Adversarial and Uncertain Reasoning for Adaptive Cyber Defense: Building the Scientific Foundation

Sushil Jajodia
George Mason University

IEEE International 5G Summit, Reston, Virginia
August 19, 2017

Outline

- Motivation
  - Current cyber defense landscape & open questions
- Pro-active Defense via Adaptation
  - Adaption Techniques
  - Scientific Challenges
- Research Highlights

Motivation

Today's Cyber Defenses are Static

- Today's approach to cyber defense is *governed by slow and deliberative processes* such as
  - Security patch deployment, testing, episodic penetration exercises, and human-in-the-loop monitoring of security events
- Adversaries can greatly benefit from this situation
  - They can *continuously and systematically probe targeted networks* with the confidence that those networks will change *slowly if at all*
  - They have the time to engineer reliable exploits and pre-plan their attacks
- Additionally, once an attack succeeds, adversaries persist for long times inside compromised networks and hosts
  - Hosts, networks, software, and services do not reconfigure, adapt, or regenerate except in deterministic ways to support maintenance and uptime requirements

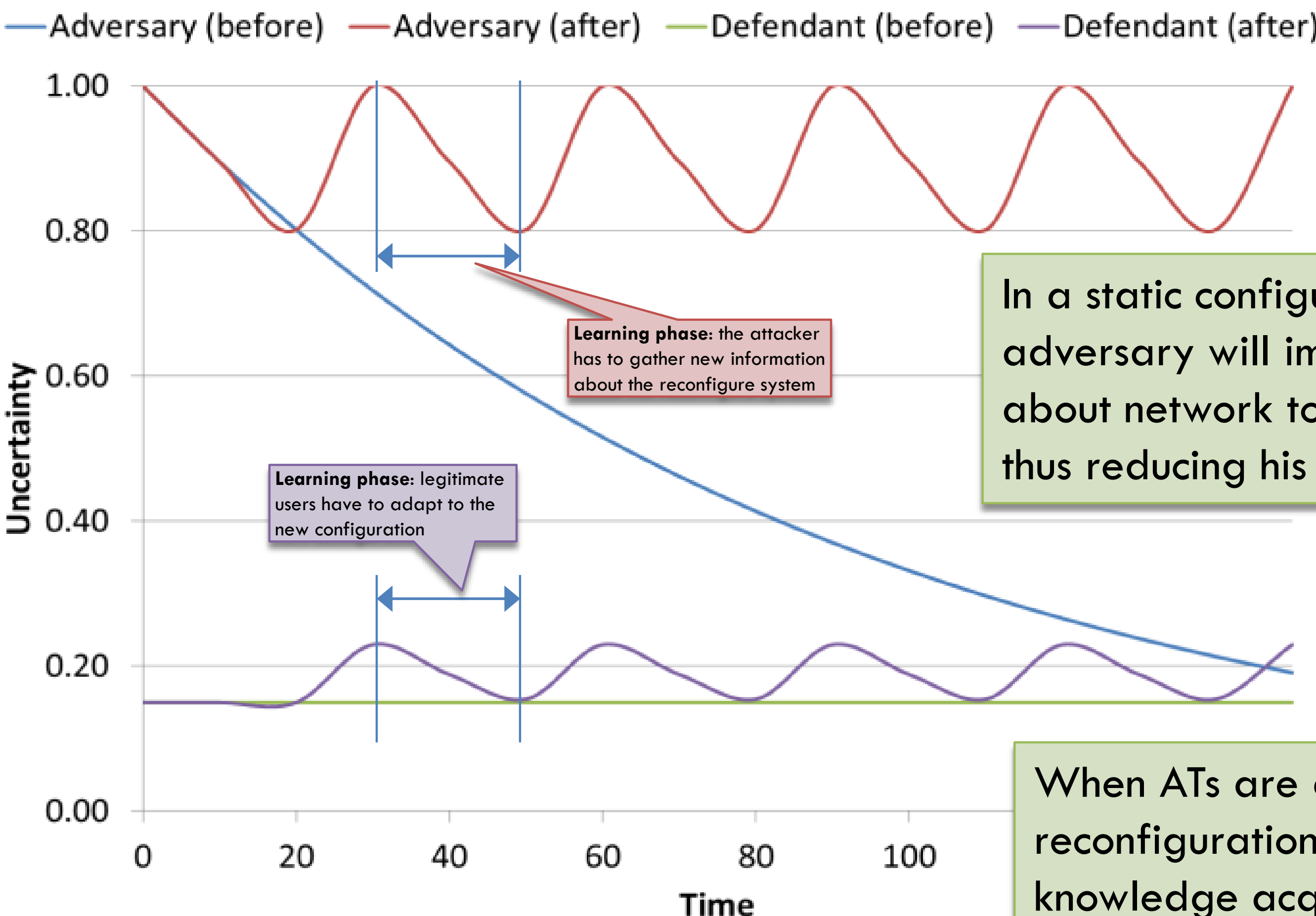
5

Pro-active Defense via Adaptation

Security through adaptation: A paradigm shift

- Adaptation Techniques (AT) consist of engineering systems that have **homogeneous functionalities** but **randomized manifestations**
  - These techniques *make networked information systems less homogeneous and less predictable*
  - ***Examples:*** Moving Target Defenses (MTD), artificial diversity, and bio-inspired defenses
- **Homogeneous functionality** allows authorized use of networks and services in predictable, standardized ways
- **Randomized manifestations** make it difficult for attackers to engineer exploits remotely, or reuse the same exploit for successful attacks against a multiplicity of hosts

