Co-Evolving Attacker and Defender Agents for Engineered System Cybersecurity (CEADAESC)

Daniel Tauritz, Interim Director, Auburn Cyber Research Center

COMP 5660/6660
Auburn University

November 3, 2021

Part I: Engineered Systems & Security

What is an Engineered System?

NSF's Engineering Research Center website defines engineered systems as:

"a combination of components that work in synergy to collectively perform a useful function. The engineered system could, for example, wholly or in part constitute a new technology for a new product line a new manufacturing process, a technology to improve the delivery of a service, or an infrastructure system."

What is an Engineered System?

NSF's Engineering Research Center website defines engineered systems as:

> "a combination of components that work in synergy to collectively perform a useful function. The engineered system could, for example, wholly or in part constitute a new technology for a new product line a new manufacturing process, a technology to improve the delivery of a service, or an infrastructure system."

Examples:

- Modern Planes, Trains, and Automobiles
- Industry 4.0: Chemical Plant, Biotechnology, Agriculture
- Modern Utilities: Electric, Water, Gas, Oil
- Satellite Constellations (e.g., Starlink)
- Internet, Enterprise Computer Networks, Cloud Computing

Critical Infrastructure Sectors

DHS' Cybersecurity and Infrastructure Security Agency (CISA) lists 16 critical infrastructure sectors:

- Chemical
- Commercial Facilities
- Communications
- Critical Manufacturing
- Dams
- Defense Industrial Base
- Emergency Services
- Energy
- Financial Services
- Food and Agriculture
- Government Facilities
- Healthcare and Public Health
- Information Technology Sector
- Nuclear Reactors, Materials, and Waste
- Transportation Systems
- Water and Wastewater Systems

The Problem

- Modern engineered systems tend to be cyber/cyber-physical in nature

The Problem

- Modern engineered systems tend to be cyber/cyber-physical in nature
- Cyber & Cyber-physical engineered systems are extremely vulnerable to attack

The Problem

- Modern engineered systems tend to be cyber/cyber-physical in nature
- Cyber & Cyber-physical engineered systems are extremely vulnerable to attack
- Cyber & Cyber-physical engineered system attack surfaces tend to be astronomically large and infeasible to fully secure

The Problem

- Modern engineered systems tend to be cyber/cyber-physical in nature
- Cyber & Cyber-physical engineered systems are extremely vulnerable to attack
- Cyber & Cyber-physical engineered system attack surfaces tend to be astronomically large and infeasible to fully secure
- AI is needed to intelligently sample the combinatorially large number of unique attacks and defenses on modern engineered systems

The Problem

- Modern engineered systems tend to be cyber/cyber-physical in nature
- Cyber & Cyber-physical engineered systems are extremely vulnerable to attack
- Cyber & Cyber-physical engineered system attack surfaces tend to be astronomically large and infeasible to fully secure
- AI is needed to intelligently sample the combinatorially large number of unique attacks and defenses on modern engineered systems
- AI is needed to defend against AI which can attack faster than humans can respond

Engineered System Security as a Game

- Two or more adversaries: one defender and one or more attackers

Engineered System Security as a Game

- Two or more adversaries: one defender and one or more attackers
- Attackers range from so-called "script-kiddies" to organized crime, terrorist organizations, and adversarial nation states

Engineered System Security as a Game

- Two or more adversaries: one defender and one or more attackers
- Attackers range from so-called "script-kiddies" to organized crime, terrorist organizations, and adversarial nation states
- Attacker goals are diverse

Engineered System Security as a Game

- Two or more adversaries: one defender and one or more attackers
- Attackers range from so-called "script-kiddies" to organized crime, terrorist organizations, and adversarial nation states
- Attacker goals are diverse
- Defender needs to simulatenously defend against this wide variety of attackers

Engineered System Security as a Game

- Two or more adversaries: one defender and one or more attackers
- Attackers range from so-called "script-kiddies" to organized crime, terrorist organizations, and adversarial nation states
- Attacker goals are diverse
- Defender needs to simulatenously defend against this wide variety of attackers
- Asymetric non-zero sum game

Engineered System Security as a Game

- Two or more adversaries: one defender and one or more attackers
- Attackers range from so-called "script-kiddies" to organized crime, terrorist organizations, and adversarial nation states
- Attacker goals are diverse
- Defender needs to simulatenously defend against this wide variety of attackers
- Asymetric non-zero sum game
- Game theory allows for mathematical analysis of adversarial models

