What You Don’t Know Will Hurt You: The Relationship Between AI and Fraud Informatics

Dr. Neal Wagner

Fraud Informatics Symposium 2019

Cyber Attacks Are Ubiquitous in Government, Defense, Industry, and Critical Infrastructure

Gov't Attacks

"Cyber attacks originating from **Russia** compromised servers containing **voter data in 39 states** as well as servers at the DNC and impacted the **2016 Presidential Election**."

Defense Attacks

"**Impregnable Radar Breached** in Simulated Cyber-Attack, 10 Apr 2018. **Ethical hackers** sent a **virus-laden email** over a **naval ship's satellite link** to the captain's computer. The **virus transferred itself** to ECDIS and **then altered the vessel's position** during a night voyage, deceiving the officer of the watch."

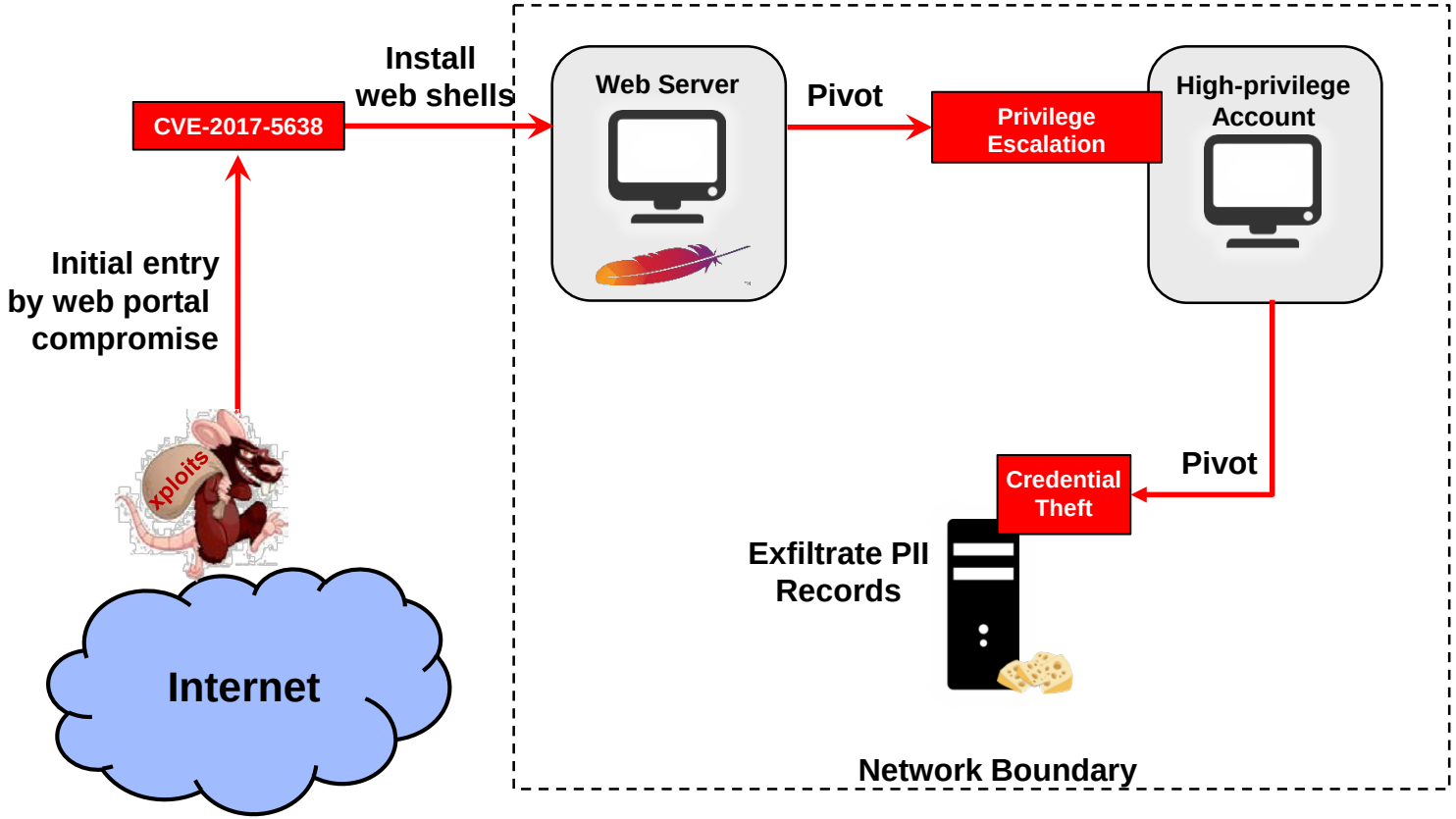
Industry Attacks

"**Trickbot** is a botnet that emerged in October 2016. Since its first appearance, it has been **targeting banks** mostly in Australia and the U.K, and in 2019, it has appeared in India, Singapore and Malesia as well."

Critical Infrastructure Attacks

"**Hackers** used firewall vulnerabilities to cause periodic **"blind spots"** for grid operators in the **western US** for about 10 hours on March 5, 2019. It's the **first known cyberattack** that has caused that kind of disruption **at a US power grid** company"

Cyber Attack Exemplar: Death by Lateral Movement



- Equifax observed suspicious traffic on 7/29/2017
  - Attacker had been in place since at least 5/13/2017
  - Exploited vulnerability to establish foothold
  - Spread laterally via allowed network communications
- Estimated cost: $1B+
- Attack duration: at least three months

Cyber breaches are commonplace in the modern, connected environment

Cyber Attack Exemplar: The Dreaded Insider

- Jiaqiang Xu – IBM software engineer
- Stole IBM source code
  - Obfuscated code's origin
- Attempted to sell the code to Chinese government agency
- Caught by undercover FBI agents posing as potential buyers
- Sentenced to 5 years in federal prison for economic espionage

Anatomy of an Insider Attack

But What Do Cyber Attacks Have To Do With Fraud?

- **Most businesses and other organizations rely on/utilize the Internet for operations**
- **Fraud via cyber means is rife!**
  - **Facilitated by the prevalence of organizations accessible via Internet**
  - **Low cost/effort required to execute a fraud attack**
- **Often difficult to track perpetrators**
  - **Criminals may be from other countries with few cybercrime laws or no extradition agreement**
  - **Can be a “no risk” crime**

Nigerian Email Scam

ADVERTISEMENT

CENTRAL BANK OF NIGERIA

PRESS STATEMENT ON ADVANCE FEE FRAUD/SCAM

DON'T BE FOOLED! MANY HAVE LOST MONEY!!