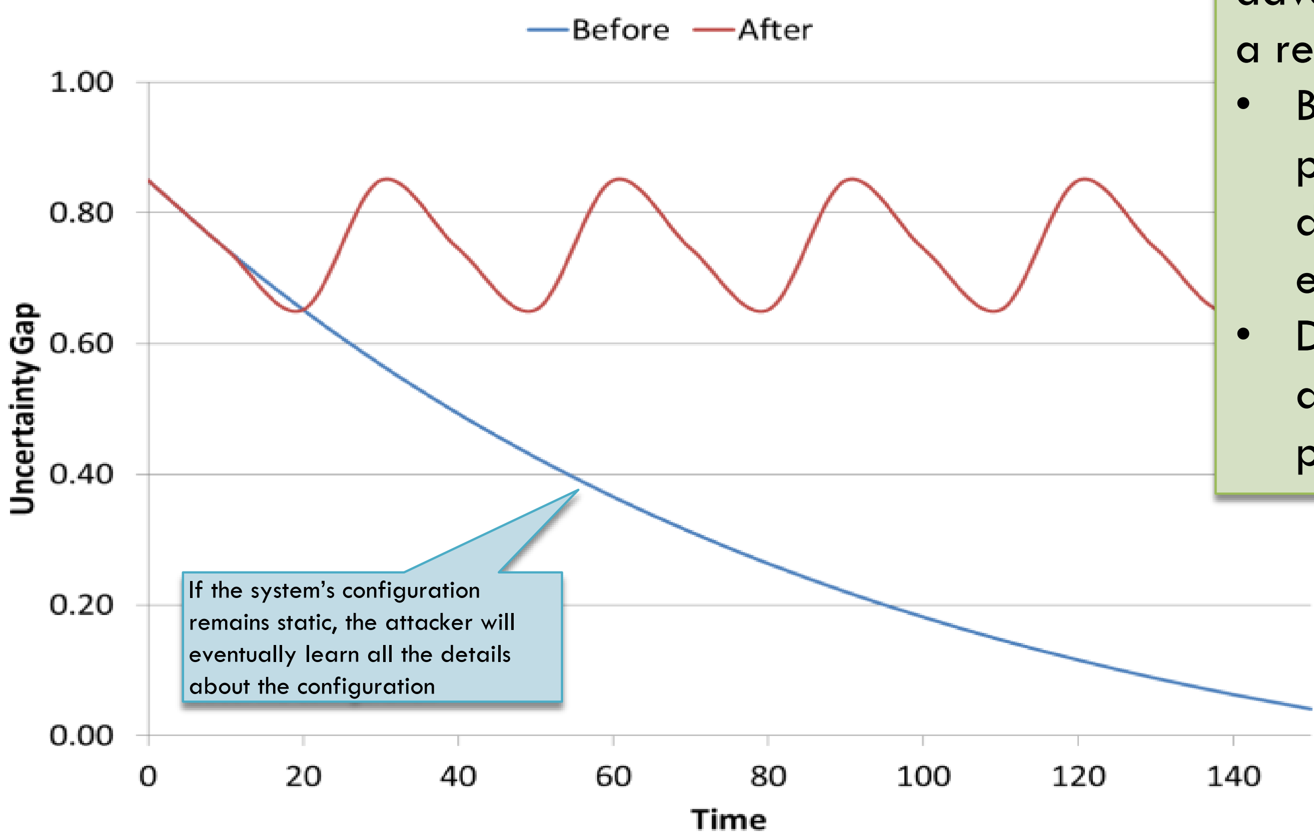
Adversary and Defender Uncertainty



In a static configuration, over time, the adversary will improve his knowledge about network topology and configuration, thus reducing his uncertainty

When ATs are deployed, each system reconfiguration will invalidate previous knowledge acquired by adversaries, thus restoring their uncertainty to higher levels

Uncertainty Gap



ATs enable us to maintain the information gap between adversaries and defenders at a relatively constant level

- Before deploying the proposed mechanisms, the defender's advantage is eroded over time
- Dynamically changing the attack surface ensures a persistent advantage

AT Benefits

- Increase **complexity, cost, and uncertainty** for attackers
- Limit exposure of vulnerabilities and opportunities for attack
- Increase system resiliency against known and unknown threats
- Offer probabilistic protection despite exposed vulnerabilities, as long as the vulnerabilities are not predictable by the adversary at the time of attack

Software-Based Adaptation

- **Address Space Layout Randomization (ASLR)**
  - Randomizes the locations of objects in memory, so that attacks depending on knowledge of the address of specific objects will fail
- **Instruction Set Randomization (ISR)**
  - A technique for preventing code injection attacks by randomly altering the instructions used by a host machine or application
- **Compiler-based Software Diversity**
  - When translating high-level source code to low-level machine code, the compiler diversifies the machine code on different targets, so that vulnerability exploits working on one target *may not work on other targets*

Network-Based Adaptation

- ID randomization
- Generation of arbitrary external attack surfaces
- VM-based dynamic virtualized network
- Phantom servers to mitigate insider and external attacks
- Proxy moving and shuffling to detect insider attacks
- Overall, these techniques aim at *giving the attacker a view of the target system that is significantly different from what the system actually is*

