifferent and accepting, which the rule prescribes because $y_k^i \notin \Omega_k^i$, is a best response. After an off-path $\rho \in \hat{\Omega}_k^{ij}$, the belief θ_1^j makes rejecting strictly better, since $c_{\text{col}}^i/d(\rho, y_k^j) > c_{\text{col}}^i/d_{ij}$. After any other off-path ρ, the belief θ_0 makes accepting optimal. The beliefs are derived from Bayes' rule on the equilibrium path, because $\bar{\rho}_k^i \neq y_k^i$ and the prior has full support.

Legitimate station. Its reference $\bar{\rho}_k^i$ is accepted and yields the cost 0, the minimum possible.

Malicious type θ_1^j. Every accepted reference lies outside $\Omega_k^i \supseteq \hat{\Omega}_k^{ij}$ and hence satisfies $d(\rho, y_k^j) \geq d_{ij}$; every rejected reference yields d_{ij}. Thus no reference yields a cost below d_{ij}, and the hover reference attains it. Hence no type has a profitable deviation. The last statement follows from (10.4.2). □

When the legitimate reference itself would move the agent closer to a detected object, the two types can pool. Xu and Zhu state that both types then pool on the hover reference y_k^i; with a continuum of references, however, the legitimate station prefers any accepted reference closer to $\bar{\rho}_k^i$ than y_k^i is. The pooling equilibrium therefore lies at the safe reference nearest to $\bar{\rho}_k^i$, which coincides with the hover reference only in a special geometry.

PROPOSITION 10.4.3 **Pooling PBNE**
Suppose $\mathcal{M}_k^i = \{j\}$, $d(\bar{\rho}_k^i, y_k^j) < d_{ij}$, $\bar{\rho}_k^i \neq y_k^j$, and the reference $\hat{\rho} \triangleq y_k^j + d_{ij}\,(\bar{\rho}_k^i - y_k^j)/d(\bar{\rho}_k^i, y_k^j)$ is admissible.

Then the following profile is a PBNE: both types send $\hat{\rho}$; the receiver uses (10.4.2); the posterior after $\hat{\rho}$ equals the prior; and the off-path beliefs are those of Theorem 10.4.2. The hover reference y_k^i is the pooling reference if and only if $\bar{\rho}_k^i$ lies on the segment from y_k^j to y_k^i.

Proof. The reference $\hat{\rho}$ is the radial projection of $\bar{\rho}_k^i$ onto the circle $d(\rho, y_k^j) = d_{ij}$; it lies outside $\Omega_k^i = \hat{\Omega}_k^{ij}$ and is the point of the closed set $\{\rho : d(\rho, y_k^j) \geq d_{ij}\}$ nearest to $\bar{\rho}_k^i$. *Receiver.* After $\hat{\rho}$, with the prior probability p of the malicious type, accepting costs pc_{col}^i/d_{ij} and rejecting costs $pc_{\text{col}}^i/d_{ij} + (1-p)\zeta^i$, so accepting is optimal; the off-path responses are verified as in Theorem 10.4.2. *Legitimate station.* An accepted reference lies in $\{\rho : d(\rho, y_k^j) \geq d_{ij}\}$, so its cost is at least $d(\hat{\rho}, \bar{\rho}_k^i)$; a rejected one costs $d(y_k^i, \bar{\rho}_k^i) \geq d(\hat{\rho}, \bar{\rho}_k^i)$, because y_k^i belongs to the same closed set. *Malicious type.* Every outcome costs at least d_{ij}, and $\hat{\rho}$ attains d_{ij}. Finally, $\hat{\rho} = y_k^i$ if and only if $\bar{\rho}_k^i - y_k^j$ is a positive multiple of $y_k^i - y_k^j$, that is, $\bar{\rho}_k^i$ lies on the segment from y_k^j to y_k^i, given $d(\bar{\rho}_k^i, y_k^j) < d_{ij}$. □

In both equilibria, the receiver accepts only references outside the full-suspicion danger zone, so no accepted reference brings an agent closer to a detected object than it already is; this is the collision-avoidance guarantee of the framework. In the separating equilibrium, the receiver learns the sender's type; in the pooling equilibrium, it does not, yet it remains safe, because safety is enforced by the geometry of the danger zone rather than by identification of the sender.

Remark 10.4.4 (Uniqueness). The separating equilibrium of Theorem 10.4.2 is not the only one. A malicious type is indifferent among all accepted references on the circle $d(\rho, y_k^j) = d_{ij}$, so any such reference can replace y_k^i without changing the outcome of the receiver or the payoffs of the senders.

Composed-game interpretation. In the language of Section 10.2, the cyber outcome is the PBNE profile with its beliefs, and the physical outcome is the MPC solution (10.4.1). The physical outcome enters the signaling game through ζ^i and the distances in the cost functions; the cyber outcome enters the controller through the accepted reference. Compared with Section 10.3, the coupling variables are beliefs, accept decisions, and MPC costs instead of transition rates and values, but the consistency requirement is the same: the trust decision must be optimal for its physical consequence, and the controller must act on the reference that the trust decision admits.

10.5 iSTRICT: Interdependent Strategic Trust in the Cloud-Enabled IoCT

Chapter 4 analyzed strategic trust between one cloud and one device as a game of games: a FlipIt game determines how often the cloud is compromised, and a signaling game determines whether the device trusts it. The iSTRICT framework of Pawlick, Chen, and Zhu [474] extends this composition to the Internet of Controlled Things (IoCT), in which a device receives data from several cloud services, each of which is contested by its own attacker and defender. The framework has three layers. The cloud layer contains one FlipIt game [577] per service; the communication layer contains one signaling game [548, 205] between the services and the device; and the physical layer contains an observer-based controller that converts trust decisions into control performance. Compromise probabilities determine trust, trust determines physical performance, and physical performance determines how much each service is worth to its attacker and defender. Figure 10.9 shows the three layers.

10.5.1 Problem Formulation

There are N_c cloud services indexed by $i \in \mathcal{I} \triangleq \{1, \ldots, N_c\}$. For service i, an attacker A_i attempts to take control of the service and a defender D_i attempts to retain or recover control. A device R receives data from all services and decides which messages to trust; it uses the trusted data in an observer and a feedback controller. The modeling goal is a steady state in which cloud security, trust, and physical performance are mutually consistent.

10.5.2 Cloud Layer: FlipIt Games

The FlipIt game of Chapter 3 is played on each service. Attacker A_i and defender D_i choose move frequencies $f_A^i, f_D^i \geq 0$, and $\bar{p}(f_A^i, f_D^i) \in [0, 1]$ denotes the long-run fraction of time during which the attacker controls the service.

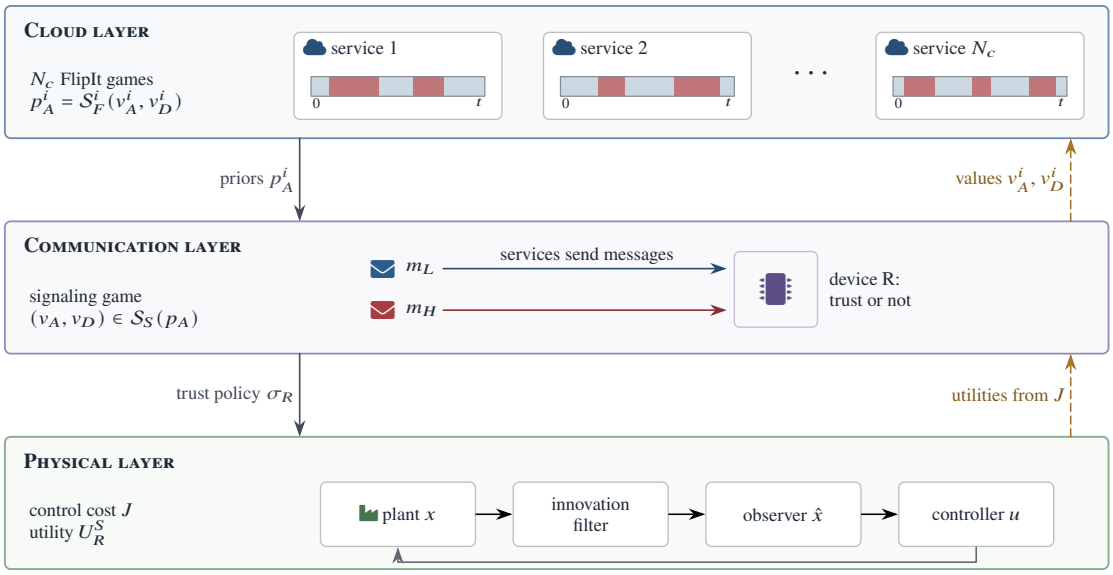


Figure 10.9: iSTRICT as a three-layer composed game. The cloud layer contains one FlipIt game per service $i \in \mathcal{I}$; each timeline shows the intervals in which the defender (light) and the attacker (dark) control the service, and the layer returns the compromise probabilities p_A^i. The communication layer uses them as priors of a signaling game, in which the services send low-risk or high-risk messages m_L or m_H and the device R decides which messages to trust; it returns the sender values (v_A^i, v_D^i) to the cloud layer. The physical layer passes the measurements through the innovation filter to the observer and the controller, converts the trust policy into the control cost J, and determines the utilities of the signaling game. Solid arrows pass outcomes down; dashed arrows pass values up.

Let $v_A^i, v_D^i \geq 0$ be the values of controlling service i per unit time, and $c_A^i, c_D^i > 0$ the costs per move. The FlipIt utilities are

$$\bar{U}_{\mathrm{A}_i}^F \triangleq v_A^i\, \bar{p}(f_A^i, f_D^i) - c_A^i f_A^i, \qquad \bar{U}_{\mathrm{D}_i}^F \triangleq v_D^i\, (1 - \bar{p}(f_A^i, f_D^i)) - c_D^i f_D^i.$$

Definition 10.5.1 (FlipIt map)**.** For values (v_A^i, v_D^i), let $(f_A^{i\star}, f_D^{i\star})$ be a Nash equilibrium of the FlipIt game of service i. The *FlipIt map* $\mathcal{S}_F^i$ assigns to the values the equilibrium compromise probability $p_A^i = \mathcal{S}_F^i(v_A^i, v_D^i) \triangleq \bar{p}(f_A^{i\star}, f_D^{i\star})$.

The values v_A^i and v_D^i are not intrinsic constants of the service; they are the equilibrium payoffs of the communication layer.

10.5.3 Communication Layer: Signaling Game

The type of service i is $\theta_i \in \{\theta_A, \theta_D\}$, according to whether the attacker or the defender controls it, with prior $\mathbb{P}(\theta_i = \theta_A) = p_A^i$ supplied by the cloud layer. Each service sends a risk label $m_i \in \{m_L, m_H\}$, low or high, and the device chooses $a_i \in \{a_T, a_N\}$, trust or not trust. The sender strategies are σ_A^i and σ_D^i, the receiver strategy is σ_R, and the posterior $\mu_i(\theta_A \mid m_i)$ is computed by Bayes' rule wherever possible.

Definition 10.5.2 (Signaling map)**.** For a prior vector $p_A = (p_A^1, \ldots, p_A^{N_c})$, let $\mathcal{S}_S(p_A)$ denote the set of vectors (v_A, v_D) of sender utilities, $v_A = (v_A^i)_{i\in\mathcal{I}}$ and $v_D = (v_D^i)_{i\in\mathcal{I}}$, attained at the selected perfect Bayesian Nash equilibria of the signaling game. These utilities are the values passed back to the FlipIt games.

10.5.4 Physical Layer: Control-Based Utilities

Appendix C reviews the observer-based control background used here. The device has the dynamics $x[k+1] = Ax[k] + Bu[k] + w[k]$, and the cloud services report measurements $y[k] = Hx[k] + \epsilon[k]$, possibly altered by bias terms injected by attacker-controlled services. Let $\tilde{y}[k]$ be the received measurement and $\hat{x}[k]$ the observer estimate. The innovation $e[k] \triangleq \tilde{y}[k] - H\hat{x}[k]$ is labeled low risk for service i if $|e_i[k]| \leq \bar{e}_i$ and high risk otherwise. The

receiver strategy σ_R acts as a *strategic innovation filter*: trusted innovations are passed to the observer, and untrusted innovations are set to zero. The observer and the feedback controller incur the finite-horizon cost

$$J \triangleq \mathbb{E}\Big[x[k_f]^\top Q_f x[k_f] + \sum_{k=0}^{k_f-1} \big(x[k]^\top Q x[k] + u[k]^\top R u[k]\big)\Big],$$

which depends on the types, messages, and trust actions. A decreasing function maps the cost to a receiver utility, for example

$$U_R^S(\theta, m, a) \triangleq \underline{v}_R + \big(\overline{v}_R - \underline{v}_R\big) e^{-b_R J(\theta,m,a)}, \qquad b_R > 0,$$

so that good tracking yields high utility. The sender utilities are similarly tied to whether their messages are trusted and to the physical effect of the trusted data. The physical layer thus determines what it costs to trust or reject a message, and trust is evaluated by its effect on controlled performance.

10.5.5 Gestalt Nash Equilibrium

Definition 10.5.3 (Gestalt Nash equilibrium [474])**.** A triple $(p_A^\dagger, v_A^\dagger, v_D^\dagger)$ is a *Gestalt Nash equilibrium* (GNE) of iSTRICT if

$$p_A^{i\dagger} = \mathcal{S}_F^i\big(v_A^{i\dagger}, v_D^{i\dagger}\big) \quad \text{for every } i \in \mathcal{I}, \qquad (v_A^\dagger, v_D^\dagger) \in \mathcal{S}_S\big(p_A^\dagger\big). \tag{10.5.1}$$

The first condition states that the FlipIt equilibria reproduce the compromise probabilities that the signaling game uses as priors; the second states that the signaling equilibrium reproduces the values that the FlipIt games use. In the notation of Section 10.2, Definition 10.5.3 is a Gestalt equilibrium with $K = N_c + 1$ components, in which the physical layer enters through the utilities of the signaling component.

10.5.6 Key Results

The following results are due to Pawlick, Chen, and Zhu [474]. They rest on five assumptions on the signaling utilities: (A1) an untrusted message gives its sender zero utility; (A2) the attacker-controlled and defender-controlled senders have opposite preferences between trusted high-risk and trusted low-risk messages; (A3) the receiver prefers not to trust high-risk messages from attacker-controlled services; (A4) the receiver prefers to trust low-risk messages from defender-controlled services; and (A5) under normal operation, trusted low-risk messages are preferred to trusted high-risk messages.

THEOREM 10.5.4 **Structure of selected signaling equilibria [474]**
Under (A1)–(A5), every GNE satisfies $v_A^{i\dagger} > 0$ and $v_D^{i\dagger} > 0$ for each $i \in \mathcal{I}$. Every pure-strategy signaling equilibrium with positive sender utilities is pooling, and the equilibria that survive the selection criteria of [474] pool on the low-risk message m_L, with the receiver trusting or rejecting m_L according to whether p_A^i lies below or above a threshold.

Proof sketch. Suppose a GNE had $v_A^{i\dagger} = 0$. Any positive attack frequency then has a positive cost and no benefit, so the attacker's FlipIt best response is $f_A^i = 0$, and $p_A^{i\dagger} = 0$. With the prior $p_A^i = 0$, the receiver trusts low-risk messages by (A4), so an attacker-controlled sender that imitates the defender-controlled sender would be trusted and would obtain a positive utility, which contradicts $v_A^{i\dagger} = 0$. The argument for $v_D^{i\dagger}$ is analogous. For the structure, consider a separating profile. The receiver identifies the attacker-controlled sender and, by (A3), does not trust it, so that sender obtains zero utility by (A1) and gains by imitating the trusted message of the defender-controlled sender; hence no separating profile with positive utilities is an equilibrium. Among pooling profiles, (A5) and the selection criteria eliminate pooling on m_H. On m_L, the receiver's expected utility from trusting is affine in p_A^i, and the threshold is its zero. □

Since revealing compromise would lead to distrust, the attacker-controlled sender imitates the defender-controlled sender, and the receiver cannot identify the type. Its defense is a trust rule that limits the value of compromise. For fixed behavior of the other services, let $p_A^{i\diamond}$ be the threshold of Theorem 10.5.4 at which the receiver is indifferent between trusting and rejecting low-risk messages from service i.

COROLLARY 10.5.5 **Worst-case compromise probability [474]**
At every GNE, $p_A^{i\dagger} \le p_A^{i\diamond}$ for each $i \in \mathcal{I}$. The bound holds independently of the FlipIt costs c_A^i and c_D^i.

Proof. If $p_A^{i\dagger} > p_A^{i\diamond}$, the selected signaling equilibrium rejects low-risk messages from service i. The attacker-controlled sender is then never trusted and, by (A1), obtains $v_A^{i\dagger} = 0$, which contradicts Theorem 10.5.4. □

The bound does not state that the cloud is secure. It states that the device's willingness to stop trusting a risky service removes the attacker's incentive to raise the compromise probability above the threshold, whatever the cost of attacking.

THEOREM 10.5.6 **Existence of GNE [474]**
Under (A1)–(A5), iSTRICT admits a Gestalt Nash equilibrium.

Proof sketch. By Theorem 10.5.4, the values at a GNE are positive, and the FlipIt equilibrium depends on them only through the ratios $v_{AD}^i \triangleq v_A^i / v_D^i$. In these variables, the FlipIt layer defines continuous maps $\widetilde{\mathcal{S}}_F^i$, and the selected signaling equilibria, with mixed receiver strategies at the thresholds, define a correspondence $\widetilde{\mathcal{S}}_S$ on a compact convex set of ratio vectors with nonempty convex values and a closed graph. The composition satisfies the hypotheses of Kakutani's theorem [307], as in Theorem 10.2.7, and its fixed points are GNE. □

10.5.7 Computation by Iteration

A GNE can be sought by alternating the layers: starting from compromise probabilities p_A, solve the signaling game for the sender values and the trust policy, solve the FlipIt games for the new compromise probabilities, and repeat. By Proposition 10.2.9, the procedure converges to the unique GNE whenever the composed map is a contraction. The following example verifies this condition for one service.

EXAMPLE 10.4 **Fixed-point iteration for one cloud service**

▸ Code: `code/istrict_fixed_point.py`

A hospital infusion pump receives dosage-relevant measurements from one cloud analytics service. The signaling and physical layers are summarized by a selected single-valued branch of $\mathcal{S}_S$: at the prior p, the pump trusts low-risk messages with probability $q(p) = 1/\big(1 + e^{4(p-0.45)}\big)$, and the induced sender values are $v_A(p) = 1 + 5q(p)$ and $v_D(p) = 3 + q(p)$. The attacker's value rises more steeply with trust, because a trusted compromised service affects the dosage. For the cloud layer, the example replaces the periodic FlipIt game of Chapter 3 by its contest form $\bar{p}(f_A, f_D) = f_A/(f_A + f_D)$ with unit move costs. The first-order conditions $v_A f_D = (f_A + f_D)^2 = v_D f_A$ of the two players give $f_A/f_D = v_A/v_D$, so the FlipIt map is $\mathcal{S}_F(v_A, v_D) = v_A/(v_A + v_D)$, and the composed map is

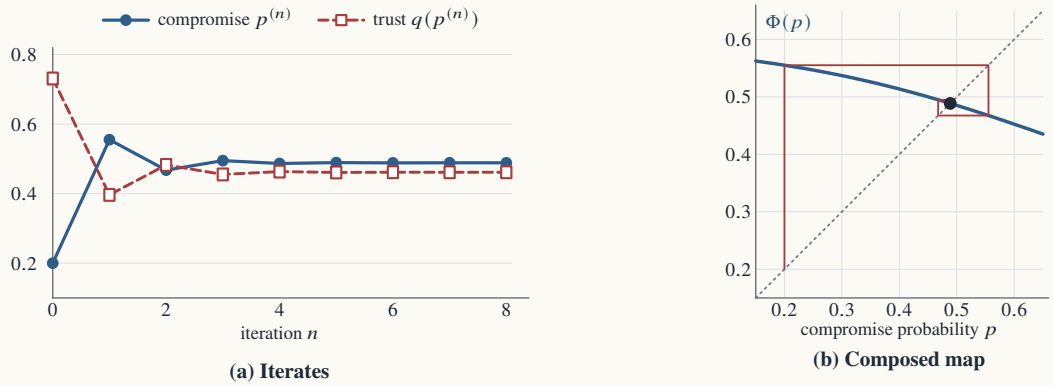
$$\Phi(p) \triangleq \mathcal{S}_F\big(v_A(p), v_D(p)\big) = \frac{1 + 5q(p)}{4 + 6q(p)}.$$

Iteration. Starting from $p^{(0)} = 0.20$, the pump trusts with probability $q(0.20) = 1/(1 + e^{-1}) = 0.731$, so $v_A = 4.655$, $v_D = 3.731$, and $p^{(1)} = 4.655/8.386 = 0.555$. High trust has made control of the service valuable, and the compromise probability rises. At $p^{(1)}$, $q = 0.396$, $v_A = 2.982$, $v_D = 3.396$, and $p^{(2)} = 0.468$: distrust has lowered the attacker's value, and the compromise probability falls. The iterates are

n	0	1	2	3	4	5	6
$p^{(n)}$	0.200	0.555	0.468	0.495	0.487	0.489	0.488
$q\big(p^{(n)}\big)$	0.731	0.396	0.482	0.455	0.463	0.461	0.462

Contraction. By the chain rule, $\Phi'(p) = \dfrac{14}{(4 + 6q)^2}\, q'(p)$ with $q'(p) = -4q(1 - q)$, so $|\Phi'(p)| = 56\,\Psi(q)$ with $\Psi(q) \triangleq q(1 - q)/(4 + 6q)^2$. The derivative $\Psi'(q)$ vanishes where $(1 - 2q)(4 + 6q) = 12q(1 - q)$, that

is, at $q = 2/7$, and $\Psi(2/7) = 1/160$. Hence $|\Phi'(p)| \le 56/160 = 0.35$ for all $p \in [0,1]$, Φ is a contraction with modulus 0.35, and Proposition 10.2.9 gives a unique GNE and convergence from every initial prior. The fixed point is $p_A^\dagger = 0.4886$, with $q^\dagger = 0.4615$, $v_A^\dagger = 3.307$, and $v_D^\dagger = 3.461$. The a priori bound $|p^{(n)} - p_A^\dagger| \le 0.35^n\,|p^{(1)} - p^{(0)}|/0.65$ guarantees an error below 10^{-3} after $n = 7$ iterations. Figure 10.10 shows the iterates and the map Φ.



(a) Iterates **(b) Composed map**

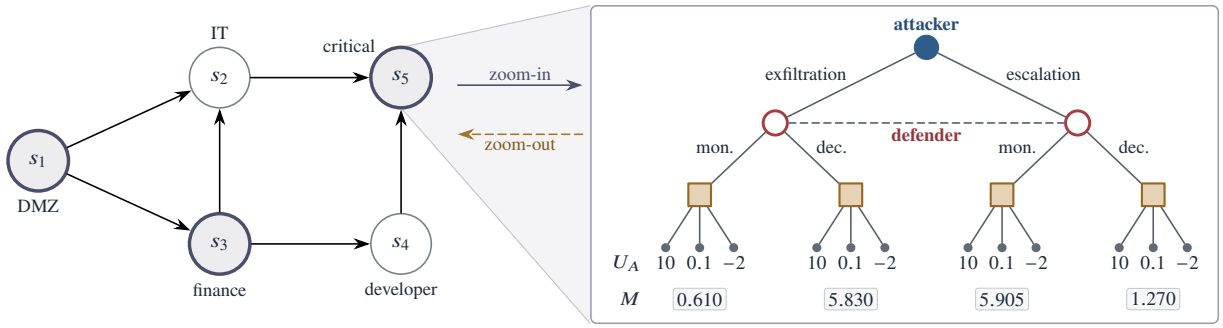
Figure 10.10: The fixed-point iteration of Example 10.4. (a) The compromise probability $p^{(n)}$ (discs, solid) and the trust probability $q(p^{(n)})$ (squares, dashed) for $n = 0, \dots, 8$; both oscillate and converge. (b) The composed map Φ (solid) and the diagonal (dotted); the staircase traces the iteration from $p^{(0)} = 0.20$, and the disc marks the unique fixed point $p_A^\dagger = 0.4886$.

Observation. A trust policy designed with the cloud layer held at its initial estimate $p = 0.20$ trusts the service with probability 0.731 and thereby raises the attacker's equilibrium share of control to 0.555. The consistent policy trusts with probability 0.4615 and holds the compromise probability at 0.489. The oscillation of the first iterates shows the two feedback effects: trust makes compromise valuable, and compromise erodes trust. The contraction modulus 0.35 is small because the trust response is flat away from the threshold and the FlipIt map is insensitive to proportional changes in the two values.

In the autonomous-vehicle application of [474], two vehicles use observer-based feedback control and receive measurements from several sources, some of them cloud services that may be compromised. The strategic innovation filter rejects high-risk innovations, so a strategic attacker shifts to low-risk messages, which are more likely to be trusted but have smaller physical impact. In the hardest parameter regime, mixed trust strategies serve as a last-resort countermeasure that enforces the worst-case compromise probability of Corollary 10.5.5 while preserving enough information for control.

10.6 Multi-Resolution Dynamic Games for Cyber Campaigns

The compositions of the preceding sections couple games that describe different functions of a system. This section couples games that describe the same function at different resolutions, following the multi-resolution framework of Yang and Zhu [610] for cross-echelon decisions in cyber campaigns. The tactical echelon needs high-resolution models of concrete techniques, such as those cataloged in MITRE ATT&CK [568]; the strategic echelon needs a low-resolution model of movement across an enterprise network toward mission-critical assets. A game tree of the entire enterprise is too large to solve or interpret, while a purely abstract Markov game can miss the tactical mechanisms that decide whether a movement succeeds. The framework links a *macro-strategic game* (MSG), a Markov game whose states are network components [528, 198], with *micro-base games* (MBGs), extensive-form games that describe the engagement inside a component [337, 205]. A *zoom-in* operation uses the macro value to define the terminal payoffs of a micro game, and a *zoom-out* operation uses the outcome distribution of the solved micro game to update the macro game. Figure 10.11 shows the two resolutions.



(a) Macro-strategic game $\Gamma_{\mathcal{G}}$ **(b) Micro-base game Γ^{s_5}**

Figure 10.11: Two resolutions of the same enterprise. (a) The macro-strategic game on the network graph: a DMZ web server s_1, an IT site s_2, a finance site s_3, a developer site s_4, and a critical server s_5; the thick circles are the components zoomed in Example 10.5. (b) The micro-base game Γ^{s_5}, opened from s_5 as through a lens: the attacker (disc) chooses a tactic, the defender (rings) responds at an information set (dashed) that does not reveal the tactic, and a chance node (squares) selects the outcome database, decoy, or block, with attacker payoffs $U_A = (10, 0.1, -2)$. The boxed numbers are the reduced payoffs $M^{s_5}(a, \ell)$. Zoom-in passes the macro value V^{Π} down as terminal payoffs U_A^s; zoom-out (dashed) passes the outcome distribution τ^s up as a new macro row.

10.6.1 Micro Layer: Extensive-Form Tactical Games

The players are the attacker P_A and the defender P_D. Inside a network component, they interact through tactical moves such as reconnaissance, exploitation, deception, isolation, and recovery.

Definition 10.6.1 (Micro-base game)**.** For a component s, the *micro-base game* is $\Gamma^s \triangleq \langle \mathcal{E}^s, \sigma_c^s, U_A^s, U_D^s \rangle$ with the game form

$$\mathcal{E}^s \triangleq \langle \mathsf{T}^s, \prec^s;\ \{\mathrm{P}_A, \mathrm{P}_D\},\ I^s;\ \mathcal{A}^s, \alpha^s;\ \mathcal{H}^s \rangle.$$

Here $(\mathsf{T}^s, \prec^s)$ is a finite game tree with terminal nodes $\mathcal{Z}^s \subset \mathsf{T}^s$; I^s assigns the decision nodes to players or to chance; $\mathcal{A}^s$ and α^s describe the feasible actions at each node; $\mathcal{H}^s$ is the partition of the decision nodes into information sets; σ_c^s is the distribution of the chance moves; and $U_A^s, U_D^s : \mathcal{Z}^s \to \mathbb{R}$ are the terminal payoffs.

A behavior strategy of P_i, $i \in \{A, D\}$, maps each information set $h \in \mathcal{H}_i^s$ of that player to a distribution $\sigma_i^s(h) \in \Delta(\mathcal{A}^s(h))$. A profile $\sigma^s \triangleq (\sigma_A^s, \sigma_D^s, \sigma_c^s)$ induces the outcome distribution $\tau^s(\cdot \mid \sigma^s) \in \Delta(\mathcal{Z}^s)$ and the expected payoffs $U_i^s(\sigma^s) \triangleq \sum_{z \in \mathcal{Z}^s} \tau^s(z \mid \sigma^s)\, U_i^s(z)$. The profile is a Nash equilibrium of Γ^s if no player can increase U_i^s by changing its own behavior strategy, and a subgame perfect equilibrium if this holds in every subgame. The micro games are assumed to have perfect recall, so that, by Kuhn's theorem [337], behavior and mixed strategies are equivalent; every finite MBG then has a Nash equilibrium in behavior strategies, and a subgame perfect one [205].

10.6.2 Macro Layer: Markov Strategic Game

Let $\mathcal{V}$ be the set of network components and $E \subseteq \mathcal{V} \times \mathcal{V}$ a directed edge set of feasible movements. The graph $\mathcal{G} = \langle \mathcal{V}, E \rangle$ is the strategic abstraction of the enterprise, and each $s \in \mathcal{V}$ corresponds to an MBG Γ^s.

Definition 10.6.2 (Macro-strategic game)**.** The *macro-strategic game* is the zero-sum Markov game $\Gamma_{\mathcal{G}} \triangleq \langle \{\mathrm{P}_A, \mathrm{P}_D\}, \mathcal{V}, \mathcal{A}_A^M, \mathcal{A}_D^M, T, U_A, \beta \rangle$. At state s, the attacker chooses an edge $a_A = (s, \hat{s}) \in E$, an attempt to move from s to $\hat{s}$, and the defender chooses an edge $a_D = (s, \check{s}) \in E$ to protect. With attacker capability $\psi_A \in [0, 1]$, the transition kernel is

$$T\big(s' \mid s, (s, \hat{s}), (s, \check{s})\big) = \begin{cases} 1, & s' = s \text{ and } (\hat{s} = s \text{ or } \hat{s} = \check{s}), \\ \psi_A, & s' = \hat{s} \text{ and } \hat{s} \notin \{s, \check{s}\}, \\ 1 - \psi_A, & s' = s \text{ and } \hat{s} \notin \{s, \check{s}\}, \\ 0, & \text{otherwise.} \end{cases}$$

A protected edge blocks the move, and an unprotected attempt succeeds with probability ψ_A. The attacker receives $U_A(s, a_A, a_D, s') = \underline{U} < 0$ if $s' = s$ and the vertex reward $\bar{U}(s') > 0$ if $s' \neq s$; the defender receives $-U_A$. Future rewards are discounted by $\beta \in (0, 1)$.

For stationary strategies $\pi_A : \mathcal{V} \to \Delta(\mathcal{A}_A^M)$ and $\pi_D : \mathcal{V} \to \Delta(\mathcal{A}_D^M)$ with profile $\Pi = (\pi_A, \pi_D)$, the attacker's value satisfies the Bellman equation

$$V^{\Pi}(s) = \sum_{a_A} \pi_A(a_A \mid s) \sum_{a_D} \pi_D(a_D \mid s) \sum_{s' \in \mathcal{V}} T(s' \mid s, a_A, a_D)\big[U_A(s, a_A, a_D, s') + \beta V^{\Pi}(s')\big]. \tag{10.6.1}$$

A stationary profile $\Pi^\star$ is a saddle point of the MSG if $V^{(\pi_A, \pi_D^\star)}(s) \le V^{\Pi^\star}(s) \le V^{(\pi_A^\star, \pi_D)}(s)$ for every $s \in \mathcal{V}$ and all stationary π_A and π_D.

PROPOSITION 10.6.3 **Two solution layers**
Every finite MBG with perfect recall has a Nash equilibrium in behavior strategies. The MSG has a unique value function $V^\star$ and a stationary saddle point.

Proof. The strategic form of a finite MBG is a finite game, so Nash's theorem [418] gives a mixed equilibrium, and Kuhn's theorem converts it into an outcome-equivalent behavior profile. For the MSG, the Shapley operator $(\mathcal{T}V)(s) \triangleq \operatorname{val}\big[\sum_{s'} T(s' \mid s, a_A, a_D)(U_A + \beta V(s'))\big]_{a_A, a_D}$ is a β-contraction in the supremum norm, by the argument of Theorem 10.3.4; its fixed point is $V^\star$, and optimal strategies of the stage matrix games form a stationary saddle point [528, 198]. □

10.6.3 Resolution Configurations and Zoom Operations

Definition 10.6.4 (Resolution configuration)**.** A *resolution configuration* selects, for each $s \in \mathcal{V}$, either the abstract state s or the micro game Γ^s. The set of configurations is $C_{\mathcal{V}} \triangleq \prod_{s \in \mathcal{V}} \{s, \Gamma^s\}$, so $|C_{\mathcal{V}}| = 2^{|\mathcal{V}|}$. The configuration in which every component is abstract is the *completely abstracted game*.

Changing the configuration does not change the system; it changes how much of the system is modeled explicitly. The two operations below move between configurations and transfer information across resolutions.

Definition 10.6.5 (Zoom-in)**.** Given a macro profile $\Pi = (\pi_A, \pi_D)$ with value V^{Π}, the zoom-in operation $\mathrm{OP}_{\mathrm{in}}$ replaces the abstract state s by Γ^s and assigns to a terminal outcome $z \in \mathcal{Z}^s$ that corresponds to the attempted move to s' the payoff

$$U_A^s(z) \triangleq \sum_{a_D} \pi_D(a_D \mid s) \sum_{\bar{s} \in \mathcal{V}} T\big(\bar{s} \mid s, (s, s'), a_D\big)\big[U_A(s, (s, s'), a_D, \bar{s}) + \beta V^{\Pi}(\bar{s})\big], \tag{10.6.2}$$

with $U_D^s(z) \triangleq -U_A^s(z)$.

The terminal payoff of a micro outcome is thus the strategic value of the macro consequence that the outcome produces. When the micro game itself resolves whether the move is blocked, as in Example 10.5 below, the macro blocking in (10.6.2) is removed and the payoff of a successful move to s' reduces to $\bar{U}(s') + \beta V^{\Pi}(s')$.

Definition 10.6.6 (Zoom-out)**.** Let Γ^s be solved with the profile σ^s. The zoom-out operation $\mathrm{OP}_{\mathrm{out}}$ replaces Γ^s by the abstract state s and updates the macro attack strategy at s by the micro outcome distribution:

$$\pi_A\big((s, s'(z)) \mid s\big) = \tau^s(z \mid \sigma^s), \qquad z \in \mathcal{Z}^s, \tag{10.6.3}$$

where $s'(z)$ is the component reached at the outcome z.

A local tactical solution thereby becomes a macro transition tendency. The two operations define a two-component composed game at each zoomed state: the MSG, with context the row τ^s and outcome V^{Π}, and the MBG Γ^s, with context the terminal payoffs U^s and outcome (σ^s, τ^s).

PROPOSITION 10.6.7 **Resolution consistency**
Fix a state s and the macro strategies at all other states. A pair (V^{Π}, σ^s) is a Gestalt equilibrium of the two-component composed game defined by zoom-in (10.6.2) and zoom-out (10.6.3) if and only if σ^s is an equilibrium of Γ^s with the terminal payoffs computed from V^{Π}, and V^{Π} solves (10.6.1) with the row at s given by $\tau^s(\cdot \mid \sigma^s)$. Repeated application of $\mathrm{OP}_{\mathrm{in}}$ and $\mathrm{OP}_{\mathrm{out}}$ at s is the fixed-point iteration of this composed game.

Proof. The zoom-in map is the coupling $C_{\mathrm{micro}}(V^{\Pi}, \sigma^s) = U^s$, and the zoom-out map is the coupling $C_{\mathrm{macro}}(V^{\Pi}, \sigma^s) = \tau^s(\cdot \mid \sigma^s)$. The two conditions of the statement are $\sigma^s \in \mathcal{S}_{\mathrm{micro}}(C_{\mathrm{micro}})$ and $V^{\Pi} \in \mathcal{S}_{\mathrm{macro}}(C_{\mathrm{macro}})$, which is (10.2.2). Applying the two operations in turn evaluates $\mathcal{F}$ once. □

Definition 10.6.8 (Operation plan)**.** A *multi-resolution operation plan* is a finite sequence

$$\mathcal{P} \triangleq \big((\mathrm{OP}_1, s_1), \ldots, (\mathrm{OP}_L, s_L)\big)$$

with $\mathrm{OP}_n \in \{\mathrm{OP}_{\mathrm{in}}, \mathrm{OP}_{\mathrm{out}}\}$, in which every zoom-out on a state is preceded by a zoom-in on the same state.

Algorithm 25 executes an operation plan. It starts from the completely abstracted game, zooms into tactically important states, solves their micro games, and propagates the results back to the macro game. Because only selected states are refined, the defender can inspect local engagements in detail, as in purple-team exercises, without expanding the entire strategic model.

Algorithm 25 Zoom-in and zoom-out computation

Require: Macro game $\Gamma_{\mathcal{G}}$, micro games $\{\Gamma^s\}_{s \in \mathcal{V}}$, initial macro profile Π^0, operation plan $\mathcal{P}$
Evaluate (10.6.1) under Π^0 to obtain V^{Π^0}
for $n = 1, \ldots, L$ **do**
  Read the operation (OP_n, s_n) from $\mathcal{P}$
  if $\mathrm{OP}_n = \mathrm{OP}_{\mathrm{in}}$ **then**
    Assign $U_A^{s_n}(z)$ by (10.6.2) and $U_D^{s_n} = -U_A^{s_n}$
    Solve Γ^{s_n} and record σ^{s_n} and $\tau^{s_n}(\cdot \mid \sigma^{s_n})$
  else
    Update $\pi_A(\cdot \mid s_n)$ by (10.6.3) and re-evaluate (10.6.1) to obtain V^{Π}
  end if
end for
return the macro profile Π, the value V^{Π}, and the micro solutions $\{\sigma^{s_n}, \tau^{s_n}\}$

EXAMPLE 10.5 **Zooming into three components of an enterprise**

▸ Code: `code/multiresolution_zoom.py`

Consider the five-component enterprise of Figure 10.11(a) with $\beta = 0.9$. The current macro evaluation gives $V^{\Pi}(s_2) = 5$, $V^{\Pi}(s_3) = 3$, $V^{\Pi}(s_4) = 8$, and $V^{\Pi}(s_5) = 10$. The defender executes the plan

$$\mathcal{P} = \big((\mathrm{OP}_{\mathrm{in}}, s_1), (\mathrm{OP}_{\mathrm{out}}, s_1), (\mathrm{OP}_{\mathrm{in}}, s_3), (\mathrm{OP}_{\mathrm{out}}, s_3), (\mathrm{OP}_{\mathrm{in}}, s_5), (\mathrm{OP}_{\mathrm{out}}, s_5)\big),$$

zooming into the DMZ server, the finance site, and the critical server, where deception can change the course of the campaign.

Zoom-in. Each zoomed component has three terminal outcomes: a successful move, a diverted move, and a block. A successful move to s' earns $\bar{U}(s') = 1$ and the continuation value, so its payoff is $1 + 0.9\,V^{\Pi}(s')$. A block evicts the attacker into an absorbing state of value 0 and pays $\underline{U} = -2$. At s_1, the outcomes reach s_2, reach s_3, or are blocked, with payoffs $U_A^{s_1} = (5.5,\ 3.7,\ -2)$. At s_3, they reach s_4, reach s_2, or are blocked, with $U_A^{s_3} = (8.2,\ 5.5,\ -2)$. At s_5, the attacker reaches the real database, which ends the campaign with payoff 10;

reaches a decoy, which reveals it and pays 0.1; or is blocked, with $U_A^{s_5} = (10,\ 0.1,\ -2)$.

Micro games. In each Γ^s, the attacker chooses one of two tactics, the defender chooses between monitoring and deception at an information set that does not reveal the tactic, and a chance node selects the outcome with probability $P^s(z \mid a, \ell)$. The outcome distribution (10.6.4) and the attacker's expected payoff (10.6.5) are

$$\tau^s(z \mid \sigma^s) = \sum_a \sum_\ell \sigma_A^s(a)\, \sigma_D^s(\ell)\, P^s(z \mid a, \ell), \tag{10.6.4}$$

$$U_A^s(\sigma^s) = \sum_a \sum_\ell \sigma_A^s(a)\, \sigma_D^s(\ell)\, M^s(a, \ell), \qquad M^s(a, \ell) \triangleq \sum_{z \in \mathcal{Z}^s} P^s(z \mid a, \ell)\, U_A^s(z), \tag{10.6.5}$$

so each micro game reduces to the zero-sum matrix game M^s. Table 10.3 gives the chance probabilities at the critical server s_5, where the tactics are exfiltration through a service account and privilege escalation.

Solution at s_5. The pure security levels are $\max_a \min_\ell M^{s_5} = 1.270$ and $\min_\ell \max_a M^{s_5} = 5.830$, so the game has no pure saddle point. Let g be the probability of exfiltration and q the probability of monitoring. Equalization gives $0.610g + 5.905(1-g) = 5.830g + 1.270(1-g)$, so $g^\star = 4.635/9.855 = 0.470$, and $0.610q + 5.830(1-q) = 5.905q + 1.270(1-q)$, so $q^\star = 4.560/9.855 = 0.463$. The value is $U_A^{s_5} = 3.415$, and (10.6.4) gives $\tau^{s_5} = (0.397,\ 0.309,\ 0.294)$ over (database, decoy, block). The micro games at s_1 and s_3 are solved in the same way; Table 10.4 collects the results. In every zoomed component, the defender randomizes almost evenly between monitoring and deception, because each response is effective against one tactic and ineffective against the other.