IF IT SOUNDS TOO GOOD TO BE TRUE THEN IT IS NOT TRUE!!!

1 The publicity campaigns by the Central Bank of Nigeria (CBN) and the Government of the Federal Republic of Nigeria have proved successful in sensitising the public about the menace of advance fee fraud and the falsehood of claims that easy money could be made in Nigeria. Consequently, the reported incidence of advance fee fraud (A.K.A. '419'), has declined significantly. Nevertheless, there are still some people who have continued to fall victim to the solicitations of advance fee fraudsters. This warning is, therefore, specifically intended for the benefit of those misguided people who, in the quest to make easy money at the expense of Nigeria, are defrauded by international fraudsters.

2 The advance fee fraud is perpetrated by enticing the victim with a bogus 'business' proposal which promises millions of US dollars as a reward. The scam

Sometimes, the 'victims' are invited to Nigeria where they are given red-carpet reception and attended to by the fraudsters, posing as Nigerian Government officials. Quite often the fraudsters invent bogus Government committees purported to have cleared the payments. Also, it is not unusual for them to contrive fake publications in the newspapers evidencing purported approvals to transfer non-existent funds.

4 To consummate the transaction, the 'victim' would be required to pay 'advance fees' for various purposes: e.g. processing fees, unforeseen taxes, licence fees, registration fees, signing/legal fees, fees for National Economic Recovery Fund, VAT, audit fees, insurance coverage fees, etc. The collection of these 'advance fees' is actually the real objective of the scam!

Chambers' of Commerce and Industry, Manufacturers' Associations of Nigeria, Federal Ministries of Commerce and Industry, Nigerian Missions in their countries of origin, their embassies or High Commissions in Nigeria for proper briefing and advice.

7 The Central Bank and indeed, the Federal Government of Nigeria cannot and should not be held responsible for bogus and shady deals transacted with criminal intentions. As a responsible corporate body, the Central Bank of Nigeria is once again warning all recipients of fraudulent letters on bogus deals, that there are no contract payments trapped in the bank's vaults. They are once again put on notice that all documents appertaining to the payment, claims, or transfers purportedly issued by the bank, its senior executives or the Government of the Federal Republic of Nigeria for the various purposes described above are all forgeries, bogus and fraudulent.

8 Please join the Central Bank and the Federal Government of Nigeria to fight the criminal syndicates who play on the gullibility and greed of their

GoFundMe Scam

Netflix Fake Sign In Scam

But What Do Cyber Attacks Have To Do With AI?

- **Whatever humans can do AI can often do faster, better, and/or with less effort**
- **AI is being used across all areas of society**
- **AI-enhanced cyber attacks are emergent**
  - **Botnet attacks**
  - **Malware signature morphing**
  - **“Smart” phishing email construction**
  - **User behavior learning and imitation**

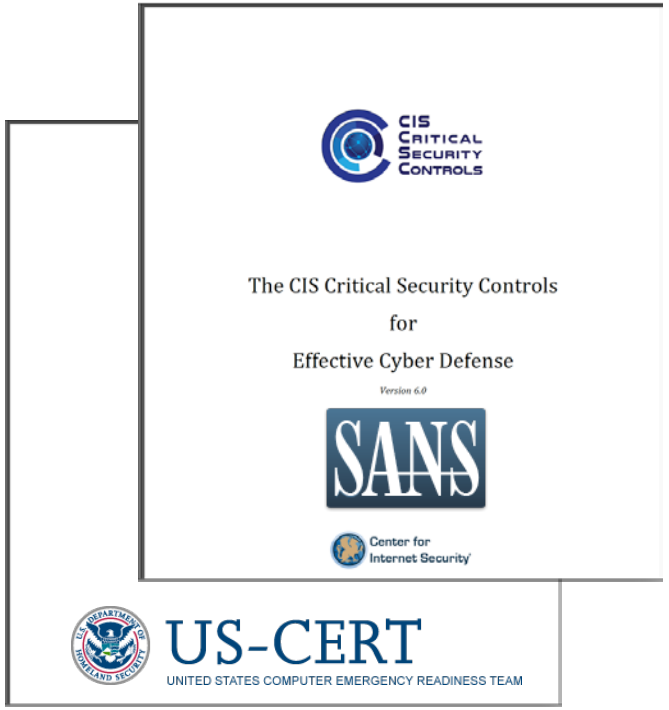
So We're All Doomed, Right?

- **Yep, that's right better just accept it…..**

- **No! If AI can be used for cyber offense, it can be used for cyber defense**
- **Commonly known offensive advantage: "Defender must defend entire system, attacker needs to find just one weakpoint"**
- **Less-commonly known defensive advantage: "Defender always has the home-court"**