Engineered System Security as a Game

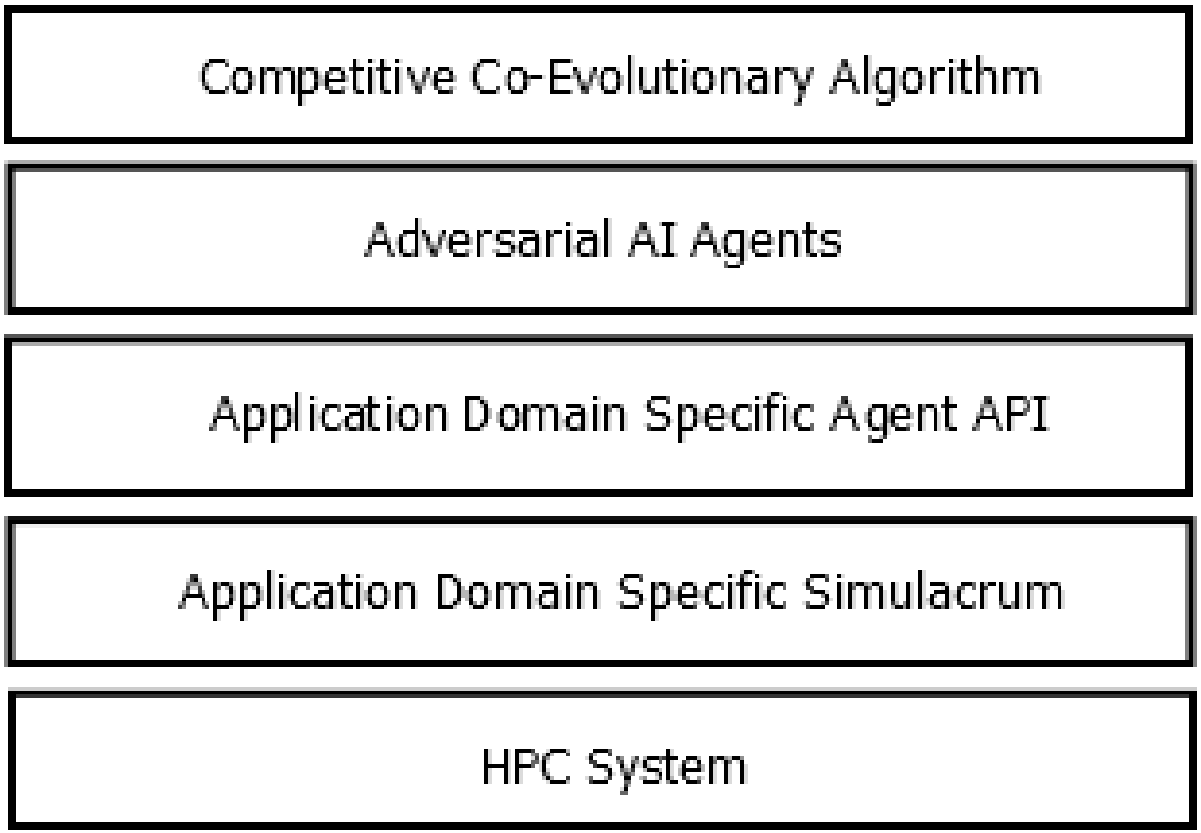
- Two or more adversaries: one defender and one or more attackers
- Attackers range from so-called "script-kiddies" to organized crime, terrorist organizations, and adversarial nation states
- Attacker goals are diverse
- Defender needs to simulatenously defend against this wide variety of attackers
- Asymetric non-zero sum game
- Game theory allows for mathematical analysis of adversarial models
- Classic game theory does not scale to complex, real-world systems