Zoom-out. The coarse MSG had used the abstract rows $(0.55, 0.25, 0.20)$ at s_1, $(0.65, 0.20, 0.15)$ at s_3, and $(0.70, 0, 0.30)$ at s_5, whose values under the same payoffs are 3.550, 6.130, and 6.400. Zoom-out replaces them by τ^{s_1}, τ^{s_3}, and τ^{s_5}. At s_5, for example, the macro attack strategy becomes $\pi_A((s_5, \text{database}) \mid s_5) = 0.397$, $\pi_A((s_5, \text{decoy}) \mid s_5) = 0.309$, and $\pi_A((s_5, \text{block}) \mid s_5) = 0.294$. Figure 10.12 compares the values.

Observation. The total attacker value at the three zoomed components falls from 16.080 in the abstract model to 9.037 after zoom-out, a reduction of 7.043. The largest reduction, 2.985, occurs at the critical server, where the abstract model cannot represent the decoy and assigns probability 0.70 to reaching the database; the micro game lowers this probability to 0.397. Refining three of the five components suffices to correct the campaign-level values, and the corrected rows enter the next macro evaluation through (10.6.1).

Attacker tactic	Defender response	database	decoy	block	$M^{s_5}(a, \ell)$
exfiltration	monitor	0.20	0.10	0.70	0.610
exfiltration	deceive	0.60	0.30	0.10	5.830
escalation	monitor	0.65	0.05	0.30	5.905
escalation	deceive	0.15	0.70	0.15	1.270

Table 10.3: Chance probabilities $P^{s_5}(z \mid a, \ell)$ and the reduced payoffs $M^{s_5}(a, \ell)$ of the micro game at the critical server in Example 10.5.

State	Tactics	σ_A^s	σ_D^s	τ^s	U_A^s	Abstract
s_1	exploit, credential reuse	$(0.460, 0.540)$	$(0.447, 0.553)$	$(0.375, 0.227, 0.397)$	2.110	3.550
s_3	phishing, VPN reuse	$(0.500, 0.500)$	$(0.517, 0.483)$	$(0.375, 0.225, 0.400)$	3.513	6.130
s_5	exfiltration, escalation	$(0.470, 0.530)$	$(0.463, 0.537)$	$(0.397, 0.309, 0.294)$	3.415	6.400

Table 10.4: Saddle points of the three micro games of Example 10.5. The defender's strategy is over (monitor, deceive), the outcome distribution τ^s is over (success, diversion or decoy, block), and the last column is the value of the abstract row.

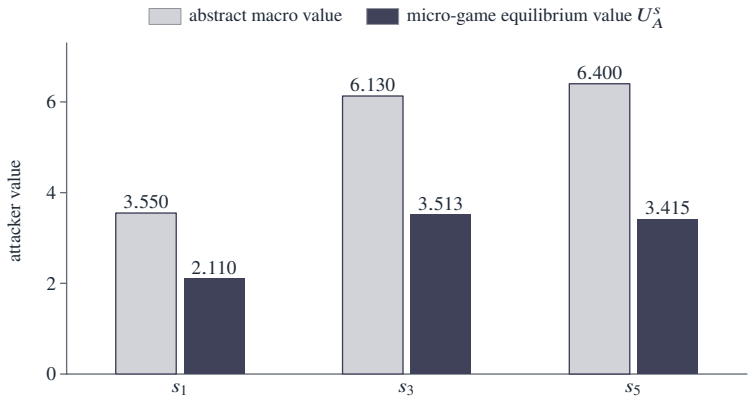


Figure 10.12: Attacker values at the three zoomed components of Example 10.5. Light bars are the values of the abstract rows of the coarse macro game; dark bars are the equilibrium values U_A^s of the micro games after zoom-out. The printed numbers are the values.

10.7 Discussion, Summary, and Further Reading

The four compositions of this chapter use different component games and different coupling variables: transition rates and values in the games-in-games principle, beliefs and MPC costs in impact-aware signaling, compromise probabilities and sender values in iSTRICT, and terminal payoffs and outcome distributions in multi-resolution games. They share one structure. Each component is solved by the solution concept appropriate to it, the coupling maps carry its outcome into the contexts of the others, and the Gestalt equilibrium requires that local solutions and global contexts reproduce one another. The modeling discipline that follows is to identify the components and the coupling variables before writing any equilibrium equation.

Selection and multiplicity. Gestalt equilibria need not be unique: the components may have several solutions, and the fixed-point condition may admit several consistent combinations, as Exercise 10.2 shows. Selection among Gestalt equilibria is an open design question. The iSTRICT selection criteria, which retain low-risk pooling equilibria, are one answer; stability under the iteration of Proposition 10.2.9 is another; and the welfare of a designer who can choose the coupling maps, as in the mechanism-design view of Chapter 9, is a third.

Agentic AI and autonomous cyber defense. Defensive systems increasingly delegate monitoring, triage, deception, and response to AI agents. Each agent may solve a local game well, yet the defense as a whole requires consistency among the agents' contexts: an alert-triage agent that assumes a stationary attacker and a deception agent that changes the attacker's behavior are coupled components. A composed-game model identifies when locally optimal agent policies support or undermine mission-level security. A Gestalt game-theoretic framework for designing agentic AI workflows in cyber deception is proposed in [7].

Multi-agent language-model systems. Language-model assistants act as planners, analysts, code reviewers, penetration testers, and incident responders. Their outputs form a closed loop: the planner's decomposition determines the analyst's evidence, the evidence determines the responder's action, and the action changes the attacker's next move. Gestalt equilibrium provides a consistency criterion for such pipelines, and the contraction condition gives a sufficient condition for their iterative coordination to settle.

Cognitive and campaign-level security. Chapters 11 and 13 model bounded rationality, attention, and deception. Cognitive modules can enter a composed game as components whose outcomes are beliefs, attention allocations, or conjectures, coupled to the cyber state and the physical consequence. At the other end of the scale, the multi-resolution framework of Section 10.6 can be extended upward to connect tactical games with operational campaigns, mission objectives, deterrence, and geopolitical constraints [610].

Connections to later chapters. Chapter 11 supplies behavioral models of the players in the components, which change the solution correspondences $\mathcal{S}_r$ but not the fixed-point structure. Chapter 12 applies cross-layer compositions to

industrial control systems, where the enterprise, operational, and process layers of Section 10.1 are modeled explicitly. Chapter 13 couples deception games with the attacker's learning, Chapter 14 composes device-level and network-level games in the Internet of Things, Chapter 15 composes the games of the tiers of a supply chain, and Chapter 16 couples physical-layer and network-layer games in next-generation wireless networks.

Connections to appendices. Appendix A supplies the convex-analysis background of the fixed-point arguments, including the properties of correspondences used in Theorem 10.2.7. Appendix B supplies the Markov-chain background of the jump process and the uniformization in Section 10.3. Appendix C supplies the linear-quadratic, Riccati, observer, and model-predictive-control background of the physical components in Sections 10.3, 10.4, and 10.5.

Further reading. The component models are classical: FlipIt [577], signaling games and sequential equilibrium [548, 336], stochastic games [528, 198], extensive-form games [337], differential games [294, 46], and minimax control [43]; the fixed-point theorem is Kakutani's [307]. The games-in-games principle for cross-layer resilient control is due to Zhu and Başar [660]; Pan and Zhu [461] extend it to regime-switching jump-diffusions, in which an outer game modulates the transition intensities of the mode process, the hierarchy yields coupled HJI equations, and linear-quadratic and exponential-affine games admit semi-closed-form solutions. Strategic trust in cloud-enabled control and the Gestalt Nash equilibrium are developed in the iSTRICT framework of Pawlick, Chen, and Zhu [474]. The coupling of signaling with model predictive control for autonomous agents is due to Xu and Zhu [603], with MPC background in Mayne et al. [388] and Rawlings, Mayne, and Diehl [498]. The micro–macro meta-game for penetration testing is MEGA-PT [214], and the multi-resolution dynamic game for cross-echelon cyber decisions is due to Yang and Zhu [610], who also outline a meta-game-theoretic paradigm for multi-echelon cyber warfare that spans policy, strategy, operations, tactics, and technical implementation [612]. Recent compositional formalisms include game Petri nets, which attach local games to the places of a colored Petri net, characterize the equilibria of sequential nets as equilibria of local games augmented by continuation values, and use an intrusion campaign derived from multi-resolution games as a running example [647], and Bayesian holonic systems, whose equilibrium requires consistency between agent-level rationality and system-level outcomes and is computed by a two-time-scale algorithm that separates intra-holon coordination from inter-holon belief adaptation [459]. Alpcan and Başar [12] provide the broader context of network-security games.

Summary. A composed game consists of component games, solution correspondences, and coupling maps, and its Gestalt equilibrium is a fixed point of the correspondence that solves every component under the context generated by the joint outcome. Existence follows from Kakutani's theorem when the outcome spaces are convex, which may require mixed outcomes even when every component has a pure solution; uniqueness and the convergence of iterative computation follow from a contraction condition. The games-in-games principle couples a cyber stochastic game with a physical differential game through transition rates and values; impact-aware signaling couples trust with model predictive control through danger zones; iSTRICT couples FlipIt, signaling, and control through compromise probabilities and values; and multi-resolution games couple tactical and strategic games through zoom operations. In each worked example, the composed solution differs materially from the prediction of a model that holds the other components fixed.

10.8 Exercises

The exercises below apply the concepts of this chapter to concrete security settings. Each exercise states its setting and the results to be derived. An exercise marked with an asterisk (*) is an open-ended design problem.

Exercise 10.1. **A costlier verification.** In Example 10.1, the shuttle's verification cost rises to $c_V = 1$ and the attacker's exploitation cost rises to $c_x = 1.5$; all other data are unchanged.

(a) Derive the trust correspondence $\mathcal{S}_T(p)$ and its threshold.

(b) Show that monitoring still strictly dominates hardening for every $W \in [0, 7]$, and derive $\mathcal{S}_C(W)$.

(c) Compute the unique Gestalt equilibrium $(\bar{p}_C^\star, q^\star, g^\star)$.

(d) Compare the vehicle's expected loss at the equilibrium with the loss predicted and the loss realized by a siloed design that solves the cyber game with $W = 0$.

Exercise 10.2. **Multiple cross-layer equilibria.** Modify Example 10.2 so that $A_{\mathrm{N}} = 0.9$, $A_{\mathrm{C}} = 1.3$, and the probability of entering C is

$$p_{\mathrm{NC}}(q, g) = 0.12 - 0.04q + 0.48g - 0.32qg.$$

(a) Compute $P(\ell, a)$ for the four pure cyber profiles, and the optimal gain $\kappa^\star(p)$ and value $Z^\star(p)$ for each.

(b) Show that attacking strictly dominates for every gain $\kappa < 1.1$, and that against an attack hardening is strictly better if and only if $c_h < 0.288(1.1 - \kappa)$.

(c) For $c_h = 0.06$, find all Gestalt equilibria in pure cyber strategies.

(d) For $c_h = 0.16$, show that both (harden, attack) and (light, attack) are Gestalt equilibria, and find a third equilibrium in which the operator hardens with probability $q \in (0, 1)$.

(e) Explain why the coupling, and not either component alone, creates the multiplicity.

Exercise 10.3. **Coupled Riccati equations for a compromise mode.** A scalar plant $\dot{x} = A_i x + u$ operates in a normal mode 1 with $A_1 = -1$ and a compromised mode 2 with $A_2 = 0.5$. Mode 2 is absorbing, and the rate from mode 1 to mode 2 is $\lambda = 0.1$ if the operator hardens and $\lambda = 0.5$ otherwise. The cost rate is $x^2 + u^2$, and there is no disturbance.

(a) Specialize (10.3.7) to this system and show that $Z_2 = A_2 + \sqrt{A_2^2 + 1}$ and $Z_1 = (A_1 - \lambda/2) + \sqrt{(A_1 - \lambda/2)^2 + 1 + \lambda Z_2}$.

(b) Compute Z_2, and Z_1 for both values of λ.

(c) From the initial state $x_0 = 1$, find the largest hardening cost for which hardening is optimal.

(d) Explain which coupling term of (10.3.7) transmits the cyber decision to the physical value.

Exercise 10.4. **Contraction in an iSTRICT loop.** A water-treatment controller trusts low-risk messages of one cloud service with probability $q(p) = 1/\big(1 + e^{6(p-0.4)}\big)$, and the sender values are $v_A(p) = 1 + 4q(p)$ and $v_D(p) = 2 + 2q(p)$. The cloud layer uses the contest-form FlipIt map $\mathcal{S}_F(v_A, v_D) = v_A/(v_A + v_D)$.

(a) Write the composed map Φ as a function of q.

(b) Starting from $p^{(0)} = 0.10$, compute $p^{(1)}$, $p^{(2)}$, and $p^{(3)}$.

(c) Show that $|\Phi'(p)| \le 1/3$ for all $p \in [0, 1]$, and identify the trust probability at which the bound is attained.

(d) Compute the GNE to three decimals, and use Proposition 10.2.9 to bound the number of iterations needed for an error below 10^{-4}.

Exercise 10.5. **Zooming into a jump server.** In a macro game with $\beta = 0.9$, an attacker at a jump server s_0 can reach the domain controller s_1, with $V^{\Pi}(s_1) = 8$, or a quarantined subnet s_2, with $V^{\Pi}(s_2) = -1$, or be blocked. A successful move earns $\bar{U} = 1$, and a block evicts the attacker with payoff -2. The attacker chooses a stealthy or a fast tactic, and the defender monitors or deceives without observing the tactic. The outcome probabilities $P(z \mid a, \ell)$ over (reach s_1, reach s_2, block) are $(0.20, 0.30, 0.50)$ for (stealthy, monitor), $(0.50, 0.30, 0.20)$ for (stealthy, deceive), $(0.60, 0.20, 0.20)$ for (fast, monitor), and $(0.10, 0.60, 0.30)$ for (fast, deceive).

(a) Compute the zoom-in payoffs $U_A^{s_0}(z)$ and the reduced matrix M^{s_0}.

(b) Show that M^{s_0} has no pure saddle point, and compute the mixed saddle point and the value.

(c) Compute the outcome distribution τ^{s_0} and write the zoom-out update of $\pi_A(\cdot \mid s_0)$.

(d) An abstract model assigns probability 0.70 to reaching s_1 and 0.30 to a block. Compare its value with the micro-informed value, and explain the difference.

Exercise 10.6. *** A composed-game design memo.** Choose one domain: autonomous vehicles, power-grid defense, cloud-enabled medical devices, enterprise penetration testing, or agentic AI defense.

(a) Name at least two component games, the outcome of each, and the solution concept used for each.

(b) Specify the coupling maps and state the Gestalt equilibrium condition (10.2.2) for the architecture.

(c) Give a numerical instance with at least two actions per player and one scalar coupling variable, and compute one complete round of the composed correspondence.

(d) State whether Assumption 10.2.6 or the contraction condition of Proposition 10.2.9 holds for the instance, and justify the answer.

Part III

Applications of Game Theory in Cyber-Physical Systems

Chapter 13

Cyber Deception

Cyber deception studies how defenders and attackers intentionally shape what an adversary observes, infers, and believes. In cyberspace, deception is not an exceptional tactic but a structural feature of conflict: attackers spoof identities, hide intent, and stage multi-step campaigns, while defenders deploy honeypots, honey accounts, decoys, moving-target defense, obfuscation, deceptive routing, and selective disclosure. A firewall blocks traffic and an intrusion detector raises an alert, but a planted credential, a believable honeypot, or a rotating address acts on a different object, the adversary's information environment, and it has value only through the attacker's strategic response. A decoy that is recognized at once has no effect, a false signal can backfire when the attacker reads it as evidence of monitoring, and a high-interaction honeypot that collects intelligence also creates operational risk.

The chapter treats cyber deception as a game of incomplete information in which the defender designs signals, environments, and timing, and the attacker forms beliefs and best responds [299, 479]. Section 13.1 defines deception and its objectives. Section 13.2 analyzes a given deception environment as a signaling game with evidence, whose detector determines when pooling survives and when the equilibrium is partially separating, and then designs an overt information policy by Bayesian persuasion and concavification. Section 13.3 reverses the direction of reasoning and designs information, incentives, and trust jointly through generator–modulator–manipulator mechanisms and duplicity games. Section 13.4 adds time, so that beliefs become state variables and engagement competes with containment, and Section 13.5 treats moving-target defense as deception through controlled change of the attack surface, with saddle-point and soft-min feedback policies. Examples include honey files, an insider repository, a deceptive probe, and a randomized web service.

The chapter builds on the signaling games of Chapter 3, the leaky-deception model of Chapter 4, the learning dynamics of Chapter 7, and the mechanism design of Chapter 9, and on the game-theoretic literature on information design and strategic learning [625, 267]; Chapter 12 develops the multi-stage game against advanced persistent threats in full. The main results are proved in the text.

Chapter Overview

1. Introduction to cyber deception (Section 13.1)
   - deception as planned manipulation of observations, beliefs, and actions;
   - defensive objectives, common techniques, the deception chain, and AI-enabled deception.
2. Game-theoretic analysis (Section 13.2)
   - signaling games with evidence, detector-dependent belief thresholds, pooling and partially separating equilibria, and the equilibrium eye;
   - Bayesian persuasion, Bayesian plausibility, and concavification of the defender's value.
3. Mechanism design for deception (Section 13.3)
   - generator–modulator–manipulator mechanisms and duplicity games;
   - incentive margins, enforceable action profiles, trust margins, and manageability of users.

4. Dynamic deception (Section 13.4)
   - dynamic Bayesian games, beliefs as state variables, and engagement versus containment;
   - multi-stage deception against advanced persistent threats and adaptive deception policies.
5. Moving-target defense (Section 13.5)
   - layer games over attack surfaces, saddle points, and exposure misalignment;
   - feedback learning with the soft-min update and its rest points.
6. Discussion, summary, further reading, and exercises (Sections 13.6 and 13.7)

Chapter Summary

Deception pays only when the information it creates or withholds changes an action that the adversary would otherwise take. In a signaling game with evidence, the detector's quality and aggressiveness partition the attacker's prior into five regimes. Pooling survives exactly where the attacker's response ignores the evidence; in the middle regime, a leaky deception is sustained only by randomization on both sides, and the area of the equilibrium eye measures how far the two defender types can behave differently. When the defender can commit to an overt policy, Bayesian plausibility is the only constraint on the posteriors it can induce, and its optimal value is the concave closure of its payoff. The best policy is then neither silence nor full revelation but a controlled distribution of suspicion, in which some artifacts are just credible enough to invite engagement. Incentive margins and best-response regions determine which type-contingent behavior any mechanism can enforce, and information alone helps the defender only when its interests and the user's are neither aligned nor opposed. Over time, a probe or a period of engagement is worth its cost only if it moves a belief across a decision threshold, and a myopic defender engages too long because the cost of an escape route appears only in the continuation value. Moving-target defense lowers the worst-case risk by randomizing the attack surface, and its soft-min update, whose completely mixed rest points are saddle points of the layer game, recovers from a change in the threat at the price of an exploration floor. Throughout, hard safety constraints bound what a deception may do, however favorable its value.

Notation and Scope

Book conventions. The chapter uses the common notation of the book, introduced in Chapter 3 and listed in Appendix F.
Key symbols. In the signaling models, $\theta \in \Theta$ is the hidden system state with prior ρ and $p \triangleq \rho(1)$, m a message, e detector evidence with alarm probabilities α, β, μ the attacker's posterior, and $\pi(m \mid \theta)$ an information policy. The mechanism model adds the user's type $q \in Q$ and the GMM mechanism $\mathcal{M} = (\pi, \kappa, b_U)$. Dynamic deception uses stages $k \leq K$, technical states x^k, defender and attacker actions d^k, a^k, and beliefs b_i^k. Moving-target defense has configurations $c \in C_l$, attack surfaces $\mathcal{S}_l(c)$, a risk matrix R with attack (row) index i and configuration index j, randomizations f, g, and the soft-min map $\mathcal{B}_\tau$.
Notation specific to this chapter: mixed and behavioral strategies are σ, while π denotes an information policy; in the mechanism model the user's type is q and θ is the hidden system state; in the dynamic models the state is x^k, the transition law P^k, and beliefs b_i^k rather than μ; J denotes detector quality rather than a cost; and $\mathcal{S}_l(c)$ is an attack surface rather than a state space.

The chapter assumes the normal-form, extensive-form, and signaling games of Chapter 3, the Bayesian games of Chapter 4, the Markov games of Chapter 5, and the learning dynamics of Chapter 7. It is restricted to models with finitely many states, types, messages, and actions; the main results are proved in the text, and the convergence of the coupled learning dynamics of moving-target defense is cited. Its aim is not to catalog deception technologies but to show how each technology becomes a formal model: what is hidden, what is shown, which belief and which action change, and which security value and which new risk result.

Learning Objectives

After reading this chapter, the reader should be able to:

(1) define cyber deception and distinguish it from hiding, denial, prevention, and detection;

(2) formulate a one-shot deception interaction as a signaling game with evidence, compute the attacker's belief thresholds, and characterize its pooling and partially separating equilibria;

(3) formulate an overt deception policy as Bayesian persuasion, verify Bayesian plausibility, and compute an optimal information policy by concavification;

(4) translate a desired, type-contingent user response into incentive-compatibility constraints, and design a generator that enforces it;

(5) compute the value of engagement and of a deceptive probe in a dynamic deception problem by backward induction and Bayes' rule;

(6) compute saddle-point and soft-min policies for moving-target defense, and explain why the rest points of the soft-min update are saddle points.

13.1 Introduction to Cyber Deception

Cyber deception is the deliberate use of information, system appearance, timing, and interaction to shape an adversary's perception. It differs from hiding. A defender who hides an asset prevents the adversary from observing it; a defender who deceives constructs an information environment in which the adversary observes something, interprets it, and then chooses an action that is less harmful to the defender.

Definition 13.1.1 (Cyber deception)**.** *Cyber deception* consists of planned defensive actions that mislead, confuse, delay, redirect, or selectively inform an adversary, so that the adversary takes, delays, or avoids actions in a way that improves the security of the defended system.

The definition has three components. Deception is *intentional*: it is designed rather than the product of accidental ambiguity. It is *adversary-centered*: it acts through the adversary's observations and inferences. It is *goal-directed*: its purpose may be deterrence, delay, detection, attribution, intelligence collection, containment, or resilience.

13.1.1 Why Cyber Defense Needs Deception

Most security controls are built around prevention and detection. Prevention stops unauthorized access, and detection recognizes malicious activity after it begins. Both remain essential, but neither removes the strategic problem: attackers probe, adapt, hide their intent, and use partial observations to decide where to move next. Cyber deception acts on this decision problem directly [299, 10]. Six defensive objectives can be distinguished, each with an observable measure of success.

(1) *Increase uncertainty.* Decoy hosts, synthetic files, and address rotation make the attacker less confident about which hosts, credentials, or routes are genuine; the measure is the attacker's posterior belief or reconnaissance error.

(2) *Increase cost.* Moving-target defense, obfuscation, and deceptive routing force the attacker to verify, probe, or retool before acting; the measures are delay, number of probes, and tool failures.

(3) *Redirect behavior.* Honeypots, honeynets, and decoy databases steer the attacker toward controlled environments; the measure is the fraction of malicious activity captured by decoys.

(4) *Reveal information.* Honey credentials, honeytokens, and engagement systems expose tools, infrastructure, and persistence methods; the measures are intelligence value and attribution confidence.

(5) *Deter action.* Plausible monitoring cues and a credible density of decoys lead the attacker to conclude that continuing is too risky; the measure is the withdrawal rate.

(6) *Buy time.* Delayed responses and deceptive paths slow lateral movement while the defender investigates and isolates; the measure is the reduction in dwell time.

Deception complements sound security engineering and does not replace it. A decoy credential does not repair weak access control, a honeypot does not replace patching, and moving-target defense does not remove the need for monitoring.

Deception is also distinct from denial. *Denial* prevents the adversary from observing true information, whereas *deception* causes the adversary to form a false or strategically useful interpretation [554]. The two usually appear together: address rotation, topology hiding, and metadata suppression deny accurate observation, while fake banners, honey credentials, decoy databases, and deceptive routes create misleading observations. In the same vein, *cryptic* deception makes the true system harder to identify, and *mimetic* deception creates a plausible false appearance.

Every deception mechanism studied in this chapter can be described by six elements: the *hidden state* that the attacker does not know, the *signal* that the attacker observes, the *evidence* that may reveal the deception, the *belief* that the attacker forms, the *action* that the belief induces, and the *value and risk* that the induced action creates for the defender. Example 13.1 identifies the six elements in a cloud deployment.

EXAMPLE 13.1 **A decoy database endpoint in a cloud network**

An enterprise web application runs in a cloud virtual private network. The production database is reachable only from the application subnet. The defender adds a decoy endpoint `orders-db-read.internal` with a valid internal certificate, a PostgreSQL banner, realistic table names, aged order records, and a few planted customer identifiers. The decoy is not connected to production data, and its security group admits connections only from hosts that never query databases in normal operation, such as bastion hosts and development containers. Figure 13.1 shows the topology.

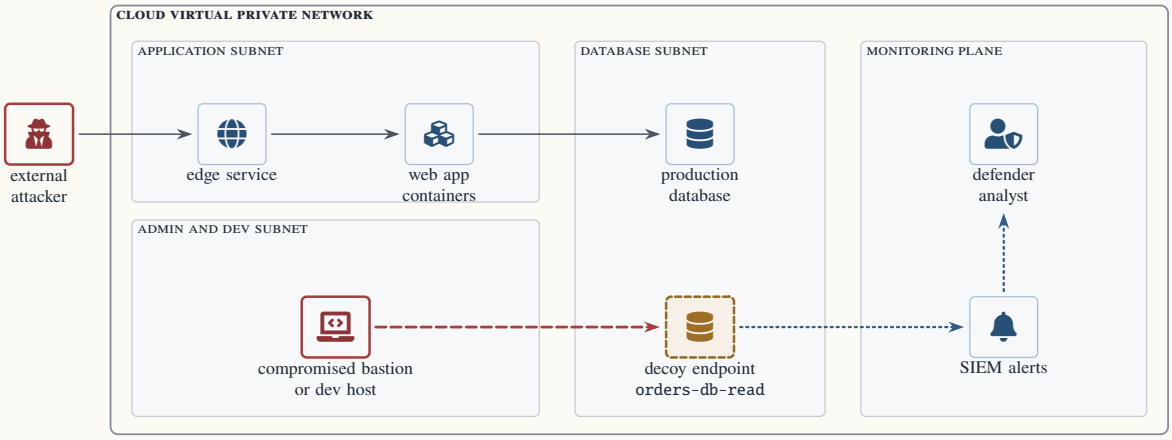


Figure 13.1: Network topology of Example 13.1. Solid arrows trace the normal path from the edge service through the application containers to the production database. The dashed arrow is a direct database probe from a compromised bastion or development host, which can reach only the decoy endpoint. The dotted arrows carry the decoy's telemetry to the security information and event management (SIEM) system and to the analyst.

The six model elements are as follows. The *hidden state* is whether the endpoint is a real read replica or a monitored decoy. The *signal* is the combination of DNS name, certificate, banner, schema, row counts, and response timing. The *evidence* available to the attacker consists of implementation clues, such as unusually uniform timing or records that never change. The attacker's *belief* is its probability that the endpoint is a decoy, and its *action* is to connect, dump tables, test credentials, or move on. For the defender, the *value* is a high-confidence alert, because no legitimate workflow touches the endpoint, together with the intelligence carried by the attacker's queries, client host, cloud role, and container image; the *risk* is that the decoy becomes a pivot point, leaks real data, or misleads legitimate operators.

Observation. The decoy does not block the attack. It changes the attacker's decision problem and converts any interaction into evidence. A single `SELECT` query is weak evidence of intent, whereas a credential test followed by a bulk dump is strong evidence. A deception design is therefore a model of hidden state, signal, belief, action, value, and risk, not merely a fake object.

13.1.2 Common Cyber-Deception Techniques

Cyber deception appears at every layer of a system. Table 13.1 lists the main families of techniques together with the signal each one creates and the belief it shapes [299, 10, 475]. The table is not a complete taxonomy; its purpose is to show that the same logic recurs across technically different mechanisms. Each technique is evaluated by the behavior it changes: a honey credential is valuable if it changes the attacker's action or reveals a compromise, a decoy if it redirects, delays, deters, or informs, and a moving-target defense if it invalidates reconnaissance or lowers the attacker's expected payoff.

Table 13.1: Families of cyber-deception techniques, the signals they present, the beliefs they shape, and the value and risk they create for the defender.

Technique	Observed signal	Belief shaped	Value and risk
Honeypots, honeynets	Host, service, or network that appears useful	Production or decoy; monitored or not	Detection and engagement; containment risk
Honeytokens, honey credentials	File, record, key, or account that appears usable	Genuine or planted; whether use exposes the attacker	High-confidence alerts; false-positive management
Decoy content	Plausible documents, repositories, or database entries	Which data are valuable; which paths are worth exploring	Redirection; intelligence on attacker interests
Obfuscation	Code, configuration, or metadata that is hard to interpret	True structure, vulnerabilities, and dependencies	Higher analysis cost; maintenance burden
Moving-target defense	Addresses, routes, or services that change over time	Whether reconnaissance is still valid	Disrupted reconnaissance; operational complexity
Deceptive responses	Errors, delays, banners, or timing returned selectively	Vulnerable, monitored, real, or emulated	Belief shaping; credibility risk
Engagement environments	Controlled space that resembles a reachable system	Whether continued interaction pays	Intelligence and delay; legal and safety limits

13.1.3 The Cyber-Deception Chain

The six model elements are linked in a chain that runs from defender design to attacker action and back. The defender designs an artifact or policy; the attacker observes part of it as a signal, combines the signal with prior knowledge and evidence, and acts on the resulting belief; the action produces a security outcome and new evidence for the defender, who updates the design. Figure 13.2 shows the chain as a feedback loop [299, 479]. A signal that changes no attacker decision on this chain may be camouflage or noise, but it is not strategically useful deception.

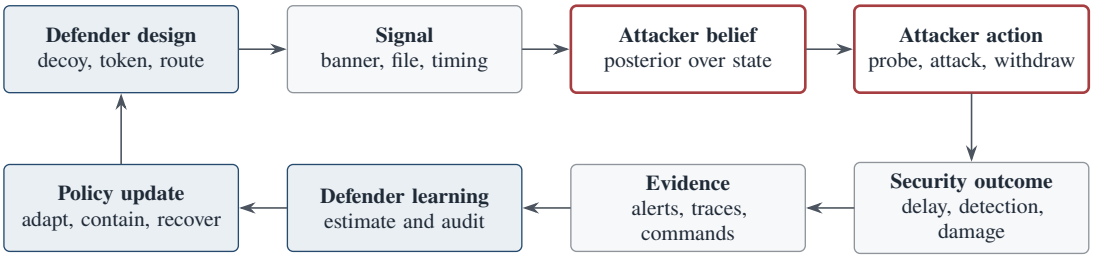


Figure 13.2: The cyber-deception chain. The upper row is the attacker's inference: the defender's design is exposed as a signal, the signal and the evidence determine the attacker's belief, and the belief determines the attacker's action. The lower row is the defender's learning: the action produces a security outcome and observable evidence, from which the defender estimates the attacker's behavior and updates the next design.

EXAMPLE 13.2 **A honey credential in a CI/CD repository**

A private repository builds container images for a software team. The defender places a realistic cloud access key in a stale deployment file, `deploy/legacy-prod.env`, surrounded by plausible region, bucket, registry, and service-account values. The key belongs to a monitored identity whose policy denies every action except a harmless identity check, and every use is tagged with a unique canary identifier in the cloud audit log. The internal secret scanner recognizes the canary, so only external use raises an alert.

In the six-element model, the hidden state is whether the credential is genuine or planted; the signal is the repository context, including file path, commit history, naming convention, and format validity; the attacker's actions are to ignore the key, test it with a low-noise call, or use it from rented infrastructure; and the defender's evidence is the audit trace. The actions generate evidence of different strength. A local read of the file is consistent with accidental developer access, an external `GetCallerIdentity` call from an unfamiliar network is strong evidence of compromise, and a sequence of failed privilege-escalation calls reveals both intent and tooling.

Observation. If the attacker believes the key is genuine, the defender obtains a high-confidence alert before any production credential is used. If the attacker suspects a planted key, it must slow down, verify, or abandon the path, which raises its cost. Either belief benefits the defender, provided the canary identity is tightly permissioned and documented so that responders do not mistake it for a real secret.

13.1.4 AI and Cyber Deception

Artificial intelligence enters cyber deception in three roles [466, 92, 246]. First, an AI system can be the deceptive actor: it can produce misleading claims, conceal information, or imitate a trusted identity, so that the hidden state becomes whether an output is reliable, generated, or strategically selected. Second, AI amplifies human deception. Generative models personalize phishing; synthesize documents, voices, and profiles; and adapt lures to observed responses. Agentic tools chain reconnaissance, lure generation, and credential testing, which makes mimetic deception cheaper and more adaptive. The same capabilities serve the defender, who can generate decoy content, synthetic users, and adaptive honeypot dialogues, at the price of stricter requirements on containment and auditability [359]. Third, AI systems are themselves targets of deception through prompt injection, poisoned context, adversarial examples, and misleading tool outputs.

These roles enlarge the scope of the models without changing their structure. The receiver may be a human attacker, an analyst, a malware classifier, or a language-model agent; the signal may be a generated message, a synthetic credential, or an adaptive conversation. The six model elements remain, and they matter more because AI generates more signals, personalizes them faster, and sustains longer interactions.

13.1.5 Why Scientific Foundations Are Needed

Deception fails when it is treated as a collection of clever tricks. A decoy may be too easy to recognize, too expensive to maintain, or too risky to expose; a warning banner may deter one attacker and alert another; a high-interaction honeypot collects intelligence but creates containment and legal risk. A scientific foundation for cyber deception therefore requires six components [299]: *concepts* that delimit deception, denial, signal, belief, and credibility; *taxonomies* that classify techniques by objective, layer, and adversary; *metrics* for delay, detection probability, attacker cost, intelligence value, false alarms, and exposure; *models* of how attackers infer, respond, and adapt; *experiments*, such as testbeds, red-team exercises, and human-subject studies, that test techniques against realistic attackers [195]; and *governance*, that is, rules of engagement, privacy controls, and safety constraints.

Game theory supplies the modeling component, because deception is strategic: the defender chooses what to reveal or hide while anticipating the attacker's inference, and the attacker chooses whether to trust, ignore, verify, or exploit what it observes. The remainder of the chapter develops this component in four steps: the analysis of a given deception environment in Section 13.2, the design of mechanisms in Section 13.3, dynamic deception in Section 13.4, and moving-target defense in Section 13.5 [625, 267, 479].

13.2 Game-Theoretic Analysis of Cyber Deception

Section 13.1 described cyber deception as the deliberate shaping of an adversary's observations, beliefs, and actions. This section turns the description into formal models and asks when a deceptive signal changes the attacker's best response.

Definition 13.2.1 (Cyber-deception game)**.** A *cyber-deception game* is a game in which a defender chooses information, signals, decoys, or system configurations in order to influence the beliefs and actions of an attacker who does not observe the true state of the system.

The attacker in such a game is not a passive sensor. It asks whether an observation is genuine, whether a system is monitored, whether a credential is planted, and whether continuing is worth the risk. The same technical object can therefore play different strategic roles: a honeypot can detect a scan, delay an intruder, reveal malware, or deter further probing, and which role it plays depends on the attacker's beliefs and payoffs. The section first analyzes a given deception environment as a signaling game with evidence, following the leaky-deception model of Chapter 4 [476], and then lets the defender design the information environment by Bayesian persuasion [310].

13.2.1 Signaling Games with Evidence

Chapter 3 introduced signaling games through sender types, messages, and Bayes-consistent receiver beliefs, and Chapter 4 added a detector that may expose a deceptive message. In cyber deception the defender is the sender and the attacker is the receiver. The detector models the clues that a deceptive artifact leaks through timing, configuration, virtual-machine fingerprints, file realism, or monitoring traces [476, 625, 479].

Definition 13.2.2 (Signaling game with evidence)**.** A *cyber-deception signaling game with evidence* consists of finite sets Θ of system states, M of messages, E of evidence values, and $\mathcal{A}_A$ of attacker actions; a prior $\rho \in \Delta(\Theta)$; an evidence kernel $\lambda(e \mid \theta, m)$; and utilities $u_D, u_A : \Theta \times M \times E \times \mathcal{A}_A \to \mathbb{R}$. Nature draws $\theta \sim \rho$; the defender observes θ and sends m with probability $\sigma_D(m \mid \theta)$; the detector produces e with probability $\lambda(e \mid \theta, m)$; the attacker observes (m, e), forms a belief $\mu(\cdot \mid m, e) \in \Delta(\Theta)$, and chooses a with probability $\sigma_A(a \mid m, e)$.

The state θ is, for example, production or honeypot; the message m is the defender-controlled appearance, such as an activity level, a banner, or a response pattern; the evidence e is an alarm or an implementation clue; and the action a is to attack, withdraw, verify, or continue. Figure 13.3 shows the game tree of the binary instance studied in Section 13.2.2.

The expected utility of a defender of type θ is

$$U_D(\sigma_D, \sigma_A \mid \theta) \triangleq \sum_{m \in M} \sigma_D(m \mid \theta) \sum_{e \in E} \lambda(e \mid \theta, m) \sum_{a \in \mathcal{A}_A} \sigma_A(a \mid m, e)\, u_D(\theta, m, e, a),$$

and the attacker's expected utility from a mixed action $\xi \in \Delta(\mathcal{A}_A)$ after observing (m, e) is

$$V_A(\xi \mid m, e) \triangleq \sum_{\theta \in \Theta} \mu(\theta \mid m, e) \sum_{a \in \mathcal{A}_A} \xi(a)\, u_A(\theta, m, e, a).$$

Definition 13.2.3 (Perfect Bayesian equilibrium with evidence)**.** A profile $(\sigma_D^*, \sigma_A^*, \mu^*)$ is a *perfect Bayesian Nash equilibrium* (PBNE) of the signaling game with evidence if the following three conditions hold.

(1) For every $\theta \in \Theta$, $\sigma_D^*(\cdot \mid \theta) \in \arg\max_{\sigma_D(\cdot \mid \theta) \in \Delta(M)} U_D(\sigma_D, \sigma_A^* \mid \theta)$.

(2) For every $(m, e) \in M \times E$, $\sigma_A^*(\cdot \mid m, e) \in \arg\max_{\xi \in \Delta(\mathcal{A}_A)} V_A(\xi \mid m, e)$, where V_A is evaluated at μ^*.

(3) Whenever $\sum_{\theta'} \lambda(e \mid \theta', m) \sigma_D^*(m \mid \theta') \rho(\theta') > 0$,

$$\mu^*(\theta \mid m, e) = \frac{\lambda(e \mid \theta, m)\, \sigma_D^*(m \mid \theta)\, \rho(\theta)}{\sum_{\theta' \in \Theta} \lambda(e \mid \theta', m)\, \sigma_D^*(m \mid \theta')\, \rho(\theta')};$$

at observations of probability zero, $\mu^*(\cdot \mid m, e)$ is an arbitrary element of $\Delta(\Theta)$.

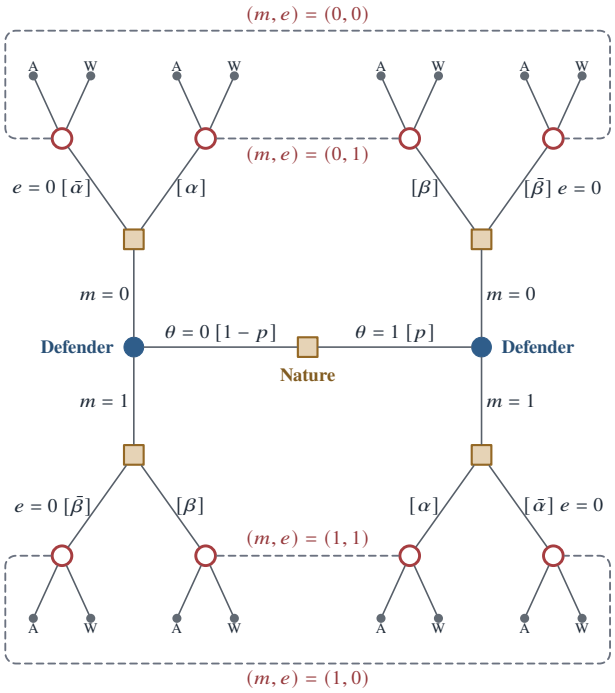


Figure 13.3: Game tree of the binary signaling game with evidence. Nature (square, center) draws a production system, $\theta = 0$, or a honeypot, $\theta = 1$. The defender (solid discs) sends the active-looking message $m = 0$ (upper half) or the inactive-looking message $m = 1$ (lower half). Bracketed numbers are branch probabilities; the detector (squares) raises an alarm, $e = 1$, with probability α after a truthful and β after a deceptive message. The attacker (rings) observes (m, e) but not θ; each dashed line joins the two nodes of one information set, labeled by (m, e), and at each node the attacker chooses to attack, A, or withdraw, W.

The belief in condition (3) differs from that of a signaling game without evidence by the factor $\lambda(e \mid \theta, m)$: the attacker updates not only on the message but also on how likely the observed evidence would be if the message were truthful or deceptive.