Cyber Defense: Attack Mitigation

- **Recommended by:**



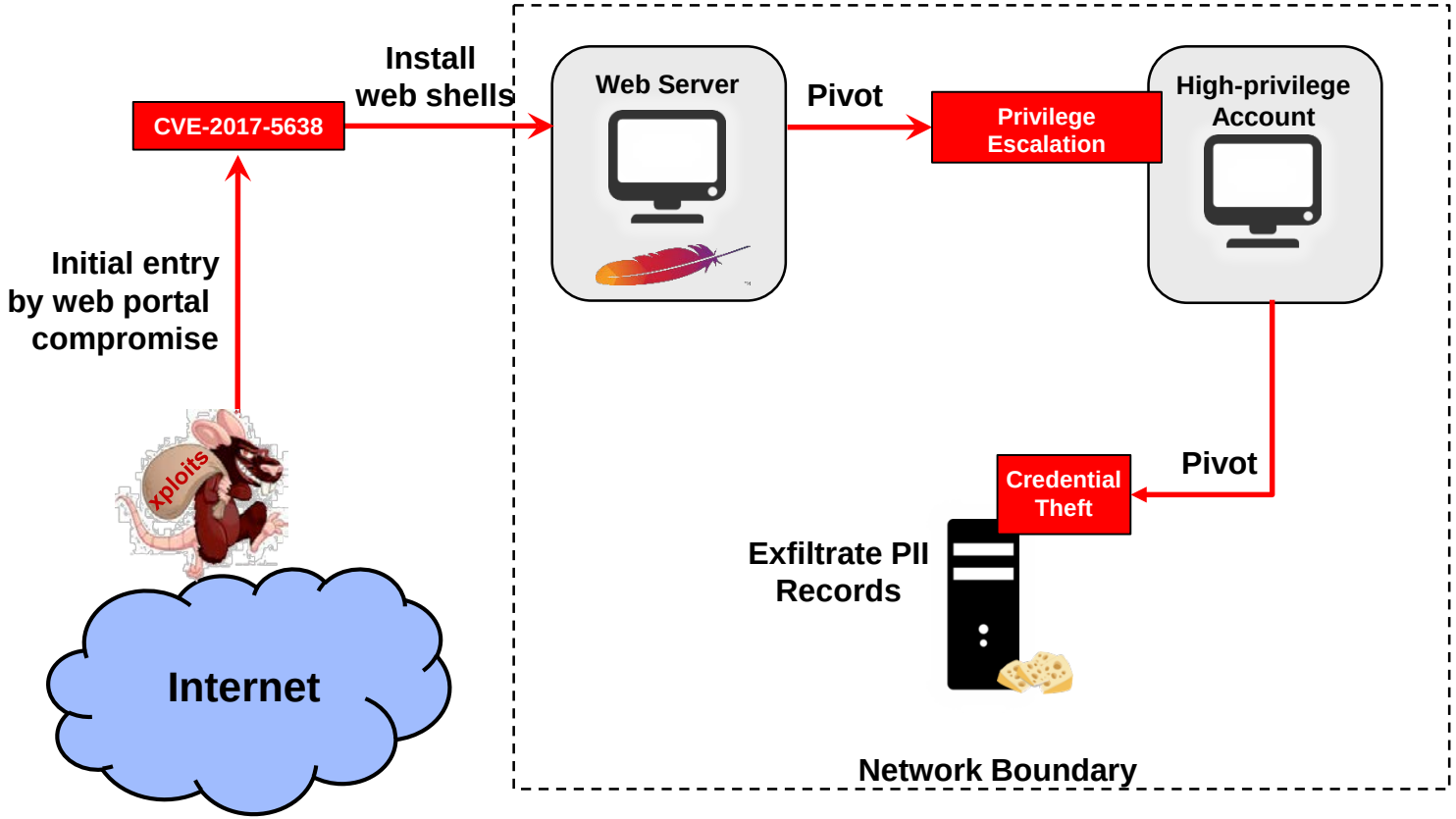
IAD's Top 10 Information Assurance Mitigation Strategies

- Anti-Exploitation (EMET)
- Limit Workstation-to-Workstation Comm.
- Application Whitelisting
- Host Intrusion Prevention System
- Network Segmentation
- Administrator Privilege Control
- Antivirus File Reputation Services
- Secure Baseline Configuration
- DNS Reputation
- Software Improvements

Mitigation: A tool, technique, control, or policy that serves to prevent or reduce the damage due to cyber attack

Cyber security authorities recommend top cyber risk mitigations

Cyber Attack Exemplar (Again): Death by Lateral Movement

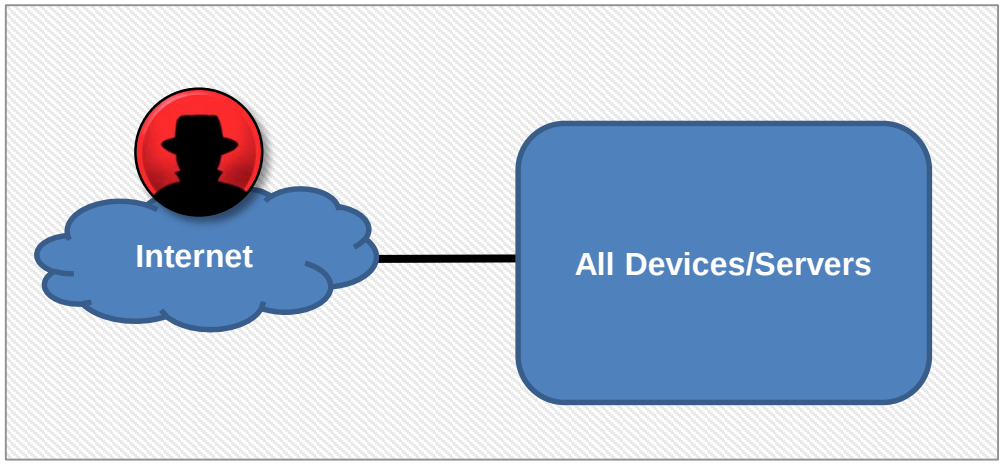


- Equifax observed suspicious traffic on 7/29/2017
  - Attacker had been in place since at least 5/13/2017
  - Exploited vulnerability to establish foothold
  - Spread laterally via allowed network communications
- Estimated cost: $1B+
- Attack duration: at least three months

Network segmentation is designed to make lateral movement difficult

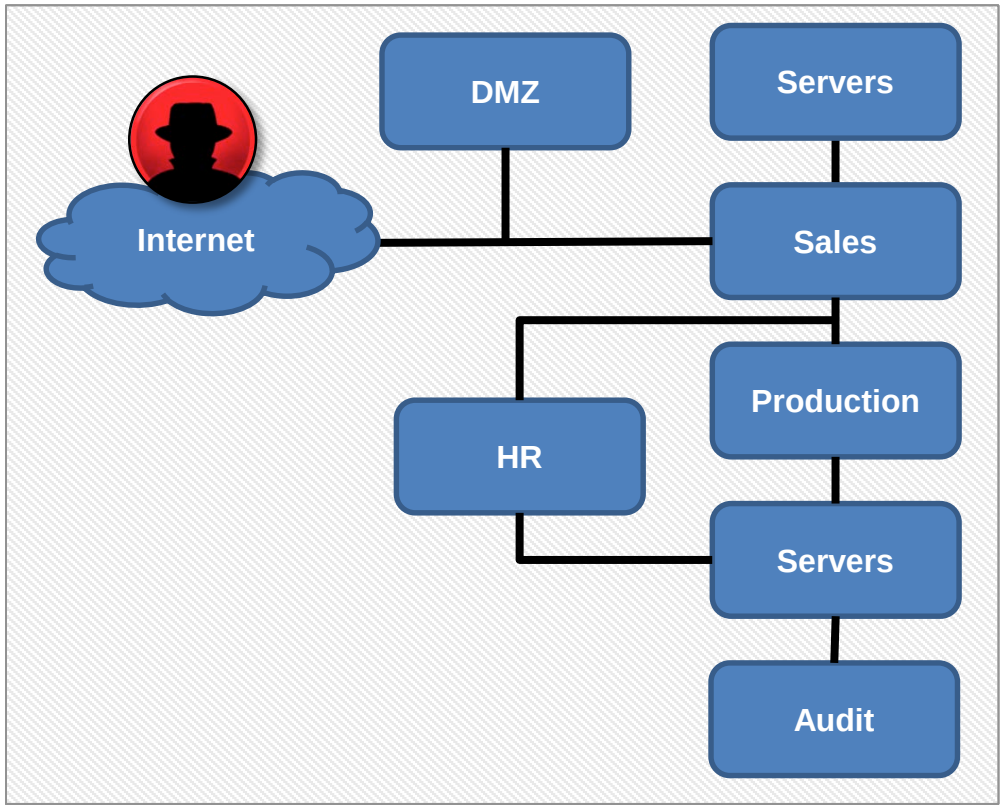
A Critical Cyber Decision Problem: How to Segment the Network?