Engineered System Security as a Game

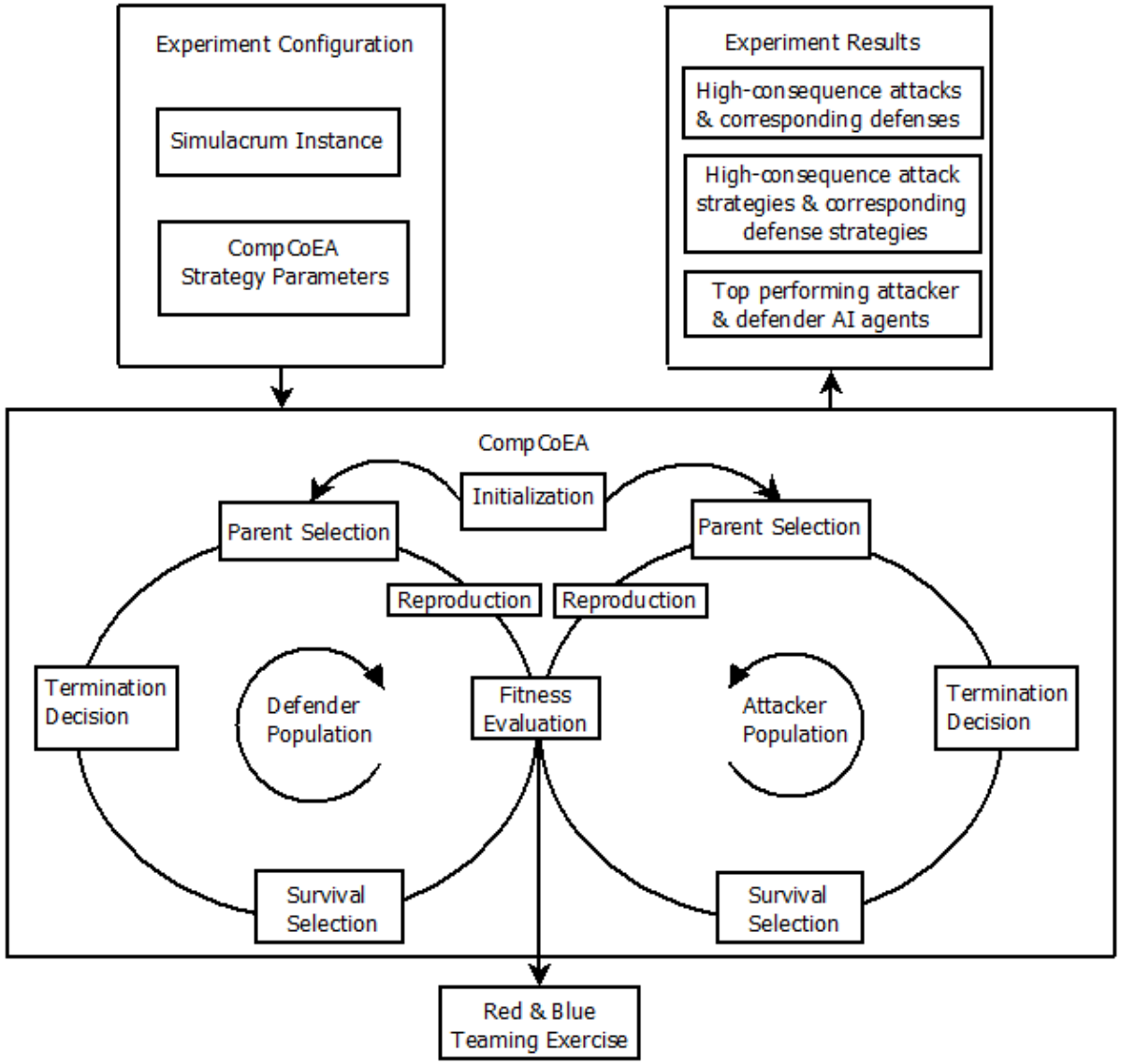
- Two or more adversaries: one defender and one or more attackers
- Attackers range from so-called "script-kiddies" to organized crime, terrorist organizations, and adversarial nation states
- Attacker goals are diverse
- Defender needs to simulatenously defend against this wide variety of attackers
- Asymetric non-zero sum game
- Game theory allows for mathematical analysis of adversarial models
- Classic game theory does not scale to complex, real-world systems
- Computational game theory achieves scalability by approximating Nash equilibria

Part II: Engineered System Security through AI Armsraces

CEADAESC system diagram



CEADAESC CompCoEA operation



Outcomes

Coevolving attacker and defender AI agents can produce three distinct capabilities:

Outcomes

Coevolving attacker and defender AI agents can produce three distinct capabilities:

Attacks & Defenses Automated identification of vulnerabilities and candidate mitigations that are already tested against a large set of attacks.

Outcomes

Coevolving attacker and defender AI agents can produce three distinct capabilities:

Attacks & Defenses Automated identification of vulnerabilities and candidate mitigations that are already tested against a large set of attacks.

Attack & Defense Strategies Automated wargaming in order to identify high-consequence attack strategies and corresponding defense strategies.

Outcomes

Coevolving attacker and defender AI agents can produce three distinct capabilities:

Attacks & Defenses Automated identification of vulnerabilities and candidate mitigations that are already tested against a large set of attacks.

Attack & Defense Strategies Automated wargaming in order to identify high-consequence attack strategies and corresponding defense strategies.

Attacker & Defender AI Agents Automated generation of highly-trained AI agents that can be deployed in live systems to augment human operators, or even autonomously engage in real-time with adversaries, both human and AI.

How to Apply CEADAESC to an Engineered System

- Create simulacrum

How to Apply CEADAESC to an Engineered System

- Create simulacrum
- Design representation for AI agent actions

How to Apply CEADAESC to an Engineered System

- Create simulacrum
- Design representation for AI agent actions
- Create AI controller logic including sensory inputs

How to Apply CEADAESC to an Engineered System

- Create simulacrum
- Design representation for AI agent actions
- Create AI controller logic including sensory inputs
- Define attacker & defender fitness functions

How to Apply CEADAESC to an Engineered System

- Create simulacrum
- Design representation for AI agent actions
- Create AI controller logic including sensory inputs
- Define attacker & defender fitness functions
- Execute AI Armsrace