13.2.2 The Binary Evidence Game

► Code: `code/binary_evidence_eye.py`

The binary instance of Pawlick, Colbert, and Zhu [476] has $\Theta = M = E = \mathcal{A}_A = \{0, 1\}$. The state $\theta = 0$ is a production system and $\theta = 1$ a honeypot; the message $m = 0$ is an active-looking system and $m = 1$ an inactive-looking one, so that the message is *truthful* when $m = \theta$ and *deceptive* when $m \neq \theta$; the evidence $e = 1$ is an alarm and $e = 0$ its absence; and the action $a = 0$ is to attack and $a = 1$ to withdraw. The detector raises an alarm with probability α after a truthful message, a false positive, and with probability β after a deceptive one, a true positive:

$$\lambda(1 \mid \theta, m) = \begin{cases} \alpha, & m = \theta, \\ \beta, & m \neq \theta, \end{cases} \qquad \lambda(0 \mid \theta, m) = 1 - \lambda(1 \mid \theta, m).$$

Write $\bar{\alpha} \triangleq 1 - \alpha$, $\bar{\beta} \triangleq 1 - \beta$, and $p \triangleq \rho(1)$ for the prior probability of a honeypot. The *detector quality* is $J \triangleq \beta - \alpha$ and the *detector aggressiveness* is $G \triangleq \alpha + \beta - 1 = \beta - \bar{\alpha}$. The detector is *conservative* if $G < 0$, so that it raises few alarms, and *aggressive* if $G > 0$.

Assumption 13.2.4 (Opposing preferences and cheap talk)**.** The detector is informative and nondegenerate, $0 < \alpha < \beta < 1$, so that $J > 0$. Utilities depend only on (θ, a), and the players have opposing preferences over the attacker's

action:

$$\Delta_0^A \triangleq u_A(0,0) - u_A(0,1) > 0, \quad \Delta_1^A \triangleq u_A(1,1) - u_A(1,0) > 0, \quad u_D(0,1) > u_D(0,0), \quad u_D(1,0) > u_D(1,1).$$

Thus the attacker wants to attack a production system and to withdraw from a honeypot, whereas a production defender wants the attacker to withdraw and a honeypot defender wants it to attack. Write $\omega \triangleq \Delta_1^A/\Delta_0^A$ and $\bar{\mu} \triangleq \Delta_0^A/(\Delta_0^A + \Delta_1^A)$. The two players' decisions reduce to simple comparisons.

LEMMA 13.2.5 **Reduced decision rules**
Under Assumption 13.2.4, the following hold.

(1) After any observation with honeypot posterior $\mu = \mu(1 \mid m, e)$, the attacker's expected gain from attacking rather than withdrawing is $(1-\mu)\Delta_0^A - \mu\Delta_1^A$. Attack is optimal if $\mu \le \bar{\mu}$, withdrawal is optimal if $\mu \ge \bar{\mu}$, and the attacker is indifferent exactly at $\mu = \bar{\mu}$.

(2) Let $W_\theta(m) \triangleq \sum_e \lambda(e \mid \theta, m)\, \sigma_A(1 \mid m, e)$ be the withdrawal probability that message m induces for a defender of type θ. The production defender, $\theta = 0$, prefers the message with the larger $W_0(m)$, and the honeypot defender, $\theta = 1$, prefers the message with the smaller $W_1(m)$.

Proof. (1) The expected utility of attacking is $(1-\mu)u_A(0,0) + \mu u_A(1,0)$ and that of withdrawing is $(1-\mu)u_A(0,1) + \mu u_A(1,1)$; their difference is $(1-\mu)\Delta_0^A - \mu\Delta_1^A$, which is nonnegative if and only if $\mu \le \bar{\mu}$. (2) Since u_D depends only on (θ, a), $U_D(\sigma_D, \sigma_A \mid \theta) = \sum_m \sigma_D(m \mid \theta)\big[u_D(\theta,0) + W_\theta(m)\big(u_D(\theta,1) - u_D(\theta,0)\big)\big]$. The bracket is increasing in $W_0(m)$ for $\theta = 0$ and decreasing in $W_1(m)$ for $\theta = 1$. □

When both defender types send the same message m, the posterior after (m, e) depends only on the evidence. Its comparison with $\bar{\mu}$ is a comparison of the prior p with a threshold, and the four observations (m, e) produce four thresholds.

PROPOSITION 13.2.6 **Detector-dependent belief thresholds**
Let both defender types send message m, and let $\Lambda_\theta \triangleq \lambda(e \mid \theta, m)$. After observation (m, e), the attacker attacks if and only if $p \le T_{me} \triangleq \Lambda_0\Delta_0^A/(\Lambda_0\Delta_0^A + \Lambda_1\Delta_1^A)$, that is,

$$T_{00} = \frac{\bar{\alpha}}{\bar{\alpha} + \bar{\beta}\,\omega}, \qquad T_{01} = \frac{\alpha}{\alpha + \beta\,\omega}, \qquad T_{10} = \frac{\bar{\beta}}{\bar{\beta} + \bar{\alpha}\,\omega}, \qquad T_{11} = \frac{\beta}{\beta + \alpha\,\omega}.$$

Under Assumption 13.2.4, a conservative detector gives $T_{01} < T_{10} < T_{00} < T_{11}$, and an aggressive detector gives $T_{10} < T_{01} < T_{11} < T_{00}$.

Proof. The posterior is $\mu(1 \mid m, e) = \Lambda_1 p/(\Lambda_1 p + \Lambda_0(1-p))$, and by Lemma 13.2.5 the attacker attacks if and only if $\Lambda_1 p\Delta_1^A \le \Lambda_0(1-p)\Delta_0^A$, which is $p \le T_{me}$. The pair (Λ_0, Λ_1) equals $(\bar{\alpha}, \bar{\beta})$, (α, β), $(\bar{\beta}, \bar{\alpha})$, and (β, α) for $(m, e) = (0,0), (0,1), (1,0), (1,1)$, which gives the four formulas. Each T_{me} is increasing in the likelihood ratio Λ_0/Λ_1, and the four ratios are $\bar{\alpha}/\bar{\beta} > 1$, $\alpha/\beta < 1$, $\bar{\beta}/\bar{\alpha} < 1$, and $\beta/\alpha > 1$. Moreover, $\alpha\bar{\alpha} - \beta\bar{\beta} = (\alpha - \beta)(1 - \alpha - \beta) = JG$. For a conservative detector $JG < 0$, so $\alpha/\beta < \bar{\beta}/\bar{\alpha}$ and $\bar{\alpha}/\bar{\beta} < \beta/\alpha$, which gives the first ordering; for an aggressive detector both inequalities reverse. □

Definition 13.2.7 (Prior regimes)**.** The thresholds of Proposition 13.2.6 partition the prior interval $(0, 1)$ into five regimes, listed from left to right: *zero-dominant* (ZD), *zero-heavy* (ZH), *middle* (M), *one-heavy* (OH), and *one-dominant* (OD). For a conservative detector, the middle regime is (T_{10}, T_{00}); for an aggressive detector, it is (T_{01}, T_{11}).

Near $p = 0$ production systems dominate and attack is optimal after every observation; near $p = 1$ withdrawal is optimal after every observation. In between, the attacker's response to a pooled message may depend on the evidence. The next result shows that such evidence dependence is exactly what destroys pooling.

PROPOSITION 13.2.8 **Pooling equilibria**
Let Assumption 13.2.4 hold, let $G \neq 0$, and let p lie in the interior of a regime. Pooling on message m is supported by a PBNE if and only if the attacker's best response after $(m, 0)$ and $(m, 1)$ is the same action. Consequently:

(1) in the zero-dominant and one-dominant regimes, pooling on either message is supported, with the attacker attacking, respectively withdrawing, after every observation;

(2) in the zero-heavy regime, pooling is supported only on $m = 1$ for a conservative detector and only on $m = 0$ for an aggressive one; in the one-heavy regime, only on $m = 0$ for a conservative detector and only on $m = 1$ for an aggressive one;

(3) in the middle regime, no pooling PBNE exists.

The negative conclusions, that pooling on a given message is not supported, hold for every specification of the off-path beliefs.

Proof. Sufficiency. Suppose the attacker plays the same action $\bar{a}$ after $(m, 0)$ and $(m, 1)$. Set the off-path beliefs after $(1 - m, e)$ to the point mass on $\theta = 0$ if $\bar{a} = 0$ and on $\theta = 1$ if $\bar{a} = 1$; by Lemma 13.2.5, $\bar{a}$ is then optimal after $(1 - m, e)$ for both e. Both messages induce the same withdrawal probability for each type, so neither type gains by deviating.

Necessity. After a pooled message $m = 0$, an alarm raises the honeypot posterior (the likelihood ratio β/α exceeds one) and its absence lowers it ($\bar{\beta}/\bar{\alpha} < 1$), so the only evidence-dependent response is to attack after $(0, 0)$ and to withdraw after $(0, 1)$. On the path, $W_0(0) = \alpha$ and $W_1(0) = \beta$. Let $\sigma_A(1 \mid 1, e) \in [0, 1]$ be the withdrawal probability after the off-path observation $(1, e)$; every value of $\sigma_A(1 \mid 1, e)$ is a best response to some off-path belief. Deviation to $m = 1$ yields $W_0(1) = \bar{\beta}\sigma_A(1 \mid 1, 0) + \beta\sigma_A(1 \mid 1, 1)$ and $W_1(1) = \bar{\alpha}\sigma_A(1 \mid 1, 0) + \alpha\sigma_A(1 \mid 1, 1)$, and by Lemma 13.2.5 pooling requires

$$\text{(i)}\ \ \bar{\beta}\sigma_A(1 \mid 1, 0) + \beta\sigma_A(1 \mid 1, 1) \leq \alpha, \qquad \text{(ii)}\ \ \bar{\alpha}\sigma_A(1 \mid 1, 0) + \alpha\sigma_A(1 \mid 1, 1) \geq \beta.$$

For a conservative detector, $\bar{\alpha} > \beta$. Subtracting $\bar{\beta}/\bar{\alpha}$ times (ii) from (i) gives $\big((\beta - \alpha)/\bar{\alpha}\big)\sigma_A(1 \mid 1, 1) \leq (\alpha\bar{\alpha} - \beta\bar{\beta})/\bar{\alpha} = JG/\bar{\alpha} < 0$, which contradicts $\sigma_A(1 \mid 1, 1) \geq 0$. For an aggressive detector, $\alpha + \beta > 1$. Subtracting (i) from β/α times (ii) gives $\big((\beta - \alpha)/\alpha\big)\sigma_A(1 \mid 1, 0) \geq (\beta^2 - \alpha^2)/\alpha$, that is, $\sigma_A(1 \mid 1, 0) \geq \alpha + \beta > 1$, which contradicts $\sigma_A(1 \mid 1, 0) \leq 1$. Hence no off-path response deters both types. The relabeling $\theta \mapsto 1 - \theta$, $m \mapsto 1 - m$, $a \mapsto 1 - a$ preserves the detector and the structure of Assumption 13.2.4, exchanging Δ_0^A and Δ_1^A, and it maps pooling on $m = 1$ into pooling on $m = 0$; the argument therefore also covers $m = 1$.

Consequences. Applying Proposition 13.2.6 regime by regime shows that the response to a pooled message is evidence-independent for both messages in ZD and OD, for exactly the message listed in (2) in ZH and OH, and for neither message in M. □

COROLLARY 13.2.9 **No pooling in the middle regime**
If $G \neq 0$, the middle regime of the binary evidence game admits no pooling PBNE.

Without evidence, pooling can usually be sustained by suitable off-path beliefs. With evidence, a deviating defender still generates evidence whose distribution depends on its type, so that no off-path response deters both types at once. In the middle regime the equilibrium must therefore be partially separating.

THEOREM 13.2.10 **Partially separating PBNE [476]**
Let Assumption 13.2.4 hold, let p lie in the middle regime, and write $o \triangleq p/(1 - p)$, $s_0 \triangleq \sigma_D^*(1 \mid 0)$, and $s_1 \triangleq \sigma_D^*(1 \mid 1)$. For a conservative detector, the profile

$$s_0 = \frac{\bar{\alpha}^2 - \bar{\alpha}\bar{\beta}\,\omega\,o}{\bar{\alpha}^2 - \bar{\beta}^2}, \qquad s_1 = \frac{\bar{\alpha}\bar{\beta}/(\omega\,o) - \bar{\beta}^2}{\bar{\alpha}^2 - \bar{\beta}^2},$$

$$\sigma_A^*(1 \mid 0,0) = \frac{1-\alpha-\beta}{2-\alpha-\beta}, \quad \sigma_A^*(1 \mid 0,1) = 1, \quad \sigma_A^*(1 \mid 1,0) = \frac{1}{2-\alpha-\beta}, \quad \sigma_A^*(1 \mid 1,1) = 0,$$

with beliefs given by Bayes' rule, is a PBNE. For an aggressive detector, the profile

$$s_0 = \frac{\alpha\beta\,\omega\,o - \alpha^2}{\beta^2 - \alpha^2}, \qquad s_1 = \frac{\beta^2 - \alpha\beta/(\omega\,o)}{\beta^2 - \alpha^2},$$

$$\sigma_A^*(1 \mid 0,0) = 0, \quad \sigma_A^*(1 \mid 0,1) = \frac{1}{\alpha+\beta}, \quad \sigma_A^*(1 \mid 1,0) = 1, \quad \sigma_A^*(1 \mid 1,1) = \frac{\alpha+\beta-1}{\alpha+\beta},$$

with beliefs given by Bayes' rule, is a PBNE. In both cases the defender and the attacker mix.

Proof. Consider the conservative case. *Feasibility.* Since $\bar{\alpha} > \bar{\beta}$, s_0 is strictly decreasing in o, and s_1 is strictly increasing in $1/o$; both are therefore strictly decreasing in p. Substituting $o = \bar{\beta}/(\bar{\alpha}\omega)$, which corresponds to $p = T_{10}$, gives $s_0 = s_1 = 1$, and substituting $o = \bar{\alpha}/(\bar{\beta}\omega)$, which corresponds to $p = T_{00}$, gives $s_0 = s_1 = 0$. Hence $s_0, s_1 \in (0,1)$ on the middle regime. The attacker's mixing probabilities lie in $(0,1)$ because $\alpha + \beta < 1$.

Attacker optimality. The two equations

$$p\,(1-s_1)\,\bar{\beta}\,\Delta_1^A = (1-p)(1-s_0)\,\bar{\alpha}\,\Delta_0^A, \qquad p\,s_1\,\bar{\alpha}\,\Delta_1^A = (1-p)\,s_0\,\bar{\beta}\,\Delta_0^A$$

state that $\mu(1 \mid 0,0) = \bar{\mu}$ and $\mu(1 \mid 1,0) = \bar{\mu}$; solving them for (s_0, s_1) yields the displayed formulas, so the attacker is indifferent after both no-alarm observations and may mix. After $(0,1)$, the first equation gives $(1-p)(1-s_0)\alpha\Delta_0^A = p(1-s_1)\Delta_1^A\,\alpha\bar{\beta}/\bar{\alpha} \le p(1-s_1)\beta\Delta_1^A$, because $\alpha\bar{\beta} \le \beta\bar{\alpha}$; hence $\mu(1 \mid 0,1) \ge \bar{\mu}$ and withdrawal is optimal. After $(1,1)$, the second equation gives $p\,s_1\alpha\Delta_1^A \le (1-p)s_0\beta\Delta_0^A$ by the same inequality, so $\mu(1 \mid 1,1) \le \bar{\mu}$ and attack is optimal.

Defender optimality. Write $X \triangleq 1/(2-\alpha-\beta)$. The production type obtains $W_0(0) = \bar{\alpha}(1-\alpha-\beta)X + \alpha = \bar{\beta}X$ and $W_0(1) = \bar{\beta}X$, and the honeypot type obtains $W_1(0) = \bar{\beta}(1-\alpha-\beta)X + \beta = \bar{\alpha}X$ and $W_1(1) = \bar{\alpha}X$. Both types are indifferent between the messages, so by Lemma 13.2.5 any mixture is optimal. The aggressive case is identical after exchanging the roles of the alarm and no-alarm observations: s_0 and s_1 solve $\mu(1 \mid 0,1) = \mu(1 \mid 1,1) = \bar{\mu}$, they equal 0 at $p = T_{01}$ and 1 at $p = T_{11}$, and the attacker's mixing gives $W_0(0) = W_0(1) = \alpha/(\alpha+\beta)$ and $W_1(0) = W_1(1) = \beta/(\alpha+\beta)$. □

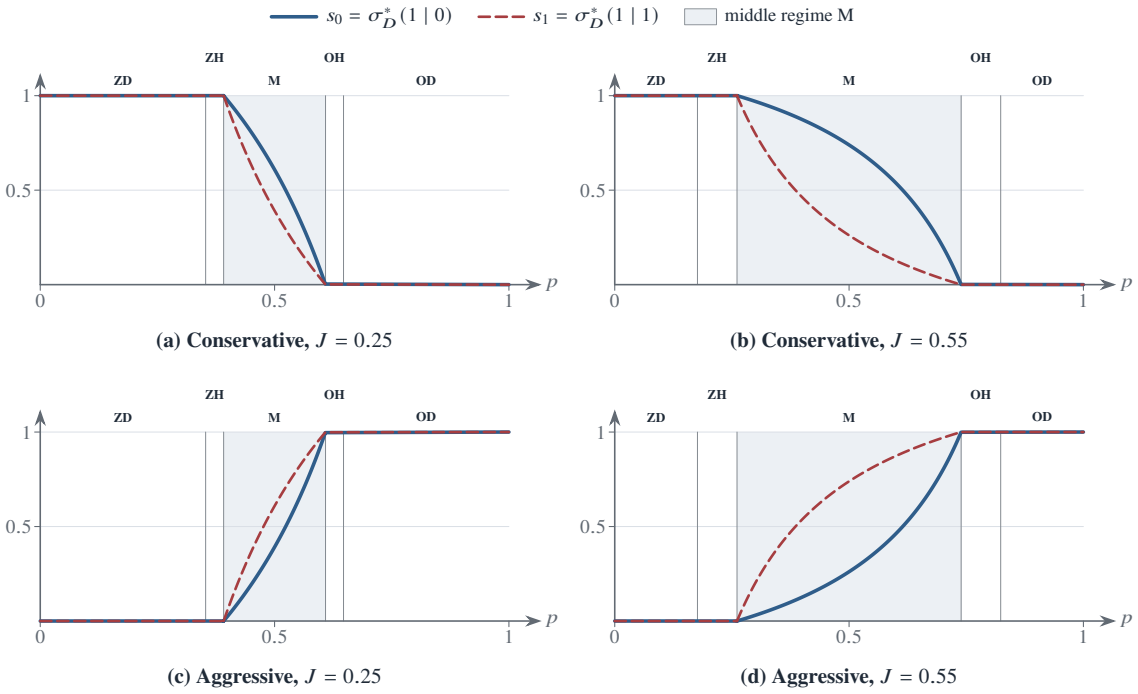
The theorem is an indifference result. The defender's mixing makes the attacker indifferent after the two observations whose meaning is ambiguous, the no-alarm observations for a conservative detector and the alarm observations for an aggressive one; the attacker's mixing makes both defender types indifferent between the two messages. Deception that is partly detectable is therefore sustained by randomization on both sides: the defender randomizes the observable activity level, and the attacker randomizes between attack and withdrawal after selected observations. The proof also shows that the partially separating branch joins the pooling branches continuously at both ends of the middle regime.

Figure 13.4 plots this continuous selection of equilibria over the whole prior interval for symmetric payoffs, $\omega = 1$. Outside the middle regime the curves follow the pooling equilibria of Proposition 13.2.8; inside it they follow Theorem 13.2.10. A higher detector quality J widens the middle regime, and the aggressiveness G determines the direction of the curves: under a conservative detector the probability of the inactive-looking message decreases as the honeypot prior rises, and under an aggressive detector it increases.

For an aggressive detector, the two curves of Theorem 13.2.10 enclose a lens-shaped region, the *equilibrium eye*, which measures the range of priors and the extent to which the two defender types behave differently.

COROLLARY 13.2.11 **Area of the equilibrium eye**
Let Assumption 13.2.4 hold with an aggressive detector, $G > 0$. On the middle regime $[p_L, p_R]$, with $p_L \triangleq T_{01} = \alpha/(\alpha+\beta\omega)$ and $p_R \triangleq T_{11} = \beta/(\beta+\alpha\omega)$, the width $w(p) \triangleq s_1(p) - s_0(p)$ is

$$w(p) = \frac{\alpha^2+\beta^2}{\beta^2-\alpha^2} - \frac{\alpha\beta}{\beta^2-\alpha^2}\left(\omega\,\frac{p}{1-p} + \frac{1}{\omega}\,\frac{1-p}{p}\right) \ge 0,$$



(a) Conservative, $J = 0.25$ **(b) Conservative, $J = 0.55$**

(c) Aggressive, $J = 0.25$ **(d) Aggressive, $J = 0.55$**

Figure 13.4: Equilibrium defender strategies $s_0 = \sigma_D^*(1 \mid 0)$ (solid line) and $s_1 = \sigma_D^*(1 \mid 1)$ (dashed line) as functions of the honeypot prior p in the binary evidence game with $\omega = 1$. Panels (a) and (b) use conservative detectors with $G = -0.15$ and quality $J = 0.25$ and $J = 0.55$; panels (c) and (d) use aggressive detectors with $G = 0.15$ and the same qualities. Thin vertical lines mark the thresholds of Proposition 13.2.6, which separate the regimes ZD, ZH, M, OH, and OD; the shaded band is the middle regime. Outside the band the curves coincide on a pooling equilibrium; in the dominant regimes the opposite pooling message is also an equilibrium.

with equality at p_L and p_R, and the area of the eye, $A_{\text{eye}} \triangleq \int_{p_L}^{p_R} w(p)\,dp$, is

$$A_{\text{eye}} = \left[\frac{\alpha^2+\beta^2}{\beta^2-\alpha^2}\,p + \frac{\alpha\beta\,\omega}{\beta^2-\alpha^2}\big(p+\log(1-p)\big) - \frac{\alpha\beta}{(\beta^2-\alpha^2)\,\omega}\big(\log p - p\big)\right]_{p_L}^{p_R}.$$

For symmetric payoffs, $\omega = 1$, and with $\tilde{J} \triangleq J/(1+G) \in (0,1)$,

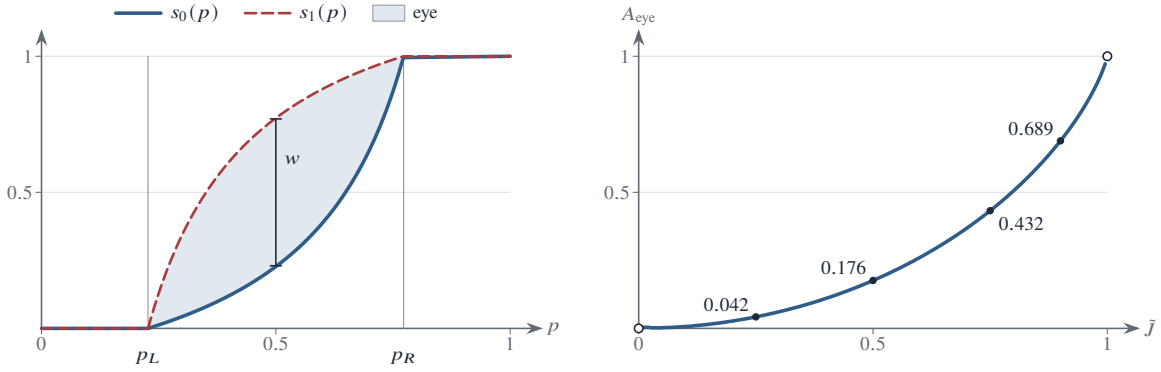
$$A_{\text{eye}} = 1 + \frac{2\alpha\beta}{\beta^2-\alpha^2}\log\frac{\alpha}{\beta} = 1 - \frac{1-\tilde{J}^2}{\tilde{J}}\operatorname{artanh}(\tilde{J}),$$

which is strictly increasing in $\tilde{J}$, satisfies $0 < A_{\text{eye}} < 1$, and tends to 0 as $\tilde{J} \to 0^+$ and to 1 as $\tilde{J} \to 1^-$.

Proof. The expression for w follows by subtracting s_0 from s_1 in Theorem 13.2.10. As a function of $o = p/(1-p)$, w is a constant minus positive multiples of o and $1/o$, so it is concave in o; it vanishes at p_L and p_R because $s_0 = s_1$ at both ends, and it is therefore nonnegative between them. The antiderivative follows from $\int p/(1-p)\,dp = -p - \log(1-p)$ and $\int (1-p)/p\,dp = \log p - p$. For $\omega = 1$, $p_L = \alpha/(\alpha+\beta)$ and $p_R = \beta/(\alpha+\beta)$; evaluating the antiderivative gives $(\alpha^2+\beta^2)/(\alpha+\beta)^2 + 2\alpha\beta/(\alpha+\beta)^2 + \big(2\alpha\beta/(\beta^2-\alpha^2)\big)\log(\alpha/\beta)$, and the first two terms sum to 1. Since $\alpha = (1+G-J)/2$ and $\beta = (1+G+J)/2$, $\alpha/\beta = (1-\tilde{J})/(1+\tilde{J})$, and $\log\big((1-\tilde{J})/(1+\tilde{J})\big) = -2\operatorname{artanh}(\tilde{J})$; moreover $2\alpha\beta/(\beta^2-\alpha^2) = (1-\tilde{J}^2)/(2\tilde{J})$, which gives the second form. With $Y \triangleq \alpha/\beta \in (0,1)$, $A_{\text{eye}} = \Psi(Y)/(1-Y^2)$ with $\Psi(Y) \triangleq 1 - Y^2 + 2Y\log Y$; since $\Psi(1) = 0$ and $\Psi'(Y) = 2(1 - Y + \log Y) < 0$ on $(0,1)$, $\Psi > 0$, and $\log Y < 0$ gives $A_{\text{eye}} < 1$. Finally, $dA_{\text{eye}}/d\tilde{J} = \big((1+\tilde{J}^2)\operatorname{artanh}(\tilde{J}) - \tilde{J}\big)/\tilde{J}^2$, whose numerator vanishes at $\tilde{J} = 0$ and has derivative $2\tilde{J}\operatorname{artanh}(\tilde{J}) + 2\tilde{J}^2/(1-\tilde{J}^2) > 0$. The limits follow from $\operatorname{artanh}(\tilde{J}) = \tilde{J} + \tilde{J}^3/3 + O(\tilde{J}^5)$ as $\tilde{J} \to 0^+$ and from

$(1 - \tilde{J}^2)\,\operatorname{artanh}(\tilde{J}) \to 0$ as $\tilde{J} \to 1^-$. □

The normalized separation $\tilde{J} = J/(1+G)$ equals the width $p_R - p_L$ of the middle regime when $\omega = 1$. A large eye means that the two defender types can sustain distinct behavior over a broad range of attacker priors; a small eye means that equilibrium play collapses quickly to pooling, so that the attacker acts almost independently of the evidence. Figure 13.5 shows the eye for one detector and the area as a function of $\tilde{J}$. The area approaches one only in the limit of a perfect detector, $\tilde{J} \to 1$.



(a) Eye for $\alpha = 0.25, \beta = 0.85$ **(b) Area of the eye**

Figure 13.5: The equilibrium eye for an aggressive detector with symmetric payoffs. (a) For $\alpha = 0.25$ and $\beta = 0.85$, the curves $s_0(p)$ (solid) and $s_1(p)$ (dashed) of Theorem 13.2.10 meet at $p_L = 0.227$ and $p_R = 0.773$ and enclose the shaded eye of area $A_{\text{eye}} = 0.212$; the bar marks the width $w(p)$ at $p = 0.5$. (b) The area A_{eye} as a function of the normalized separation $\tilde{J} = J/(1+G)$, with the values at $\tilde{J} = 0.25$, 0.5, 0.75, and 0.9 printed; the open circles mark the limits at $\tilde{J} = 0$ and $\tilde{J} = 1$.

13.2.3 Bayesian Persuasion for Cyber Deception

The signaling model analyzes a given deception environment in which the defender's message is a strategy chosen type by type. *Bayesian persuasion* goes one step further toward design: the defender commits, before the state is realized, to an information policy, that is, to the statistical relationship between the hidden state and the observable artifact [310, 309, 57]. Canonical persuasion is *overt*: the policy, or at least the commitment to an information experiment, is known to the attacker. This fits warnings, disclosures, security labels, risk scores, and announced monitoring policies. When the defender hides its policy, persuasion remains a useful benchmark but no longer describes the interaction exactly; covert deception is better described by the signaling model of Section 13.2.1.

Definition 13.2.12 (Information policy)**.** An *information policy* is a family $\pi = \{\pi(\cdot \mid \theta)\}_{\theta \in \Theta}$ of distributions on M; the defender commits to π, Nature draws $\theta \sim \rho$, the message $m \sim \pi(\cdot \mid \theta)$ is displayed, and the attacker, who knows π, observes m and acts. The policy is chosen from an *admissible set* Π that encodes implementation constraints.

The admissible set distinguishes cyber persuasion from an abstract communication experiment. It may contain support constraints $\pi(m \mid \theta) = 0$ for artifacts that cannot be displayed in state θ, a budget $\sum_\theta \rho(\theta) \sum_m \pi(m \mid \theta)\, \Xi(\theta, m) \le \bar{\Xi}$ on the expected cost $\Xi(\theta, m)$ of producing artifacts, and safety constraints that forbid signals that steer the attacker toward protected assets. Under π, message m occurs with probability $\nu_m \triangleq \sum_\theta \rho(\theta)\pi(m \mid \theta)$, and when $\nu_m > 0$ the attacker's posterior is

$$\mu_m(\theta) \triangleq \mu(\theta \mid m) = \frac{\rho(\theta)\,\pi(m \mid \theta)}{\nu_m}.$$

The attacker then chooses $a^*(\mu_m) \in \arg\max_{a \in \mathcal{A}_A} \sum_\theta \mu_m(\theta) u_A(\theta, a)$; ties are broken in the defender's favor, the standard convention in information design, since the defender can break a tie by an arbitrarily small change of the policy. Write $\widetilde{u}_D(\mu) \triangleq \sum_\theta \mu(\theta) u_D(\theta, a^*(\mu))$ for the defender's expected payoff when the attacker acts on belief μ. The

persuasion problem is

$$\max_{\pi \in \Pi} \sum_{m \in M} \nu_m \, \widetilde{u}_D(\mu_m).$$

PROPOSITION 13.2.13 **Bayesian plausibility and implementation**
For finite Θ and M, every information policy induces message probabilities and posteriors satisfying the *Bayesian-plausibility* condition

$$\sum_{m \in M} \nu_m \, \mu_m = \rho.$$

Conversely, if $\nu_m \geq 0$, $\sum_m \nu_m = 1$, $\mu_m \in \Delta(\Theta)$, and $\sum_m \nu_m \mu_m = \rho$, then the policy $\pi(m \mid \theta) \triangleq \nu_m \mu_m(\theta)/\rho(\theta)$ for $\rho(\theta) > 0$, completed arbitrarily for $\rho(\theta) = 0$, induces exactly these message probabilities and, for every message with $\nu_m > 0$, these posteriors.

Proof. By the definition of μ_m, $\sum_m \nu_m \mu_m(\theta) = \sum_m \rho(\theta)\pi(m \mid \theta) = \rho(\theta)$. Conversely, the specified $\pi(\cdot \mid \theta)$ is nonnegative and satisfies $\sum_m \pi(m \mid \theta) = \rho(\theta)^{-1} \sum_m \nu_m \mu_m(\theta) = 1$. It induces the message probability $\sum_\theta \rho(\theta)\pi(m \mid \theta) = \nu_m \sum_\theta \mu_m(\theta) = \nu_m$ and, when $\nu_m > 0$, by Bayes' rule, the posterior $\rho(\theta)\pi(m \mid \theta)/\nu_m = \mu_m(\theta)$. Types with zero prior affect neither quantity. □

Bayesian plausibility is the feasibility discipline of deception: the defender can redistribute suspicion across messages, but the average posterior must equal the prior. Proposition 13.2.13 also separates belief design from implementation. The analyst may first ask which posteriors would change the attacker's action and only afterwards whether banners, files, warnings, or timing patterns can implement them within Π.

Problem 13.2.14 (Posterior-form persuasion)**.** With Π unconstrained, the persuasion problem is equivalent to

$$\max_{\{\nu_m, \mu_m\}_{m \in M}} \quad \sum_{m \in M} \nu_m \, \widetilde{u}_D(\mu_m)$$
$$\text{s.t.} \quad \nu_m \geq 0, \quad \sum_m \nu_m = 1, \quad \mu_m \in \Delta(\Theta), \quad \sum_m \nu_m \mu_m = \rho.$$

The plausibility constraint is the information constraint; the attacker's incentive constraint is built into $\widetilde{u}_D$, which evaluates each posterior at the action it induces. The solution has a geometric form. The *concave closure* $\operatorname{cav} \widetilde{u}_D$ is the smallest concave function on $\Delta(\Theta)$ that is pointwise at least $\widetilde{u}_D$.

THEOREM 13.2.15 **Concavification [310]**
Let Θ, M, and $\mathcal{A}_A$ be finite with $|\Theta| \leq |M|$, and let ties be broken in the defender's favor. The optimal value of Problem 13.2.14 equals $\operatorname{cav} \widetilde{u}_D(\rho)$, and it is attained by a policy that uses at most $|\Theta|$ messages.

Proof. For any feasible $\{\nu_m, \mu_m\}$, concavity of $\operatorname{cav} \widetilde{u}_D$ and Bayesian plausibility give $\sum_m \nu_m \widetilde{u}_D(\mu_m) \leq \sum_m \nu_m \operatorname{cav} \widetilde{u}_D(\mu_m) \leq \operatorname{cav} \widetilde{u}_D(\rho)$. Conversely, $\operatorname{cav} \widetilde{u}_D(\rho)$ is the supremum of $\sum_i \nu_i \widetilde{u}_D(\mu_i)$ over finite convex combinations with $\sum_i \nu_i \mu_i = \rho$. The function $\widetilde{u}_D$ is upper semicontinuous: for each action a, the set of beliefs at which a is optimal is closed, and $\widetilde{u}_D$ is the maximum over a of the affine function $\sum_\theta \mu(\theta) u_D(\theta, a)$ restricted to that set. For fixed posteriors, the weights solve a linear program with $|\Theta|$ equality constraints, so by Carathéodory's theorem in the $(|\Theta| - 1)$-dimensional simplex an optimal vertex has at most $|\Theta|$ positive weights. The supremum can therefore be taken over combinations of $|\Theta|$ points, and upper semicontinuity on the compact set of such combinations ensures that it is attained. Proposition 13.2.13 implements the optimal posteriors with $|\Theta|$ messages. □

EXAMPLE 13.3 **An overt warning policy for a honey file**

▸ Code: `code/honey_file_persuasion.py`

An attacker finds `customer_backup.zip` in a compromised development workspace. The file is either a monitored honey file, $\theta = H$, or a sanitized real file, $\theta = N$, with prior $\rho(H) = 0.6$. The defender has announced

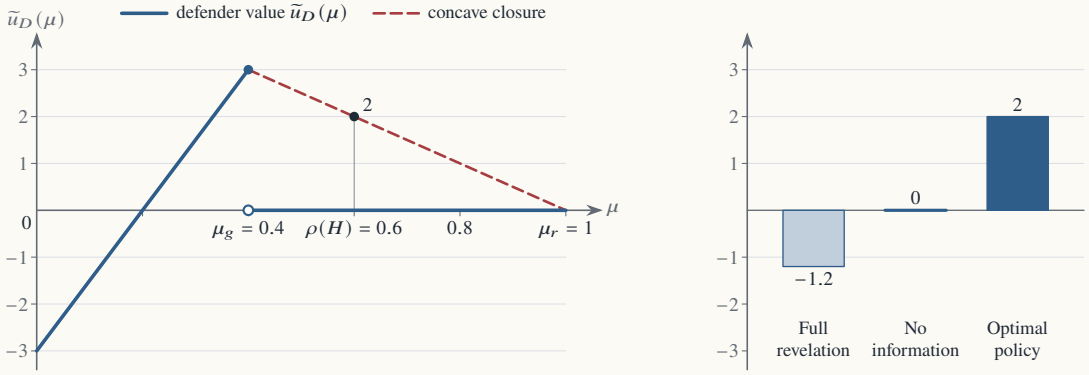
a labeling policy with a green label m_g and a red label m_r. The attacker opens the file or leaves; opening a real file gives it 10, opening a honey file gives -15, and leaving gives 0. With $\mu = \mu(H \mid m)$, the attacker opens if and only if $10(1-\mu) - 15\mu \geq 0$, that is, $\mu \leq \bar{\mu} = 0.4$. The defender wants controlled engagement: opening a honey file yields intelligence worth 12, opening a real file costs 3, and leaving yields 0. Hence

$$\widetilde{u}_D(\mu) = \begin{cases} 12\mu - 3(1-\mu) = 15\mu - 3, & \mu \leq 0.4, \\ 0, & \mu > 0.4. \end{cases}$$

At the prior $0.6 > 0.4$ the attacker leaves, so the defender's value without information is 0. Figure 13.6 shows $\widetilde{u}_D$ and its concave closure. The closure at $\rho(H) = 0.6$ is attained on the chord between the points $(0.4, 3)$ and $(1, 0)$: the green label makes the file as suspicious as possible while still inducing opening, $\mu_g = 0.4$, and the red label absorbs the remaining suspicion, $\mu_r = 1$. Bayesian plausibility, $0.4\,\nu_g + 1 \cdot (1-\nu_g) = 0.6$, gives $\nu_g = 2/3$ and $\nu_r = 1/3$, and the value is $\frac{2}{3}(15 \cdot 0.4 - 3) = 2$. By Proposition 13.2.13, the policy is

$$\pi(m_g \mid H) = \frac{(2/3)(0.4)}{0.6} = \frac{4}{9}, \qquad \pi(m_g \mid N) = \frac{(2/3)(0.6)}{0.4} = 1 :$$

every real file and $4/9$ of the honey files are labeled green, and the remaining $5/9$ of the honey files are labeled red.



(a) Defender value and concave closure **(b) Defender value**

Figure 13.6: Concavification for the honey-file warning policy of Example 13.3. (a) The solid line is the defender's value $\widetilde{u}_D(\mu)$ as a function of the attacker's honeypot posterior μ: the attacker opens the file for $\mu \leq 0.4$ and leaves for $\mu > 0.4$. The dashed line is the concave closure where it differs from $\widetilde{u}_D$. At the prior $\rho(H) = 0.6$, the optimal policy splits the prior into the posteriors $\mu_g = 0.4$ and $\mu_r = 1$, with probabilities $2/3$ and $1/3$, and raises the defender's value from 0 to 2. (b) The bars compare this value with those of full revelation and of no information.

Observation. The optimal policy is neither silence nor full revelation. Full revelation would induce the attacker to open every real file and leave every honey file, giving $0.4 \cdot (-3) = -1.2$, which is worse than silence. The optimal policy pools all real files with a controlled fraction of honey files under the green label, so that green is just credible enough to invite opening. If strict incentives are required, $\mu_g = 0.4 - \varepsilon$ gives a value arbitrarily close to 2.

The example illustrates the general structure of binary persuasion. When the attacker attacks at posteriors below a threshold $\bar{\mu}$, the prior exceeds $\bar{\mu}$, and the defender's value from an attack at $\bar{\mu}$ exceeds its value from withdrawal at the prior, as in the example, an optimal policy places one posterior exactly at $\bar{\mu}$ and the other at certainty. Cyber deception therefore does not need to make every artifact equally suspicious: a controlled distribution of suspicion, in which some observations invite engagement and others deter, is consistent with the attacker's prior and dominates uniform realism. Persuasion also describes the labels that platforms attach to user-generated misinformation; Yang, Li, and Zhu [607] couple it with a branching-process model of misinformation spread and find that transparent tagging is optimal even under misdetection.

Bayesian persuasion changes only information. It does not change the attacker's payoffs, its participation options,

or its initial trust in the defender's reports. Section 13.3 widens the design space so that the defender also chooses incentives and trust.

13.3 Mechanism Design Theory of Cyber Deception

Section 13.2 analyzed how an attacker responds to a given deception environment. Mechanism design, developed in general form in Chapter 9, reverses the direction of reasoning: the defender starts from a desired security outcome and asks which information, incentives, and operational rules make that outcome rational for the attacker.

Definition 13.3.1 (Cyber-deception mechanism)**.** A *cyber-deception mechanism* is a designed environment that maps hidden system states, user types, and defender choices into the observations, incentives, beliefs, and feasible actions of the user.

A mechanism may be a randomized feature pattern, a rule for placing honey credentials, a moving-target schedule, an authentication delay, a warning banner, a selective-disclosure rule, or a trust-management policy. It succeeds only if the response it induces improves security within the defender's safety and cost constraints.

13.3.1 Reverse Game-Theoretic Reasoning

Game-theoretic analysis takes the mechanism as given, whether a decoy, a monitoring policy, or a moving-target schedule, and predicts the attacker's inference, best response, and the resulting equilibrium. Mechanism design starts from a behavioral target, such as withdrawal from a critical asset, entry into a contained decoy, delayed lateral movement, or disclosure of tools, and searches for an information structure, an incentive structure, and a trust environment under which the target is a best response. The defender cannot command the attacker; it can only make the desired action optimal. Figure 13.7 contrasts the two directions.

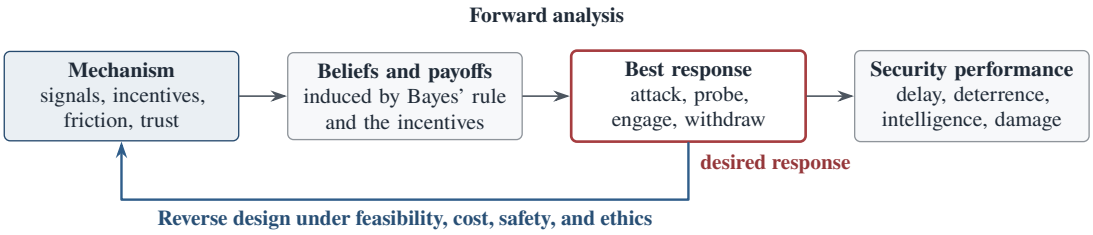


Figure 13.7: Forward analysis and reverse design. Game-theoretic analysis (upper arrows) maps a given mechanism to the induced beliefs and payoffs, the attacker's best response, and the resulting security performance. Mechanism design (lower arrow) starts from the desired response and searches, within the feasibility, cost, safety, and ethical constraints, for a mechanism whose induced best response is that response.