No Network Segmentation



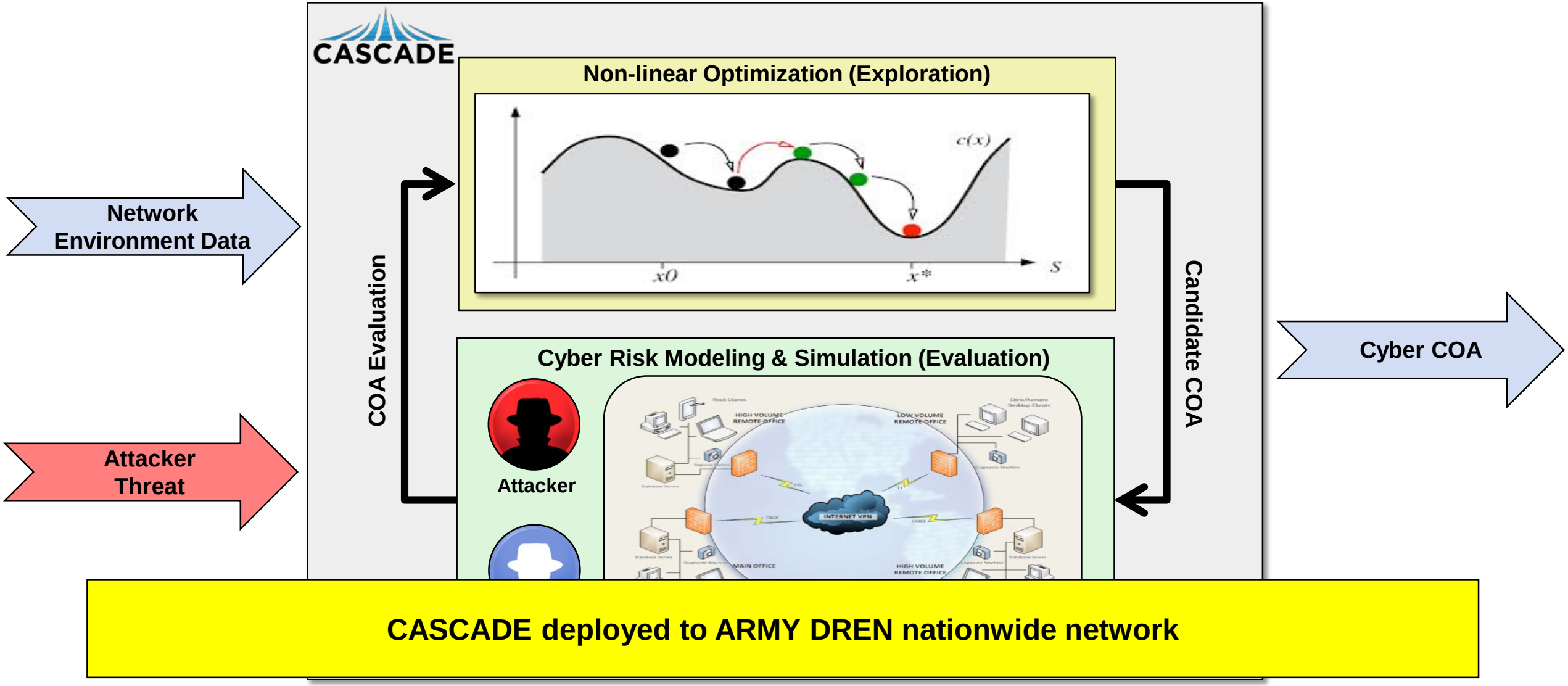
VS.

Network Segmentation



- **Current best practices offer only vague guidance**
  - **E.g.: Segmentation via organizational functional units or principle of least privilege**

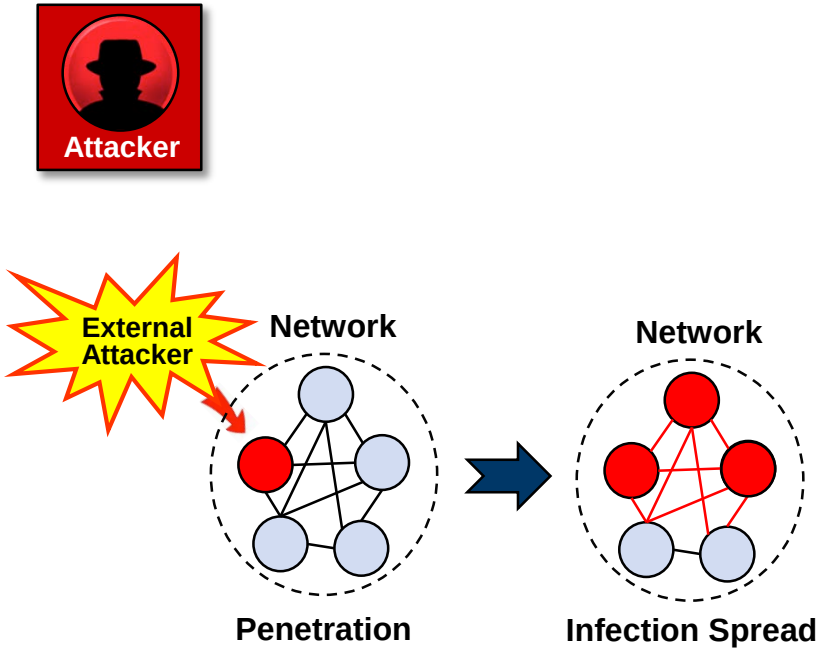
CASCADE: A General-purpose Cyber AI Decision Engine