But there are Many ACD Ideas…

ESC-EN-HA-TR-2012-109

Technical Report
1166

Survey of Cyber Moving Targets

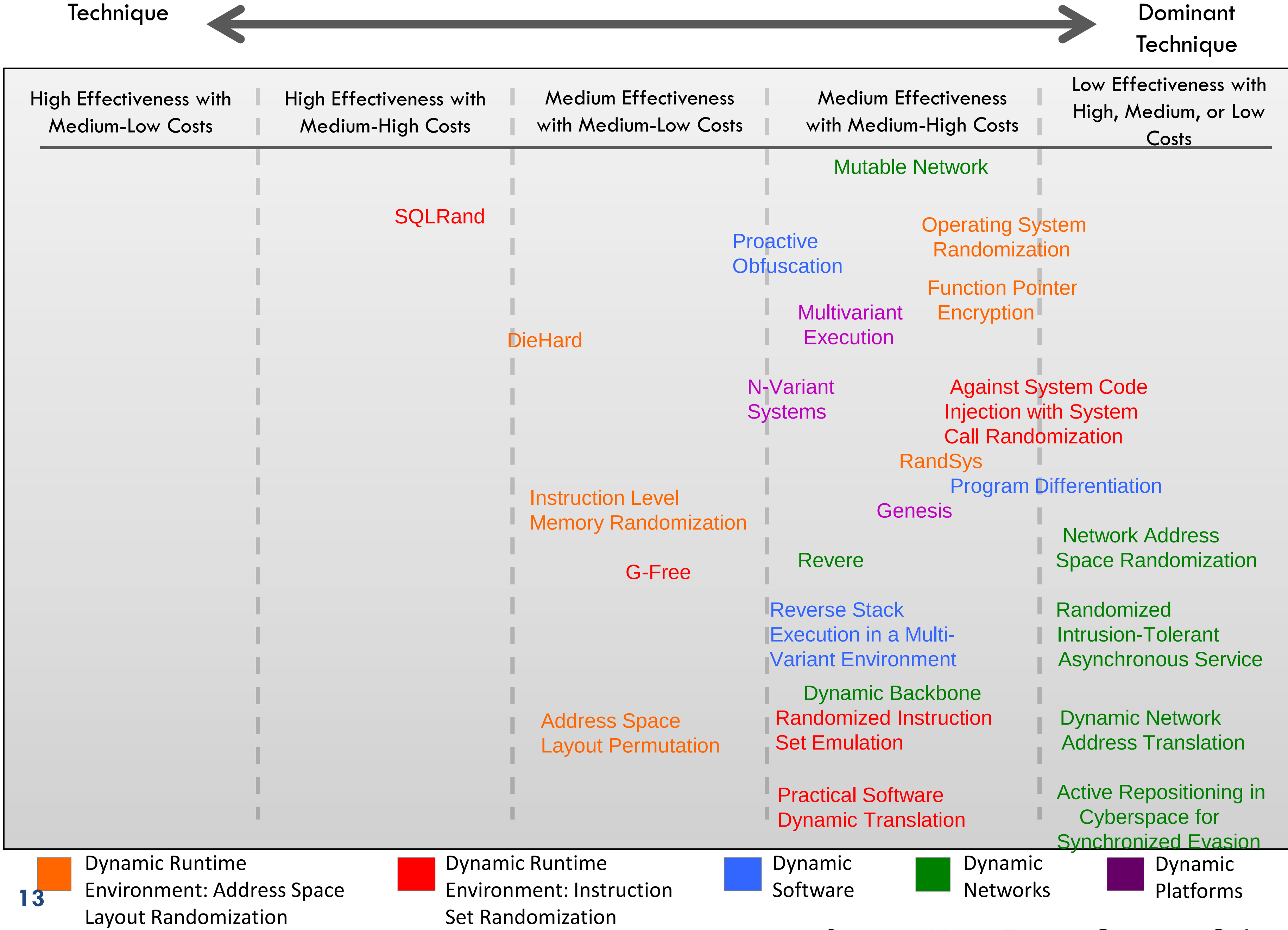
H. Okhravi
M.A. Rabe
T.J. Mayberry
W.G. Leonard
T.R. Hobson
D. Bigelow
W.W. Streilein

At least 39 documented in this 2013 MIT Lincoln Labs Report

>50 today?

How can we compare them?

Spectrum of Moving Target Defense Techniques



Source: Kate Ferris, George Cybenko

Limitations of Current Approaches

- The *contexts in which ATs are useful and their added cost* (in terms of performance and maintainability) to the defenders can vary significantly
  - Most ATs aim at preventing a specific type of attack
- The focus of existing approaches is on *developing new techniques*, not on understanding overall operational costs, when they are most useful, and what their possible interrelationships might be
- While each AT might have some engineering rigor, the *overall discipline is largely ad hoc* when it comes to understanding the totality of AT methods and their optimized application
- AT approaches assume *non-adversarial, environments*

Adaptive Cyber Defense (ACD)

- We need to *understand*
  - the overall operational costs of these techniques
  - when they are most useful
  - their possible inter-relationships
- Propose new classes of techniques that force adversaries to *continually re-assess and re-plan their cyber operations*
- Present adversaries with *optimally changing attack surfaces* and system configurations

Adaptive Cyber Defense (ACD)

Advanced Persistent Threats (APTs) have the time and technology to easily exploit our systems now

Attack Phase	**Reconnaissance** Identify the attack surface	**Access** Compromise a targeted component	**Persistence** Maintain presence and exploitation
Possible Adaptation Techniques (AT)	Randomized network addressing and layout; Obfuscated OS types and services.	Randomized instruction set and memory layout; Just-in-time compiling and decryption.	Dynamic virtualization; Workload and service migration; System regeneration.