13.3.2 A General Deception-Mechanism Model

Let $\theta \in \Theta$ be the hidden state of the system, for example honeypot or production server, real or honey credential, monitored or unmonitored route, with true distribution $\rho \in \Delta(\Theta)$. Let $q \in Q$ be the private type of the user, for example careless insider, adversarial insider, automated scanner, or careful operator, and let $a \in \mathcal{A}_A$ be the user's action, such as access, verify, attack, or drop out. The defender's belief about types is $b_D(q \mid \theta)$.

Definition 13.3.2 (GMM deception mechanism [270])**.** A *generator–modulator–manipulator* (GMM) mechanism is a triple $\mathcal{M} = (\pi, \kappa, b_U)$, where

(1) the *generator* $\pi(m \mid \theta)$ chooses an observable message, or security policy, $m \in M$ conditional on the hidden state;

(2) the *modulator* κ shifts the user's payoffs, so that a type-q user who takes action a in state θ receives $u_U(\theta, q, a; \kappa)$ and the defender receives $u_D(\theta, q, a; \kappa)$, for example $u_U(\theta, q, a; \kappa) = u_U(\theta, q, a) - \kappa(a)$ for an action-dependent friction cost $\kappa(a) \geq 0$;

(3) the *manipulator* $b_U(\cdot \mid q) \in \Delta(\Theta)$ is the prior over states that a type-q user perceives.

The three components are three design levers. The generator maps the state into evidence and hence into posterior beliefs; operationally it consists of feature patterns, banners, files, routes, credentials, response timing, and honeypot configurations. The modulator represents authentication friction, delay, monitoring intensity, rate limits, penalties, and access restrictions, all of which shift the thresholds at which actions become attractive. The manipulator represents the trust environment in which signals are interpreted, such as the reported density of honeypots, the credibility of warnings, and the apparent reliability of detectors. Given $\mathcal{M}$, a type-q user who observes m forms the posterior

$$\mu_q(\theta \mid m) \triangleq \frac{b_U(\theta \mid q)\,\pi(m \mid \theta)}{\sum_{\theta' \in \Theta} b_U(\theta' \mid q)\,\pi(m \mid \theta')}$$

and chooses

$$a_q^*(m) \in \arg\max_{a \in \mathcal{A}_A} \sum_{\theta \in \Theta} \mu_q(\theta \mid m)\, u_U(\theta, q, a; \kappa),$$

with ties broken in the defender's favor.

Problem 13.3.3 (Cyber-deception mechanism design)**.** Let $\mathfrak{F}$ be the set of feasible mechanisms, those satisfying the containment, cost, usability, legal, and safety constraints. The defender solves

$$\max_{(\pi, \kappa, b_U) \in \mathfrak{F}} \sum_{\theta \in \Theta} \rho(\theta) \sum_{m \in M} \pi(m \mid \theta) \sum_{q \in Q} b_D(q \mid \theta)\, u_D\big(\theta, q, a_q^*(m); \kappa\big).$$

If only π can be chosen, Problem 13.3.3 is an information-design problem; if κ can also be chosen, it includes incentive design; if b_U can be shaped, it includes trust design. Bayesian persuasion is the special case with a single type, a fixed modulator, and $b_U = \rho$; Section 13.3.5 returns to it.

PROPOSITION 13.3.4 **Existence of an optimal mechanism**
Let Θ, Q, M, and $\mathcal{A}_A$ be finite, and let $\mathfrak{F}$ be a nonempty compact subset of a Euclidean space on which u_U and u_D are continuous. If ties are broken in the defender's favor, Problem 13.3.3 has an optimal solution.

Proof. For a message m and type q, write $\bar{u}_{U,a}(\mathcal{M}) \triangleq \sum_\theta b_U(\theta \mid q)\pi(m \mid \theta)u_U(\theta, q, a; \kappa)$, which is proportional to the user's posterior value of a, and $\bar{u}_{D,a}(\mathcal{M}) \triangleq \sum_\theta \rho(\theta)\pi(m \mid \theta)b_D(q \mid \theta)u_D(\theta, q, a; \kappa)$, the defender's contribution when a is chosen. Both are continuous on $\mathfrak{F}$. The set $F_a \triangleq \{\mathcal{M} : \bar{u}_{U,a} \geq \bar{u}_{U,a'}\ \forall a'\}$ on which a is a best response is closed. The defender-favorable contribution of (m, q) is $\max_a \hat{u}_{D,a}$, where $\hat{u}_{D,a} \triangleq \bar{u}_{D,a}$ on F_a and $\hat{u}_{D,a} \triangleq -\infty$ off F_a; each $\hat{u}_{D,a}$ is upper semicontinuous, and so are the finite maximum and the finite sum over (m, q). An upper semicontinuous function attains its maximum on a nonempty compact set. □

13.3.3 Incentive Compatibility and Enforceability

A mechanism is implemented through the user's choices, not through the defender's intentions. Suppose the defender wants type q to choose $a^\dagger$ after message m. The *incentive margin* of this assignment is

$$\Gamma_q(m, a^\dagger) \triangleq \min_{a \in \mathcal{A}_A \setminus \{a^\dagger\}} \sum_{\theta \in \Theta} \mu_q(\theta \mid m)\big[u_U(\theta, q, a^\dagger; \kappa) - u_U(\theta, q, a; \kappa)\big].$$

PROPOSITION 13.3.5 **Incentive-margin test**
Fix a mechanism $\mathcal{M} = (\pi, \kappa, b_U)$, a type q, and a message m that type q observes with positive probability. The action $a^\dagger$ is a best response of type q after m if and only if $\Gamma_q(m, a^\dagger) \geq 0$. If $\Gamma_q(m, a^\dagger) < 0$, the minimizing action gives type q a strictly higher expected utility than $a^\dagger$.

Proof. The action $a^\dagger$ is a best response if and only if each difference in the definition of Γ_q is nonnegative, which holds if and only if their minimum is nonnegative. If the minimum is negative, the minimizing action yields a strictly higher expected utility than $a^\dagger$. □

A positive margin makes the assignment robust to small errors in beliefs and payoffs; a zero margin leaves the user indifferent; a negative margin identifies the alternative action that defeats the design. The levers act on the margin in different ways: the generator changes the posterior, the modulator changes the payoff differences, and the manipulator changes the prior from which the posterior is formed. If no feasible combination makes the margin nonnegative, the desired behavior is not enforceable and the defender must choose a different objective. Participation is a special case. If the defender wants a user to continue into a decoy, engagement must beat dropping out; if it wants withdrawal, the drop-out action must beat attack, verification, and lateral movement.

13.3.4 Duplicity Games

Duplicity games make the GMM design space explicit for settings with two-sided asymmetric information [270]. The defender knows or controls the state θ, and the user privately knows its type $q \in Q = \{q_1, \dots, q_n\}$. The action set contains a drop-out action a_{DO} that represents nonparticipation. Since every message induces a type-contingent action profile, messages can be labeled by the profiles they are intended to induce: $m_{(a_1,\dots,a_n)}$ denotes a security policy under which type q_ℓ is intended to choose a_ℓ. With $|\mathcal{A}_A|$ actions there are at most $|\mathcal{A}_A|^n$ such labels.

Problem 13.3.6 (GMM duplicity design)**.** Choose a feasible mechanism (π, κ, b_U), with π supported on the labels $m_{(a_1,\dots,a_n)}$, to solve

$$\begin{aligned} \max_{\pi,\kappa,b_U} \quad & \sum_{\theta\in\Theta} \rho(\theta) \sum_{(a_1,\dots,a_n)} \pi\big(m_{(a_1,\dots,a_n)} \mid \theta\big) \sum_{\ell=1}^{n} b_D(q_\ell \mid \theta)\, u_D(\theta, q_\ell, a_\ell; \kappa) \\ \text{s.t.} \quad & \sum_{\theta\in\Theta} \big[u_U(\theta, q_\ell, a_\ell; \kappa) - u_U(\theta, q_\ell, a; \kappa)\big]\, b_U(\theta \mid q_\ell)\, \pi\big(m_{(a_1,\dots,a_n)} \mid \theta\big) \ge 0 \quad \forall a,\ \ell,\ (a_1,\dots,a_n), \\ & \kappa(a_{\mathrm{DO}}) = 0, \qquad (\pi, \kappa, b_U) \in \mathfrak{F}. \end{aligned}$$

The inequality constraint is the incentive-margin condition of Proposition 13.3.5 multiplied by the positive normalizing constant of the posterior; it is linear in π for fixed (κ, b_U), so that the generator problem is a linear program whenever $\mathfrak{F}$ imposes linear constraints on π. The normalization $\kappa(a_{\mathrm{DO}}) = 0$ prevents the modulator from acting as an unconstrained transfer.

THEOREM 13.3.7 **Action-profile enforceability**
Fix a modulator κ. For each type q and action a, the best-response region

$$\mathcal{R}_a^q(\kappa) \triangleq \Big\{\mu \in \Delta(\Theta) : \sum_{\theta\in\Theta} \mu(\theta)\big[u_U(\theta, q, a; \kappa) - u_U(\theta, q, a'; \kappa)\big] \ge 0 \ \ \forall a' \in \mathcal{A}_A\Big\}$$

is a closed convex polytope, possibly empty, and the regions $\{\mathcal{R}_a^q(\kappa)\}_{a\in\mathcal{A}_A}$ cover $\Delta(\Theta)$. A label $m_{(a_1,\dots,a_n)}$ is enforceable if and only if, for every ℓ, the posterior of type q_ℓ after the label lies in $\mathcal{R}_{a_\ell}^{q_\ell}(\kappa)$. In particular, if all types share the perceived prior, the profile $(a_1, \dots, a_n)$ can be induced by some generator only if $\bigcap_\ell \mathcal{R}_{a_\ell}^{q_\ell}(\kappa) \neq \emptyset$.

Proof. Each defining inequality is affine in μ, so $\mathcal{R}_a^q(\kappa)$ is the intersection of the simplex with finitely many closed half-spaces. Every μ has at least one maximizing action, so the regions cover the simplex. The enforceability statement restates Proposition 13.3.5 for each type. When all types share the perceived prior, they also share the posterior after any label, and that common posterior must lie in every $\mathcal{R}_{a_\ell}^{q_\ell}(\kappa)$. □

The geometry of the best-response regions determines which behavior patterns are implementable: although there are $|\mathcal{A}_A|^n$ type-contingent profiles, only those with suitable belief regions can appear in a mechanism. The regions also determine whether types can be told apart by their behavior. Two types are behaviorally separable at a posterior if their best responses there differ. If their utilities are related by $u_U(\theta, q', a) = \Omega\, u_U(\theta, q, a) + t(\theta)$ with $\Omega > 0$, they choose the same actions at every posterior and, if they perceive the same prior, no information policy separates them; if $\Omega < 0$, their preferences are reversed and they are easier to separate. When types perceive different priors, the manipulator can place them at different posteriors after the same label, which is how Example 13.4 separates them.

EXAMPLE 13.4 A GMM mechanism for an insider-honeypot repository

▸ Code: `code/gmm_repository.py`

A company protects a sensitive source-code repository. The hidden state is whether a valuable-looking repository is the production repository, N, or a contained honeypot mirror, H, with $\rho(N) = 0.65$ and $\rho(H) = 0.35$. The user is either a careless insider q_C, who browses out of convenience but wants to avoid an explicit policy violation, or an adversarial insider q_A, who accepts risk for valuable information; the defender believes $b_D(q_A \mid \theta) = 0.3$ in both states. The actions are to drop out and request authorization, a_{DO}, to verify through an approved channel, a_V, and to enter and enumerate the repository, a_E. The defender wants the profile (a_{DO}, a_E): the careless type drops out, and the adversarial type enters the honeypot, where its tools and intent are observed.

The three levers are concrete. The generator decides when a repository is exposed with an engagement profile: plausible project name, commit history, stale deployment notes, and planted but monitored secrets. The modulator adds step-up authentication, visible audit logging, and throttled downloads. The manipulator is the trust environment: training and documentation have convinced careless insiders that half of such repositories are instrumented, $b_U(H \mid q_C) = 0.5$, whereas adversarial insiders believe $b_U(H \mid q_A) = 0.2$. The user payoffs, written as (value in N, value in H), are

$$\begin{aligned} q_C: &\quad u_U(a_{\mathrm{DO}}) = (0,0), \quad u_U(a_V) = (1,-2), \quad u_U(a_E) = (4,-12), \\ q_A: &\quad u_U(a_{\mathrm{DO}}) = (0,0), \quad u_U(a_V) = (2,-2), \quad u_U(a_E) = (10,-6). \end{aligned}$$

The defender's payoff is 1 when a careless user drops out, 0 when an adversarial user drops out, 20 when an adversarial user enters the honeypot, and -15 when it enters the production repository.

Labels. The defender uses two labels. The protected label m_S carries the profile $(a_{\mathrm{DO}}, a_{\mathrm{DO}})$ and the modulator $\kappa(a_V) = \kappa(a_E) = 12$, a friction cost subtracted from the user's payoff of each of these actions, which makes a_{DO} optimal for both types at every posterior, because no action is worth more than 10 to either type. The engagement label m_E carries the profile (a_{DO}, a_E) and no friction. Let $\pi_N \triangleq \pi(m_E \mid N)$ and $\pi_H \triangleq \pi(m_E \mid H)$. Safety limits production exposure to $\pi_N \le 0.12$, and honeypot capacity limits $\pi_H \le 0.75$.

Incentive constraints. With honeypot posterior μ, the careless type prefers a_{DO} to a_V if $1 - 3\mu \le 0$ and to a_E if $4 - 16\mu \le 0$, that is, $\mu \ge 1/3$. After m_E its posterior is $\mu_C = 0.5\pi_H/(0.5\pi_H + 0.5\pi_N)$, so the constraint is $\pi_H \ge \pi_N/2$. The adversarial type prefers a_E to a_{DO} if $10 - 16\mu \ge 0$ and to a_V if $8 - 12\mu \ge 0$, that is, $\mu \le 0.625$; its posterior is $\mu_A = 0.2\pi_H/(0.2\pi_H + 0.8\pi_N)$, so the constraint is $\pi_H \le \frac{20}{3}\pi_N$. Figure 13.8 shows the two belief regions.

Linear program. Under m_S the defender's expected payoff is 0.7 in either state; under m_E it is $0.7 \cdot 1 + 0.3 \cdot (-15) = -3.8$ in state N and $0.7 \cdot 1 + 0.3 \cdot 20 = 6.7$ in state H. The design problem is therefore

$$\max_{\pi_N, \pi_H} \; 0.7 - 2.925\,\pi_N + 2.1\,\pi_H \quad \text{s.t.} \quad \tfrac{1}{2}\pi_N \le \pi_H \le \tfrac{20}{3}\pi_N, \quad 0 \le \pi_N \le 0.12, \quad 0 \le \pi_H \le 0.75.$$

For fixed π_H the objective decreases in π_N, so the adversarial constraint binds, $\pi_N = \frac{3}{20}\pi_H$. On this boundary the objective increases at the rate $2.1 - 2.925 \cdot 0.15 > 0$ in π_H, so the optimum is $\pi_H^* = 0.75$ and $\pi_N^* = 0.75 \cdot \frac{3}{20} = 0.1125$, with value 1.946. The induced posteriors after m_E are $\mu_C = 20/23 \approx 0.870 \ge 1/3$ and $\mu_A = 0.625$, so the careless type drops out and the adversarial type is just willing to enter. Figure 13.9 shows the feasible set and the optimal generator.

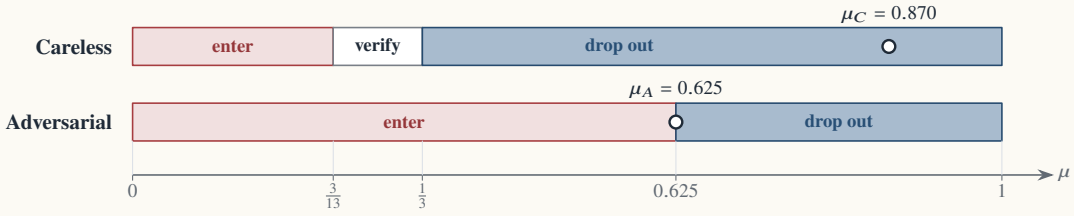
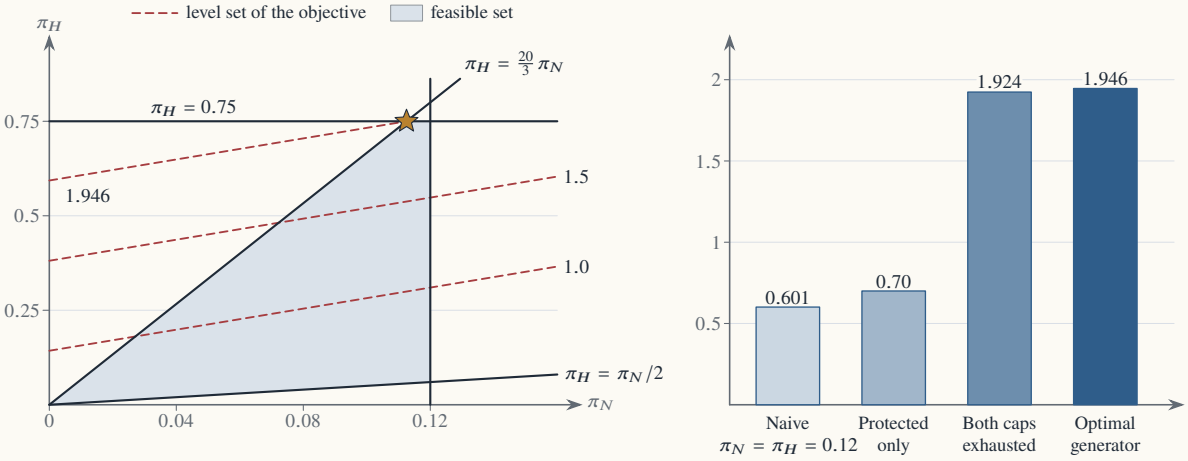


Figure 13.8: Best-response regions of the two insider types in Example 13.4 on the honeypot-posterior axis. The careless type enters for $\mu \le 3/13$, verifies for $3/13 \le \mu \le 1/3$, and drops out for $\mu \ge 1/3$; the adversarial type enters for $\mu \le 0.625$ and drops out above. The markers show the posteriors induced by the engagement label at the optimum: $\mu_C = 0.870$ lies in the careless type's drop-out region, and $\mu_A = 0.625$ lies on the boundary of the adversarial type's entry region. The two types reach different posteriors from the same label because they perceive different priors.



(a) Feasible generators **(b) Defender value**

Figure 13.9: The generator design problem of Example 13.4. (a) The feasible set (shaded) of $\pi_N = \pi(m_E \mid N)$ and $\pi_H = \pi(m_E \mid H)$ is bounded by the careless type's constraint $\pi_H \ge \pi_N/2$, the adversarial type's constraint $\pi_H \le \frac{20}{3}\pi_N$, the safety cap $\pi_N \le 0.12$, and the capacity cap $\pi_H \le 0.75$; dashed lines are level sets of the defender's objective, and the star marks the optimum $(0.1125, 0.75)$. (b) The defender's value under four policies: the naive policy $\pi_N = \pi_H = 0.12$, the protected label alone, the policy that exhausts both caps, and the optimal generator.

Observation. The optimal generator exposes the production repository with the engagement profile as rarely as the adversarial type's incentive constraint allows. The protected policy alone yields 0.70, exhausting both caps yields 1.924, and the optimum yields 1.946. A naive policy that shows the engagement profile equally often in both states, $\pi_N = \pi_H = 0.12$, yields only 0.601, less than the protected policy, because it exposes production as often as it engages the adversary.

The example also shows how the framework diagnoses failure. If the engagement profile looks too artificial, or the perceived density of honeypots is too high, the adversarial posterior exceeds 0.625 and the adversary drops out, so the defender loses intelligence. If the profile looks too genuine and friction is too weak, the careless posterior falls below $1/3$ and careless users no longer drop out: they verify, or, below $3/13$, enter and produce false alarms. If the profile is ever deployed on production without containment, the mechanism benefits the attacker rather than the defender.

13.3.5 Information, Incentives, and Trust

In the information-only case, a single type perceives the true prior, $b_U = \rho$, and the modulator is fixed, so Problem 13.3.6 reduces to Bayesian persuasion, and Theorem 13.2.15 gives its value as $\operatorname{cav} \widetilde{u}_D(\rho)$, where $\widetilde{u}_D(\mu)$ is the defender's value when the user acts on belief μ. The gain from information has a name.

Definition 13.3.8 (Trust margin and manageability)**.** In the information-only case with prior ρ, the *trust margin* of a generator π is its value minus $\widetilde{u}_D(\rho)$, the value of the uninformative generator. A user is *manageable* at ρ if the largest trust margin over feasible generators is positive.

PROPOSITION 13.3.9 **Aligned and opposed utilities**
Consider the information-only case for a single type q with unconstrained generators, and suppose that

$$u_D(\theta, q, a) = \Omega\, u_U(\theta, q, a) + t(\theta)$$

for a constant Ω and a function t that does not depend on the action. If $\Omega \le 0$, no generator has a positive trust margin, and the uninformative generator is optimal. If $\Omega > 0$, full revelation of θ is optimal and its trust margin is nonnegative; it is positive whenever no single action is a best response of the user in every state in the support of ρ.

Proof. The user's value $V_U(\mu) \triangleq \max_a \sum_\theta \mu(\theta) u_U(\theta, q, a)$ is a maximum of affine functions of μ, hence convex. Because the user's chosen action maximizes u_U, $\widetilde{u}_D(\mu) = \Omega V_U(\mu) + \sum_\theta \mu(\theta) t(\theta)$. If $\Omega \le 0$, $\widetilde{u}_D$ is concave, so $\operatorname{cav} \widetilde{u}_D = \widetilde{u}_D$ and, by Theorem 13.2.15, no generator beats $\widetilde{u}_D(\rho)$. If $\Omega > 0$, $\widetilde{u}_D$ is convex; for any Bayes-plausible split, Jensen's inequality applied to each posterior $\mu_m = \sum_\theta \mu_m(\theta) \mathbf{1}_{\{\theta\}}$, together with Bayesian plausibility, gives $\sum_m \nu_m \widetilde{u}_D(\mu_m) \le \sum_\theta \rho(\theta) \widetilde{u}_D(\mathbf{1}_{\{\theta\}})$, which is the value of full revelation. Convexity gives $\widetilde{u}_D(\rho) \le \sum_\theta \rho(\theta) \widetilde{u}_D(\mathbf{1}_{\{\theta\}})$, with strict inequality when no single action is optimal in every state of the support of ρ, because each affine piece $\sum_\theta \mu(\theta) u_U(\theta, q, a)$ meets the chord $\sum_\theta \mu(\theta) V_U(\mathbf{1}_{\{\theta\}})$ at ρ only if a is optimal in every such state. □

The proposition makes the phrase "the user can be managed by deception" precise. When the defender's and the user's interests are opposed, $\Omega < 0$, information only helps the user choose what the defender dislikes, and silence is optimal. When they are aligned, the defender should reveal everything, and deception is unnecessary. Deception is valuable in the intermediate cases, as in Example 13.3, where the defender values some of the user's actions in some states and not in others.

PROPOSITION 13.3.10 **Modulator and manipulator benchmarks**
The modulator κ changes the coefficients of the inequalities that define $\mathcal{R}_a^q(\kappa)$ and can therefore make an action profile enforceable or unenforceable. If, in the information-only case, an unconstrained manipulator could choose any common perceived prior $b_U \in \Delta(\Theta)$ at no cost, the joint design of generator and manipulator would attain at least $\max\{\sum_\theta \rho(\theta) u_D(\theta, q, a; \kappa) : a \in \mathcal{A}_A,\ \mathcal{R}_a^q(\kappa) \neq \emptyset\}$, a value achieved by an uninformative generator, and at most $\sum_\theta \rho(\theta) \max_{a \in \mathcal{A}_A} u_D(\theta, q, a; \kappa)$.

Proof. The first statement follows from the definition of $\mathcal{R}_a^q(\kappa)$, whose coefficients are differences of $u_U(\cdot; \kappa)$. For the second, if $\mathcal{R}_a^q(\kappa) \neq \emptyset$, choosing $b_U \in \mathcal{R}_a^q(\kappa)$ and sending no information makes a a best response in every state, so that, with ties broken in the defender's favor, the defender receives at least $\sum_\theta \rho(\theta) u_D(\theta, q, a; \kappa)$; the defender's payoff is evaluated under the true prior ρ, not under b_U. The upper bound holds because in each state θ the defender receives at most $\max_a u_D(\theta, q, a; \kappa)$. □

The lower bound can exceed the persuasion value by a wide margin. In Example 13.3, a perceived prior $b_U(H) \le 0.4$ makes the attacker open the file in both states and yields $0.6 \cdot 12 + 0.4 \cdot (-3) = 6$, three times the persuasion value 2.

The benchmark is not a recommendation. Real manipulation of trust is limited by law, ethics, auditability, and the attacker's own learning, and under these limits the three levers are complements: the generator changes what is observed, the modulator changes the payoff comparison, and the manipulator changes the prior from which observations are read. A successful design is not merely a realistic decoy. It requires identifying the enforceable action profiles, deciding whether the types are behaviorally separable, verifying that information yields a positive trust margin, and checking that the friction imposed on legitimate users is acceptable. Comparable levers appear in language-mediated deception: in the semantic signaling game of Zhu [651], a language model generates the message from the sender's semantic control, the receiver's awareness is a type that determines which linguistic features it perceives, and reshaping awareness, penalizing deceptive controls, and changing the receiver population serve as mechanisms that induce benign pooling equilibria.

13.4 Dynamic Cyber Deception

Cyber deception is usually dynamic. An attacker probes a decoy, updates its beliefs, moves laterally, retreats, escalates, or returns with new tools; the defender adapts the deceptive environment, rotates services, reveals or conceals information,

engages the attacker, or isolates hosts. Static deception asks whether one observation changes one action; dynamic deception asks how deception should evolve over time. Time matters for three reasons. Attacks unfold in stages, from reconnaissance through initial access, privilege escalation, and lateral movement to exfiltration or impact, and a deception that is useful in one stage may be useless in another. Both sides learn: the attacker about the defender's environment, and the defender about the attacker's type, goal, and capability. Finally, engagement creates a trade-off, because keeping an attacker inside a controlled environment produces intelligence and delay but also risk.

Definition 13.4.1 (Dynamic Bayesian deception game)**.** A *dynamic Bayesian deception game* with horizon K consists of a defender D with private type $\theta_D \in \Theta_D$, for example primitive or sophisticated monitoring; a user or attacker A with private type $\theta_A \in \Theta_A$, for example legitimate user or advanced persistent threat; and, for each stage $k = 0, 1, \dots, K$, a finite set $\mathcal{X}^k$ of technical states, finite action sets $\mathcal{D}^k$ and $\mathcal{A}^k$, a transition law $P^k(x^{k+1} \mid x^k, d^k, a^k, \theta_D, \theta_A)$, and stage utilities u_D^k, u_A^k of $(x^k, d^k, a^k, \theta_D, \theta_A)$. The types are drawn from a common prior; at stage k, both players observe x^k, player i holds a belief $b_i^k \in \Delta(\Theta_j)$ about its opponent j's type, and the players choose d^k and a^k according to behavioral strategies $\sigma_D^k(\cdot \mid x^k, \theta_D)$ and $\sigma_A^k(\cdot \mid x^k, \theta_A)$.

The technical state records the part of the environment that affects future evolution: exposed services, attacker location, privilege level, decoy status, alert history, and response posture. The stage utilities measure security value, attack value, service cost, deception cost, intelligence gain, and risk. Figure 13.10 shows the timing of one stage. Each player knows its own type and maintains a belief about the other's; beliefs and state determine the strategies, the strategies determine the actions, the actions move the state, and the state transition provides the evidence for the next belief update. The defender's policy changes what the attacker believes, and the attacker's policy changes what the defender observes next; this two-way coupling is what distinguishes dynamic deception from repeated control. If the attacker's behavior is fixed in advance, the model reduces to a partially observed Markov decision process for the defender [266].

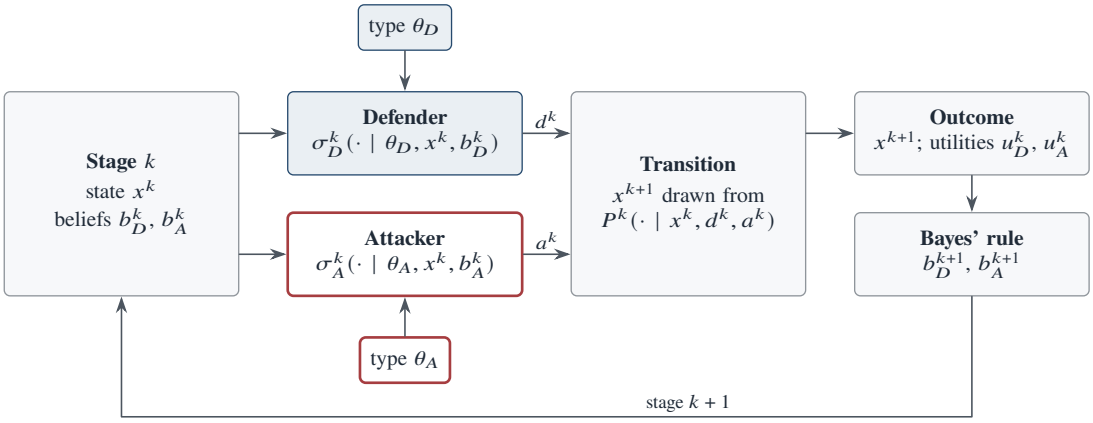


Figure 13.10: One stage of the dynamic Bayesian deception game. The private types θ_D and θ_A condition the players' strategies. At stage k the state x^k and the beliefs b_D^k, b_A^k determine the behavioral strategies, which generate the actions d^k and a^k; the transition P^k produces the next state x^{k+1} and the stage utilities, and Bayes' rule converts the observed transition into the beliefs b_D^{k+1}, b_A^{k+1} used at stage $k+1$.

The expected utility of player i from stage k_0 onward is

$$U_i^{k_0}(\sigma_D, \sigma_A \mid x^{k_0}, \theta_i) \triangleq \mathbb{E}\Big[\sum_{k=k_0}^{K} u_i^k(x^k, d^k, a^k, \theta_D, \theta_A) \,\Big|\, x^{k_0}, \theta_i\Big],$$

where the expectation is taken over the opponent's type according to player i's belief, over both players' mixed actions, and over the state transitions.

Definition 13.4.2 (PBNE and ε-PBNE)**.** A profile $\big((\sigma_D, \sigma_A), (b_D, b_A)\big)$ of strategies and beliefs is a *perfect Bayesian Nash equilibrium* if (1) the beliefs are obtained by Bayes' rule from the prior and the strategies whenever the observed history has positive probability, and (2) at every stage, state, and type, no player can raise its expected continuation

utility U_i^k by a unilateral deviation, given its beliefs. If condition (2) holds up to an additive loss $\varepsilon \geq 0$ and condition (1) holds on the equilibrium path, the profile is an *ε-PBNE*. If the beliefs are fixed exogenously and only condition (2) is imposed, the strategy profile is a *dynamic Bayesian Nash equilibrium under those beliefs*.

13.4.1 State, Actions, and Rewards

The decision state of player i at stage k is the triple (x^k, θ_i, b_i^k): what the player observes about the system, what it privately knows about itself, and what it believes about its opponent. The defender's action d^k may engage the attacker, rotate an address, reveal or conceal a cue, add a decoy, isolate a host, or restore a service. The attacker's action a^k may probe, exploit, move laterally, exfiltrate, wait, verify, or withdraw. The transition is both technical and informational: a rotation changes which services are exposed, engagement changes what the defender learns, a probe changes the defender's belief about the attacker's intent, and containment changes the actions available later.

For a fixed belief sequence and a fixed continuation strategy of the opponent, player i's continuation value satisfies the recursion

$$V_i^k(x^k, \theta_i) = \max_{\sigma_i^k} \mathbb{E}\Big[u_i^k(x^k, d^k, a^k, \theta_D, \theta_A) + V_i^{k+1}(x^{k+1}, \theta_i)\Big], \qquad V_i^{K+1} \equiv 0, \tag{13.4.1}$$

where the expectation is over the opponent's type, both players' mixed actions, and $x^{k+1} \sim P^k(\cdot \mid x^k, d^k, a^k, \theta_D, \theta_A)$. The recursion is game-theoretic because the opponent's continuation strategy is part of the equilibrium rather than a given of nature.

THEOREM 13.4.3 **Sequential-rationality recursion**
Fix a finite horizon and a belief sequence $\{b_i^k\}_{k=0}^K$. A strategy profile σ^* is a dynamic Bayesian Nash equilibrium under these beliefs if and only if, for each player i, stage k, state x^k, and type θ_i, the stage strategy $\sigma_i^{*,k}$ attains the maximum in (13.4.1) against $\sigma_j^{*,k}$, with V_i^{k+1} computed from $\sigma^{*,k+1:K}$.

Proof. The argument is backward induction, as for the subgame-perfect equilibria of Chapter 3. At stage K the continuation value is zero, and (13.4.1) is the one-stage Bayesian best-response condition. Suppose the claim holds from stage $k+1$ onward. A unilateral deviation at stage k changes the expected stage utility and the distribution of x^{k+1}, after which the continuation value is V_i^{k+1} by the induction hypothesis; no deviation is profitable at stage k if and only if $\sigma_i^{*,k}$ maximizes the bracket in (13.4.1). By the one-shot deviation principle for finite-horizon games, the absence of profitable one-stage deviations at every decision state is equivalent to sequential rationality. □

13.4.2 Beliefs as State Variables

In deception the state is rarely fully observed, and the incomplete information is two-sided. The defender may not know whether a user is legitimate, whether an attacker is automated, or whether it has discovered the honeypot; the attacker may not know whether the defender is sophisticated, whether a host is a decoy, or whether a quiet action has already triggered monitoring. Beliefs convert observations into strategic state variables. Let player i, having taken its own action d_i^k (which is d^k for the defender and a^k for the attacker), observe the transition from x^k to x^{k+1}. The likelihood of this transition under opponent type θ_j is

$$\Lambda_i^k(\theta_j) \triangleq \sum_{d_j \in \mathcal{A}_j^k} \sigma_j^k(d_j \mid x^k, \theta_j)\, P^k\big(x^{k+1} \mid x^k, d^k, a^k, \theta_D, \theta_A\big),$$

where $\mathcal{A}_j^k$ is the opponent's action set ($\mathcal{D}^k$ if $j = D$ and $\mathcal{A}^k$ if $j = A$), the opponent's action d_j takes the place of d^k or a^k, respectively, and player i updates by

$$b_i^{k+1}(\theta_j) = \frac{\Lambda_i^k(\theta_j)\, b_i^k(\theta_j)}{\sum_{\theta_j' \in \Theta_j} \Lambda_i^k(\theta_j')\, b_i^k(\theta_j')}. \tag{13.4.2}$$

The likelihood is induced by the opponent's strategy and the transition law, not by a sensor model alone: the next state is informative because different types choose different actions with different probabilities.

PROPOSITION 13.4.4 **Bayes-consistent Markov belief**
If the denominator of (13.4.2) is positive, then b_i^{k+1} is a probability distribution on Θ_j and equals player i's conditional distribution of the opponent's type given the observed transition from x^k to x^{k+1}, its own action, and its previous belief. An opponent type that assigns zero likelihood to the observed transition receives posterior probability zero, and if all types assign the same positive likelihood, the posterior equals the previous belief b_i^k.

Proof. The denominator $Z \triangleq \sum_{\theta_j'} \Lambda_i^k(\theta_j') b_i^k(\theta_j')$ is the total probability of the observed transition under player i's belief. When $Z > 0$, the right-hand side of (13.4.2) is nonnegative and sums to one, and it is Bayes' formula for the conditional probability of θ_j. The two remaining claims follow from the numerator. □

A defender's belief moves toward an adversarial type after a suspicious command sequence, unusual lateral movement, or repeated access to decoy files; an attacker's belief moves toward a sophisticated defender after delayed access, suppressed alerts, or inconsistent reconnaissance. Each posterior enters the next decision: a defender who believes it faces an advanced persistent threat (APT) may shorten engagement and contain, and an attacker who believes the defender is sophisticated may slow down or choose a stealthier path. An alert is an observation; a belief is a strategic interpretation of a sequence of observations. Dynamic deception policies should depend on beliefs, not on the latest raw alert, a point also made by the IDS-alert example of Appendix B.

13.4.3 Engagement Versus Containment

► Code: `code/dynamic_engagement.py`

The central dynamic trade-off is between engagement and containment. *Engagement* lets the adversary continue inside a controlled environment so that the defender observes its tools, intent, infrastructure, and decision making; *containment* ends or sharply restricts the interaction by isolating a host, revoking a credential, blocking a session, or redirecting traffic. Engagement creates intelligence, delay, and attribution evidence, but also exposure, collateral damage, legal risk, and the possibility that the adversary recognizes the deception. Chapter 7 studies the engagement problem as a semi-Markov decision process (Section 7.4.1) learned by reinforcement learning [264]; here the trade-off is isolated in its simplest form.

At stage k, let b^k be the defender's belief about the attacker's type and intent, z^k the operational exposure, and h^k the probability that the attacker has recognized the decoy. Engagement yields intelligence I^k and delay value δ^k and incurs a risk cost $C^k(z^k, h^k)$.

PROPOSITION 13.4.5 **Dynamic engagement threshold**
Let V be the defender's continuation value, let $\gamma \in (0, 1]$ be the discount factor, and define

$$V_E \triangleq I^k + \delta^k - C^k(z^k, h^k) + \gamma\, \mathbb{E}\big[V(b^{k+1}, z^{k+1}, h^{k+1}) \mid \text{engage}\big], \qquad V_C \triangleq \gamma\, \mathbb{E}\big[V(b^{k+1}, z^{k+1}, h^{k+1}) \mid \text{contain}\big].$$

If engagement and containment both satisfy the hard safety constraints, engagement is optimal at stage k if and only if its surplus $S^k \triangleq V_E - V_C$ is nonnegative.

Proof. By the recursion (13.4.1), with the continuation value discounted by γ, the defender chooses the feasible action with the larger value of stage utility plus discounted continuation value, and V_E and V_C are these values for engagement and containment. □

The rule differs from the myopic rule $I^k + \delta^k - C^k \geq 0$ through the continuation terms. A myopic defender continues too long when engagement raises the probability of a later escape or discovery, because the cost appears only in the continuation value. Hard safety constraints are not part of the surplus: no access to production data, no outbound attack traffic from the decoy, no uncontrolled malware execution, and no engagement beyond a protected boundary, however favorable the value calculation.

EXAMPLE 13.5 **Myopic and forward-looking engagement**

A defender can engage an intruder in a decoy for at most three periods, with $\gamma = 1$. Each period of engagement yields intelligence $I = 2$, and the risk costs are $C^1 = 0.5$, $C^2 = 1.5$, and $C^3 = 4$. The myopic net values are 1.5, 0.5, and -2, so a myopic defender engages in periods 1 and 2 and contains before period 3, for a total of 2. Suppose now that engagement in period 2 lets the intruder stage an escape route, which causes an expected loss $L = 3$ in period 3 whether or not the defender engages then. By backward induction, containment is optimal in period 3; in period 2 the surplus is $S^2 = 2 - 1.5 - 3 = -2.5 < 0$, so the defender contains; and in period 1 the surplus is $1.5 > 0$.

Observation. The forward-looking defender engages only in period 1 and obtains 1.5, whereas the myopic defender engages in periods 1 and 2 and obtains $1.5 + 0.5 - 3 = -1$. Period 2 looks locally profitable, but its cost appears only in the continuation value.

Engagement is therefore an experiment under risk. A further period of engagement may distinguish a commodity scanner from a human operator, reveal command-and-control infrastructure, or expose a planned lateral path; such observations are valuable only to the extent that they change future containment, recovery, attribution, or hardening decisions. Operational policies are typically layered: early suspicious behavior is redirected to a high-interaction decoy, privilege escalation triggers stronger isolation while observation continues, approach to production assets triggers immediate containment, and evidence that the attacker has recognized the deception triggers ejection or a switch to moving-target defense, since continued engagement would teach the attacker to recognize future decoys.

13.4.4 Multi-Stage Bayesian Deception Against APTs

An advanced persistent threat proceeds through a sequence of stages in which the attacker gathers intelligence, gains initial access, escalates privileges, moves laterally, and may finally affect a physical or mission process. Huang and Zhu model this interaction as a multi-stage Bayesian game with two-sided incomplete information [266]: the defender does not know whether a user is legitimate or adversarial, and the attacker does not know whether the defender is primitive or sophisticated. *Defensive deception* keeps the defender's own type uncertain and so induces conservative attacker behavior; *adversarial deception* disguises malicious behavior as legitimate activity. Chapter 12 develops this model in full, with its Tennessee Eastman case study; this subsection extracts the computational structure that dynamic deception requires.

In the full Bayesian game, a player can update on the entire action history, which grows with every stage. The Markov representation replaces the history by the state x^k, chosen to retain the information that matters for future payoffs and belief updates, and the belief is then updated by (13.4.2). Backward computation evaluates the optimal strategies for a given belief sequence, and forward updating recomputes the belief sequence that those strategies induce; an equilibrium is a fixed point of the two operations.

THEOREM 13.4.6 **Forward-belief/backward-policy fixed point [266]**

For a candidate belief sequence $b = \{b_i^k\}$, let $\sigma^\star(b)$ be a strategy profile obtained from the recursion (13.4.1) under b, and let $\Phi(b)$ be the belief sequence generated from the prior by (13.4.2) under $\sigma^\star(b)$, with the off-path beliefs of b retained. If $b = \Phi(b)$, then $\big(\sigma^\star(b), b\big)$ is a PBNE. If $\sigma^\star(b)$ satisfies (13.4.1) within ε at every decision state and b coincides with $\Phi(b)$ on the equilibrium path, then $\big(\sigma^\star(b), b\big)$ is a $(K+1)\varepsilon$-PBNE.