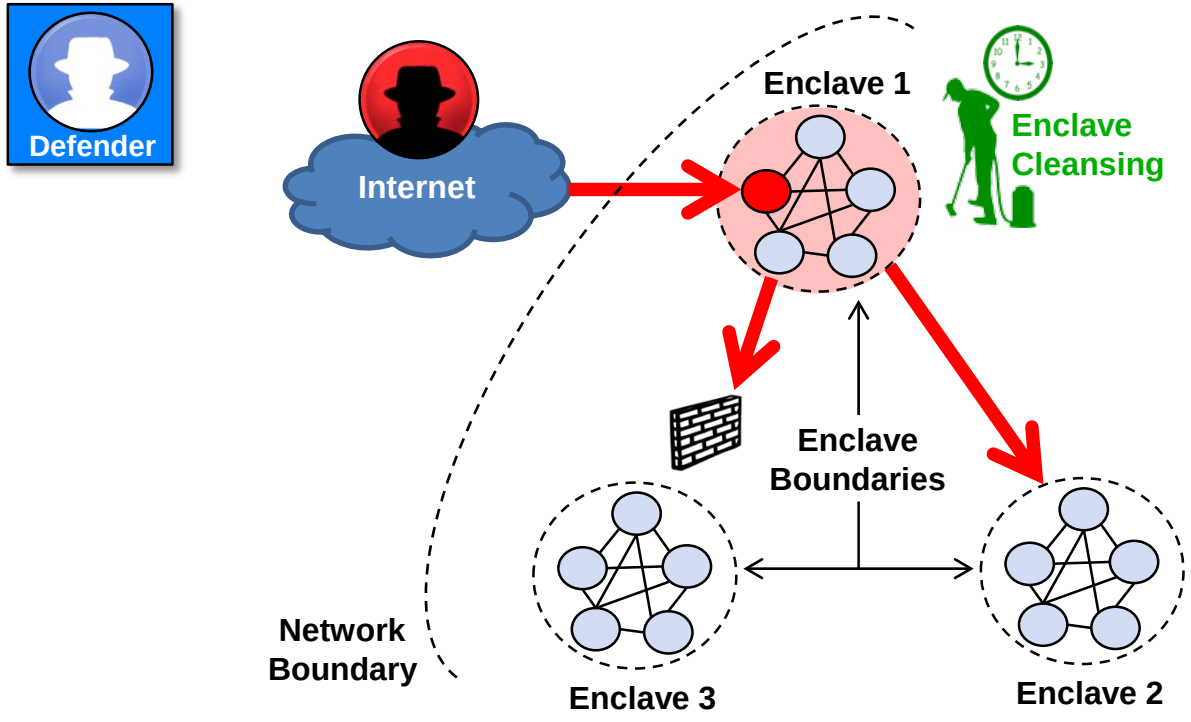


CASCADE deployed to ARMY DREN nationwide network

CASCADE: Cyber Adversarial SCenario modeling and Automated Decision Engine
COA = Course of Action

Network Segmentation Use Case: Cyber Risk Modeling and Simulation Component





Attacker

- Exploit vulnerability to penetrate network
- Pivot and spread throughout network

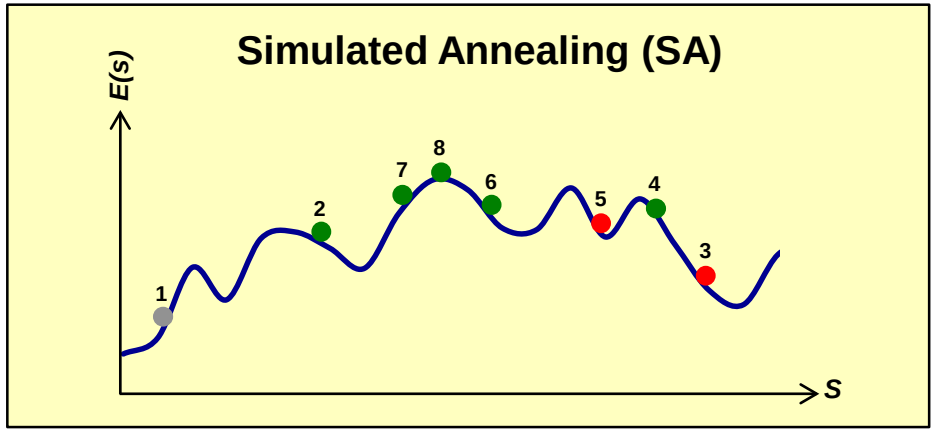
Defender

- Network protected by segmentation architecture
- Communications restricted
- Compromised enclaves periodically cleansed

Network Segmentation Use Case: Non-linear Optimization Component

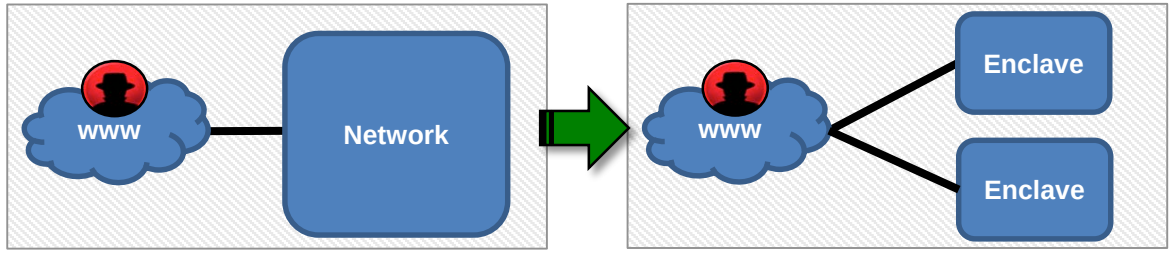
- **Optimization component suggests candidate segmentation architectures to be evaluated via modeling and simulation**
- **Uses evaluation to guide the search**

S **= Space of all possible solution states (i.e. segmentation architectures)**
E(s) **= Evaluation of segmentation architecture** $s \in S$

