A Perspective on AI and Security

Ravi Sandhu
Ram Krishnan

Cisco Research
Generative AI and Security Summit
October 3, 2023
The Big Picture
Security and AI

AI Attackers vs Human Defenders Hopeless
The Big Picture
Security and AI

AI Attackers vs Human Defenders  Hopeless

AI Attackers vs AI Defenders  Only Hope
The Big Picture

Security and AI

AI Attackers vs Human Defenders

Hopeless

AI Attackers vs AI Defenders

Only Hope

Traditional Position: attackers have asymmetric advantage

Fundamental Challenge: *how to flip the asymmetric advantage*
Asymmetric Attacker Advantage

➢ Traditional argument:
  Attackers need to exploit ONE vulnerability
  Defenders need to defend ALL weaknesses including ZERO DAY ones
Asymmetric Attacker Advantage

➢ Traditional argument:
   Attackers need to exploit ONE vulnerability
   Defenders need to defend ALL weaknesses including ZERO DAY ones

➢ Modern AI argument:
   Good AI is about good DATA and good TRAINING

   Assumption: TRAINING is symmetric and confers no benefit to either side (beyond the traditional asymmetry argument)

Corollary: Asymmetry impact will flow from good DATA
The Big Picture
Security and AI

Al Attackers vs Human Defenders
Hopeless

Al Attackers vs Al Defenders
Only Hope

Traditional Position: attackers have asymmetric advantage
Fundamental Challenge: how to flip the asymmetric advantage

Corollary: Cannot flip without good DATA for AI Defenders
The Good Data Recognition Problem

- First we have the defender’s data-poverty problem
- But even with data-abundance we will have the good-data-recognition problem
First we have the defender’s data-poverty problem

But even with data-abundance we will have the good-data-recognition problem

Brandolini's law (BS asymmetry principle, 2013):

The amount of energy needed to refute BS is an order of magnitude bigger than that needed to produce it.
The Good Data Recognition Problem

- First we have the defender’s data-poverty problem
- But even with data-abundance we will have the good-data-recognition problem

Brandolini's law (BS asymmetry principle, 2013:)
reject bad data
The amount of energy needed to refute BS is an order of magnitude bigger than that needed to produce it.
produce bad data
The Big Picture
Holistic Security

RESILIENCE
Assume Breach
Attack Aware
Measured Response

ZERO-TRUST
Beyond Static Deterministic Decisions
Dynamic Score-Based Decisions
Continuous Authorization
Obligations: Pre, Ongoing, Post

AI/AUTOMATION
Machine Speed
Machine Scale
Smart Escalation to Stakeholders
Rapid Policy Adjustment
Traditional Attack-Oblivious Multi-Layer Security

Goals
informal articulation

Policy
specification

Enforcement
architecture

Implementation
executable code
Traditional Attack-Oblivious Multi-Layer Security

**Goals**
- Informal articulation

**Policy**
- Specification

**Enforcement**
- Architecture

**Implementation**
- Executable code

- Minimize repeated authentication for legitimate users
- Enable one-hop lateral movement without authentication
- Configure firewall rules to authorize one-hop links
- Cache credentials to enable lateral movement without authentication
Attack-Aware Multi-Layer Security

**Goals**
informal articulation

**Policy**
specification

**Enforcement**
architecture

**Implementation**
executable code

- Minimize repeated authentication for legitimate users
- Enable one-hop lateral movement without authentication
- Configure firewall rules to authorize one-hop links
- Cache credentials to enable lateral one-hop moves without authentication

Attacker somehow acquires credentials for one user account
Attacker expands reach by harvesting cached credentials to move laterally
Attack-Resilient Multi-Layer Security

- **Goals**
  - informal articulation

- **Policy**
  - specification

- **Enforcement**
  - architecture

- **Implementation**
  - executable code

Minimize repeated authentication for legitimate users

Enable one-hop lateral movement without authentication

Configure firewall rules to authorize one-hop links

Cache credentials to enable lateral movement without authentication

Attacker somehow acquires credentials for one user account

Attacker expands reach by harvesting cached credentials to move laterally

**Measured response for resilience**
Multi-Layer Attack-Resilient Security (MLAR Security)

Goals (G)
informal articulation

Policy (P)
specification

Enforcement (E)
architecture

Implementation (I)
executable code

Need a holistic framework

- 3 players: Attackers, Defenders, Users
- Attacks exploit vulnerabilities at all layers
- Defenders defend/respond at all layers
- AI/Automation needed at all layers and cross-layer
Multi-Layer Attack-Resilient Security (MLAR Security)

Goals (G)
informal articulation

Policy (P)
specification

Enforcement (E)
arbitrary architecture

Implementation (I)
executable code

Need a holistic framework

- 3 players: Attackers, Defenders, Users
- Attacks exploit vulnerabilities at all layers
- Defenders defend/respond at all layers
- AI/Automation needed at all layers and cross-layer
- Existing literature focus is almost exclusively on the I layer
Asymmetric advantage to AI defenders requires solving:
- The data-poverty problem
- The good-data-recognition problem (even with data-abundance)

We lack a scientific discipline to engineer multi-layer attack-resilient cyber systems