Proof. By Theorem 13.4.3, $\sigma^\star(b)$ is sequentially rational under b. The condition $b = \Phi(b)$ states that b is generated by Bayes' rule under $\sigma^\star(b)$ wherever Bayes' rule applies, which is the belief-consistency condition. The second statement follows by replacing exact maximization by ε-maximization in the same argument; the per-stage losses accumulate over at most $K + 1$ stages, which gives the bound $(K + 1)\varepsilon$. □

Algorithm 31 alternates the two directions until they agree within the tolerance. The iteration is not guaranteed to converge in general; when it does, Theorem 13.4.6 certifies the limit. Example 13.6 carries out one backward and one forward pass by hand.

Algorithm 31 Forward-belief/backward-policy computation for multi-stage deception

Require: Prior beliefs, state spaces $\mathcal{X}^k$, action sets, transition laws P^k, utilities u_i^k, horizon K, tolerance ε_{tol}

Initialize a belief sequence $b = \{b_i^k\}$
repeat
  Set $V_i^{K+1} \equiv 0$; for $k = K, K-1, \dots, 0$, compute σ^k from (13.4.1) under b
  For $k = 0, 1, \dots, K-1$, compute b_i^{k+1} from (13.4.2) under σ, giving $\Phi(b)$
  $\Delta \leftarrow \|\Phi(b) - b\|$; $b \leftarrow \Phi(b)$
until $\Delta \le \varepsilon_{\text{tol}}$
return strategies σ and beliefs b

EXAMPLE 13.6 **The value of a deceptive probe**

A defender holds the belief $b = \mathbb{P}(\theta_A = \text{malicious})$ about a user; in this example, subscripts on b and V index the stage. At the last stage, it chooses between continued engagement E and containment C, with values $V_2(\text{E}; b) = 4 - 12b$ and $V_2(\text{C}; b) = -3$: engagement yields intelligence but causes damage if the user is malicious. Hence $V_2(b) = \max\{4 - 12b, -3\}$, and engagement is optimal if and only if $b \le 7/12$. One stage earlier, with $b_1 = 0.35$ and $\gamma = 0.9$, the defender can use ordinary monitoring M_o at no cost or deploy a deceptive probe D_p, such as a decoy credential, at cost 0.2. The probe raises an alert $e = 1$ with probability 0.8 for a malicious user and 0.2 for a legitimate one; ordinary monitoring raises an alert with probabilities 0.5 and 0.3.

Backward pass. Under D_p, $\mathbb{P}(e = 1) = 0.8 \cdot 0.35 + 0.2 \cdot 0.65 = 0.41$, and Bayes' rule gives $b_2(1) = 0.28/0.41 = 0.683$ and $b_2(0) = 0.07/0.59 = 0.119$. Since $V_2(0.683) = -3$ and $V_2(0.119) = 2.576$,

$$V_1(\text{D}_\text{p}) = -0.2 + 0.9\,\big[0.41 \cdot (-3) + 0.59 \cdot 2.576\big] = 0.061.$$

Under M_o, $\mathbb{P}(e = 1) = 0.37$, $b_2(1) = 0.473$, and $b_2(0) = 0.278$; both posteriors lie below $7/12$, so engagement follows either observation, and $V_1(\text{M}_\text{o}) = 0.9\,\big[0.37 \cdot (-1.676) + 0.63 \cdot 0.667\big] = -0.18$, which equals the value $0.9\,V_2(0.35)$ of acting on the prior. The defender deploys the probe.

Forward pass. If the alert occurs, the belief jumps to 0.683 and the defender contains; otherwise it falls to 0.119 and the defender continues to engage. Figure 13.11 shows both computations.

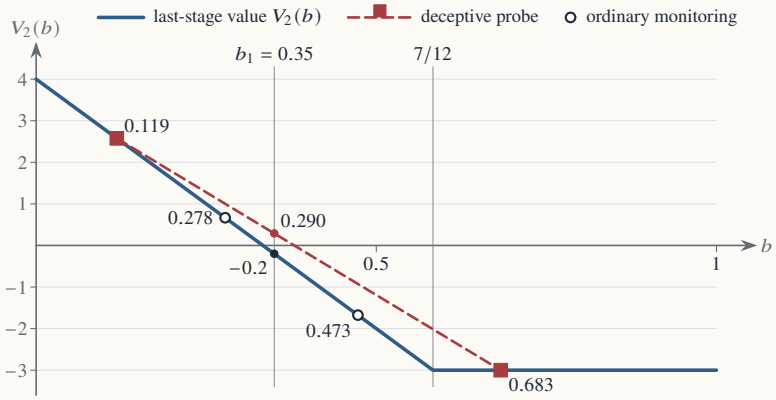


Figure 13.11: The value of a deceptive probe in Example 13.6. The solid line is the last-stage value $V_2(b) = \max\{4 - 12b, -3\}$, with a kink at $b = 7/12$. From the prior $b_1 = 0.35$, the deceptive probe splits the belief into 0.119 and 0.683 (squares); the expected continuation value lies on the chord between them (dashed line) and equals 0.290 before discounting. Ordinary monitoring splits the belief into 0.278 and 0.473 (circles), which lie on the same linear piece of V_2, so its expected continuation value -0.2 equals the value of acting on the prior.

Observation. Ordinary monitoring is informative but worthless here, because neither of its posteriors crosses the containment threshold $7/12$. The deceptive probe is worth its cost because one of its posteriors does. Deception pays only when the information it produces changes a future action.

The lesson for APT defense is that late defense is often too late. If the attacker reaches the final stage with high privilege and good intelligence, the defender has few safe options. Proactive deception acts earlier: fake profiles poison reconnaissance, training and warnings change phishing outcomes, monitoring affects privilege escalation, and uncertainty about the defender's sophistication makes the attacker more conservative. Partial observability is therefore not only a sensing problem: legitimate users generate suspicious actions, adversarial users imitate legitimate behavior, and each side observes the other through delayed and ambiguous evidence.

13.4.5 Learning for Active, Adaptive, and Autonomous Defense

Modern defense should be *active*, interacting with the attacker before damage is complete; *adaptive*, updating beliefs and policies as the attacker probes and learns; and *autonomous*, reducing reliance on perfect supervision and perfect prior knowledge [267]. Learning enters at four levels. *Parameter learning* estimates the attacker's type, skill, intent, or the reliability of a detector. *Payoff learning* estimates which defensive actions actually change attacker behavior. *Environmental learning* estimates transition probabilities when the system model is incomplete. *Strategic learning* estimates how the attacker is learning the defense; evidence that a honeypot has been recognized should change engagement, isolation, and disclosure decisions. Chapter 7 develops the corresponding algorithms. In each case learning has value only when it changes a future action, which Algorithm 32 makes explicit.

Algorithm 32 Adaptive dynamic deception loop

Require: Prior belief b^0, initial exposure z^1, policy σ_D mapping beliefs and exposure to actions, safety constraints
1: **for** $k = 1, 2, \ldots$ **do**
2: Observe alerts, commands, file access, timing, and lateral-movement indicators y^k
3: Update the belief b^k about the attacker's type, stage, and awareness of the deception
4: Select $d^k = \sigma_D(b^k, z^k)$ among the actions that satisfy the safety constraints
5: Engage, rotate, reveal, conceal, isolate, or contain; update the exposure z^{k+1}
6: Record intelligence value, operating cost, and risk
7: **end for**

EXAMPLE 13.7 **A threshold policy for adaptive deception**

A defender monitors a session with prior belief $b^0 = 0.18$ that the user is malicious. Each period it observes $y^k \in \{0, 1\}$, where a suspicious observation $y^k = 1$ has probability 0.70 under malicious and 0.15 under benign behavior, and it updates the belief by Bayes' rule. The policy σ_D uses three belief thresholds and an exposure limit: it contains if $b^k \geq 0.9$ or $z^k \geq 0.7$; otherwise it rotates the exposed services if $b^k \geq 0.7$, engages if $b^k \geq 0.3$, and only monitors if $b^k < 0.3$. The exposure starts at $z^1 = 0.2$; engagement raises it by 0.15, plus 0.05 after a suspicious observation, rotation lowers it by 0.25 to no less than 0.1, and monitoring leaves it unchanged. For the observation sequence $(0, 1, 0, 1, 1, 0, 1, 1)$ the policy produces

Period k	1	2	3	4	5	6	7	8
Observation y^k	0	1	0	1	1	0	1	1
Belief b^k	0.072	0.266	0.113	0.373	0.735	0.495	0.821	0.955
Exposure z^k	0.20	0.20	0.20	0.20	0.40	0.15	0.30	0.10
Action	monitor	monitor	monitor	engage	rotate	engage	rotate	contain

Figure 13.12 plots the belief and exposure trajectories.

Observation. A single suspicious observation is not enough to engage, because the low prior absorbs it (period 2). Engagement begins only after a second suspicious observation, rotation reduces the exposure that engagement creates, and containment follows when the belief exceeds 0.9. The exposure limit was never binding here; it is the safety constraint that would force containment if engagement continued at a moderate belief.

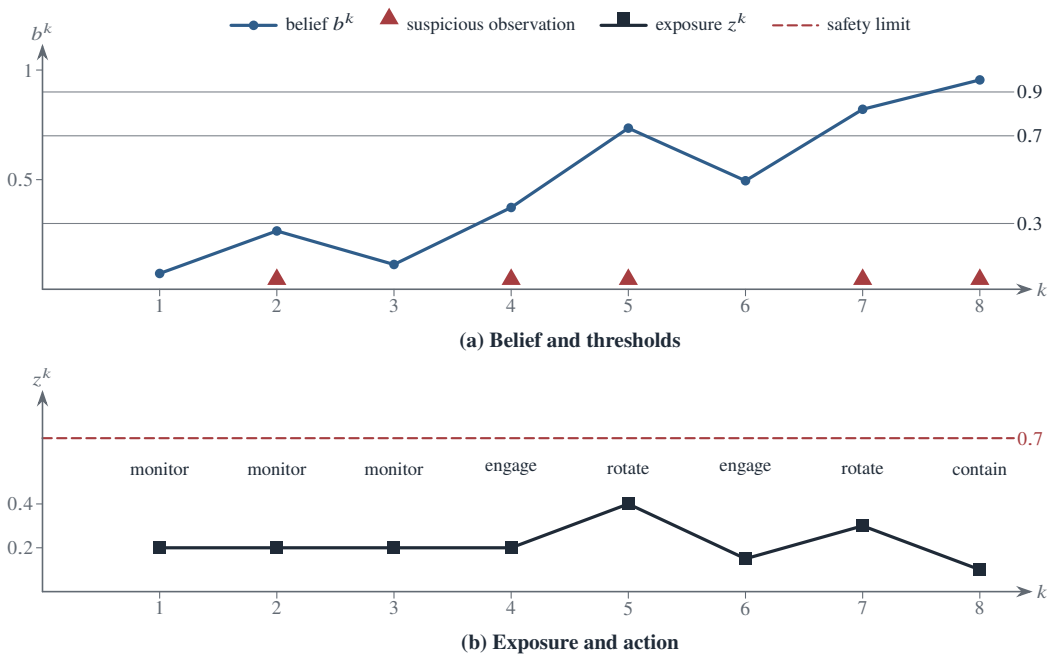


Figure 13.12: The threshold policy of Example 13.7. (a) The belief b^k (solid line with dots) against the engagement, rotation, and containment thresholds 0.3, 0.7, and 0.9 (thin lines); triangles on the axis mark suspicious observations. (b) The exposure z^k (solid line with squares) against the safety limit 0.7 (dashed line), with the action taken in each period printed above the curve.

13.4.6 Designing Dynamic Deception Policies

A dynamic deception policy is designed by aligning stages, observations, beliefs, actions, rewards, and safety triggers. The analyst first specifies the attack stages that the policy addresses and the state variables that distinguish them, and then the observations that separate the hidden types or stages, such as alerts, commands, file access, timing, lateral movement, and interaction with decoys. These observations determine the belief variables: attacker type, stage, capability, goal, and awareness of the deception. The action model specifies what the defender can change safely, since engagement, rotation, concealment, isolation, recovery, and containment carry different risks. The reward model assigns value to delay, intelligence, protected assets, service continuity, deception cost, response cost, and escape risk. Finally, *safety triggers* specify when engagement must stop even though its information value remains positive, and *adaptation rules* specify how the policy changes when the attacker appears to have detected the deception. Because deceptions are costly and lose effectiveness over time, their timing is also a design variable; in the cross-layer Stackelberg game of Yang and Zhu [615], a strategic layer solves a stopping-time problem that selects when to deploy which technique, and a tactical layer describes the engagement under the chosen deception as a one-sided information Markov game.

Remark 13.4.7 (Safety before cleverness). A policy that produces intelligence while exposing production systems to unacceptable risk is not a good defensive policy. Mathematically, operational safety enters the optimization as hard constraints on the feasible actions or as penalties large enough that unsafe states are never optimal.

13.5 Moving-Target Defense

Moving-target defense (MTD) is dynamic deception through controlled change. Instead of leaving the attack surface fixed and trying to detect every malicious action, the defender changes addresses, routes, software versions, configurations, credentials, service locations, or protocol appearances over time [118, 526]. The objective is to make reconnaissance stale, to misalign the attacker's chosen exploit with the vulnerabilities that the system currently exposes, and to raise the cost of learning the system. The feedback-driven multi-stage model of Zhu and Başar [658] formalizes this idea as a game in which the defender shifts the attack surface while observing risk over time.

13.5.1 Layer Game and Attack Surface

The system is decomposed into layers, for example network addressing, service tier, and host software. At layer l, let $\mathcal{V}_l = \{v_{l,1}, \dots, v_{l,|\mathcal{V}_l|}\}$ be the set of vulnerabilities and $C_l = \{c_{l,1}, \dots, c_{l,|C_l|}\}$ the set of feasible configurations. The *attack surface* $\mathcal{S}_l(c) \subseteq \mathcal{V}_l$ is the set of vulnerabilities exposed by configuration c. Attack i exploits vulnerability $v_{l,i}$ and causes the loss $L_{l,i} \geq 0$ if $v_{l,i}$ is exposed and nothing otherwise, so the *risk matrix* has entries

$$R_{l,ij} \triangleq L_{l,i}\, B_{l,ij}, \qquad B_{l,ij} \triangleq \begin{cases} 1, & v_{l,i} \in \mathcal{S}_l(c_{l,j}), \\ 0, & \text{otherwise.} \end{cases}$$

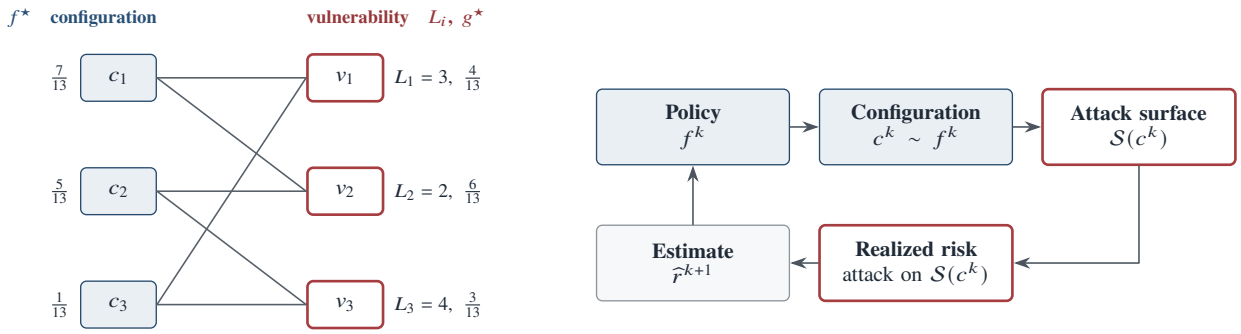
If the defender randomizes over configurations with $f_l \in \Delta(C_l)$ and the attacker over attacks with $g_l \in \Delta(\mathcal{V}_l)$, the expected risk is $\bar{R}_l(g_l, f_l) \triangleq g_l^\top R_l f_l$, which the defender minimizes and the attacker maximizes. The *layer game* is therefore a zero-sum matrix game in the sense of Chapter 3, with the roles of rows and columns exchanged.

THEOREM 13.5.1 **Static MTD saddle point**
Each layer game has a value $V_l^\star$ and mixed strategies $f_l^\star \in \Delta(C_l)$ and $g_l^\star \in \Delta(\mathcal{V}_l)$ such that

$$g_l^\top R_l f_l^\star \leq V_l^\star \leq (g_l^\star)^\top R_l f_l \qquad \text{for all } f_l \in \Delta(C_l),\ g_l \in \Delta(\mathcal{V}_l),$$

and $V_l^\star = \min_{f_l} \max_{g_l} g_l^\top R_l f_l = \max_{g_l} \min_{f_l} g_l^\top R_l f_l$.

Proof. The strategy sets are simplexes, which are compact and convex, and $g_l^\top R_l f_l$ is bilinear. The minimax theorem of von Neumann [585] gives the equality of the two values and the existence of optimal strategies; the saddle inequalities restate their optimality. □



(a) Attack surfaces

(b) Feedback structure

Figure 13.13: Moving-target defense. (a) The attack surfaces of Example 13.8: each configuration (left) is connected to the vulnerabilities it exposes (right). The defender's saddle-point probabilities $f^\star$ are printed to the left of the configurations, and the losses L_i and the attacker's saddle-point probabilities $g^\star$ to the right of the vulnerabilities. (b) The feedback structure of moving-target defense: the randomized policy f^k selects a configuration, the configuration exposes an attack surface, the attack realizes a risk, the risk updates the estimate $\widehat{r}^{k+1}$, and the estimate updates the policy.

EXAMPLE 13.8 **Randomized configuration of a web service**

▸ Code: `code/mtd_softmin.py`

A web service can run in three configurations: c_1 exposes the vulnerabilities $\{v_1, v_2\}$, c_2 exposes $\{v_2, v_3\}$, and c_3 exposes $\{v_1, v_3\}$. The losses are $L_1 = 3$ for a database injection, $L_2 = 2$ for a session hijack, and $L_3 = 4$ for

remote code execution. With rows indexed by attacks and columns by configurations,

$$R = \begin{bmatrix} 3 & 0 & 3 \\ 2 & 2 & 0 \\ 0 & 4 & 4 \end{bmatrix}.$$

Every pure configuration has a worst-case risk of at least 3, attained by attacking the most damaging exposed vulnerability. The equalizer strategies

$$f^\star = \left(\tfrac{7}{13}, \tfrac{5}{13}, \tfrac{1}{13}\right), \qquad g^\star = \left(\tfrac{4}{13}, \tfrac{6}{13}, \tfrac{3}{13}\right)$$

satisfy $Rf^\star = \frac{24}{13}\mathbf{1}$ and $R^\top g^\star = \frac{24}{13}\mathbf{1}$, so each is a best response to the other and the value is $V^\star = 24/13 \approx 1.85$. Figure 13.13(a) shows the attack surfaces and both strategies.

Observation. Randomization lowers the worst-case risk from 3 for the best static configuration, and from $8/3 \approx 2.67$ for uniform rotation, to 1.85. The saddle-point policy rarely uses c_3, which exposes the two most damaging vulnerabilities, and the attacker's saddle-point strategy concentrates on the session hijack, the only vulnerability exposed by both of the frequently used configurations.

13.5.2 Exposure Distribution and Misalignment

The defender's randomization induces an *exposure distribution* over vulnerabilities,

$$\eta_{l,i}(f_l) \triangleq \frac{(B_l f_l)_i}{\sum_{i'} (B_l f_l)_{i'}},$$

the normalized probability that $v_{l,i}$ is exposed. A vulnerability exposed in many likely configurations remains exposed often, and reconnaissance about a rarely exposed vulnerability is of little use. Zhu and Başar measure the misalignment between the defender's exposure distribution and the attacker's attack distribution by the Kullback–Leibler divergence [658]

$$\mathrm{Mis}(f_l, g_l) \triangleq \mathrm{KL}\big(\eta_l \,\|\, g_l\big) = \sum_{i=1}^{|\mathcal{V}_l|} \eta_{l,i} \log \frac{\eta_{l,i}}{g_{l,i}}.$$

A large divergence means that the vulnerabilities the attacker is prepared to exploit are poorly aligned with those the moving system exposes. The divergence is not a complete security score, because losses and mission context still matter, but it turns the movement of the attack surface into a quantitative design object.

PROPOSITION 13.5.2 **Nonnegative misalignment**
If the support of η_l is contained in that of g_l, then $\mathrm{Mis}(f_l, g_l) \geq 0$, with equality if and only if $\eta_l = g_l$.

Proof. This is Gibbs' inequality. Since $\log x \leq x - 1$ for $x > 0$, with equality only at $x = 1$,

$$-\mathrm{Mis}(f_l, g_l) = \sum_{i:\eta_{l,i}>0} \eta_{l,i} \log \frac{g_{l,i}}{\eta_{l,i}} \leq \sum_{i:\eta_{l,i}>0} \big(g_{l,i} - \eta_{l,i}\big) \leq 0,$$

and equality holds throughout if and only if $g_{l,i} = \eta_{l,i}$ for every i in the support of η_l and g_l puts no mass outside it, that is, $\eta_l = g_l$. □

13.5.3 Feedback Learning and Entropic Reconfiguration

Moving too rarely lets reconnaissance remain valid; moving too often disrupts legitimate service, complicates management, and invites operational mistakes. The feedback-driven model captures this trade-off by charging a

reconfiguration cost for changing the randomized policy from f^k to f^{k+1}:

$$\tau\,\mathrm{KL}\big(f^{k+1}\,\|\,f^k\big) = \tau\sum_j f_j^{k+1}\log\frac{f_j^{k+1}}{f_j^k},$$

where the layer index is dropped and $\tau > 0$ is the switching-cost weight. A small τ lets the defender react sharply to new risk information; a large τ makes the policy smooth and conservative.

The defender does not know the risk of each configuration and estimates it from observations. If configuration c^k is used at stage k and the realized risk is r^k, the estimate is updated by

$$\widehat{r}^{\,k+1}(c) = \widehat{r}^{\,k}(c) + \zeta\,\mathbf{1}_{\{c^k=c\}}\big(r^k - \widehat{r}^{\,k}(c)\big),$$

with step size $\zeta \in (0,1]$. The policy is then updated, using the new estimate, by the one-step regularized problem

$$f^{k+1} \in \operatorname*{arg\,min}_{f\in\Delta(C)}\Big\{\sum_j f_j\,\widehat{r}^{\,k+1}(c_j) + \tau\sum_j f_j\log\frac{f_j}{f_j^k}\Big\}. \tag{13.5.1}$$

THEOREM 13.5.3 **Soft-min MTD update**
If $f_j^k > 0$ for every j and $\tau > 0$, the problem (13.5.1) has the unique solution

$$f_j^{k+1} = \big[\mathcal{B}_\tau(\widehat{r}^{\,k+1}, f^k)\big]_j \triangleq \frac{f_j^k\exp\big(-\widehat{r}^{\,k+1}(c_j)/\tau\big)}{\sum_{j'} f_{j'}^k\exp\big(-\widehat{r}^{\,k+1}(c_{j'})/\tau\big)}.$$

Proof. Write $\widehat{r}_j \triangleq \widehat{r}^{\,k+1}(c_j)$. The objective is continuous on the compact simplex and strictly convex, because the relative entropy is strictly convex in f, so it has a unique minimizer. Its partial derivative with respect to f_j tends to $-\infty$ as $f_j \downarrow 0$, so the minimizer is interior and satisfies the Lagrange condition $\widehat{r}_j + \tau\big(\log(f_j/f_j^k) + 1\big) + \psi = 0$ for a multiplier ψ of the constraint $\sum_j f_j = 1$. Hence $f_j = f_j^k\exp(-\widehat{r}_j/\tau)\exp(-1-\psi/\tau)$; the last factor is common to all j and is fixed by normalization. □

Configurations with a higher estimated risk receive less probability, while the previous policy and the switching cost prevent overreaction to a single observation. As $\tau \downarrow 0$, the update concentrates on the configurations of minimum estimated risk; as $\tau \to \infty$, $f^{k+1} \to f^k$. Figure 13.14 evaluates the update for three values of τ. The attacker's counterpart, a regularized maximization with a learning cost, yields a soft-max update over attacks, the Boltzmann–Gibbs map of Chapter 7 (Proposition 7.1.7).

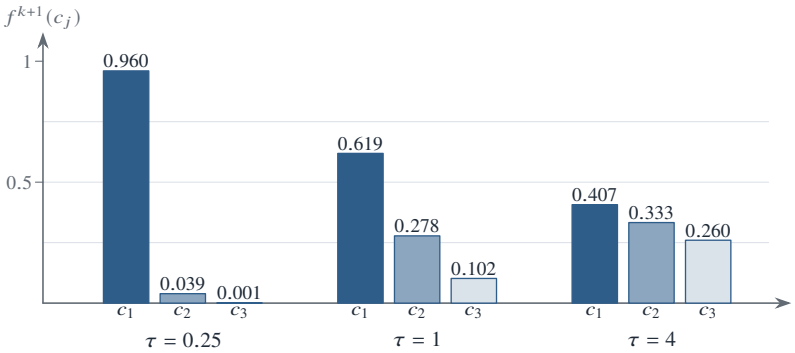


Figure 13.14: The soft-min update of Theorem 13.5.3 from the uniform policy $f^k = (\frac{1}{3},\frac{1}{3},\frac{1}{3})$ with estimated risks $\widehat{r}^{\,k+1} = (0.2, 1.0, 2.0)$. Each group of bars gives f^{k+1} for one switching-cost weight τ, with the bars for c_1, c_2, and c_3 shaded from dark to light and the probabilities printed. A small weight concentrates the policy on the configuration of least estimated risk; a large weight keeps it near the previous policy.

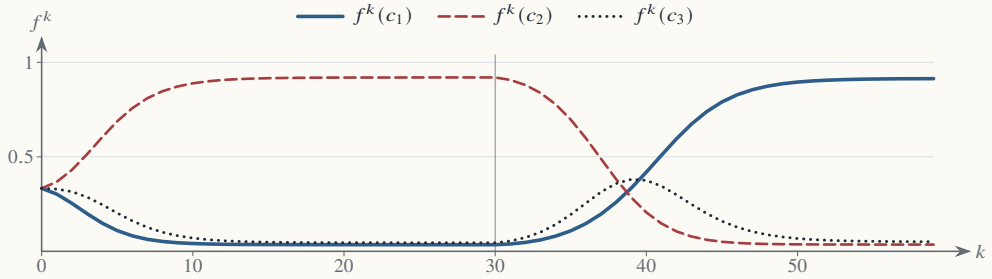
The rest points of the update have a game-theoretic meaning.

PROPOSITION 13.5.4 **Rest points of the soft-min update**
A policy f satisfies $f = \mathcal{B}_\tau(r, f)$ if and only if r is constant on the support of f. Consequently, if the defender uses the risks $r = R^\top g$ induced by an attacker strategy g and the attacker uses the values Rf, then every pair (f, g) of completely mixed strategies that is a rest point of both updates is a saddle point of the layer game.

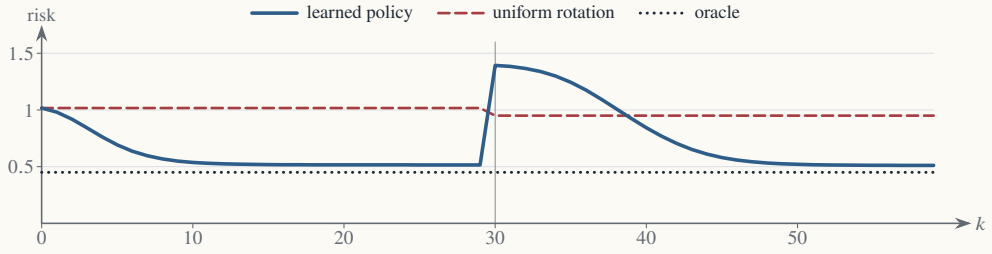
Proof. For j in the support of f, $f = \mathcal{B}_\tau(r, f)$ means $\exp(-r_j/\tau) = \sum_{j'} f_{j'} \exp(-r_{j'}/\tau)$, a constant; hence r_j is constant on the support, and the converse is immediate. If f and g are completely mixed rest points, then $R^\top g$ and Rf are constant vectors. Every configuration then yields the same risk against g, so f is a best response to g, and every attack yields the same risk against f, so g is a best response to f. A pair of mutual best responses in a zero-sum game is a saddle point. □

In Example 13.8, the saddle point $(f^\star, g^\star)$ satisfies $R^\top g^\star = Rf^\star = \frac{24}{13}\mathbf{1}$ and is therefore, by the first part of Proposition 13.5.4, a rest point of both updates.

EXAMPLE 13.9 **Feedback MTD after a change in the threat**
A service has three configurations whose true risks are $(1.55, 0.45, 1.05)$ for $k < 30$, so that c_2 is safest, and $(0.45, 1.45, 0.95)$ for $30 \le k < 60$, so that c_1 becomes safest, for example because a new exploit for c_2 appears. The defender starts from the uniform policy and the estimate $\widehat{r}^0 = (1, 1, 1)$. To obtain a reproducible run, the stochastic updates are replaced by their expectations: the estimate of c_j moves toward its true risk with the step $0.45 f_j^k + 0.08(1 - f_j^k)$, where the second term represents background telemetry on unused configurations. The policy moves toward the soft-min update with $\tau = 0.6$, with relaxation 0.55, and is then mixed with the uniform policy with weight 0.05, an exploration floor. Figure 13.15 shows the run.



(a) Learned policy



(b) Expected risk

Figure 13.15: Feedback moving-target defense in Example 13.9. (a) The probabilities $f^k(c_1)$ (solid), $f^k(c_2)$ (dashed), and $f^k(c_3)$ (dotted) of the learned policy; the vertical line marks the change of the true risks at $k = 30$. (b) The expected risk of the learned policy (solid) compared with uniform rotation (dashed) and an oracle that always uses the safest configuration (dotted).

Observation. The learned policy concentrates on c_2 within about ten stages, with $f^{29} \approx (0.03, 0.92, 0.05)$; after the change it shifts to c_1, which receives more than half of the probability from $k = 41$ on, and it ends at

$f^{59} \approx (0.91, 0.04, 0.05)$. Its average expected risk over the 60 stages is 0.694, against 0.983 for uniform rotation and 0.450 for the oracle. The spike at $k = 30$ is unavoidable, because feedback cannot anticipate an unobserved change; the residual gap to the oracle in steady state, 0.515 before the change and 0.511 after it against 0.45, is the price of the exploration floor that makes the recovery possible.

13.5.4 Learning Dynamics

Zhu and Başar further study the coupled dynamics in which the defender updates risk estimates and configuration policies while an intelligent attacker updates attack estimates and attack strategies [658]. A common relaxed form is

$$f^{k+1} = (1 - \chi_k) f^k + \chi_k \, \mathcal{B}_\tau(\tilde{r}^{k+1}, f^k),$$

with step sizes $\chi_k \in (0, 1]$, and the attacker uses the analogous soft-max map. Under the standard stochastic-approximation conditions on the step sizes, the iterates track an ordinary differential equation of the kind analyzed in Chapter 7 (Theorem 7.1.4), and Proposition 13.5.4 identifies its completely mixed rest points with the completely mixed saddle points of the layer game. The static saddle point is thus the target of the game analysis, and feedback learning is an implementable way to approach good randomized policies when the risks must be estimated online.

This view explains why moving-target defense belongs to cyber deception. It does not merely hide a fixed asset: it changes the evidence available to the attacker, degrades the value of past reconnaissance, and forces the attacker to reason about an information structure that is itself changing.

13.6 Discussion, Summary, and Further Reading

The chapter has treated cyber deception as a change in the adversary's decision problem. A decoy, warning, honey credential, moving-target policy, or engagement environment matters only when it changes what the adversary believes and therefore what it does. The directions below extend the models of the chapter.

Choosing the model. The appropriate model depends on the question. A threshold model suffices when the issue is whether suspicion crosses an attack-or-withdraw boundary. A signaling game with evidence is appropriate when the defender's deception is covert and may be exposed by implementation clues. Bayesian persuasion fits overt information policies, such as labels, warnings, and announced monitoring, for which the defender's commitment is credible. Mechanism design is needed when the defender can also shape incentives, friction, trust, or participation, and dynamic Bayesian games are needed when deception unfolds over stages and today's engagement changes tomorrow's risk.

Credibility, measurement, and safety. A deception policy must be credible, measurable, safe, and lawful. Credibility asks whether the artifact remains plausible after the adversary reasons strategically, which is the content of the equilibrium conditions of Section 13.2. Measurement asks whether the policy creates delay, deterrence, intelligence, attribution, or reduced damage. Safety asks whether engagement, generated content, decoys, and moving targets can be contained, and legal and ethical constraints determine which data may be collected, which users may be affected, and when engagement must stop. Controlled experiments with human red teams have shown that both the presence of decoys and the attackers' belief in their presence change attacker behavior [195], and such experiments are as important to the field as equilibrium calculations.

Deception, cognition, and AI agents. Many deceptions act on the attention and cognition of human attackers and analysts rather than on purely rational inference. Cognitive security models attention, bounded rationality, and feint attacks that exhaust human analysts [272, 271, 263]; human risks motivate cognition-inspired adaptive deception [605]; and bi-level planning of deception at the strategic, operational, and tactical levels lets a defender exploit an APT attacker's cognitive biases, such as rational inattention and confirmation bias, as a form of cognitive arbitrage [609]. Language-model agents both enlarge the attack surface and offer new defensive capabilities: they can generate and sustain deceptive interactions at scale, and game models can be used to keep their behavior consistent with the incentives of the other players [359, 639] and to design agentic deception workflows [7]. Language models are also targets: Han and Zhu [234] model jailbreaking as an extensive-form game in which an attacker searches the prompt space and

the model accepts, rejects, or redirects. Trust links these roles: Ge and Zhu [216] treat trust as both an output and a requirement of AI in networked systems.

Misinformation and information attacks. Deception also targets populations and algorithms. Tags on user-generated misinformation persuade a population of users, whose responses in turn nudge the content provider toward authentic content [607]. On the attack side, Yang, Lei, and Zhu [606] model an attacker who fabricates travel demand as a Stackelberg leader that steers a navigation recommender system, and show that a trust mechanism based on users' confidence in the recommender reduces the impact of such attacks.

Connections to later chapters. Chapter 14 turns to the Internet of Things, where remote attestation is an inspection game and hidden device security creates a lemons problem that labels and certification must resolve; both are signaling and verification questions of the kind studied here, now at the scale of device populations. Chapter 15 treats supply chains as inherited attack surfaces, in which verification and counterfeit detection face adversaries who disguise compromised components as genuine ones. Chapter 16 studies jamming, eavesdropping, and AI-native network control in next-generation wireless networks, where primary-user emulation is an attacker-side deception and randomized channel and relay selection plays the role of the moving targets of this chapter.

Connections to appendices. Appendix A supplies the linear programming behind the duplicity-design and saddle-point computations and the convex duality behind the soft-min update. Appendix B supplies Bayes' rule, the belief filtering of hidden Markov models used for belief-based deception, and the stochastic-approximation results behind the learning dynamics of moving-target defense. Appendix D describes the attack lifecycle and attack surfaces that the dynamic and moving-target models formalize.

Further reading. The edited volume of Jajodia et al. [299] surveys cyber deception, denial, adversary modeling, honeypots, and virtual attack surfaces, and the volume edited by Zhu et al. [672] addresses the modeling, analysis, design, and human factors of cyber deception; Almeshekah and Spafford [10], Stech et al. [554], and Wang and Lu [589] give accessible introductions. Pawlick et al. [475] classify defensive deception into perturbation, moving-target defense, obfuscation, mixing, honey-x, and attacker engagement, and pair each class with a game model; Pawlick and Zhu [479] give a book-length game-theoretic treatment; tutorials are given by Zhu [635] and, for deception in game theory and control, by Vamvoudakis et al. [576]. Signaling games with evidence originate in [477, 476] and are developed further in [625]. For information design, Kamenica and Gentzkow [310] is the starting point, and [309, 57] survey the field; the duplicity-game framework is due to Huang and Zhu [270]. Honeypot allocation as a security game is studied in [482, 322], and against lateral movement as a partially observable stochastic game in [255]; AI and game-theoretic models for deception in critical networks are reviewed in [206]. Dynamic deception against advanced persistent threats is developed in [266, 673], adaptive honeypot engagement in [264], strategic learning for active, adaptive, and autonomous defense in [267], and zero-trust authentication against lateral movement in [213]. The feedback model of moving-target defense is due to Zhu and Başar [658], and Cho et al. [118] and Sengupta et al. [526] survey the field.

Summary. Cyber deception shapes the adversary's information environment so that its best response becomes less harmful. In a signaling game with evidence, the detector's quality and aggressiveness determine five prior regimes; pooling survives exactly where the attacker's response ignores the evidence, and in the middle regime the equilibrium is partially separating, with randomization on both sides. Bayesian persuasion designs the information policy itself, subject to Bayesian plausibility, and its value is the concave closure of the defender's payoff. Mechanism design adds incentives and trust, and its feasibility is governed by incentive margins and best-response regions. Dynamic deception treats beliefs as state variables and decides engagement by forward-looking surpluses, and moving-target defense randomizes the attack surface, with saddle points as the static benchmark and soft-min learning as an implementable feedback policy. Throughout, a deception is valuable only when the information it creates or withholds changes an action, and it is acceptable only when it remains within the defender's safety constraints.

13.7 Exercises

The exercises below apply the concepts of this chapter to concrete security settings. Each exercise states its setting and the results to be derived.

Exercise 13.7.1 (Posterior suspicion and a production banner)**.** A server is either production, N, or a honeypot, H, with prior $\rho(H) = 0.3$. A production-like banner m_B appears on production systems with probability 0.9 and on honeypots with probability 0.6. Attacking production gives the attacker 10, attacking a honeypot gives -5, and withdrawing gives 0.

(a) Compute the posterior $\mu(H \mid m_B)$.

(b) Compute the posterior threshold $\bar{\mu}$ below which the attacker attacks, and decide whether it attacks after observing m_B.

(c) Holding the banner probabilities fixed, find the smallest prior $\rho(H)$ at which the attacker withdraws after observing m_B.

Exercise 13.7.2 (Equilibria of a conservative evidence game)**.** In the binary evidence game of Section 13.2.2, let $\alpha = 0.2$, $\beta = 0.6$, and $\Delta_0^A = \Delta_1^A = 1$.

(a) Compute J and G, and classify the detector.

(b) Compute the thresholds T_{00}, T_{01}, T_{10}, and T_{11}, and identify the five prior regimes.

(c) For $p = 0.5$, compute the partially separating equilibrium of Theorem 13.2.10, and verify that the attacker is indifferent after the observation $(m, e) = (0, 0)$.

(d) For $p = 0.3$, determine which pooling equilibria exist, and explain why pooling on the other message fails.

Exercise 13.7.3 (Persuasion for deterrence)**.** A defender labels systems with a warning m_w or a quiet label m_s under an announced policy. Each system is a honeypot, H, or a production system, N, and the prior probability of a honeypot is $\rho(H) = 0.2$. The attacker withdraws if and only if its honeypot posterior is at least 0.4, and the defender receives 1 if the attacker withdraws and 0 otherwise.

(a) Compute the defender's value without information and under full revelation.

(b) Use concavification to find the optimal posteriors, the probability of the warning label, and the optimal value.

(c) Compute the policy $\pi(m_w \mid H)$ and $\pi(m_w \mid N)$ that implements the optimal posteriors.

(d) Explain, using Bayesian plausibility, why no policy deters the attacker with probability greater than $1/2$.

Exercise 13.7.4 (Enforceable profiles in a duplicity mechanism)**.** Two user types, an automated scanner q_1 and a human operator q_2, share the perceived prior $b_U(H \mid q_1) = b_U(H \mid q_2) = 0.25$ that a reachable service is a honeypot, H, rather than a production service, N. Each type either engages, a_E, or drops out, a_{DO}, which yields 0; there is no friction cost unless stated otherwise. Engagement yields $(6, -2)$ to q_1 and $(3, -6)$ to q_2, written as (value on production, value on honeypot).

(a) Compute the best-response regions $\mathcal{R}^q_{a_E}$ and $\mathcal{R}^q_{a_{\mathrm{DO}}}$ for both types on the honeypot-posterior axis.

(b) Determine which of the four action profiles can be induced by some generator, and explain why one of them cannot.

(c) For the label $m_{(a_E, a_{\mathrm{DO}})}$ with $\pi_N = \pi(m \mid N)$ and $\pi_H = \pi(m \mid H)$, derive the incentive constraints as bounds on π_H/π_N.

(d) Explain how a friction cost $\kappa(a_E) > 0$ moves the regions of part (a).

Exercise 13.7.5 (How accurate must a deceptive probe be?)**.** Consider the setting of Example 13.6, with $V_2(b) = \max\{4 - 12b, -3\}$, prior $b_1 = 0.35$, $\gamma = 0.9$, and probe cost 0.2, but let the probe raise an alert with probability φ for a malicious user and $1 - \varphi$ for a legitimate one, where $\varphi \in (\frac{1}{2}, 1)$.

(a) Find the smallest φ for which the posterior after an alert reaches the containment threshold $7/12$.

(b) For φ above this value, show that the expected continuation value $\mathbb{E}[V_2(b_2)]$ equals $6.3\,\varphi - 4.75$.

(c) Find the smallest φ for which deploying the probe is better than acting on the prior, and explain why a probe with φ below the value found in part (a) has no value however cheap it is.

Exercise 13.7.6 (Saddle point and soft-min update for a moving target)**.** A service has three configurations: c_1 exposes $\{v_1, v_2\}$, c_2 exposes $\{v_2, v_3\}$, and c_3 exposes $\{v_3\}$. The losses are $L_1 = 4$, $L_2 = 2$, and $L_3 = 3$.

(a) For $f = (\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$, compute the exposure probability of each vulnerability and the exposure distribution $\eta(f)$.

(b) Write the risk matrix R, show that c_2 is weakly dominated for the defender, and compute a saddle point and the value of the layer game. Decide whether the defender's saddle-point strategy is unique.