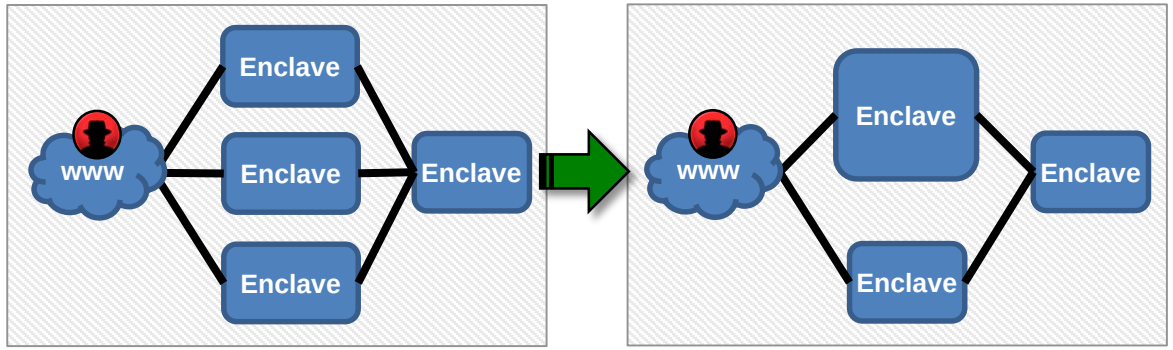


SA algorithm progressively adapts initial solution to find better performing solutions

Designated Search Operations

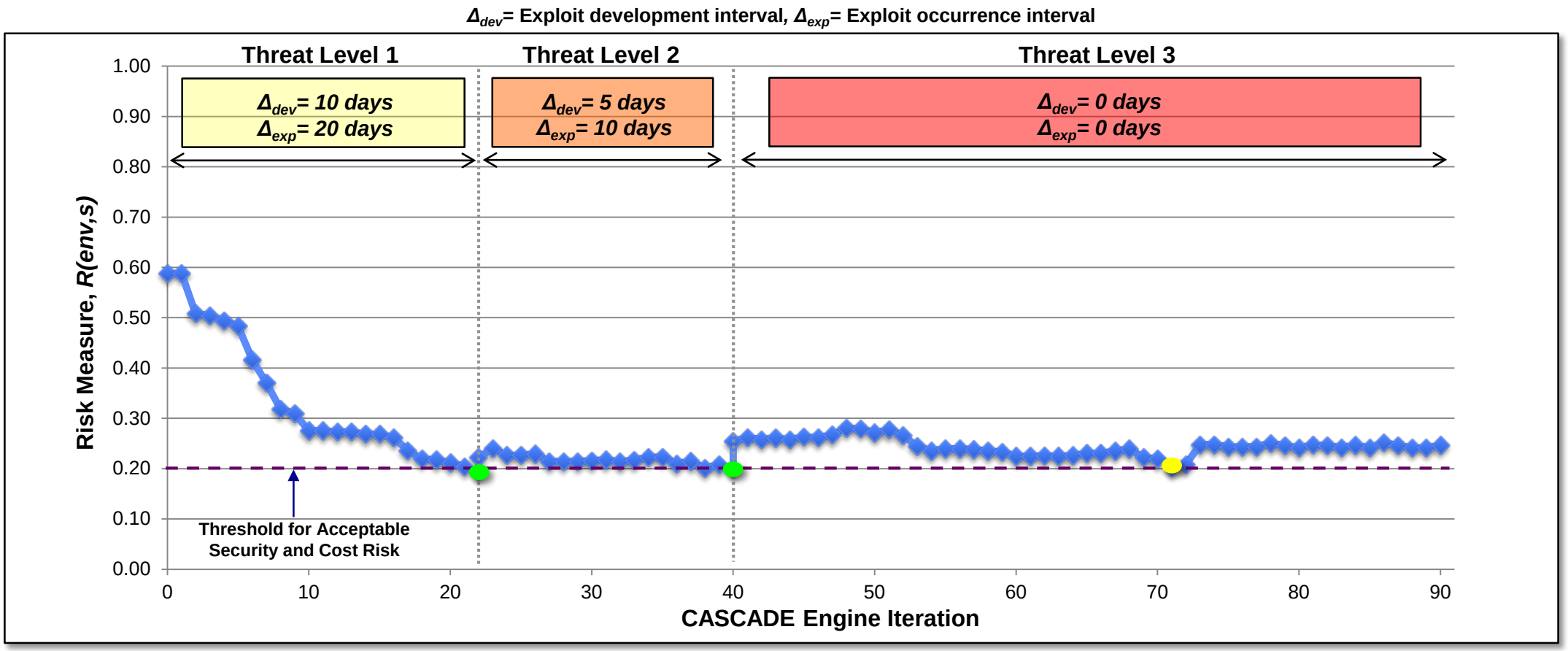


Split Enclave Operation



Merge Enclaves Operation

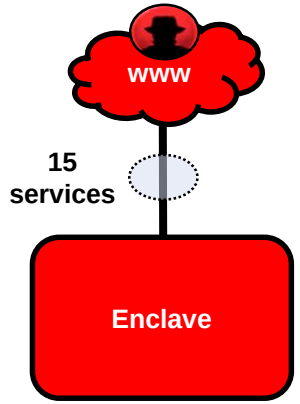
Results



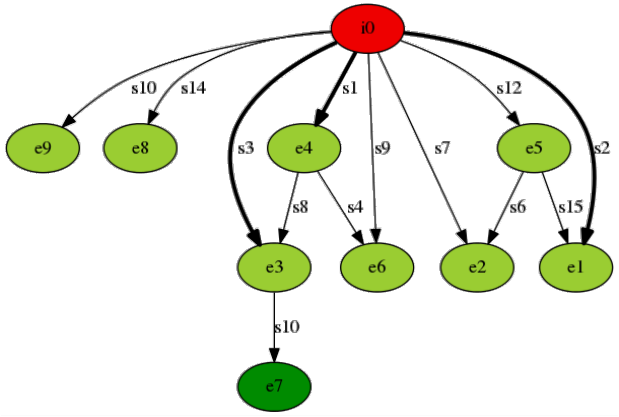
- **CASCADE is able to automatically generate segmentation architecture to meet threshold for acceptable risk to security and cost**
- **Adapts to changing threat levels to generate new architectures that meet threshold**