(c) An attacker concentrates on v_2. Compute the configuration risks $r = R^{\top} g$ and the soft-min update from $f = (\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$ with $\tau = 1$, and explain the direction of the change.

Bibliography

[1] 116th United States Congress. Internet of Things Cybersecurity Improvement Act of 2020. Public Law 116-207, 2020.

[2] 3rd Generation Partnership Project. Security architecture and procedures for 5G system. 3GPP TS 33.501 V19.6.0 (Release 19), Mar. 2026.

[3] M. Abdallah, S. Bagchi, S. D. Bopardikar, K. Chan, X. Gao, M. Kantarcioglu, C. Li, P. Liu, and Q. Zhu. Game theory in distributed systems security: Foundations, challenges, and future directions. *IEEE Security & Privacy*, 23(1):64–74, 2025.

[4] D. Abshari and M. Sridhar. Cyber-physical systems security: A comprehensive review of anomaly detection techniques. *Internet of Things*, 37:101938, 2026. Preprint: arXiv:2502.13256 (2025).

[5] D. Acemoglu, A. Malekian, and A. Ozdaglar. Network security and contagion. *Journal of Economic Theory*, 166:536–585, 2016.

[6] G. A. Akerlof. The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3):488–500, 1970.

[7] M. A. Al Bari and Q. Zhu. A Gestalt game-theoretic framework for designing agentic AI workflows in cyber deception. In *Game Theory and AI for Security (GameSec 2025)*, Lecture Notes in Computer Science, pages 228–248. Springer, Cham, 2025.

[8] R. Albert, H. Jeong, and A.-L. Barabási. Error and attack tolerance of complex networks. *Nature*, 406:378–382, 2000.

[9] A. AlHarmali, S. Ali, W. Aman, and O. Hussain. Cyber risk assessment for cyber-physical systems: A review of methodologies and recommendations for improved assessment effectiveness. In *Proceedings of the 12th International Conference on Computational Science and Engineering (CSE 2024)*, volume 14 of *Computer Science & Information Technology (CS & IT)*, pages 77–94. AIRCC Publishing Corporation, 2024.

[10] M. H. Almeshekah and E. H. Spafford. Cyber security deception. In S. Jajodia, V. S. Subrahmanian, V. Swarup, and C. Wang, editors, *Cyber Deception: Building the Scientific Foundation*, pages 23–50. Springer, 2016.

[11] T. Alpcan and T. Başar. A game theoretic approach to decision and analysis in network intrusion detection. In *Proceedings of the 42nd IEEE Conference on Decision and Control*, volume 3, pages 2595–2600, 2003.

[12] T. Alpcan and T. Başar. *Network Security: A Decision and Game-Theoretic Approach*. Cambridge University Press, Cambridge, 2010.

[13] O. Alrawi, C. Lever, M. Antonakakis, and F. Monrose. SoK: Security evaluation of home-based IoT deployments. In *Proceedings of the IEEE Symposium on Security and Privacy (S&P)*, 2019.

[14] E. Altman, K. Avrachenkov, and A. Garnaev. Jamming in wireless networks under uncertainty. In *Proceedings of the 7th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, 2009.

[15] J. Alvarez. Stuxnet: The world's first cyber weapon. `https://cisac.fsi.stanford.edu/news/stuxnet`, Feb. 2015. Center for International Security and Cooperation (CISAC), Stanford University, news, 3 February 2015.

[16] M. Ambrosin, M. Conti, R. Lazzeretti, M. M. Rabbani, and S. Ranise. Collective remote attestation at the Internet of Things scale: State-of-the-art and future challenges. *IEEE Communications Surveys & Tutorials*, 22(4), 2020.

[17] S. Amini, F. Pasqualetti, and H. Mohsenian-Rad. Dynamic load altering attacks against power system stability: Attack models and protection schemes. *IEEE Transactions on Smart Grid*, 9(4):2862–2872, 2018.

[18] L. An, A. Chakrabortty, and A. Duel-Hallen. A stackelberg security investment game for voltage stability of power systems. In *Proceedings of the 59th IEEE Conference on Decision and Control*, pages 3359–3364, 2020.

[19] B. D. O. Anderson and J. B. Moore. *Optimal Filtering*. Prentice-Hall, Englewood Cliffs, NJ, 1979.

[20] B. D. O. Anderson and J. B. Moore. *Optimal Control: Linear Quadratic Methods*. Prentice-Hall, Englewood Cliffs, NJ, 1990.

[21] J. P. Anderson. Computer security threat monitoring and surveillance. Technical report, James P. Anderson Co., Fort Washington, PA, 1980.

[22] R. Anderson. *Security Engineering: A Guide to Building Dependable Distributed Systems*. Wiley, Indianapolis, IN, 3 edition, 2020.

[23] R. Anderson and T. Moore. The economics of information security. *Science*, 314(5799):610–613, 2006.

[24] R. Anderson and T. Moore. The economics of information security. *Science*, 314(5799):610–613, 2006.

[25] M. Antonakakis, T. April, M. Bailey, M. Bernhard, E. Bursztein, J. Cochran, Z. Durumeric, J. A. Halderman, L. Invernizzi, M. Kallitsis, D. Kumar, C. Lever, Z. Ma, J. Mason, D. Menscher, C. Seaman, N. Sullivan, K. Thomas, and Y. Zhou. Understanding the Mirai botnet. In *Proceedings of the 26th USENIX Security Symposium*, pages 1093–1110, 2017.

[26] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.

[27] M. R. Asghar, Q. Hu, and S. Zeadally. Cybersecurity in industrial control systems: Issues, technologies, and challenges. *Computer Networks*, 165:106946, 2019.

[28] J. Aspnes, K. Chang, and A. Yampolskiy. Inoculation strategies for victims of viruses and the sum-of-squares partition problem. *Journal of Computer and System Sciences*, 72(6):1077–1093, 2006.

[29] R. J. Aumann. Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics*, 1(1):67–96, 1974.

[30] R. J. Aumann. Agreeing to disagree. *Annals of Statistics*, 4(6):1236–1239, 1976.

[31] R. J. Aumann and A. Brandenburger. Epistemic conditions for Nash equilibrium. *Econometrica*, 63(5):1161–1180, 1995.

[32] T. Aven. *Risk Analysis*. Wiley, Chichester, 2 edition, 2015.

[33] R. Avenhaus, B. von Stengel, and S. Zamir. Inspection games. In R. J. Aumann and S. Hart, editors, *Handbook of Game Theory with Economic Applications*, volume 3, chapter 51, pages 1947–1987. Elsevier, 2002.

[34] R. Avenhaus, B. von Stengel, and S. Zamir. Inspection games. In R. J. Aumann and S. Hart, editors, *Handbook of Game Theory with Economic Applications*, volume 3, pages 1947–1987. Elsevier, 2002.

[35] S. Axelsson. The base-rate fallacy and the difficulty of intrusion detection. *ACM Transactions on Information and System Security*, 3(3):186–205, 2000.

[36] P. Baguer, G. M. Yilma, E. Municio, G. Garcia-Aviles, A. Garcia-Saavedra, M. Liebsch, and X. Costa-Pérez. Attacking O-RAN interfaces: Threat modeling, analysis and practical experimentation. *IEEE Open Journal of the Communications Society*, 5:4559–4577, 2024.

[37] C. Bakker, A. Bhattacharya, S. Chatterjee, and D. L. Vrabie. Learning and information manipulation: Repeated hypergames for cyber-physical security. *IEEE Control Systems Letters*, 4(2):295–300, 2020.

[38] N. Bakshi and P. Kleindorfer. Co-opetition and investment for supply-chain resilience. *Production and Operations Management*, 18(6):583–603, 2009.

[39] V. Bala and S. Goyal. A noncooperative model of network formation. *Econometrica*, 68(5):1181–1229, 2000.

[40] C. Ballester, A. Calvó-Armengol, and Y. Zenou. Who's who in networks. Wanted: The key player. *Econometrica*, 74(5):1403–1417, 2006.

[41] S. Banach. Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. *Fundamenta Mathematicae*, 3:133–181, 1922.

[42] J. Barreiro-Gomez and Q. Zhu. Coalitional stochastic differential games for networks. *IEEE Control Systems Letters*, 6:2707–2712, 2022.

[43] T. Başar and P. Bernhard. *H_∞-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*. Birkhäuser, Boston, MA, 2nd edition, 1995.

[44] T. Başar, T. Hayakawa, H. Ishii, and Q. Zhu. Cooperative and noncooperative paradigms for game-theoretic control of socio-technical systems, 2026. arXiv:2605.17886.

[45] T. Başar and G. J. Olsder. *Dynamic noncooperative game theory*. SIAM, 1998.

[46] T. Başar and G. J. Olsder. *Dynamic Noncooperative Game Theory*, volume 23 of *Classics in Applied Mathematics*. SIAM, Philadelphia, PA, 2nd edition, 1999.

[47] T. Başar and Q. Zhu. Prices of anarchy, information, and cooperation in differential games. *Dynamic Games and Applications*, 1(1):50–73, 2011.

[48] T. Başar and Q. Zhu. Lecture 15: Mechanism design theory. Lecture notes on static and dynamic game theory, 2026. Draft lecture notes.

[49] D. Basin, J. Dreier, L. Hirschi, S. Radomirović, R. Sasse, and V. Stettler. A formal analysis of 5G authentication. In *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security (CCS)*, pages 1383–1396, 2018.

[50] A. Beck. *First-Order Methods in Optimization*. MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2017.

[51] D. E. Bell and L. J. LaPadula. Secure computer systems: Mathematical foundations. Technical Report MTR-2547, Volume I, MITRE Corporation, 1973.

[52] B. Beltz, P.-Y. Chen, J. Doty, Y. Fonken, N. Gurney, H.-W. Hsing, S. Hirschmann, B. Israelsen, N. Lau, M. Li, S. Marsella, M. Murray, J. Oh, A. Sliva, K. Srivastava, S. Trent, P. Wu, Y.-T. Yang, Q. Zhu, et al. Guarding against malicious biased threats (GAMBiT): Experimental design of cognitive sensors and triggers with behavioral impact analysis, 2025. arXiv:2512.00098v1.

[53] B. Beltz, J. Doty, Y. Fonken, N. Gurney, B. Israelsen, N. Lau, S. Marsella, R. Thomas, S. Trent, P. Wu, Y.-T. Yang, and Q. Zhu. Guarding against malicious biased threats (GAMBiT) datasets: Revealing cognitive bias in human-subjects red-team cyber range operations. *Data in Brief*, 65:112476, 2026.

[54] M. Benaïm. Dynamics of stochastic approximation algorithms. In *Séminaire de Probabilités XXXIII*, volume 1709 of *Lecture Notes in Mathematics*, pages 1–68. Springer, Berlin, 1999.

[55] P. G. Bennett. Toward a theory of hypergames. *Omega*, 5(6):749–751, 1977.

[56] A. Bensoussan, M. Kantarcioglu, and S. C. Hoe. A game-theoretical approach for finding optimal strategies in a botnet defense model. In *Decision and Game Theory for Security (GameSec 2010)*, volume 6442 of *Lecture Notes in Computer Science*, pages 135–148. Springer, 2010.

[57] D. Bergemann and S. Morris. Information design: A unified perspective. *Journal of Economic Literature*, 57(1):44–95, 2019.

[58] A. R. Bergen and D. J. Hill. A structure preserving model for power system stability analysis. *IEEE Transactions on Power Apparatus and Systems*, PAS-100(1):25–35, 1981.

[59] R. H. Berk. Limiting behavior of posterior distributions when the model is incorrect. *The Annals of Mathematical Statistics*, 37(1):51–58, 1966.

[60] P. Bernhard. Linear-quadratic, two-person, zero-sum differential games: Necessary and sufficient conditions. *Journal of Optimization Theory and Applications*, 27(1):51–69, 1979.

[61] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, MA, 2nd edition, 1999.

[62] D. P. Bertsekas. *Convex Optimization Algorithms*. Athena Scientific, Belmont, MA, 2015.

[63] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, MA, 3 edition, 2016.

[64] D. P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, Belmont, MA, 4th edition, 2017.

[65] D. P. Bertsekas. *Rollout, Policy Iteration, and Distributed Reinforcement Learning*. Athena Scientific, Belmont, MA, 2020.

[66] D. P. Bertsekas and J. N. Tsitsiklis. Neuro-dynamic programming: an overview. In *Proceedings of 1995 34th IEEE conference on decision and control*, volume 1, pages 560–564. IEEE, 1995.

[67] S. Bhunia, M. S. Hsiao, M. Banga, and S. Narasimhan. Hardware Trojan attacks: Threat analysis and countermeasures. *Proceedings of the IEEE*, 102(8):1229–1247, 2014.

[68] S. Bi and Y. J. Zhang. Graphical methods for defense against false-data injection attacks on power system state estimation. arXiv:1304.4151, 2013.

[69] V. M. Bier and M. N. Azaiez. *Game Theoretic Risk Analysis of Security Threats*. Springer, New York, 2009.

[70] B. Biggio and F. Roli. Wild patterns: Ten years after the rise of adversarial machine learning. *Pattern Recognition*, 84:317–331, 2018.

[71] P. Billingsley. *Probability and Measure*. Wiley, New York, 3 edition, 1995.

[72] Z. W. Birnbaum. On the importance of different components in a multicomponent system. In P. R. Krishnaiah, editor, *Multivariate Analysis II*, pages 581–592. Academic Press, New York, 1969.

[73] M. Bishop. *Computer Security: Art and Science*. Addison-Wesley, Boston, 2003.

[74] L. Blume, D. Easley, J. Kleinberg, R. Kleinberg, and É. Tardos. Network formation in the presence of contagious risk. *ACM Transactions on Economics and Computation*, 1(2):6:1–6:20, 2013.

[75] P. Bolton and M. Dewatripont. *Contract Theory*. MIT Press, Cambridge, MA, 2005.

[76] K. C. Border. *Fixed Point Theorems with Applications to Economics and Game Theory*. Cambridge University Press, Cambridge, 1985.

[77] É. Borel. The theory of play and integral equations with skew symmetric kernels. *Econometrica*, 21(1):97–100, 1953. English translation of the 1921 note by L. J. Savage.

[78] T. Börgers. *An Introduction to the Theory of Mechanism Design*. Oxford University Press, Oxford, 2015.

[79] V. S. Borkar. Stochastic approximation with two time scales. *Systems & Control Letters*, 29(5):291–294, 1997.

[80] V. S. Borkar. *Stochastic Approximation: A Dynamical Systems Viewpoint*. Hindustan Book Agency, New Delhi, 2008.

[81] V. S. Borkar. *Stochastic Approximation: A Dynamical Systems Viewpoint*. Texts and Readings in Mathematics. Springer, Singapore, 2 edition, 2023.

[82] S. Boyd, S.-J. Kim, L. Vandenberghe, and A. Hassibi. A tutorial on geometric programming. *Optimization and Engineering*, 8(1):67–127, 2007.

[83] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, 2004.

[84] J. Boyens, A. Smith, N. Bartol, K. Winkler, A. Holbrook, and M. Fallon. Cybersecurity supply chain risk management practices for systems and organizations. NIST Special Publication 800-161 Rev. 1, National Institute of Standards and Technology, 2022.

[85] S. J. Bradtke and M. O. Duff. Reinforcement learning methods for continuous-time Markov decision problems. In G. Tesauro, D. S. Touretzky, and T. K. Leen, editors, *Advances in Neural Information Processing Systems 7*, pages 393–400, Cambridge, MA, 1995. MIT Press.

[86] D. Braess. Über ein paradoxon aus der verkehrsplanung. *Unternehmensforschung*, 12:258–268, 1968.

[87] Y. Bramoullé and R. Kranton. Public goods in networks. *Journal of Economic Theory*, 135(1):478–494, 2007.

[88] L. E. J. Brouwer. Über Abbildung von Mannigfaltigkeiten. *Mathematische Annalen*, 71(1):97–115, 1911.

[89] G. Brown, M. Carlyle, J. Salmerón, and K. Wood. Defending critical infrastructure. *Interfaces*, 36(6):530–544, 2006.

[90] G. G. Brown and L. A. T. Cox, Jr. How probabilistic risk assessment can mislead terrorism risk analysts. *Risk Analysis*, 31(2):196–204, 2011.

[91] G. W. Brown. Iterative solution of games by fictitious play. In T. C. Koopmans, editor, *Activity Analysis of Production and Allocation*, pages 374–376. John Wiley and Sons, New York, 1951.

[92] M. Brundage, S. Avin, J. Clark, H. Toner, P. Eckersley, B. Garfinkel, A. Dafoe, P. Scharre, T. Zeitzoff, B. Filar, H. Anderson, H. Roff, G. C. Allen, J. Steinhardt, C. Flynn, S. Ó hÉigeartaigh, S. J. Beard, H. Belfield, S. Farquhar, C. Lyle, R. Crootof, O. Evans, M. Page, J. Bryson, R. Yampolskiy, and D. Amodei. The malicious use of artificial intelligence: Forecasting, prevention, and mitigation, 2018.

[93] M. Bruneau, S. E. Chang, R. T. Eguchi, G. C. Lee, T. D. O'Rourke, A. M. Reinhorn, M. Shinozuka, K. Tierney, W. A. Wallace, and D. von Winterfeldt. A framework to quantitatively assess and enhance the seismic resilience of communities. *Earthquake Spectra*, 19(4):733–752, 2003.

[94] E. Budish. Trust at scale: The economic limits of cryptocurrencies and blockchains. *Quarterly Journal of Economics*, 140(1):1–62, 2025.

[95] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin. Catastrophic cascade of failures in interdependent networks. *Nature*, 464:1025–1028, 2010.

[96] S. Caltagirone, A. Pendergast, and C. Betz. The Diamond Model of intrusion analysis. Technical report, Center for Cyber Intelligence Analysis and Threat Research, 2013.

[97] C. F. Camerer, T.-H. Ho, and J.-K. Chong. A cognitive hierarchy model of games. *Quarterly Journal of Economics*, 119(3):861–898, 2004.

[98] A. A. Cárdenas, S. Amin, B. Sinopoli, A. Giani, A. Perrig, and S. Sastry. Challenges for securing cyber physical systems. In *Workshop on Future Directions in Cyber-Physical Systems Security*, Newark, NJ, July 2009. U.S. Department of Homeland Security.

[99] D. A. Cerdeiro, M. Dziubiński, and S. Goyal. Individual security, contagion, and network design. *Journal of Economic Theory*, 170:182–226, 2017.

[100] N. Cesa-Bianchi and G. Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006.

[101] C.-T. Chen. *Linear System Theory and Design*. Oxford University Press, New York, 4th edition, 2013.

[102] J. Chen, Y. Huang, R. Zhang, and Q. Zhu. Optimal curing strategy for competing epidemics spreading over complex networks. *IEEE Transactions on Signal and Information Processing over Networks*, 7:294–308, 2021.

[103] J. Chen, C. Touati, and Q. Zhu. A dynamic game approach to strategic design of secure and resilient infrastructure network. *IEEE Transactions on Information Forensics and Security*, 15:462–474, 2020.

[104] J. Chen, C. Touati, and Q. Zhu. Optimal secure two-layer IoT network design. *IEEE Transactions on Control of Network Systems*, 7(1):398–409, 2020.

[105] J. Chen, C. Touati, and Q. Zhu. A dynamic game approach to designing secure interdependent IoT-enabled infrastructure network. *IEEE Transactions on Network Science and Engineering*, 8(3):2601–2612, 2021.

[106] J. Chen and Q. Zhu. Interdependent network formation games. SIAM Workshop on Network Science, 2015. arXiv:1602.07745, 2016.

[107] J. Chen and Q. Zhu. Resilient and decentralized control of multi-level cooperative mobile networks to maintain connectivity under adversarial environment. In *Proceedings of the 55th IEEE Conference on Decision and Control*, pages 5183–5188, 2016.

[108] J. Chen and Q. Zhu. A game-theoretic framework for resilient and distributed generation control of renewable energies in microgrids. *IEEE Transactions on Smart Grid*, 8(1):285–295, 2017.

[109] J. Chen and Q. Zhu. Security as a service for cloud-enabled Internet of controlled things under advanced persistent threats: A contract design approach. *IEEE Transactions on Information Forensics and Security*, 12(11):2736–2750, 2017.

[110] J. Chen and Q. Zhu. *A Game- and Decision-Theoretic Approach to Resilient Interdependent Network Analysis and Design*. Springer, Cham, 2019.

[111] J. Chen and Q. Zhu. Interdependent strategic security risk management with bounded rationality in the Internet of Things. *IEEE Transactions on Information Forensics and Security*, 14(11):2958–2971, 2019.

[112] J. Chen and Q. Zhu. Control of multilayer mobile autonomous systems in adversarial environments: A games-in-games approach. *IEEE Transactions on Control of Network Systems*, 7(3):1056–1068, 2020.

[113] J. Chen and Q. Zhu. A cross-layer design approach to strategic cyber defense and robust switching control of cyber-physical wind energy systems. *IEEE Transactions on Automation Science and Engineering*, 20(1):624–635, 2023.

[114] J. Chen, Q. Zhu, and T. Başar. Dynamic contract design for systemic cyber risk management of interdependent enterprise networks. *Dynamic Games and Applications*, 11(2):294–325, 2021.

[115] Z. S. Chen and Q. Zhu. A theory of multilevel interactive equilibrium in NeuroAI, 2026. arXiv:2605.10505.

[116] A. Cherepanov. Win32/Industroyer: A new threat for industrial control systems. Technical report, ESET, 2017.

[117] I.-K. Cho and D. M. Kreps. Signaling games and stable equilibria. *The Quarterly Journal of Economics*, 102(2):179–221, 1987.

[118] J.-H. Cho, D. P. Sharma, H. Alavizadeh, S. Yoon, N. Ben-Asher, T. J. Moore, D. S. Kim, H. Lim, and F. F. Nelson. Toward proactive, adaptive defense: A survey on moving target defense. *IEEE Communications Surveys & Tutorials*, 22(1):709–745, 2020.

[119] J.-H. Cho, M. Zhu, and M. P. Singh. Modeling and analysis of deception games based on hypergame theory. In *Autonomous Cyber Deception*, pages 49–74. Springer, Cham, 2019.

[120] J. Chod, N. Trichakis, G. Tsoukalas, H. Aspegren, and M. Weber. On the financing benefits of supply chain transparency and blockchain adoption. *Management Science*, 66(10):4378–4396, 2020.

[121] E. H. Clarke. Multipart pricing of public goods. *Public Choice*, 11:17–33, 1971.

[122] W. H. Clohessy and R. S. Wiltshire. Terminal guidance system for satellite rendezvous. *Journal of the Aerospace Sciences*, 27(9):653–658, 1960.

[123] R. Cohen, K. Erez, D. ben Avraham, and S. Havlin. Resilience of the Internet to random breakdowns. *Physical Review Letters*, 85(21):4626–4628, 2000.

[124] V. Conitzer and T. Sandholm. Computing the optimal strategy to commit to. In *Proceedings of the ACM Conference on Electronic Commerce*, pages 82–90, 2006.

[125] J. Conlisk. Why bounded rationality? *Journal of Economic Literature*, 34(2):669–700, 1996.

[126] Control Engineering Staff. Throwback attack: An insider releases 265,000 gallons of sewage on the Maroochy Shire. `https://www.controleng.com/throwback-attack-an-insider-releases-265000-gallons-of-sewage-on-the-maroochy-shire/`, Nov. 2021. Control Engineering, Throwback Attack series.

[127] O. L. V. Costa, M. D. Fragoso, and M. G. Todorov. *Continuous-Time Markov Jump Linear Systems*. Springer, Berlin, 2013.

[128] A. Costin, J. Zaddach, A. Francillon, and D. Balzarotti. A large-scale analysis of the security of embedded firmwares. In *Proceedings of the 23rd USENIX Security Symposium*, 2014.

[129] P. Courty and H. Li. Sequential screening. *The Review of Economic Studies*, 67(4):697–717, 2000.

[130] P. Courty and H. Li. Sequential screening. *The Review of Economic Studies*, 67(4):697–717, 2000.

[131] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley, Hoboken, 2nd edition, 2006.

[132] L. A. T. Cox, Jr. Game theory and risk analysis. *Risk Analysis*, 29(8):1062–1068, 2009.

[133] Cybersecurity and Infrastructure Security Agency. Cyber-attack against Ukrainian critical infrastructure. ICS Alert IR-ALERT-H-16-056-01, `https://www.cisa.gov/news-events/ics-alerts/ir-alert-h-16-056-01`, 2016. Released February 25, 2016.

[134] Cybersecurity and Infrastructure Security Agency. Compromise of U.S. water treatment facility. Cybersecurity Advisory AA21-042A, `https://www.cisa.gov/news-events/cybersecurity-advisories/aa21-042a`, 2021. Released February 11, 2021.

[135] Cybersecurity and Infrastructure Security Agency and Federal Bureau of Investigation. DarkSide ransomware: Best practices for preventing business disruption from ransomware attacks. Joint Cybersecurity Advisory AA21-131A, `https://www.cisa.gov/news-events/cybersecurity-advisories/aa21-131a`, 2021. Released May 11, 2021.

[136] Cybersecurity and Infrastructure Security Agency, Federal Bureau of Investigation, and Department of Energy. Tactics, techniques, and procedures of indicted state-sponsored Russian cyber actors targeting the energy sector. Joint Cybersecurity Advisory AA22-083A, `https://www.cisa.gov/news-events/cybersecurity-advisories/aa22-083a`, 2022. Released March 24, 2022; covers the TRITON/TRISIS safety-system intrusion.

[137] Cybersecurity and Infrastructure Security Agency, Federal Bureau of Investigation, National Security Agency, Environmental Protection Agency, and Israel National Cyber Directorate. IRGC-affiliated cyber actors exploit PLCs in multiple sectors, including US water and wastewater systems facilities. Joint Cybersecurity Advisory AA23-335A, `https://www.cisa.gov/news-events/cybersecurity-advisories/aa23-335a`, 2023. Released December 1, 2023.

[138] Cybersecurity and Infrastructure Security Agency, National Security Agency, and Federal Bureau of Investigation. PRC state-sponsored actors compromise and maintain persistent access to U.S. critical infrastructure. Joint Cybersecurity Advisory AA24-038A, `https://www.cisa.gov/news-events/cybersecurity-advisories/aa24-038a`, 2024. Released February 7, 2024.

[139] Cybersecurity for the Operational Technology Environment (CyOTE). CyOTE case study: Colonial pipeline. Technical report, Idaho National Laboratory for the U.S. Department of Energy, Office of Cybersecurity, Energy Security, and Emergency Response (CESER), Idaho Falls, ID, 2022. Case study of the May 2021 DarkSide ransomware attack.

[140] L. Dai. *Singular Control Systems*, volume 118 of *Lecture Notes in Control and Information Sciences*. Springer, Berlin, 1989.

[141] P. Dasgupta and J. B. Collins. A survey of game theoretic approaches for adversarial machine learning in cybersecurity tasks. *AI Magazine*, 40(2):31–43, 2019.

[142] M. H. A. Davis. *Markov Models and Optimization*. Chapman and Hall, London, 1993.

[143] I. De Oliveira Nunes, S. Jakkamsetti, N. Rattanavipanon, and G. Tsudik. On the TOCTOU problem in remote attestation. In *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2021.

[144] E. Debenedetti, J. Zhang, M. Balunović, L. Beurer-Kellner, M. Fischer, and F. Tramèr. AgentDojo: A dynamic environment to evaluate prompt injection attacks and defenses for LLM agents. In *Advances in Neural Information Processing Systems 37 (NeurIPS 2024), Datasets and Benchmarks Track*, pages 82895–82920, 2024.

[145] E. Delavari, J. Farooq, M. M. Butt, and Q. Zhu. Process-aware cross-layer adaptation for O-RAN-enabled industrial systems, 2026. arXiv:2608.13372.

[146] Z. Deng, Y. Guo, C. Han, W. Ma, J. Xiong, S. Wen, and Y. Xiang. AI agents under threat: A survey of key security challenges and future pathways. *ACM Computing Surveys*, 57(7):1–36, 2025.

[147] D. E. Denning. An intrusion-detection model. *IEEE Transactions on Software Engineering*, SE-13(2):222–232, 1987.

[148] A. Der Kiureghian and O. Ditlevsen. Aleatory or epistemic? Does it matter? *Structural Safety*, 31(2):105–112, 2009.

[149] Z. DeSmit, A. Kulkarni, and C. Wernz. Enhancing cyber-physical security in manufacturing through game-theoretic analysis. *Cyber-Physical Systems*, 4(4):232–259, 2018.

[150] S. M. Dibaji, M. Pirani, D. B. Flamholz, A. M. Annaswamy, K. H. Johansson, and A. Chakrabortty. A systems and control perspective of CPS security. *Annual Reviews in Control*, 47:394–411, 2019.

[151] C. T. Do, N. H. Tran, C. Hong, C. A. Kamhoua, K. A. Kwiat, E. Blasch, S. Ren, N. Pissinou, and S. S. Iyengar. Game theory for cyber security and privacy. *ACM Computing Surveys*, 50(2):1–37, 2017.

[152] E. J. Dockner, S. Jørgensen, N. V. Long, and G. Sorger. *Differential Games in Economics and Management Science*. Cambridge University Press, Cambridge, 2000.

[153] Dover Microsystems. German steel mill cyberattack. `https://www.dovermicrosystems.com/case-study/german-steel-mill-cyberattack/`, 2021. Case study of the 2014 incident; accessed June 2025.

[154] J. J. Downs and E. F. Vogel. A plant-wide industrial process control problem. *Computers & Chemical Engineering*, 17(3):245–255, 1993.

[155] J. C. Doyle. Guaranteed margins for LQG regulators. *IEEE Transactions on Automatic Control*, 23(4):756–757, 1978.

[156] M. Dresher. *Games of Strategy: Theory and Applications*. Prentice-Hall, Englewood Cliffs, NJ, 1961.

[157] Y. Du, S. Li, A. Torralba, J. B. Tenenbaum, and I. Mordatch. Improving factuality and reasoning in language models through multiagent debate. *arXiv preprint arXiv:2305.14325*, 2023.

[158] P. Dupuis and A. Nagurney. Dynamical systems and variational inequalities. *Annals of Operations Research*, 44:9–42, 1993.

[159] K. Durkota, V. Lisý, C. Kiekintveld, K. Horák, B. Bošanský, and T. Pevný. Optimal strategies for detecting data exfiltration by internal and external attackers. In *International Conference on Decision and Game Theory for Security*, pages 171–192. Springer, 2017.

[160] R. Durrett. *Probability: Theory and Examples*. Cambridge University Press, Cambridge, 5 edition, 2019.

[161] M. Dziubiński and S. Goyal. Network design and defence. *Games and Economic Behavior*, 79:30–43, 2013.

[162] K. Eldefrawy, G. Tsudik, A. Francillon, and D. Perito. SMART: Secure and minimal architecture for (establishing a dynamic) root of trust. In *Proceedings of the Network and Distributed System Security Symposium (NDSS)*, 2012.

[163] P. Emami-Naeini, Y. Agarwal, L. F. Cranor, and H. Hibshi. Ask the experts: What should be on an IoT privacy and security label? In *Proceedings of the IEEE Symposium on Security and Privacy (S&P)*, 2020.

[164] J. C. Engwerda. *LQ Dynamic Optimization and Differential Games*. John Wiley and Sons, Chichester, 2005.

[165] J. C. Engwerda and Salmah. The open-loop linear quadratic differential game for index one descriptor systems. *Automatica*, 45(2):585–592, 2009.

[166] J. C. Engwerda and Salmah. Feedback Nash equilibria for linear quadratic descriptor differential games. *Automatica*, 48(4):625–631, 2012.

[167] J. D. Esary and F. Proschan. Coherent structures of non-identical components. *Technometrics*, 5(2):191–209, 1963.

[168] J. D. Esary, F. Proschan, and D. W. Walkup. Association of random variables, with applications. *The Annals of Mathematical Statistics*, 38(5):1466–1474, 1967.

[169] ESET Research. Industroyer2: Industroyer reloaded. WeLiveSecurity, `https://www.welivesecurity.com/2022/04/12/industroyer2-industroyer-reloaded/`, 2022. Published April 12, 2022.

[170] I. Esponda and D. Pouzo. Berk–Nash equilibrium: A framework for modeling agents with misspecified models. *Econometrica*, 84(3):1093–1130, 2016.

[171] I. Esponda and D. Pouzo. Berk–Nash equilibrium with exhaustive learning. *Econometrica*, 85(4):1275–1306, 2017.

[172] I. Esponda and D. Pouzo. Equilibrium in misspecified markov decision processes. *Theoretical Economics*, 16(2):717–757, 2021.

[173] S. R. Etesami and T. Başar. Dynamic games in cyber-physical security: An overview. *Dynamic Games and Applications*, 9(4):884–913, 2019.

[174] ETSI. Cyber security for consumer Internet of Things: Baseline requirements. ETSI EN 303 645 V3.1.3, 2024.

[175] ETSI. O-RAN security threat modeling and risk assessment. ETSI TR 104 106 V3.0.0, 2025.

[176] European Parliament and Council of the European Union. Directive (EU) 2022/2555 on measures for a high common level of cybersecurity across the union (NIS 2 directive). Official Journal of the European Union, L 333, 2022.

[177] European Parliament and Council of the European Union. Regulation (EU) 2024/2847 on horizontal cybersecurity requirements for products with digital elements (Cyber Resilience Act). Official Journal of the European Union, L series, 2024.

[178] European Parliament and Council of the European Union. Regulation (EU) 2024/2847 on horizontal cybersecurity requirements for products with digital elements (Cyber Resilience Act). Official Journal of the European Union, L 2024/2847, 2024.

[179] L. C. Evans and P. E. Souganidis. Differential games and representation formulas for solutions of Hamilton–Jacobi–Isaacs equations. *Indiana University Mathematics Journal*, 33(5):773–797, 1984.

[180] Executive Office of the President. Executive Order 14028: Improving the Nation's cybersecurity. Federal Register 86 FR 26633, May 2021.

[181] I. Eyal. The miner's dilemma. In *Proceedings of the IEEE Symposium on Security and Privacy*, pages 89–103, 2015.

[182] I. Eyal and E. G. Sirer. Majority is not enough: Bitcoin mining is vulnerable. In *Financial Cryptography and Data Security (FC 2014)*, volume 8437 of *Lecture Notes in Computer Science*, pages 436–454. Springer, 2014.

[183] A. Fabrikant, A. Luthra, E. Maneva, C. H. Papadimitriou, and S. Shenker. On a network creation game. In *Proceedings of the 22nd Annual ACM Symposium on Principles of Distributed Computing (PODC)*, pages 347–351, 2003.

[184] F. Facchinei and J.-S. Pang. *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer, New York, 2003.

[185] M. Fagan, K. N. Megas, K. Scarfone, and M. Smith. IoT device cybersecurity capability core baseline. NISTIR 8259A, National Institute of Standards and Technology, 2020.

[186] J. Farooq and Q. Zhu. *Resource Management for On-Demand Mission-Critical Internet of Things Applications*. Wiley–IEEE Press, Hoboken, NJ, 2021.

[187] J. Farooq and Q. Zhu. *Resource Management for On-Demand Mission-Critical Internet of Things Applications*. Wiley, Hoboken, NJ, 2021.

[188] J. Farooq and Q. Zhu. Cyber resilience in next-generation networks: Threat landscape, theoretical foundations, and design paradigms. *Foundations and Trends in Networking*, 17(1–2):1–145, 2026.

[189] M. J. Farooq and Q. Zhu. A multi-layer feedback system approach to resilient connectivity of remotely deployed mobile Internet of Things. *IEEE Transactions on Cognitive Communications and Networking*, 4(2):422–432, 2018.

[190] M. J. Farooq and Q. Zhu. On the secure and reconfigurable multi-layer network design for critical information dissemination in the Internet of Battlefield Things (IoBT). *IEEE Transactions on Wireless Communications*, 17(4):2618–2632, 2018.

[191] M. J. Farooq and Q. Zhu. Optimal dynamic contract for spectrum reservation in mission-critical UNB-IoT systems. In *Proceedings of the 16th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, pages 1–6, Shanghai, China, 2018.

[192] M. J. Farooq and Q. Zhu. Modeling, analysis, and mitigation of dynamic botnet formation in wireless IoT networks. *IEEE Transactions on Information Forensics and Security*, 14(9):2412–2426, 2019.

[193] H. Fawzi, P. Tabuada, and S. Diggavi. Secure estimation and control for cyber-physical systems under adversarial attacks. *IEEE Transactions on Automatic Control*, 59(6):1454–1467, 2014.

[194] Federal Communications Commission. Cybersecurity labeling for Internet of Things. Report and Order FCC 24-26, 2024.

[195] K. Ferguson-Walter, T. Shade, A. Rogers, E. Niedbala, M. Trumbo, K. Nauer, K. Divis, A. Jones, A. Combs, and R. Abbott. The Tularosa study: An experimental design and implementation to quantify the effectiveness of cyber deception. In *Proceedings of the 52nd Hawaii International Conference on System Sciences (HICSS)*, 2019.

[196] M. Fiedler. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23(2):298–305, 1973.

[197] J. Filar and K. Vrieze. *Competitive Markov decision processes*. Springer Science & Business Media, 2012.

[198] J. A. Filar and K. Vrieze. *Competitive Markov Decision Processes*. Springer, 1997.

[199] A. M. Fink. Equilibrium in a stochastic n-person game. *Journal of Science of the Hiroshima University, Series A-I (Mathematics)*, 28(1):89–93, 1964.

[200] W. H. Fleming and H. M. Soner. *Controlled Markov Processes and Viscosity Solutions*, volume 25 of *Stochastic Modelling and Applied Probability*. Springer, New York, 2nd edition, 2006.

[201] A. Fonseca-Morales and O. Hernández-Lerma. Potential differential games. *Dynamic Games and Applications*, 8(2):254–279, 2018.

[202] Forum of Incident Response and Security Teams. Common Vulnerability Scoring System version 3.1: Specification document, 2019. FIRST standard.

[203] Y. Freund and R. E. Schapire. Adaptive game playing using multiplicative weights. *Games and Economic Behavior*, 29(1–2):79–103, 1999.

[204] D. Fudenberg and D. K. Levine. *The Theory of Learning in Games*. MIT Press, 1998.

[205] D. Fudenberg and J. Tirole. *Game Theory*. MIT Press, Cambridge, MA, 1991.

[206] S. Fugate and K. Ferguson-Walter. Artificial intelligence and game theory models for defending critical networks with cyber deception. *AI Magazine*, 40(1):49–62, 2019.

[207] X. Gabaix. A sparsity-based model of bounded rationality. *Quarterly Journal of Economics*, 129(4):1661–1710, 2014.

[208] X. Gabaix. Behavioral inattention. In *Handbook of Behavioral Economics*, volume 2, pages 261–343. Elsevier, 2019.

[209] A. Galeotti, B. Golub, and S. Goyal. Targeting interventions in networks. *Econometrica*, 88(6):2445–2471, 2020.

[210] A. Galeotti, S. Goyal, M. O. Jackson, F. Vega-Redondo, and L. Yariv. Network games. *Review of Economic Studies*, 77(1):218–244, 2010.

[211] A. Garnaev, Y. Hayel, and E. Altman. A Bayesian jamming game in an OFDM wireless network. In *Proceedings of the 10th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, pages 41–48, Paderborn, Germany, 2012.

[212] Y. Ge, Y.-T. Yang, and Q. Zhu. Attributing responsibility in AI-induced incidents: A computational reflective equilibrium framework for accountability, 2024. arXiv:2404.16957.

[213] Y. Ge and Q. Zhu. GAZETA: GAme-theoretic ZEro-Trust authentication for defense against lateral movement in 5G IoT networks. *IEEE Transactions on Information Forensics and Security*, 19:540–554, 2024.

[214] Y. Ge and Q. Zhu. MEGA-PT: A meta-game framework for agile penetration testing. arXiv:2409.14219, 2024.

[215] Y. Ge and Q. Zhu. Game-theoretic frameworks for zero-trust authentication in autonomous cyber resilience. In C. A. Kamhoua, A. Kott, Q. Zhu, and N. O. Leslie, editors, *Autonomous Cyber Resilience*, pages 105–138. Wiley–IEEE Press, Hoboken, NJ, 2026.

[216] Y. Ge and Q. Zhu. The game-theoretic symbiosis of trust and AI in networked systems. In *Foundations of Cyber Deception*, Advances in Information Security, pages 63–89. Springer, Cham, 2026.

[217] S. J. Gershman, E. J. Horvitz, and J. B. Tenenbaum. Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Behavioral and Brain Sciences*, 38:e71, 2015.

[218] A. Ghosh and S. Boyd. Growing well-connected graphs. In *Proceedings of the 45th IEEE Conference on Decision and Control*, pages 6605–6611, 2006.

[219] G. Gigerenzer and R. Selten, editors. *Bounded Rationality: The Adaptive Toolbox*. MIT Press, Cambridge, MA, 2001.

[220] I. L. Glicksberg. A further generalization of the Kakutani fixed point theorem, with application to Nash equilibrium points. *Proceedings of the American Mathematical Society*, 3(1):170–174, 1952.

[221] J. K. Goeree, C. A. Holt, and T. R. Palfrey. Regular quantal response equilibrium. *Experimental Economics*, 8(4):347–367, 2005.

[222] L. A. Gordon and M. P. Loeb. The economics of information security investment. *ACM Transactions on Information and System Security*, 5(4):438–457, 2002.

[223] S. Goyal. *Connections: An Introduction to the Economics of Networks*. Princeton University Press, Princeton, NJ, 2007.

[224] S. Goyal, S. Jabbari, M. Kearns, S. Khanna, and J. Morgenstern. Strategic network formation with attack and immunization. In *Web and Internet Economics (WINE 2016)*, volume 10123 of *Lecture Notes in Computer Science*, pages 429–443. Springer, 2016.

[225] S. Goyal and A. Vigier. Attack, defence, and contagion in networks. *Review of Economic Studies*, 81(4):1518–1542, 2014.

[226] K. Greshake, S. Abdelnabi, S. Mishra, C. Endres, T. Holz, and M. Fritz. Not what you've signed up for: Compromising real-world LLM-integrated applications with indirect prompt injection. In *Proceedings of the 16th ACM Workshop on Artificial Intelligence and Security (AISec)*, pages 79–90. ACM, 2023.

[227] O. Gross and R. Wagner. A continuous Colonel Blotto game. Research Memorandum RM-408, RAND Corporation, Santa Monica, CA, 1950.

[228] T. Groves. Incentives in teams. *Econometrica*, 41(4):617–631, 1973.

[229] C. Guestrin, D. Koller, R. Parr, and S. Venkataraman. Efficient solution algorithms for factored MDPs. *Journal of Artificial Intelligence Research*, 19:399–468, 2003.

[230] Y. Y. Haimes. *Risk Modeling, Assessment, and Management*. Wiley, Hoboken, NJ, 4 edition, 2016.

[231] K. Hammar, T. Li, R. Stadler, and Q. Zhu. Adaptive security response strategies through conjectural online learning. *IEEE Transactions on Information Forensics and Security*, 20:4055–4070, 2025.

[232] K. Hammar and R. Stadler. Finding effective security strategies through reinforcement learning and self-play. In *16th International Conference on Network and Service Management (CNSM)*, pages 1–9. IEEE, 2020.

[233] Z. Han, D. Niyato, W. Saad, T. Başar, and A. Hjørungnes. *Game Theory in Wireless and Communication Networks: Theory, Models, and Applications*. Cambridge University Press, Cambridge, 2012.

[234] Z. Han and Q. Zhu. Toward a dynamic Stackelberg game-theoretic framework for agentic AI defense against LLM jailbreaking, 2025. arXiv:2507.08207.

[235] J. Hannan. Approximation to Bayes risk in repeated play. In M. Dresher, A. W. Tucker, and P. Wolfe, editors, *Contributions to the Theory of Games, Volume III*, number 39 in Annals of Mathematics Studies, pages 97–139. Princeton University Press, Princeton, NJ, 1957.

[236] E. A. Hansen, D. S. Bernstein, and S. Zilberstein. Dynamic programming for partially observable stochastic games. In *Proceedings of the Nineteenth National Conference on Artificial Intelligence*, pages 709–715, 2004.

[237] F. Harary. The maximum connectivity of a graph. *Proceedings of the National Academy of Sciences*, 48(7):1142–1146, 1962.

[238] M. Hardt and C. Mendler-Dünner. Performative prediction: Past and future. *Statistical Science*, 40(3):417–436, 2025.

[239] J. C. Harsanyi. Games with incomplete information played by "bayesian" players, I–III: Part I. the basic model. *Management Science*, 14(3):159–182, 1967.

[240] J. C. Harsanyi. Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points. *International Journal of Game Theory*, 2(1):1–23, 1973.

[241] J. C. Harsanyi and R. Selten. A generalized Nash solution for two-person bargaining games with incomplete information. *Management Science*, 18(5, Part 2):P80–P106, 1972.

[242] S. Hart and A. Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150, 2000.

[243] J.-P. Hauet, P. Bock, R. Foley, and R. Françoise. Ukrainian power grids cyberattack: A forensic analysis based on ISA/IEC 62443. *InTech*, Mar./Apr. 2017.

[244] Y. Hayel and Q. Zhu. Evolutionary Poisson games for controlling large population behaviors. In *Proceedings of the 54th IEEE Conference on Decision and Control (CDC)*, Osaka, Japan, 2015.

[245] Y. Hayel and Q. Zhu. Epidemic protection over heterogeneous networks using evolutionary Poisson games. *IEEE Transactions on Information Forensics and Security*, 12(8):1786–1800, 2017.

[246] F. Heiding, S. Lermen, A. Kao, B. Schneier, and A. Vishwanath. Evaluating large language models' capability to launch fully automated spear phishing campaigns: Validated on human subjects, 2024.

[247] S. Herwig, K. Harvey, G. Hughey, R. Roberts, and D. Levin. Measurement and analysis of Hajime, a peer-to-peer IoT botnet. In *Proceedings of the Network and Distributed System Security Symposium (NDSS)*, 2019.

[248] R. Hiesgen, M. Nawrocki, T. C. Schmidt, and M. Wählisch. The race to the vulnerable: Measuring the Log4j shell incident. In *Proceedings of the Network Traffic Measurement and Analysis Conference (TMA)*, 2022.

[249] Y. C. Ho, A. E. Bryson, and S. Baron. Differential games and optimal pursuit-evasion strategies. *IEEE Transactions on Automatic Control*, 10(4):385–389, 1965.

[250] J. Hofbauer and W. H. Sandholm. On the global convergence of stochastic fictitious play. *Econometrica*, 70(6):2265–2294, 2002.

[251] J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge University Press, Cambridge, 1998.

[252] B. Holmström. Moral hazard and observability. *The Bell Journal of Economics*, 10(1):74–91, 1979.

[253] K. Horák, B. Bošanský, V. Kovařík, and C. Kiekintveld. Solving zero-sum one-sided partially observable stochastic games. *Artificial Intelligence*, 316:103838, 2023.

[254] K. Horák, B. Bošanský, and M. Pěchouček. Heuristic search value iteration for one-sided partially observable stochastic games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, 2017.

[255] K. Horák, B. Bošanský, P. Tomášek, C. Kiekintveld, and C. Kamhoua. Optimizing honeypot strategies against dynamic lateral movement using partially observable stochastic games. *Computers & Security*, 87:101579, 2019.

[256] K. Horák, Q. Zhu, and B. Bošanský. Manipulating adversary's belief: A dynamic game approach to deception by design for proactive network security. In *Decision and Game Theory for Security (GameSec 2017)*, volume 10575 of *Lecture Notes in Computer Science*, pages 273–294, Cham, 2017. Springer.

[257] J. Hörner, D. Rosenberg, E. Solan, and N. Vieille. On a Markov game with one-sided information. *Operations Research*, 58(4-part-2):1107–1115, 2010.

[258] J. T. House and G. Cybenko. Hypergame theory applied to cyber attack and defense. In *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense IX*, volume 7666 of *Proceedings of SPIE*, page 766604, 2010.

[259] J. Hu and M. P. Wellman. Nash q-learning for general-sum stochastic games. *Journal of machine learning research*, 4(Nov):1039–1069, 2003.

[260] B. Huang, A. A. Cárdenas, and R. Baldick. Not everything is dark and gloomy: Power grid protections against IoT demand attacks. In *Proceedings of the 28th USENIX Security Symposium*, pages 1115–1132, 2019.

[261] J. Huang and Q. Zhu. PenHeal: A two-stage LLM framework for automated pentesting and optimal remediation. In *Proceedings of the Workshop on Autonomous Cybersecurity (AutonomousCyber '24), ACM CCS*, pages 11–22, 2024.

[262] L. Huang, J. Chen, and Q. Zhu. Factored markov game theory for secure interdependent infrastructure networks. In S. Rass and S. Schauer, editors, *Game Theory for Security and Risk Management: From Theory to Practice*, Static & Dynamic Game Theory: Foundations & Applications, pages 99–126. Birkhäuser, Cham, 2018.

[263] L. Huang, S. Jia, E. Balcetis, and Q. Zhu. ADVERT: An adaptive and data-driven attention enhancement mechanism for phishing prevention. *IEEE Transactions on Information Forensics and Security*, 17:2585–2597, 2022.

[264] L. Huang and Q. Zhu. Adaptive honeypot engagement through reinforcement learning of semi-markov decision processes. In *Decision and Game Theory for Security*, volume 11836 of *Lecture Notes in Computer Science*, pages 196–216. Springer, 2019.

[265] L. Huang and Q. Zhu. Adaptive strategic cyber defense for advanced persistent threats in critical infrastructure networks. *ACM SIGMETRICS Performance Evaluation Review*, 46(2):52–56, 2019.

[266] L. Huang and Q. Zhu. A dynamic games approach to proactive defense strategies against advanced persistent threats in cyber-physical systems, 2019.

[267] L. Huang and Q. Zhu. Strategic learning for active, adaptive, and autonomous cyber defense, 2019.

[268] L. Huang and Q. Zhu. A dynamic games approach to proactive defense strategies against advanced persistent threats in cyber-physical systems. *Computers & Security*, 89:101660, 2020.

[269] L. Huang and Q. Zhu. Strategic learning for active, adaptive, and autonomous cyber defense. In S. Jajodia, G. Cybenko, V. S. Subrahmanian, V. Swarup, C. Wang, and M. Wellman, editors, *Adaptive Autonomous Secure Cyber Systems*, pages 205–230. Springer, Cham, 2020.

[270] L. Huang and Q. Zhu. Duplicity games for deception design with an application to insider threat mitigation. *IEEE Transactions on Information Forensics and Security*, 2021.

[271] L. Huang and Q. Zhu. RADAMS: Resilient and adaptive alert and attention management strategy against informational denial-of-service (IDoS) attacks. *Computers & Security*, 121:102844, 2022.

[272] L. Huang and Q. Zhu. *Cognitive Security: A System-Scientific Approach*. SpringerBriefs in Computer Science. Springer, Cham, 2023.

[273] L. Huang and Q. Zhu. ZETAR: Modeling and computational design of strategic and adaptive compliance policies. *IEEE Transactions on Computational Social Systems*, 11(3):4001–4015, 2024.

[274] S. Huang, F. Jones, N. Gurney, D. Pynadath, K. Srivastava, S. Trent, P. Wu, and Q. Zhu. PsybORG+: Modeling and simulation for detecting cognitive biases in advanced persistent threats, 2024. arXiv:2408.01310.

[275] Y. Huang, J. Chen, L. Huang, and Q. Zhu. Dynamic games for secure and resilient control system design. arXiv:1910.07510, 2019.

[276] Y. Huang, J. Chen, L. Huang, and Q. Zhu. Dynamic games for secure and resilient control system design. *National Science Review*, 7(7):1125–1141, 2020.

[277] Y. Huang, J. Chen, and Q. Zhu. Defending an asset with partial information and selected observations: A differential game framework. In *Proceedings of the 60th IEEE Conference on Decision and Control (CDC)*, pages 2366–2373, 2021.

[278] Y. Huang, L. Huang, and Q. Zhu. Reinforcement learning for feedback-enabled cyber resilience. *Annual Reviews in Control*, 53:273–295, 2022.

[279] Y. Huang and Q. Zhu. A differential game approach to decentralized virus-resistant weight adaptation policy over complex networks. *IEEE Transactions on Control of Network Systems*, 7(2):944–955, 2020.

[280] Y. Huang and Q. Zhu. Manipulating reinforcement learning: Stealthy attacks on cost signals. In C. A. Kamhoua, C. D. Kiekintveld, F. Fang, and Q. Zhu, editors, *Game Theory and Machine Learning for Cyber Security*, pages 367–388. Wiley-IEEE Press, 2021.

[281] Y. Huang and Q. Zhu. Game-theoretic frameworks for epidemic spreading and human decision-making: A review. *Dynamic Games and Applications*, 12(1):7–48, 2022.

[282] D. W. Hubbard and R. Seiersen. *How to Measure Anything in Cybersecurity Risk*. Wiley, Hoboken, NJ, 2016.

[283] A. A. Humayed, J. Lin, F. Li, and B. Luo. Cyber-physical systems security – a survey. *IEEE Internet of Things Journal*, 4(6):1802–1831, 2017.

[284] S. R. Hussain, M. Echeverria, I. Karim, O. Chowdhury, and E. Bertino. 5GReasoner: A property-directed security and privacy analysis framework for 5G cellular network protocol. In *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security (CCS)*, pages 669–684, 2019.

[285] E. M. Hutchins, M. J. Cloppert, and R. M. Amin. Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. In *Leading Issues in Information Warfare and Security Research*, volume 1, pages 80–106. Academic Publishing International, 2011.

[286] A. Ibrahim, A.-R. Sadeghi, and S. Zeitouni. SeED: Secure non-interactive attestation for embedded devices. In *Proceedings of the 10th ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec)*, 2017.

[287] O. C. Imer, S. Yüksel, and T. Başar. Optimal control of LTI systems over unreliable communication links. *Automatica*, 42(9):1429–1439, 2006.

[288] International Electrotechnical Commission. Industrial communication networks—network and system security. Technical Report 62443, IEC, 2018.

[289] International Organization for Standardization. Risk management—guidelines. International Standard 31000:2018, ISO, 2018.

[290] International Organization for Standardization. Information security, cybersecurity and privacy protection—information security management systems—requirements. Technical Report 27001:2022, ISO/IEC, 2022.

[291] International Organization for Standardization (ISO). Cyber-physical systems. `https://www.iso.org/foresight/cyber-physical-systems.html`, 2017. ISO website (accessed 2025-06).

[292] International Society of Automation. ISA/IEC 62443 series of standards. `https://www.isa.org/standards-and-publications/isa-standards/isa-iec-62443-series-of-standards`, 2026. Accessed June 30, 2026.

[293] International Society of Automation (ISA). Isa/iec 62443 series of standards. `https://www.isa.org/standards-and-publications/isa-standards/isa-iec-62443-series-of-standards`, 2024. Accessed 2025-06.

[294] R. Isaacs. *Differential Games: A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization*. John Wiley and Sons, 1965.

[295] E. Israeli and R. K. Wood. Shortest-path network interdiction. *Networks*, 40(2):97–111, 2002.

[296] M. O. Jackson. *Social and Economic Networks*. Princeton University Press, Princeton, NJ, 2008.

[297] M. O. Jackson and A. Watts. The evolution of social and economic networks. *Journal of Economic Theory*, 106(2):265–295, 2002.

[298] M. O. Jackson and A. Wolinsky. A strategic model of social and economic networks. *Journal of Economic Theory*, 71(1):44–74, 1996.

[299] S. Jajodia, V. S. Subrahmanian, V. Swarup, and C. Wang, editors. *Cyber Deception: Building the Scientific Foundation*. Springer, 2016.

[300] Y. Ji and H. J. Chizeck. Controllability, stabilizability, and continuous-time Markovian jump linear quadratic control. *IEEE Transactions on Automatic Control*, 35(7):777–788, 1990.

[301] L. Jia, Y. Xu, Y. Sun, S. Feng, and A. Anpalagan. Stackelberg game approaches for anti-jamming defence in wireless networks. *IEEE Wireless Communications*, 25(6):120–128, 2018.

[302] J. Jin, Z. Pang, J. Kua, Q. Zhu, K. H. Johansson, N. Marchenko, and D. Cavalcanti. Cloud-fog automation: The new paradigm toward autonomous industrial cyber-physical systems. *IEEE Journal on Selected Areas in Communications*, 43(9):2917–2937, 2025.

[303] B. Johnson, D. Caban, M. Krotofil, D. Scali, N. Brubaker, and C. Glyer. Attackers deploy new ICS attack framework "TRITON" and cause operational disruption to critical infrastructure. Mandiant (FireEye) Threat Intelligence, `https://cloud.google.com/blog/topics/threat-intelligence/attackers-deploy-new-ics-attack-framework-triton`, 2017. Published December 14, 2017.

[304] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1–2):99–134, 1998.

[305] D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291, 1979.

[306] T. Kailath. *Linear Systems*. Prentice-Hall, Englewood Cliffs, NJ, 1980.

[307] S. Kakutani. A generalization of Brouwer's fixed point theorem. *Duke Mathematical Journal*, 8(3):457–459, 1941.

[308] R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(1):35–45, 1960.

[309] E. Kamenica. Bayesian persuasion and information design. *Annual Review of Economics*, 11:249–272, 2019.

[310] E. Kamenica and M. Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.

[311] C. A. Kamhoua, C. D. Kiekintveld, F. Fang, and Q. Zhu, editors. *Game Theory and Machine Learning for Cyber Security*. Wiley-IEEE Press, Hoboken, NJ, 2021.

[312] C. A. Kamhoua, A. Kott, Q. Zhu, and N. O. Leslie, editors. *Autonomous Cyber Resilience*. Wiley–IEEE Press, Hoboken, NJ, 2026.

[313] C. A. Kamhoua, M. Rodriguez, and K. A. Kwiat. Testing for hardware Trojans: A game-theoretic approach. In *Decision and Game Theory for Security (GameSec 2014)*, volume 8840 of *Lecture Notes in Computer Science*, pages 360–369. Springer, 2014.

[314] S. Kaplan and B. J. Garrick. On the quantitative definition of risk. *Risk Analysis*, 1(1):11–27, 1981.

[315] D. Kar, T. H. Nguyen, F. Fang, M. Brown, A. Sinha, M. Tambe, and A. X. Jiang. Trends and applications in stackelberg security games. Book chapter, 2017.

[316] H. Kellerer, U. Pferschy, and D. Pisinger. *Knapsack Problems*. Springer, Berlin, 2004.

[317] J. G. Kemeny and J. L. Snell. *Finite Markov Chains*. Springer, New York, 1976.

[318] M. H. R. Khouzani, S. Sarkar, and E. Altman. Maximum damage malware attack in mobile wireless networks. *IEEE/ACM Transactions on Networking*, 20(5):1347–1360, 2012.

[319] M. H. R. Khouzani, S. Sarkar, and E. Altman. Optimal dissemination of security patches in mobile wireless networks. *IEEE Transactions on Information Theory*, 58(7):4714–4732, 2012.

[320] M. H. R. Khouzani, S. Sarkar, and E. Altman. Saddle-point strategies in malware attack. *IEEE Journal on Selected Areas in Communications*, 30(1):31–43, 2012.

[321] C. Kiekintveld, M. Jain, J. Tsai, J. Pita, F. Ordóñez, and M. Tambe. Computing optimal randomized resource allocations for massive security games. In *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems*, pages 689–696, 2009.

[322] C. Kiekintveld, V. Lisý, and R. Píbil. Game-theoretic foundations for the strategic use of honeypots in network security. In S. Jajodia, P. Shakarian, V. S. Subrahmanian, V. Swarup, and C. Wang, editors, *Cyber Warfare: Building the Scientific Foundation*, Advances in Information Security, pages 81–101. Springer, Cham, 2015.

[323] C. Kiennert, Z. Ismail, H. Debar, and J. Leneutre. A survey on game-theoretic approaches for intrusion detection and response optimization. *ACM Computing Surveys*, 51(5):1–31, 2018.

[324] T. Kieras, J. Farooq, and Q. Zhu. I-SCRAM: A framework for IoT supply chain risk analysis and mitigation decisions. *IEEE Access*, 9:29827–29840, 2021.

[325] T. Kieras, J. Farooq, and Q. Zhu. *IoT Supply Chain Security Risk Analysis and Mitigation: Modeling, Computations, and Software Tools*. SpringerBriefs in Computer Science. Springer, Cham, 2022.

[326] T. Kieras, M. J. Farooq, and Q. Zhu. Modeling and assessment of IoT supply chain security risks: The role of structural and parametric uncertainties. In *Proceedings of the IEEE Security and Privacy Workshops (SPW)*, 2020.

[327] T. Kieras, M. J. Farooq, and Q. Zhu. RIoTS: Risk analysis of IoT supply chain threats. In *Proceedings of the IEEE 6th World Forum on Internet of Things (WF-IoT)*, pages 1–6, 2020.

[328] Y. Kim, G. Dán, and Q. Zhu. Human-in-the-loop cyber intrusion detection using active learning. *IEEE Transactions on Information Forensics and Security*, 19:8658–8672, 2024.

[329] Y. Kim, Q. Zhu, and G. Dán. ADAPTD: Adaptive detection and proactive threat defense for autonomous APT attacks, 2026. arXiv:2608.17251.

[330] F. H. Knight. *Risk, Uncertainty and Profit*. Houghton Mifflin, Boston, 1921.

[331] C. Kolias, G. Kambourakis, A. Stavrou, and J. Voas. DDoS in the IoT: Mirai and other botnets. *Computer*, 50(7):80–84, 2017.

[332] D. Koller, N. Megiddo, and B. von Stengel. Efficient computation of equilibria for extensive two-person games. *Games and Economic Behavior*, 14(2):247–259, 1996.

[333] B. Kordy, L. Piètre-Cambacédès, and P. Schweitzer. DAG-based attack and defense modeling: Don't miss the forest for the attack trees. *Computer Science Review*, 13–14:1–38, 2014.

[334] J. Koshal, A. Nedić, and U. V. Shanbhag. Multiuser optimization: Distributed algorithms and error analysis. *SIAM Journal on Optimization*, 21(3):1046–1081, 2011.

[335] M. Kravchik and A. Shabtai. Detecting cyberattacks in industrial control systems using convolutional neural networks. arXiv:1806.08110, 2018.

[336] D. M. Kreps and R. Wilson. Sequential equilibria. *Econometrica*, 50(4):863–894, 1982.

[337] H. W. Kuhn. Extensive games and the problem of information. In H. W. Kuhn and A. W. Tucker, editors, *Contributions to the Theory of Games II*, volume 28 of *Annals of Mathematics Studies*, pages 193–216. Princeton University Press, Princeton, NJ, 1953.

[338] P. Kundur. *Power System Stability and Control*. McGraw-Hill, New York, 1994.

[339] P. Kunkel and V. Mehrmann. *Differential-Algebraic Equations: Analysis and Numerical Solution*. European Mathematical Society, Zürich, 2006.

[340] H. Kunreuther and G. Heal. Interdependent security. *Journal of Risk and Uncertainty*, 26(2–3):231–249, 2003.

[341] H. Kunreuther and G. Heal. Interdependent security. *Journal of Risk and Uncertainty*, 26(2–3):231–249, 2003.

[342] H. J. Kushner and G. G. Yin. *Stochastic Approximation and Recursive Algorithms and Applications*. Springer, New York, 2 edition, 2003.

[343] P. Ladisa, H. Plate, M. Martinez, and O. Barais. SoK: Taxonomy of attacks on open-source software supply chains. In *Proceedings of the IEEE Symposium on Security and Privacy*, pages 1509–1526, 2023.

[344] J.-J. Laffont and D. Martimort. *The Theory of Incentives: The Principal-Agent Model*. Princeton University Press, Princeton, NJ, 2002.

[345] S. Lakshminarayana, E. V. Belmega, and H. V. Poor. Moving-target defense against cyber-physical attacks in power grids via game theory. arXiv:2006.07697, 2020.

[346] B. W. Lampson. Protection. In *Proceedings of the 5th Princeton Symposium on Information Sciences and Systems*, pages 437–443, 1971.

[347] R. Langner. Stuxnet: Dissecting a cyberwarfare weapon. *IEEE Security & Privacy*, 9(3):49–51, 2011.

[348] H. Lei, Y. Ge, and Q. Zhu. ADAPT: A game-theoretic and neuro-symbolic framework for automated distributed adaptive penetration testing. In *Proceedings of the IEEE Military Communications Conference (MILCOM)*, pages 7–12, 2024.

[349] S. Leonardos, W. Overman, I. Panageas, and G. Piliouras. Global convergence of multi-agent policy gradient in Markov potential games. In *International Conference on Learning Representations (ICLR)*, 2022.

[350] N. G. Leveson. *Engineering a Safer World: Systems Thinking Applied to Safety*. MIT Press, Cambridge, MA, 2012.

[351] N. G. Leveson and J. P. Thomas. *STPA Handbook*. Massachusetts Institute of Technology, 2018.

[352] D. Levin and L. Zhang. Bridging Level-K to Nash equilibrium, 2022. Accepted at The Review of Economics and Statistics.

[353] D. A. Levin and Y. Peres. *Markov Chains and Mixing Times*. American Mathematical Society, Providence, RI, 2 edition, 2017.

[354] T. Li, J. Guevara, X. Xie, and Q. Zhu. Self-confirming transformer for belief-conditioned adaptation in offline multi-agent reinforcement learning. In *Proceedings of the 17th Adaptive and Learning Agents Workshop (ALA)*, 2025. arXiv:2310.04579.

[355] T. Li, K. Hammar, R. Stadler, and Q. Zhu. Conjectural online learning with first-order beliefs in asymmetric information stochastic games. In *Proceedings of the 63rd IEEE Conference on Decision and Control (CDC)*, pages 6780–6785, 2024.

[356] T. Li, H. Li, Y. Pan, T. Xu, Z. Zheng, and Q. Zhu. Meta Stackelberg game: Robust federated learning against adaptive and mixed poisoning attacks, 2024. arXiv:2410.17431.

[357] T. Li, Y.-T. Yang, Y. Pan, and Q. Zhu. From texts to shields: Convergence of large language models and cybersecurity, 2025. arXiv:2505.00841.

[358] T. Li and Q. Zhu. Agentic AI for cyber resilience: A new security paradigm and its system-theoretic foundations, 2025. arXiv:2512.22883.

[359] T. Li and Q. Zhu. Symbiotic game and foundation models for cyber deception operations in strategic cyber warfare. In *Foundations of Cyber Deception*, Advances in Information Security, pages 25–62. Springer, Cham, 2026.

[360] T.-Y. Li and Z. Gajić. Lyapunov iterations for solving coupled algebraic Riccati equations of Nash differential games and algebraic Riccati equations of zero-sum games. In G. J. Olsder, editor, *New Trends in Dynamic Games and Applications*, pages 333–351. Birkhäuser, Boston, MA, 1995.

[361] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong. A review of false data injection attacks against modern power systems. *IEEE Transactions on Smart Grid*, 8(4):1630–1638, 2017.

[362] D. Liberzon. *Calculus of Variations and Optimal Control Theory: A Concise Introduction*. Princeton University Press, Princeton, NJ, 2012.

[363] F. Lieder and T. L. Griffiths. Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, 43:e1, 2020.

[364] I. Linkov and A. Kott. Fundamental concepts of cyber resilience: Introduction and overview. In A. Kott and I. Linkov, editors, *Cyber Resilience of Systems and Networks*, pages 1–25. Springer, Cham, 2019.

[365] M. L. Littman. Markov games as a framework for multi-agent reinforcement learning. In *Machine Learning Proceedings 1994*, pages 157–163. Morgan Kaufmann, 1994.

[366] S. Liu and Q. Zhu. Cyber insurance for cyber resilience, 2023. arXiv:2312.02921.

[367] S. Liu and Q. Zhu. Mitigating moral hazard in insurance contracts using risk preference design. *Operations Research Letters*, 62:107322, 2025.

[368] S. Liu and Q. Zhu. Stackelberg risk preference design. *Mathematical Programming*, 209(1–2):785–823, 2025.

[369] S. Liu and Q. Zhu. Cyber insurance for cyber resilience. In *Autonomous Cyber Resilience*, pages 139–182. Wiley, 2026.

[370] X. Liu, J. Zhang, H. Shang, S. Guo, C. Yang, and Q. Zhu. Exploring prosocial irrationality for LLM agents: A social cognition view. In *International Conference on Learning Representations*, 2025. ICLR 2025.

[371] Y. Liu, Y. Jia, R. Geng, J. Jia, and N. Z. Gong. Formalizing and benchmarking prompt injection attacks and defenses. In *33rd USENIX Security Symposium*. USENIX Association, 2024.

[372] Y. Liu, P. Ning, and M. K. Reiter. False data injection attacks against state estimation in electric power grids. In *Proceedings of the 16th ACM Conference on Computer and Communications Security (CCS)*, pages 21–32. ACM, 2009.

[373] A. Lizzeri. Information revelation and certification intermediaries. *The RAND Journal of Economics*, 30(2):214–231, 1999.

[374] J. Louis Anthony Cox. What's wrong with risk matrices? *Risk Analysis*, 28(2):497–512, 2008.

[375] Y. Luo, L. Gao, and J. Huang. Spectrum reservation contract design in TV white space networks. *IEEE Transactions on Cognitive Communications and Networking*, 1(2):147–160, 2015.

[376] K.-w. Lye and J. M. Wing. Game strategies in network security. *International Journal of Information Security*, 4(1–2):71–86, 2005.

[377] M. Lyu, J. Farooq, and Q. Zhu. Mapping cyber threats in the 5G supply chain: Landscape, vulnerabilities, and risk management. *IEEE Network*, 39(1):251–260, 2025.

[378] E. F. Mageirou. Values and strategies for infinite time linear quadratic games. *IEEE Transactions on Automatic Control*, 21(4):547–550, 1976.

[379] P. K. Manadhata and J. M. Wing. An attack surface metric. *IEEE Transactions on Software Engineering*, 37(3):371–386, 2011.

[380] M. H. Manshaei, Q. Zhu, T. Alpcan, T. Başar, and J.-P. Hubaux. Game theory meets network security and privacy. *ACM Computing Surveys*, 45(3):1–39, 2013.

[381] M. H. Manshaei, Q. Zhu, T. Alpcan, T. Başar, and J.-P. Hubaux. Game theory meets network security and privacy. *ACM Computing Surveys*, 45(3):25:1–25:39, 2013.

[382] E. Maskin and J. Tirole. Markov perfect equilibrium: I. Observable actions. *Journal of Economic Theory*, 100(2):191–219, 2001.

[383] R. Masum. A review on game theory with smart grid security. arXiv:2304.11738, 2023.

[384] F. Matějka and A. McKay. Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1):272–298, 2015.

[385] MATPOWER. MATPOWER: Free, open-source tools for electric power system simulation and optimization. `https://matpower.org/`, 2026. Accessed June 30, 2026.

[386] S. Mauw and M. Oostdijk. Foundations of attack trees. In *Information Security and Cryptology – ICISC 2005*, volume 3935 of *Lecture Notes in Computer Science*, pages 186–198, Berlin, Heidelberg, 2006. Springer.

[387] J. Maynard Smith and G. R. Price. The logic of animal conflict. *Nature*, 246(5427):15–18, 1973.

[388] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert. Constrained model predictive control: Stability and optimality. *Automatica*, 36(6):789–814, 2000.

[389] R. D. McKelvey and T. R. Palfrey. An experimental study of the centipede game. *Econometrica*, 60(4):803–836, 1992.

[390] R. D. McKelvey and T. R. Palfrey. Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10(1):6–38, 1995.

[391] R. D. McKelvey and T. R. Palfrey. Quantal response equilibria for extensive form games. *Experimental Economics*, 1(1):9–41, 1998.

[392] F. Miao, Q. Zhu, M. Pajic, and G. J. Pappas. A moving-horizon hybrid stochastic game for secure control of cyber-physical systems. arXiv:1710.00288, 2017.

[393] E. Miehling, M. Rasouli, and D. Teneketzis. A POMDP approach to the dynamic defense of large-scale cyber networks. *IEEE Transactions on Information Forensics and Security*, 13(10):2490–2505, 2018.

[394] J. A. Mirrlees. An exploration in the theory of optimum income taxation. *The Review of Economic Studies*, 38(2):175–208, 1971.

[395] T. J. Misa. Computer security discourse at RAND, SDC, and NSA (1958–1970). *IEEE Annals of the History of Computing*, 38(4):12–25, 2016.

[396] MITRE. MITRE ATT&CK. `https://attack.mitre.org/`, 2026. Accessed June 11, 2026.

[397] MITRE Corporation. MITRE ATT&CK: Adversarial Tactics, Techniques, and Common Knowledge, 2024. Knowledge base.

[398] MITRE Corporation. MITRE ATLAS: Adversarial threat landscape for artificial-intelligence systems. `https://atlas.mitre.org/`, 2026. Accessed 2026-11-10.

[399] MITRE Corporation. MITRE ATT&CK for Industrial Control Systems. `https://attack.mitre.org/matrices/ics/`, 2026. Accessed June 30, 2026.

[400] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533, 2015.

[401] Y. Mo, R. Chabukswar, and B. Sinopoli. Detecting integrity attacks on SCADA systems. *IEEE Transactions on Control Systems Technology*, 22(4):1396–1407, 2014.

[402] Y. Mo and B. Sinopoli. Secure control against replay attacks. In *Proceedings of the 47th Annual Allerton Conference on Communication, Control, and Computing*, pages 911–918. IEEE, 2009.

[403] D. Monderer and L. S. Shapley. Fictitious play property for games with identical interests. *Journal of Economic Theory*, 68(1):258–265, 1996.

[404] D. Monderer and L. S. Shapley. Potential games. *Games and Economic Behavior*, 14(1):124–143, 1996.

[405] A. Mukherjee, S. A. A. Fakoorian, J. Huang, and A. L. Swindlehurst. Principles of physical layer security in multiuser wireless networks: A survey. *IEEE Communications Surveys & Tutorials*, 16(3):1550–1573, 2014.

[406] A. S. Musleh, G. Chen, and Z. Y. Dong. A survey on the detection algorithms for false data injection attacks in smart grids. *IEEE Transactions on Smart Grid*, 11(3):2218–2234, 2020.

[407] R. B. Myerson. Graphs and cooperation in games. *Mathematics of Operations Research*, 2(3):225–229, 1977.

[408] R. B. Myerson. Incentive compatibility and the bargaining problem. *Econometrica*, 47(1):61–73, 1979.

[409] R. B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981.

[410] R. B. Myerson. *Game Theory: Analysis of Conflict*. Harvard University Press, 1991.

[411] R. B. Myerson. Population uncertainty and Poisson games. *International Journal of Game Theory*, 27(3):375–392, 1998.

[412] R. B. Myerson. Large Poisson games. *Journal of Economic Theory*, 94(1):7–45, 2000.

[413] R. B. Myerson. Comments on Harsanyi's papers on games with incomplete information. *International Journal of Game Theory*, 33(2):59–73, 2004.

[414] R. Nagel. Unraveling in guessing games: An experimental study. *American Economic Review*, 85(5):1313–1326, 1995.

[415] A. Nagurney, P. Daniele, and S. Shukla. A supply chain network game theory model of cybersecurity investments with nonlinear budget constraints. *Annals of Operations Research*, 248(1):405–427, 2017.

[416] S. Nakamoto. Bitcoin: A peer-to-peer electronic cash system. `https://bitcoin.org/bitcoin.pdf`, 2008.

[417] J. Nash. Equilibrium points in *n*-person games. *Proceedings of the National Academy of Sciences*, 36(1):48–49, 1950.

[418] J. Nash. Non-cooperative games. *Annals of Mathematics*, 54(2):286–295, 1951.

[419] National Institute of Standards and Technology. Standards for security categorization of federal information and information systems. FIPS PUB 199, NIST, 2004.

[420] National Institute of Standards and Technology. Information security continuous monitoring for federal information systems and organizations. Special Publication 800-137, NIST, 2011.

[421] National Institute of Standards and Technology. Managing information security risk. Special Publication 800-39, NIST, 2011.

[422] National Institute of Standards and Technology. Guide for conducting risk assessments. Special Publication 800-30 Revision 1, NIST, 2012.

[423] National Institute of Standards and Technology. Integrating cybersecurity and enterprise risk management. NISTIR 8286, NIST, 2020.

[424] National Institute of Standards and Technology. Security and privacy controls for information systems and organizations. Special Publication 800-53 Revision 5, NIST, 2020.

[425] National Institute of Standards and Technology. IoT device cybersecurity guidance for the Federal Government. NIST Special Publication 800-213, National Institute of Standards and Technology, 2021.

[426] National Institute of Standards and Technology. Artificial intelligence risk management framework (AI RMF 1.0). NIST AI 100-1, National Institute of Standards and Technology, 2023.

[427] National Institute of Standards and Technology. Guide to operational technology security. Special Publication 800-82 Revision 3, NIST, 2023.

[428] National Institute of Standards and Technology. Artificial intelligence risk management framework: Generative artificial intelligence profile. NIST AI 600-1, National Institute of Standards and Technology, 2024.

[429] National Institute of Standards and Technology. National Vulnerability Database, 2024. Vulnerability database.

[430] National Institute of Standards and Technology. The NIST cybersecurity framework (CSF) 2.0. NIST Cybersecurity White Paper NIST CSWP 29, National Institute of Standards and Technology, 2024.

[431] National Telecommunications and Information Administration. The minimum elements for a software bill of materials (SBOM), 2021.

[432] A. Nemirovski, A. Juditsky, G. Lan, and A. Shapiro. Robust stochastic approximation approach to stochastic programming. *SIAM Journal on Optimization*, 19(4):1574–1609, 2009.

[433] Y. Nesterov. *Lectures on Convex Optimization*, volume 137 of *Springer Optimization and Its Applications*. Springer, Cham, 2 edition, 2018.

[434] Z. Newman, J. S. Meyers, and S. Torres-Arias. Sigstore: Software signing for everybody. In *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security (CCS)*, pages 2353–2367, 2022.

[435] A. Neyman and S. Sorin, editors. *Stochastic Games and Applications*, volume 570 of *NATO Science Series C: Mathematical and Physical Sciences*. Springer, Dordrecht, 2003.

[436] A. T. Nguyen and Q. Zhu. Detection-control games under hidden modes: Resilience-induced blindness phenomenon. *IEEE Control Systems Letters*, 2026. To appear; arXiv:2606.29080.

[437] A. T. Nguyen and Q. Zhu. Divergence-based safety measure for large language models via rational inattention, 2026. arXiv:2606.29081.

[438] A. T. Nguyen, Q. Zhu, and A. Teixeira. Bilateral cognitive security games in networked control systems under stealthy injection attacks. In *Proceedings of the 64th IEEE Conference on Decision and Control (CDC)*, pages 7305–7312, 2025.

[439] K. C. Nguyen, T. Alpcan, and T. Başar. Stochastic games for security in networks with interdependent nodes. In *Proceedings of the International Conference on Game Theory for Networks (GameNets)*, pages 697–703. IEEE, 2009.

[440] V.-L. Nguyen, P.-C. Lin, B.-C. Cheng, R.-H. Hwang, and Y.-D. Lin. Security and privacy for 6G: A survey on prospective technologies and challenges. *IEEE Communications Surveys & Tutorials*, 23(4):2384–2428, 2021.

[441] N. Nisan, T. Roughgarden, É. Tardos, and V. V. Vazirani, editors. *Algorithmic Game Theory*. Cambridge University Press, Cambridge, 2007.

[442] NIST. Cyber-physical systems (cps). `https://csrc.nist.gov/topics/applications/cyber-physical-systems`, 2016. Computer Security Resource Center, NIST (accessed 2025-06).

[443] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, New York, 2 edition, 2006.

[444] J. R. Norris. *Markov Chains*. Cambridge University Press, Cambridge, 1998.

[445] Y. Nugraha, A. Cetinkaya, T. Hayakawa, H. Ishii, and Q. Zhu. Dynamic resilient network games with applications to multiagent consensus. *IEEE Transactions on Control of Network Systems*, 8(1):246–259, 2021.

[446] Y. E. Nugraha, T. Hayakawa, H. Ishii, A. Cetinkaya, and Q. Zhu. Perfect Bayesian equilibria of two-player games in resilient multiagent systems. *Dynamic Games and Applications*, 16(1):198–219, 2026.

[447] C. Okafor, T. R. Schorlemmer, S. Torres-Arias, and J. C. Davis. SoK: Analysis of software supply chain security by establishing secure design properties. In *Proceedings of the ACM Workshop on Software Supply Chain Offensive Research and Ecosystem Defenses (SCORED)*, pages 15–24, 2022.

[448] Open Source Security Foundation. Supply-chain levels for software artifacts (SLSA) specification v1.0, 2023.

[449] M. J. Osborne. *An Introduction to Game Theory*. Oxford University Press, New York, 2004.

[450] M. J. Osborne and A. Rubinstein. *A Course in Game Theory*. MIT Press, Cambridge, MA, 1994.

[451] X. Ou, W. F. Boyer, and M. A. McQueen. A scalable approach to attack graph generation. In *Proceedings of the 13th ACM Conference on Computer and Communications Security*, pages 336–345, 2006.

[452] OWASP Foundation. Agentic AI: Threats and mitigations. `https://genai.owasp.org/resource/agentic-ai-threats-and-mitigations/`, 2025. OWASP GenAI Security Project, Version 1.0; accessed June 11, 2026.

[453] OWASP Foundation. OWASP Top 10 for Large Language Model Applications and Generative AI 2025. `https://genai.owasp.org/llm-top-10/`, 2025. Accessed June 11, 2026.

[454] Y. Pan, J. Li, L. Xu, S. Sun, and Q. Zhu. A game-theoretic approach for high-resolution automotive FMCW radar interference avoidance. In *Proceedings of the IEEE International Radar Conference (RADAR)*, pages 1–6, 2025.

[455] Y. Pan, J. Li, L. Xu, S. Sun, and Q. Zhu. Decentralized no-regret frequency–time scheduling for FMCW radar interference avoidance. *IEEE Transactions on Aerospace and Electronic Systems*, 62:17218–17234, 2026.

[456] Y. Pan, T. Li, and Q. Zhu. On the variational interpretation of mirror play in monotone games. In *Proceedings of the 63rd IEEE Conference on Decision and Control (CDC)*, pages 6799–6804, 2024.

[457] Y. Pan, T. Li, and Q. Zhu. Model-agnostic Hessian-free meta-policy optimization via zeroth-order estimation: A linear quadratic regulator perspective. *Dynamic Games and Applications*, 16(1):386–425, 2026.

[458] Y. Pan, T. Li, and Q. Zhu. A unified variational design of predictive mirror descent in convex games under stochastic feedback. *IEEE Control Systems Letters*, 10:1759–1764, 2026.

[459] Y. Pan and Q. Zhu. Bayesian holonic systems: Equilibrium, uniqueness, and computation, 2025. arXiv:2512.18112.

[460] Y. Pan and Q. Zhu. Extending no-regret hopping in FMCW radar interference avoidance. *ACM SIGMETRICS Performance Evaluation Review*, 53(2):131–133, 2025.

[461] Y. Pan and Q. Zhu. A games-in-games paradigm for strategic hybrid jump-diffusions: Hamilton–Jacobi–Isaacs hierarchy and spectral structure. *Nonlinear Analysis: Hybrid Systems*, 2026. To appear; arXiv:2512.18098.

[462] Y. Pan and Q. Zhu. Timing-aware two-player stochastic games with self-triggered control. In *Proceedings of the 23rd IFAC World Congress*, Busan, Republic of Korea, 2026. arXiv:2512.18109.

[463] Z. Pan and T. Başar. H_∞-control of Markovian jump systems and solutions to associated piecewise-deterministic differential games. In G. J. Olsder, editor, *New Trends in Dynamic Games and Applications*, pages 61–94. Birkhäuser, Boston, MA, 1995.