CASCADE-generated Segmentation Architectures

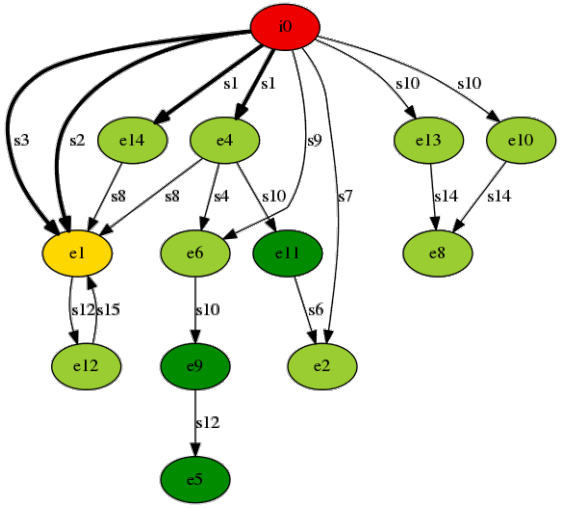
Baseline Architecture



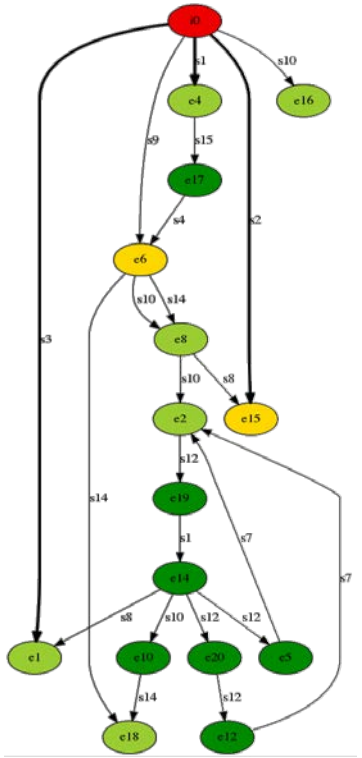
Threat Level 1 Architecture



Threat Level 2 Architecture



Threat Level 3 Architecture



- **Bold** connections denote services that are required to be present in the network.
- Enclaves are colored based on their individual Probability of Compromise (PoC).

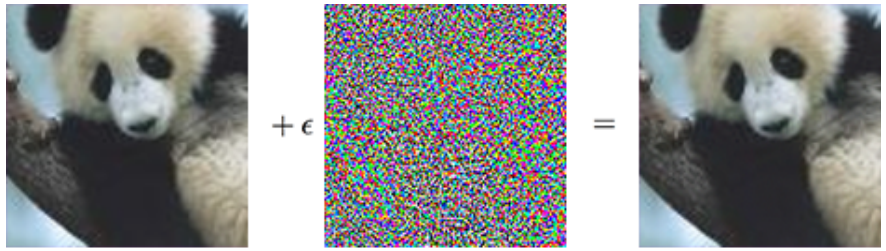
PoC < 0.2	0.4 > PoC >= 0.2	0.6 > PoC >= 0.4	0.8 > PoC >= 0.6	PoC > 0.8

CASCADE can improve baseline architecture to satisfy requirement for acceptable risk and adapt architecture in response to changing threat levels

The Future? Adversarial AI!

Artificial Intelligence

- **Recent favorable results have relied on a benign operating environment**
- **Subtle adversarial interference can confound such approaches**



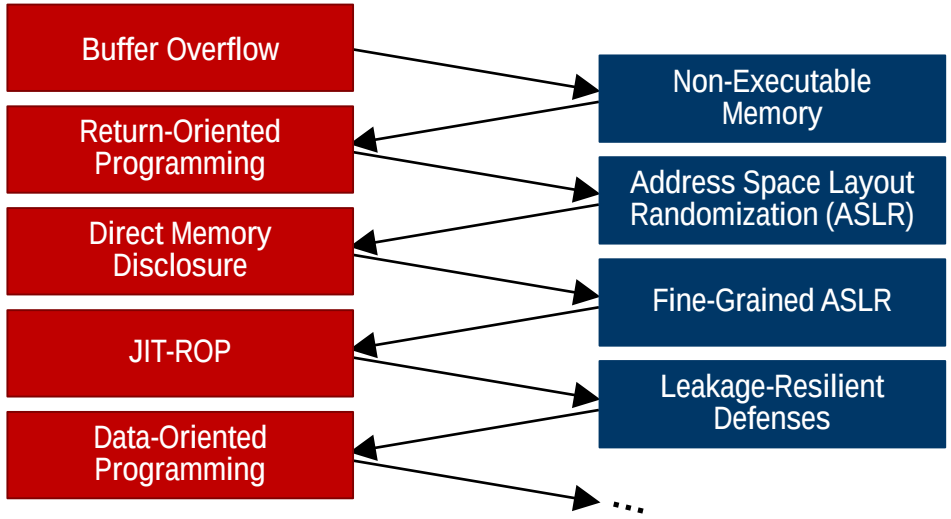
"panda"
57.7% confidence

small noise
carefully constructed

"gibbon"
99.3% confidence

Cyber

- **The cyber domain is not a benign operating environment**
- **Cyber adversaries adapt intelligently to foil each other**