[464] G. P. Papavassilopoulos, J. V. Medanić, and J. B. Cruz. On the existence of Nash strategies and solutions to coupled Riccati equations in linear-quadratic games. *Journal of Optimization Theory and Applications*, 28(1):49–76, 1979.

[465] N. Papernot, P. McDaniel, A. Sinha, and M. P. Wellman. SoK: Security and privacy in machine learning. In *2018 IEEE European Symposium on Security and Privacy (EuroS&P)*, pages 399–414. IEEE, 2018.

[466] P. S. Park, S. Goldstein, A. O'Gara, M. Chen, and D. Hendrycks. Ai deception: A survey of examples, risks, and potential solutions, 2023.

[467] T. Parthasarathy, S. H. Tijs, and O. J. Vrieze. Stochastic games with state independent transitions and separable rewards. In G. Hammer and D. Pallaschke, editors, *Selected Topics in Operations Research and Mathematical Economics*, volume 226 of *Lecture Notes in Economics and Mathematical Systems*, pages 262–271. Springer, Berlin, 1984.

[468] P. Paruchuri, J. P. Pearce, J. Marecki, M. Tambe, F. Ordóñez, and S. Kraus. Playing games with security: An efficient exact algorithm for bayesian stackelberg games. In *Proceedings of the 7th International Conference on Autonomous Agents and Multiagent Systems*, pages 895–902, 2008.

[469] F. Pasqualetti, F. Dörfler, and F. Bullo. Attack detection and identification in cyber-physical systems. *IEEE Transactions on Automatic Control*, 58(11):2715–2729, 2013.

[470] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani. Epidemic processes in complex networks. *Reviews of Modern Physics*, 87(3):925–979, 2015.

[471] R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *Physical Review Letters*, 86(14):3200–3203, 2001.

[472] R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *Physical Review Letters*, 86(14):3200–3203, 2001.

[473] A. Pavan, I. Segal, and J. Toikka. Dynamic mechanism design: A Myersonian approach. *Econometrica*, 82(2):601–653, 2014.

[474] J. Pawlick, J. Chen, and Q. Zhu. iSTRICT: An interdependent strategic trust mechanism for the cloud-enabled internet of controlled things. *IEEE Transactions on Information Forensics and Security*, 14(6):1654–1669, 2019.

[475] J. Pawlick, E. Colbert, and Q. Zhu. A game-theoretic taxonomy and survey of defensive deception for cybersecurity and privacy. *ACM Computing Surveys*, 52(4):1–28, 2019.

[476] J. Pawlick, E. Colbert, and Q. Zhu. Modeling and analysis of leaky deception using signaling games with evidence. *IEEE Transactions on Information Forensics and Security*, 14(7):1871–1886, 2019.

[477] J. Pawlick and Q. Zhu. Deception by design: Evidence-based signaling games for network defense. arXiv:1503.05458, 2015.

[478] J. Pawlick and Q. Zhu. Strategic trust in cloud-enabled cyber-physical systems with an application to glucose control. *IEEE Transactions on Information Forensics and Security*, 12(12):2906–2919, 2017.

[479] J. Pawlick and Q. Zhu. *Game Theory for Cyber Deception: From Theory to Applications*. Static & Dynamic Game Theory: Foundations & Applications. Springer, Cham, 2021.

[480] J. Perdomo, T. Zrnic, C. Mendler-Dünner, and M. Hardt. Performative prediction. In H. Daumé III and A. Singh, editors, *Proceedings of the 37th International Conference on Machine Learning (ICML)*, volume 119 of *Proceedings of Machine Learning Research*, pages 7599–7609. PMLR, 2020.

[481] C. Phillips and L. P. Swiler. A graph-based system for network-vulnerability analysis. In *Proceedings of the 1998 Workshop on New Security Paradigms*, pages 71–79, 1998.

[482] R. Píbil, V. Lisý, C. Kiekintveld, B. Bošanský, and M. Pěchouček. Game theoretic model of strategic honeypot selection in computer networks. In J. Grossklags and J. Walrand, editors, *Decision and Game Theory for Security (GameSec 2012)*, volume 7638 of *Lecture Notes in Computer Science*, pages 201–220, Berlin, Heidelberg, 2012. Springer.

[483] J. Pita, M. Jain, J. Marecki, F. Ordóñez, C. Portway, M. Tambe, C. Western, P. Paruchuri, and S. Kraus. Deployed ARMOR protection: The application of a game-theoretic model for security at the Los Angeles International Airport. In *Security and Game Theory*, pages 67–87. Cambridge University Press, Cambridge, 2011.

[484] M. Polese, L. Bonati, S. D'Oro, S. Basagni, and T. Melodia. Understanding O-RAN: Architecture, interfaces, algorithms, security, and research challenges. *IEEE Communications Surveys & Tutorials*, 25(2):1376–1411, 2023.

[485] N. Poolsappasit, R. Dewri, and I. Ray. Dynamic security risk management using Bayesian attack graphs. *IEEE Transactions on Dependable and Secure Computing*, 9(1):61–74, 2012.

[486] D. Prelec. The probability weighting function. *Econometrica*, 66(3):497–527, 1998.

[487] G. Provan and Q. Zhu. An inspection game for monitor design against GNSS timing attacks in synchronous islanding, 2026. arXiv:2611.09331.

[488] P. Przymus and T. Durieux. Wolves in the repository: A software engineering analysis of the XZ Utils supply chain attack. In *Proceedings of the IEEE/ACM 22nd International Conference on Mining Software Repositories (MSR)*, 2025.

[489] M. L. Psiaki and T. E. Humphreys. GNSS spoofing and detection. *Proceedings of the IEEE*, 104(6):1258–1270, 2016.

[490] M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley, 1994.

[491] A. Ramdas, P. Grünwald, V. Vovk, and G. Shafer. Game-theoretic statistics and safe anytime-valid inference. *Statistical Science*, 38(4):576–601, 2023.

[492] T. Rashid, M. Samvelyan, C. Schroeder de Witt, G. Farquhar, J. Foerster, and S. Whiteson. QMIX: Monotonic value function factorisation for deep multi-agent reinforcement learning. In J. Dy and A. Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 4295–4304. PMLR, 2018.

[493] S. Rass, S. König, and E. Panaousis. Cut-the-rope: A game of stealthy intrusion. In *Decision and Game Theory for Security*, volume 11836 of *Lecture Notes in Computer Science*, pages 404–416. Springer, Cham, 2019.

[494] S. Rass, S. Schauer, S. König, and Q. Zhu. *Cyber-Security in Critical Infrastructures: A Game-Theoretic Approach*. Advanced Sciences and Technologies for Security Applications. Springer, Cham, 2020.

[495] S. Rass, S. Schauer, S. König, and Q. Zhu. *Cybersicherheit in kritischen Infrastrukturen: Ein spieltheoretischer Zugang*. Springer Vieweg, Cham, 2025.

[496] S. Rass and Q. Zhu. GADAPT: A sequential game-theoretic framework for designing defense-in-depth strategies against advanced persistent threats. In *Decision and Game Theory for Security (GameSec 2016)*, Lecture Notes in Computer Science, pages 314–326, Cham, 2016. Springer.

[497] A. S. Rawat, P. Anand, H. Chen, and P. K. Varshney. Collaborative spectrum sensing in the presence of Byzantine attacks in cognitive radio networks. *IEEE Transactions on Signal Processing*, 59(2):774–786, 2011.

[498] J. B. Rawlings, D. Q. Mayne, and M. M. Diehl. *Model Predictive Control: Theory, Computation, and Design*. Nob Hill Publishing, Madison, WI, 2nd edition, 2017.

[499] S. M. Rinaldi, J. P. Peerenboom, and T. K. Kelly. Identifying, understanding, and analyzing critical infrastructure interdependencies. *IEEE Control Systems Magazine*, 21(6):11–25, 2001.

[500] H. Robbins and S. Monro. A stochastic approximation method. *The Annals of Mathematical Statistics*, 22(3):400–407, 1951.

[501] H. Robbins and D. Siegmund. A convergence theorem for non negative almost supermartingales and some applications. In J. S. Rustagi, editor, *Optimizing Methods in Statistics*, pages 233–257. Academic Press, New York, 1971.

[502] J. Robinson. An iterative method of solving a game. *Annals of Mathematics*, 54(2):296–301, 1951.

[503] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, NJ, 1970.

[504] R. T. Rockafellar and S. Uryasev. Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26(7):1443–1471, 2002.

[505] S. Rose, O. Borchert, S. Mitchell, and S. Connelly. Zero trust architecture. NIST Special Publication 800-207, National Institute of Standards and Technology, 2020.

[506] J. B. Rosen. Existence and uniqueness of equilibrium points for concave N-person games. *Econometrica*, 33(3):520–534, 1965.

[507] R. Ross, V. Pillitteri, R. Graubart, D. Bodeau, and R. McQuaid. Developing cyber-resilient systems: A systems security engineering approach. NIST Special Publication 800-160 Vol. 2 Rev. 1, National Institute of Standards and Technology, 2021.

[508] S. M. Ross. *Introduction to Probability Models*. Academic Press, London, 12th edition, 2019.

[509] M. Rothschild and J. Stiglitz. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *The Quarterly Journal of Economics*, 90(4):629–649, 1976.

[510] S. Roy, C. Ellis, S. Shiva, D. Dasgupta, V. Shandilya, and Q. Wu. A survey of game theory as applied to network security. In *Proceedings of the 43rd Hawaii International Conference on System Sciences*, pages 1–10, 2010.

[511] S. Roy, S. U. Kadir, Y. Vorobeychik, and A. Laszka. Strategic remote attestation: Testbed for Internet-of-Things devices and Stackelberg security game for optimal strategies. In *Decision and Game Theory for Security (GameSec 2021)*, volume 13061 of *Lecture Notes in Computer Science*, pages 271–290. Springer, 2021.

[512] A. Rubinstein. Finite automata play the repeated prisoner's dilemma. *Journal of Economic Theory*, 39(1):83–96, 1986.

[513] W. Saad, M. Bennis, and M. Chen. A vision of 6G wireless systems: Applications, trends, technologies, and open research problems. *IEEE Network*, 34(3):134–142, 2020.

[514] W. Saad, A. Sanjab, Y. Wang, C. A. Kamhoua, and K. A. Kwiat. Hardware Trojan detection game: A prospect-theoretic approach. *IEEE Transactions on Vehicular Technology*, 66(9):7697–7710, 2017.

[515] J. H. Saltzer and M. D. Schroeder. The protection of information in computer systems. *Proceedings of the IEEE*, 63(9):1278–1308, 1975.

[516] H. Sandberg, S. Amin, and K. H. Johansson. Cyberphysical security in networked control systems: An introduction to the issue. *IEEE Control Systems Magazine*, 35(1):20–23, 2015.

[517] W. H. Sandholm. *Population Games and Evolutionary Dynamics*. MIT Press, 2010.

[518] J. Sanz, C.-Y. Xia, S. Meloni, and Y. Moreno. Dynamics of interacting diseases. *Physical Review X*, 4(4):041005, 2014.

[519] A. Sapirshtein, Y. Sompolinsky, and A. Zohar. Optimal selfish mining strategies in Bitcoin. In *Financial Cryptography and Data Security (FC 2016)*, volume 9603 of *Lecture Notes in Computer Science*, pages 515–532. Springer, 2017.

[520] T. C. Schelling. *The Strategy of Conflict*. Harvard University Press, Cambridge, MA, 1960.

[521] K. H. Schlag. Why imitate, and if so, how? A boundedly rational approach to multi-armed bandits. *Journal of Economic Theory*, 78(1):130–156, 1998.

[522] B. Schneier. Attack trees: Modeling security threats. *Dr. Dobb's Journal*, 24(12):21–29, 1999.

[523] G. Scutari, D. P. Palomar, F. Facchinei, and J.-S. Pang. Convex optimization, game theory, and variational inequality theory. *IEEE Signal Processing Magazine*, 27(3):35–49, 2010.

[524] R. Selten. Spieltheoretische behandlung eines oligopolmodells mit nachfragetragheit. *Zeitschrift für die gesamte Staatswissenschaft*, 121:301–324 and 667–689, 1965.

[525] R. Selten. Bounded rationality. *Journal of Institutional and Theoretical Economics*, 146(4):649–658, 1990.

[526] S. Sengupta, A. Chowdhary, A. Sabur, A. Alshamrani, D. Huang, and S. Kambhampati. A survey of moving target defenses for network security. *IEEE Communications Surveys & Tutorials*, 22(3):1909–1941, 2020.

[527] C. E. Shannon and W. Weaver. *The Mathematical Theory of Communication*. University of Illinois Press, Urbana, 1949.

[528] L. S. Shapley. Stochastic games. *Proceedings of the national academy of sciences*, 39(10):1095–1100, 1953.

[529] L. S. Shapley. Stochastic games. *Proceedings of the National Academy of Sciences*, 39(10):1095–1100, 1953.

[530] L. S. Shapley. Some topics in two-person games. In M. Dresher, L. S. Shapley, and A. W. Tucker, editors, *Advances in Game Theory*, number 52 in Annals of Mathematics Studies, pages 1–28. Princeton University Press, Princeton, NJ, 1964.

[531] O. Sheyner, J. Haines, S. Jha, R. Lippmann, and J. M. Wing. Automated generation and analysis of attack graphs. In *Proceedings of the 2002 IEEE Symposium on Security and Privacy*, pages 273–284, 2002.

[532] Y. Shoham and K. Leyton-Brown. *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge University Press, 2008.

[533] P. Shukla, L. An, A. Chakrabortty, and A. Duel-Hallen. A robust stackelberg game for cyber-security investment in networked control systems. arXiv:2103.16611, 2021.

[534] M. Simaan and J. B. Cruz. On the Stackelberg strategy in nonzero-sum games. *Journal of Optimization Theory and Applications*, 11(5):533–555, 1973.

[535] H. A. Simon. A behavioral model of rational choice. *Quarterly Journal of Economics*, 69(1):99–118, 1955.

[536] H. A. Simon. Rational choice and the structure of the environment. *Psychological Review*, 63(2):129–138, 1956.

[537] H. A. Simon. *Models of Man: Social and Rational*. Wiley, New York, 1957.

[538] H. A. Simon. *Administrative Behavior*. Free Press, New York, 4th edition, 1997.

[539] J. Simon and A. Omar. Cybersecurity investments in the supply chain: Coordination and a strategic attacker. *European Journal of Operational Research*, 282(1):161–171, 2020.

[540] C. A. Sims. Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690, 2003.

[541] C. A. Sims. Rational inattention: Beyond the linear-quadratic case. *American Economic Review*, 96(2):158–163, 2006.

[542] S. K. Singh and Q. Zhu. Consistent conjectural approach to adversarial intent tracking under sensing constraints in multi-target defense differential games. In *Game Theory and AI for Security (GameSec 2025)*, Lecture Notes in Computer Science, pages 292–312. Springer, Cham, 2025.

[543] A. Sinha, F. Fang, B. An, C. Kiekintveld, and M. Tambe. Stackelberg security games: Looking beyond a decade of success. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18)*, pages 5494–5501, 2018.

[544] J. Slay and M. Miller. Lessons learned from the Maroochy water breach. In E. Goetz and S. Shenoi, editors, *Critical Infrastructure Protection*, IFIP International Federation for Information Processing, pages 73–82. Springer, Boston, MA, 2008.

[545] S. Soltan, P. Mittal, and H. V. Poor. BlackIoT: IoT botnet of high wattage devices can disrupt the power grid. In *Proceedings of the 27th USENIX Security Symposium*, pages 15–32, 2018.

[546] M. Souppaya, K. Scarfone, and D. Dodson. Secure software development framework (SSDF) version 1.1. NIST Special Publication 800-218, National Institute of Standards and Technology, 2022.

[547] J. C. Spall. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Transactions on Automatic Control*, 37(3):332–341, 1992.

[548] A. M. Spence. Job market signaling. *Quarterly Journal of Economics*, 87(3):355–374, 1973.

[549] D. O. Stahl and P. W. Wilson. Experimental evidence on players' models of other players. *Journal of Economic Behavior and Organization*, 25:309–327, 1994.

[550] D. O. Stahl and P. W. Wilson. On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior*, 10(1):218–254, 1995.

[551] W. Stallings. *Network Security Essentials: Applications and Standards*. Pearson, Boston, 6 edition, 2017.

[552] A. W. Starr and Y. C. Ho. Further properties of nonzero-sum differential games. *Journal of Optimization Theory and Applications*, 3(4):207–219, 1969.

[553] A. W. Starr and Y. C. Ho. Nonzero-sum differential games. *Journal of Optimization Theory and Applications*, 3(3):184–206, 1969.

[554] F. J. Stech, K. E. Heckman, and B. E. Strom. Integrating cyber-d&d into adversary modeling for active cyber defense. In S. Jajodia, V. S. Subrahmanian, V. Swarup, and C. Wang, editors, *Cyber Deception: Building the Scientific Foundation*, pages 1–22. Springer, 2016.

[555] L. A. Stole. Lectures on the theory of contracts. Lecture notes, 1999. First version 1993; current version September 1999.

[556] K. Stouffer, M. Pease, C. Tang, T. Zimmerman, V. Pillitteri, S. Lightman, A. Hahn, S. Saravia, A. Sherule, and M. Thompson. Guide to operational technology (ot) security. Special Publication 800-82 Rev. 3, National Institute of Standards and Technology, Gaithersburg, MD, 2023.

[557] X. Sun. Lecture notes on game theory: Theory and examples. Lecture notes, 2018.

[558] R. S. Sutton and A. G. Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

[559] C. Szepesvári and M. L. Littman. A unified analysis of value-function-based reinforcement-learning algorithms. *Neural Computation*, 11(8):2017–2060, 1999.

[560] M. Tambe. *Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned*. Cambridge University Press, Cambridge, 2011.

[561] A. Tanwani and Q. Zhu. Feedback Nash equilibrium for randomly switching differential-algebraic games. *IEEE Transactions on Automatic Control*, 65(8):3286–3301, 2020.

[562] A. Tarski. A lattice-theoretical fixpoint theorem and its applications. *Pacific Journal of Mathematics*, 5(2):285–309, 1955.

[563] P. D. Taylor and L. B. Jonker. Evolutionary stable strategies and game dynamics. *Mathematical Biosciences*, 40(1–2):145–156, 1978.

[564] M. Tehranipoor and F. Koushanfar. A survey of hardware Trojan taxonomy and detection. *IEEE Design & Test of Computers*, 27(1):10–25, 2010.

[565] A. Teixeira, G. Dán, H. Sandberg, and K. H. Johansson. A cyber security study of a SCADA energy management system: Stealthy deception attacks on the state estimator. arXiv:1011.1828, 2010.

[566] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson. A secure control framework for resource-limited adversaries. *Automatica*, 51:135–148, 2015.

[567] H. Tembine, Q. Zhu, and T. Başar. Risk-sensitive mean-field games. *IEEE Transactions on Automatic Control*, 59(4):835–850, 2014.

[568] The MITRE Corporation. MITRE ATT&CK. `https://attack.mitre.org/`, 2020. Accessed June 14, 2026.

[569] The Open Group. Open FAIR: Factor analysis of information risk, 2015. Risk taxonomy and analysis standard.

[570] S. Torres-Arias, H. Afzali, T. K. Kuppusamy, R. Curtmola, and J. Cappos. in-toto: Providing farm-to-table guarantees for bits and bytes. In *Proceedings of the 28th USENIX Security Symposium*, pages 1393–1410, 2019.

[571] W. Tushar, C. Yuen, T. K. Saha, S. Nizami, M. R. Alam, D. B. Smith, and H. V. Poor. A survey of cyber-physical systems from a game-theoretic perspective. *IEEE Access*, 11:9799–9834, 2023.

[572] A. Tversky and D. Kahneman. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323, 1992.

[573] UK Parliament. Product Security and Telecommunications Infrastructure Act 2022. c. 46, 2022.

[574] D. Umsonst, S. Sarıtaş, G. Dán, and H. Sandberg. A bayesian nash equilibrium-based moving target defense against stealthy sensor attacks. arXiv:2111.06682, 2021.

[575] U.S. Code of Federal Regulations. 47 CFR section 96.39: Citizens Broadband Radio Service device general requirements, 2026. Accessed 2026.

[576] K. G. Vamvoudakis, F. Fotiadis, T. Başar, V. Gupta, J. Poveda, M. Tang, M. Krstić, and Q. Zhu. Deception in game theory and control: A tutorial. In *Proceedings of the American Control Conference (ACC)*, pages 2146–2163, 2025.

[577] M. van Dijk, A. Juels, A. Oprea, and R. L. Rivest. FlipIt: The game of "stealthy takeover". *Journal of Cryptology*, 26(4):655–713, 2013.

[578] H. van Hasselt, A. Guez, and D. Silver. Deep reinforcement learning with double Q-learning. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)*, volume 30, pages 2094–2100. AAAI Press, 2016.

[579] P. Van Mieghem, J. Omic, and R. Kooij. Virus spread in networks. *IEEE/ACM Transactions on Networking*, 17(1):1–14, 2009.

[580] R. Vane. Advances in hypergame theory. In *Proceedings of the AAMAS Workshop on Game Theoretic and Decision Theoretic Agents*, pages 1–8, 2006.

[581] R. Vane and P. E. Lehner. Using hypergames to select plans in adversarial environments. In *Proceedings of the 1st Workshop on Game Theoretic and Decision Theoretic Agents*, pages 103–111, 1999.

[582] L. R. Varshney and X. Wu. A graph-based Stackelberg security game for trustworthy 6G disaggregated architecture. *arXiv preprint arXiv:2609.29500*, 2026.

[583] A. Vassilev, A. Oprea, A. Fordyce, H. Anderson, X. Davies, and M. Hamin. Adversarial machine learning: A taxonomy and terminology of attacks and mitigations. NIST Trustworthy and Responsible AI NIST AI 100-2e2025, National Institute of Standards and Technology, Gaithersburg, MD, 2025.

[584] W. Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of Finance*, 16(1):8–37, 1961.

[585] J. von Neumann. Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen*, 100:295–320, 1928.

[586] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, NJ, 1944.

[587] H. von Stackelberg. *Marktform und Gleichgewicht*. Julius Springer, Vienna, 1934.

[588] A. Wald. Sequential tests of statistical hypotheses. *The Annals of Mathematical Statistics*, 16(2):117–186, 1945.

[589] C. Wang and Z. Lu. Cyber deception: Overview and the road ahead. *IEEE Security & Privacy*, 16(2):80–85, 2018.

[590] L. Wang, C. Ma, X. Feng, Z. Zhang, H. Yang, J. Zhang, Z. Chen, J. Tang, X. Chen, Y. Lin, W. X. Zhao, Z. Wei, and J. Wen. A survey on large language model based autonomous agents. *Frontiers of Computer Science*, 18(6):186345, 2024.

[591] X. Wang, J. Wei, D. Schuurmans, Q. Le, E. Chi, S. Narang, A. Chowdhery, and D. Zhou. Self-consistency improves chain of thought reasoning in language models. *International Conference on Learning Representations*, 2023.

[592] Z. Wang, T. Schaul, M. Hessel, H. van Hasselt, M. Lanctot, and N. de Freitas. Dueling network architectures for deep reinforcement learning. In M. F. Balcan and K. Q. Weinberger, editors, *Proceedings of the 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pages 1995–2003, New York, NY, USA, 2016. PMLR.

[593] J. G. Wardrop. Some theoretical aspects of road traffic research. *Proceedings of the Institution of Civil Engineers*, 1(3):325–362, 1952.

[594] W. H. Ware. Security controls for computer systems: Report of defense science board task force on computer security. Technical Report R-609-1, RAND Corporation, 1979. Originally published for the Department of Defense in 1970.

[595] A. Washburn and K. Wood. Two-person zero-sum games for network interdiction. *Operations Research*, 43(2):243–251, 1995.

[596] C. J. C. H. Watkins and P. Dayan. Q-learning. *Machine Learning*, 8(3–4):279–292, 1992.

[597] J. Wei, X. Wang, D. Schuurmans, M. Bosma, B. Ichter, F. Xia, E. Chi, Q. Le, and D. Zhou. Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, 35:24824–24837, 2022.

[598] D. Williams. *Probability with Martingales*. Cambridge University Press, Cambridge, 1991.

[599] T. J. Williams. The Purdue enterprise reference architecture. *Computers in Industry*, 24(2–3):141–158, 1994.

[600] X. Wu, J. Farooq, and T. Li. Performative security games: Strategic manipulation of decision-dependent distributions in learning-based O-RAN control. In *Decision and Game Theory for Security (GameSec 2026)*, volume 17097 of *Lecture Notes in Computer Science*, pages 541–560. Springer, 2027.

[601] Y. Wu, B. Wang, K. J. R. Liu, and T. C. Clancy. Anti-jamming games in multi-channel cognitive radio networks. *IEEE Journal on Selected Areas in Communications*, 30(1):4–15, 2012.

[602] L. Xiao, T. Chen, J. Liu, and H. Dai. Anti-jamming transmission Stackelberg game with observation errors. *IEEE Communications Letters*, 19(6):949–952, 2015.

[603] Z. Xu and Q. Zhu. A cyber-physical game framework for secure and resilient multi-agent autonomous systems. In *Proceedings of the IEEE Conference on Decision and Control*, pages 5157–5162, 2015.

[604] J.-P. A. Yaacoub, O. Salman, H. N. Noura, N. Kaaniche, A. Chehab, and M. Malli. Cyber-physical systems security: Limitations, issues and future trends. *Microprocessors and Microsystems*, 77:103201, 2020.

[605] Y.-T. Yang, Y. Ge, and Q. Zhu. Human risks and cognition-inspired adaptive cyber deception. In *Foundations of Cyber Deception*, Advances in Information Security, pages 181–200. Springer, Cham, 2026.

[606] Y.-T. Yang, H. Lei, and Q. Zhu. PRADA: Proactive risk assessment and mitigation of misinformed demand attacks on navigational route recommendations. *IEEE Transactions on Information Forensics and Security*, 20:10879–10892, 2025.

[607] Y.-T. Yang, T. Li, and Q. Zhu. Transparent tagging for strategic social nudges on user-generated misinformation. *IEEE Transactions on Network Science and Engineering*, 13:1147–1161, 2026.

[608] Y.-T. Yang, T. Zhang, and Q. Zhu. Herd accountability of privacy-preserving algorithms: A Stackelberg game approach. *IEEE Transactions on Information Forensics and Security*, 20:2237–2251, 2025.

[609] Y.-T. Yang and Q. Zhu. Bi-level game-theoretic planning of cyber deception for cognitive arbitrage, 2025. arXiv:2509.05498.

[610] Y.-T. Yang and Q. Zhu. A multi-resolution dynamic game framework for cross-echelon decision-making in cyber warfare. In *Game Theory and AI for Security (GameSec 2025)*, Lecture Notes in Computer Science, pages 335–355. Springer, Cham, 2025.

[611] Y.-T. Yang and Q. Zhu. PACT: A contract-theoretic framework for pricing agentic AI services powered by large language models. In *Proceedings of the IEEE Global Communications Conference (GLOBECOM)*, pages 5677–5682, 2025.

[612] Y.-T. Yang and Q. Zhu. Toward a multi-echelon cyber warfare theory: A meta-game-theoretic paradigm for defense and dominance, 2025. arXiv:2509.08976.

[613] Y.-T. Yang and Q. Zhu. Cognitive sensing and control for arbitrage: A game-theoretic framework, 2026. Preprint.

[614] Y.-T. Yang and Q. Zhu. Internet of agentic AI: Incentive-compatible distributed teaming and workflow. In *Proceedings of the 24th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, pages 1–8, 2026.

[615] Y.-T. Yang and Q. Zhu. When to deceive: A cross-layer Stackelberg game framework for strategic timing of cyber deception. In *Proceedings of the 17th Conference on Game Theory and AI for Security (GameSec)*, 2026. To appear; arXiv:2505.21244.

[616] S. Yao, D. Yu, J. Zhao, I. Shafran, T. L. Griffiths, Y. Cao, and K. Narasimhan. Tree of thoughts: Deliberate problem solving with large language models. *Advances in Neural Information Processing Systems*, 2023.

[617] Y. Yao, J. Duan, K. Xu, Y. Cai, Z. Sun, and Y. Zhang. A survey on large language model (LLM) security and privacy: The good, the bad, and the ugly. *High-Confidence Computing*, 4(2):100211, 2024.

[618] D. Ye, M. Shi, and Z. Sun. Satellite proximate pursuit-evasion game with different thrust configurations. *Aerospace Science and Technology*, 99:105715, 2020.

[619] K. Zhang, Z. Yang, and T. Başar. Multi-agent reinforcement learning: A selective overview of theories and algorithms. In K. G. Vamvoudakis, Y. Wan, F. L. Lewis, and D. Cansever, editors, *Handbook of Reinforcement Learning and Control*, volume 325 of *Studies in Systems, Decision and Control*, pages 321–384. Springer, Cham, 2021.

[620] L. Zhang and Q. Zhu. ActivityRDI: A centralized solution framework for activity retrieval and detection intelligence based on knowledge graphs, large language models, and imbalanced learning. *Machine Learning and Knowledge Extraction*, 8(3):75, 2026.

[621] R. Zhang and Q. Zhu. Attack-aware cyber insurance of interdependent computer networks. Working Paper 16-18, NET Institute, Sept. 2016. Available at SSRN: 2848576.

[622] R. Zhang and Q. Zhu. A game-theoretic approach to design secure and resilient distributed support vector machines. *IEEE Transactions on Neural Networks and Learning Systems*, 29(11):5512–5527, 2018.

[623] R. Zhang and Q. Zhu. FlipIn: A game-theoretic cyber insurance framework for incentive-compatible cyber risk management of internet of things. *IEEE Transactions on Information Forensics and Security*, 15:2026–2041, 2020.

[624] R. Zhang, Q. Zhu, and Y. Hayel. A bi-level game approach to attack-aware cyber insurance of computer networks. *IEEE Journal on Selected Areas in Communications*, 35(3):779–794, 2017.

[625] T. Zhang, L. Huang, J. Pawlick, and Q. Zhu. Game-theoretic analysis of cyber deception: Evidence-based strategies and dynamic risk mitigation, 2019.

[626] Y. Zhao, J. Chen, Y. Lu, and Q. Zhu. Mean field control by stochastic Koopman operator via a spectral method, 2024. arXiv:2411.06180.

[627] Y. Zhao, J. Chen, and Q. Zhu. Integrated cyber-physical resiliency for power grids under IoT-enabled dynamic botnet attacks. *IEEE Transactions on Control Systems Technology*, 2024. Preprint arXiv:2401.01963.

[628] Y. Zhao, Y. Ge, and Q. Zhu. Combating ransomware in Internet of Things: A games-in-games approach for cross-layer cyber defense and security investment. In *Decision and Game Theory for Security (GameSec 2021)*, Lecture Notes in Computer Science, pages 208–228, Cham, 2021. Springer.

[629] Y. Zhao, L. Huang, C. Smidts, and Q. Zhu. Finite-horizon semi-Markov game for time-sensitive attack response and probabilistic risk assessment in nuclear power plants. *Reliability Engineering & System Safety*, 201:106878, 2020.

[630] Y. Zhao, C. Rieger, and Q. Zhu. Multi-agent learning for resilient distributed control systems. In *Power Grid Resilience*, Power Electronics and Power Systems, pages 425–458. Springer, Cham, 2025.

[631] Y. Zhao and Q. Zhu. Combating online counterfeits: A game-theoretic analysis of cyber supply chain ecosystem. In *Decision and Game Theory for Security (GameSec 2020)*, volume 12513 of *Lecture Notes in Computer Science*, pages 326–345. Springer, 2020.

[632] Y. Zhao and Q. Zhu. Autonomous and resilient control for optimal LEO satellite constellation coverage against space threats. arXiv:2203.02050, 2022.

[633] Y. Zhao and Q. Zhu. Distributed and resilient planning-control for optimal LEO satellite constellation coverage. In *Proceedings of the American Control Conference (ACC)*, pages 1841–1846, 2022.

[634] K. Zhou, J. C. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice-Hall, Upper Saddle River, NJ, 1996.

[635] Q. Zhu. Game theory for cyber deception: A tutorial. In *Proceedings of the 6th Annual Symposium on Hot Topics in the Science of Security (HotSoS)*, pages 1–3. ACM, 2019.

[636] Q. Zhu. Multilayer cyber-physical security and resilience for smart grid. In J. Stoustrup, A. Annaswamy, A. Chakrabortty, and Z. Qu, editors, *Smart Grid Control: Overview and Research Opportunities*, Power Electronics and Power Systems, pages 225–239. Springer, Cham, 2019.

[637] Q. Zhu. Foundations of cyber resilience: The confluence of game, control, and learning theories. In I. Linkov and A. Kott, editors, *Cyber Resilience: Applied Perspectives*, Risk, Systems and Decisions, pages 27–58. Springer, Cham, 2025.

[638] Q. Zhu. Game theory meets LLM and agentic AI: Reimagining cybersecurity for the age of intelligent threats, 2025. arXiv:2507.10621.

[639] Q. Zhu. Generative-conjectural LLM equilibrium for agentic AI deception with applications to spearphishing. In J. S. Baras, S. Papavassiliou, E. E. Tsiropoulou, and M. O. Sayin, editors, *Game Theory and AI for Security (GameSec 2025)*, Lecture Notes in Computer Science, pages 356–375, Cham, 2025. Springer.

[640] Q. Zhu. LLM-Stackelberg games: Conjectural reasoning equilibria and their applications to spearphishing, 2025. arXiv:2507.09407v1.

[641] Q. Zhu. Reasoning and behavioral equilibria in LLM-Nash games: From mindsets to actions, 2025. arXiv:2507.08208v1.

[642] Q. Zhu. Agentomics: Economic foundations for the valuation, attribution, and pricing of AI agents in human-AI workflows, 2026. arXiv:2606.14769.

[643] Q. Zhu. Berk–Nash equilibrium and learning for satellite orbital pursuit–evasion games. In *Proceedings of the 23rd IFAC World Congress*, Busan, Republic of Korea, 2026.

[644] Q. Zhu. Calibrated learning of sequential equilibrium in signaling games. In *Proceedings of the 65th IEEE Conference on Decision and Control (CDC)*, 2026. To appear.

[645] Q. Zhu. Exit-and-join dynamics for decentralized coalition formation, 2026. arXiv:2606.19683.

[646] Q. Zhu. From optimal transport to optimal quantization: A variational study of the Witsenhausen counterexample, 2026. arXiv:2609.18992.

[647] Q. Zhu. Game Petri nets: Compositional modeling and analysis of large-scale strategic systems, 2026. arXiv:2610.12779.

[648] Q. Zhu. Insurance of agentic AI, 2026. arXiv:2606.05449.

[649] Q. Zhu. The Internet of agentic AI: Communication, coordination, and collective intelligence at scale, 2026. arXiv:2606.12835.

[650] Q. Zhu. Internet of agentic things: Networked AI agents for closed-loop IoT orchestration, 2026. arXiv:2607.12662.

[651] Q. Zhu. LLM semantic signaling game and mechanism design: Systematic blindness, awareness shaping, and mindset dynamics, 2026. arXiv:2606.29113.

[652] Q. Zhu. Neuronic Nash equilibrium: An EEG data-driven game-theoretic framework for BCI-enabled multi-agent behaviors. *Automation*, 7(3):78, 2026.

[653] Q. Zhu. Notes on rational inattention, 2026. Unpublished lecture notes, April 7, 2026.

[654] Q. Zhu. Performative Shapley allocation for endogenous network formation. In *Proceedings of the IFAC Workshop on Cyber-Physical and Human Systems (CPHS)*, 2026. To appear.

[655] Q. Zhu and T. Başar. Dynamic policy-based IDS configuration. In *Proceedings of the 48th IEEE Conference on Decision and Control held jointly with the 28th Chinese Control Conference*, pages 8600–8605. IEEE, 2009.

[656] Q. Zhu and T. Başar. Robust and resilient control design for cyber-physical systems with an application to power systems. In *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pages 4066–4071. IEEE, 2011.

[657] Q. Zhu and T. Başar. A dynamic game-theoretic approach to resilient control system design for cascading failures. In *Proceedings of the 1st International Conference on High Confidence Networked Systems (HiCoNS)*, pages 41–46. ACM, 2012.

[658] Q. Zhu and T. Başar. Game-theoretic approach to feedback-driven multi-stage moving target defense. In *Decision and Game Theory for Security*, volume 8252 of *Lecture Notes in Computer Science*, pages 246–263. Springer, 2013.

[659] Q. Zhu and T. Başar. Multi-resolution large population stochastic differential games and their application to demand response management in the smart grid. *Dynamic Games and Applications*, 3(1):68–88, 2013.

[660] Q. Zhu and T. Başar. Game-theoretic methods for robustness, security, and resilience of cyberphysical control systems: Games-in-games principle for optimal cross-layer resilient control systems. *IEEE Control Systems*, 35(1):46–65, 2015.

[661] Q. Zhu and T. Başar. Game-theoretic methods for robustness, security, and resilience of cyberphysical control systems: Games-in-games principle for optimal cross-layer resilient control systems. *IEEE Control Systems Magazine*, 35(1):46–65, 2015.

[662] Q. Zhu and T. Başar. Disentangling resilience from robustness: Contextual dualism, interactionism, and game-theoretic paradigms. *IEEE Control Systems Magazine*, 44(3):95–103, 2024.

[663] Q. Zhu and T. Başar. Revisiting game-theoretic control in socio-technical networks: Emerging design frameworks and contemporary applications. *IEEE Control Systems Letters*, 9:74–89, 2025.

[664] Q. Zhu and T. Başar. A multistage extension of Witsenhausen's counterexample with linear signaling, 2026. arXiv:2611.11181.

[665] Q. Zhu, A. Clark, R. Poovendran, and T. Başar. Deceptive routing games. In *Proceedings of the 51st IEEE Conference on Decision and Control*, pages 2704–2711. IEEE, 2012.

[666] Q. Zhu, C. Fung, R. Boutaba, and T. Başar. GUIDEX: A game-theoretic incentive-based mechanism for intrusion detection networks. *IEEE Journal on Selected Areas in Communications*, 30(11):2220–2230, 2012.

[667] Q. Zhu and Z. Han. Learning, misspecification, and cognitive arbitrage in linear-quadratic network games. In *Proceedings of the International Symposium on Mathematical Theory of Networks and Systems (MTNS)*, 2026. arXiv:2603.17157.

[668] Q. Zhu and Z. Han. A mathematical programming approach to computing and learning Berk–Nash equilibria in infinite-horizon MDPs, 2026. arXiv:2603.13641.

[669] Q. Zhu and Z. Han. Performative scenario optimization. In *Proceedings of the 65th IEEE Conference on Decision and Control (CDC)*, 2026. To appear; arXiv:2603.29982.

[670] Q. Zhu and Z. Han. Split-merge dynamics for Shapley-fair coalition formation. In *Proceedings of the International Symposium on Mathematical Theory of Networks and Systems (MTNS)*, 2026. arXiv:2603.17153.

[671] Q. Zhu, H. Li, Z. Han, and T. Başar. A stochastic game model for jamming in multi-channel cognitive radio systems. In *Proceedings of the IEEE International Conference on Communications (ICC)*, 2010.

[672] Q. Zhu, Z. Lu, P. L. Yu, and C. Wang, editors. *Foundations of Cyber Deception: Modeling, Analysis, Design, and Human Factors*. Advances in Information Security. Springer, Cham, 2026.

[673] Q. Zhu and S. Rass. On multi-phase and multi-stage game-theoretic modeling of advanced persistent threats. *IEEE Access*, 6:13958–13971, 2018.

[674] Q. Zhu, S. Rass, B. Dieber, and V. Mayoral Vilches. Cybersecurity in robotics: Challenges, quantitative modeling, and practice. *Foundations and Trends in Robotics*, 9(1):1–129, 2021.

[675] Q. Zhu, W. Saad, Z. Han, H. V. Poor, and T. Başar. Eavesdropping and jamming in next-generation wireless networks: A game-theoretic approach. In *Proceedings of the IEEE Military Communications Conference (MILCOM)*, pages 119–124, 2011.

[676] Q. Zhu, J. B. Song, and T. Başar. Dynamic secure routing game in distributed cognitive radio networks. In *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, pages 1–6, 2011.

[677] Q. Zhu, H. Tembine, and T. Başar. Heterogeneous learning in zero-sum stochastic games with incomplete information. In *Proceedings of the 49th IEEE Conference on Decision and Control*, pages 219–224. IEEE, 2010.

[678] Q. Zhu, H. Tembine, and T. Başar. Network security configurations: A nonzero-sum stochastic game approach. In *Proceedings of the 2010 American Control Conference*, pages 1059–1064. IEEE, 2010.

[679] Q. Zhu, H. Tembine, and T. Başar. Distributed strategic learning with application to network security. In *Proceedings of the 2011 American Control Conference*, pages 4057–4062. IEEE, 2011.

[680] Q. Zhu and Z. Xu. *Cross-Layer Design for Secure and Resilient Cyber-Physical Systems: A Decision and Game Theoretic Approach*. Advances in Information Security. Springer, Cham, 2020.

Strategic reasoning for securing the systems where the cyber and physical worlds meet.

Every security decision changes the incentives of an adversary who is watching and adapting. This book develops game theory as the language for reasoning about that interdependence, and applies it to the cyber-physical systems on which power grids, factories, vehicles, and communication networks depend.

Part I builds the foundations: normal- and extensive-form games, games of incomplete information, Markov games, and differential games. Part II develops learning in games, network games, mechanism design, the composition of games, and bounded rationality. Part III applies these models to industrial control systems, cyber deception, the Internet of Things, supply chains, and next-generation wireless networks.

Assuming only linear algebra, calculus, and elementary probability, the book is written for graduate students in engineering, computer science, and operations research; for researchers entering the security literature from control, networking, machine learning, or economics; and for practitioners who design or evaluate security policy.

Features

- Sixteen chapters in three parts, from first principles to applications
- Worked examples, each closing with the modeling lesson it teaches
- Learning objectives, exercises, and annotated further reading in every chapter
- Five appendices: optimization, probability, control, network security, cyber risk