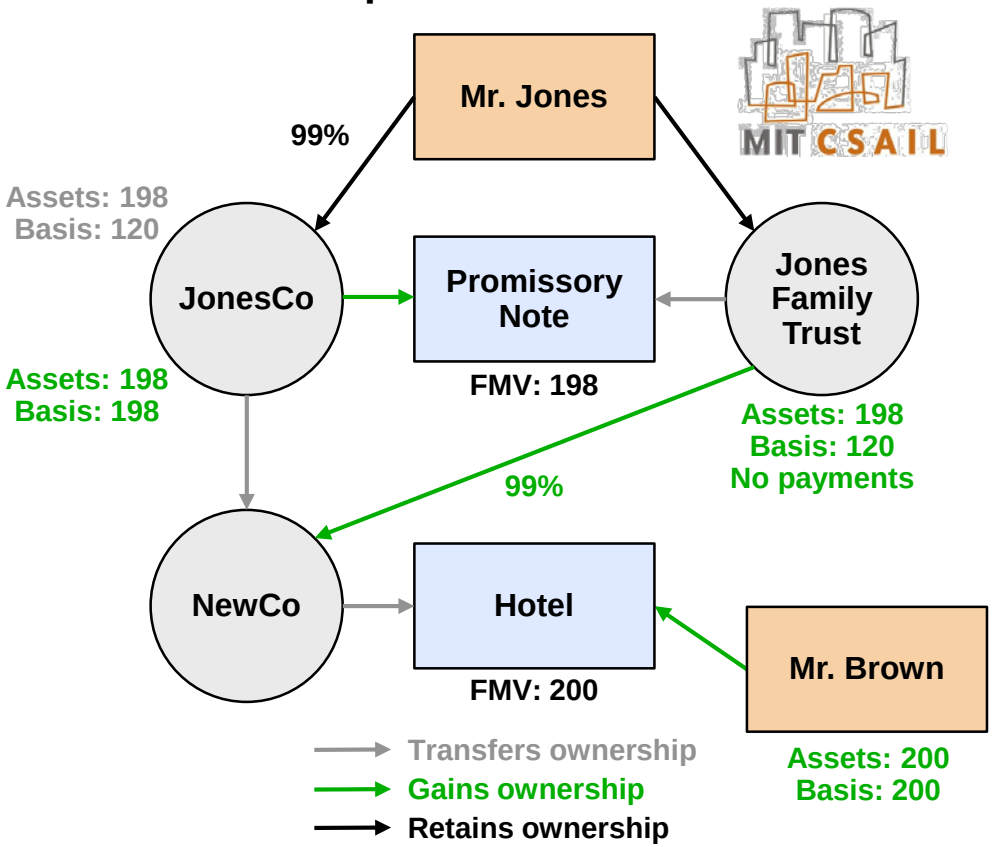
Cyber (AI) approaches are under threat unless they are robust to adversarial adaptation

IJ Goodfellow, J Shlens, C Szegedy (2015) Explaining and harnessing adversarial examples. *ICLR 2015*.

Co-evolutionary Algorithms: AI for Simulating an Adaptive Adversary

- **Attacker and defender evolve simultaneously to optimize against each other**
- **Defender strategies are tested against a range of possible attacker strategies**
  - **Search through strategies takes adversary into account**
  - **Outcome is robust to adversarial adaptation within modeled behavior scope**
- **Prior work* modeling tax evasion and audit rediscovered the iBOB† tax scheme**
  - **Researchers modeled relevant portions of U.S. tax law**
  - **Taxpayer agents learned to structure complex iBOB transactions to avoid capital gains taxes and audits**

Simple iBOB tax scheme

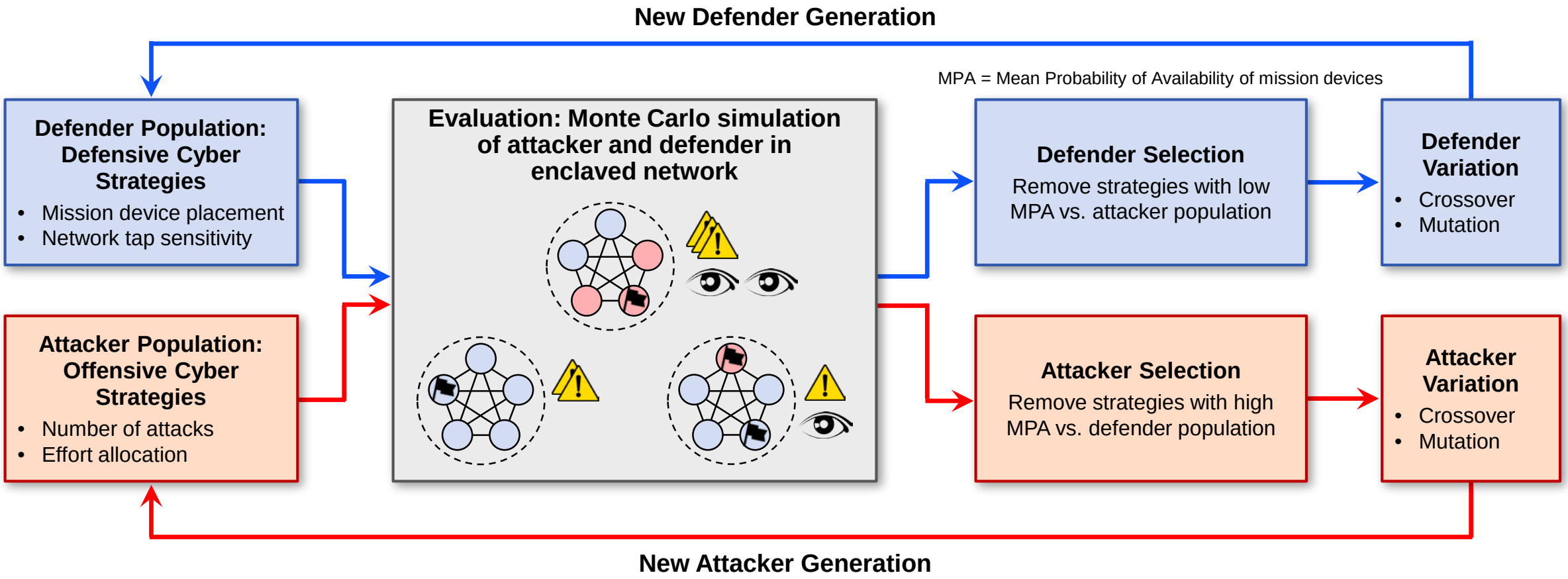


Co-evolve adversaries within a cyber simulation to model intelligent adaptation

* E Hemberg, J Rosen, G Warner, S Wijesinghe, UM O'Reilly (2016) Detecting tax evasion: a co-evolutionary approach. *Artificial Intelligence and Law* 24(2): 149–82.

† Installment-sale Bogus Optional Basis: example from https://www.gao.gov/products/GAO-10-968

CASCADE Extension: Adversarial Co-Evolution of Defender and Attacker in Segmented Network Model*



- **Can evolve defender only (→) or attacker only (→) or co-evolve both (⇄)**
- **Grammatical evolution allows for rapid integration of new cyber domain knowledge into model**

*Hemberg, E., Zipkin, J.R., Skowyra, R.W., Wagner, N. and O'Reilly, U.M., 2018, July. Adversarial co-evolution of attack and defense in a segmented computer network environment. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion* (pp. 1648-1655). ACM.

Summary and Next Steps

- **AI is here, there is no escaping it now**
- **AI, cybercrime, and fraud are joined at the hip**
- **If AI can be used for offense, it can be used for defense**
- **“The best defense (can be) a good offense…”**

Contact Information

Neal Wagner
neal.f.wagner@gmail.com