There are many possible AT options

Adaptation techniques are typically aimed at defeating different stages of possible attacks

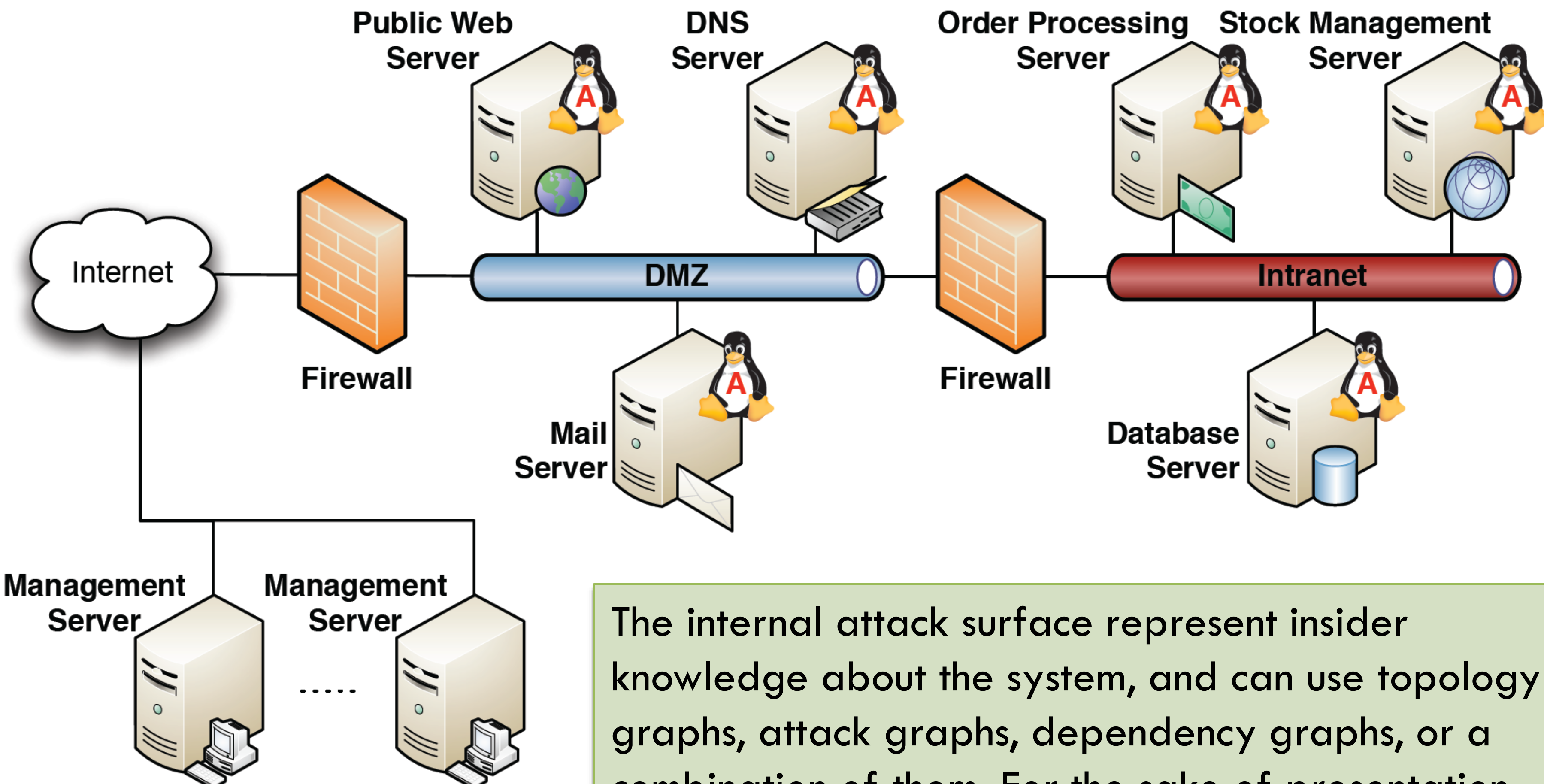
We need to develop a scientific framework for optimizing strategies for deploying adaptation techniques for different attack types, stages and underlying missions

Research Highlights

Novel Adaptive Techniques

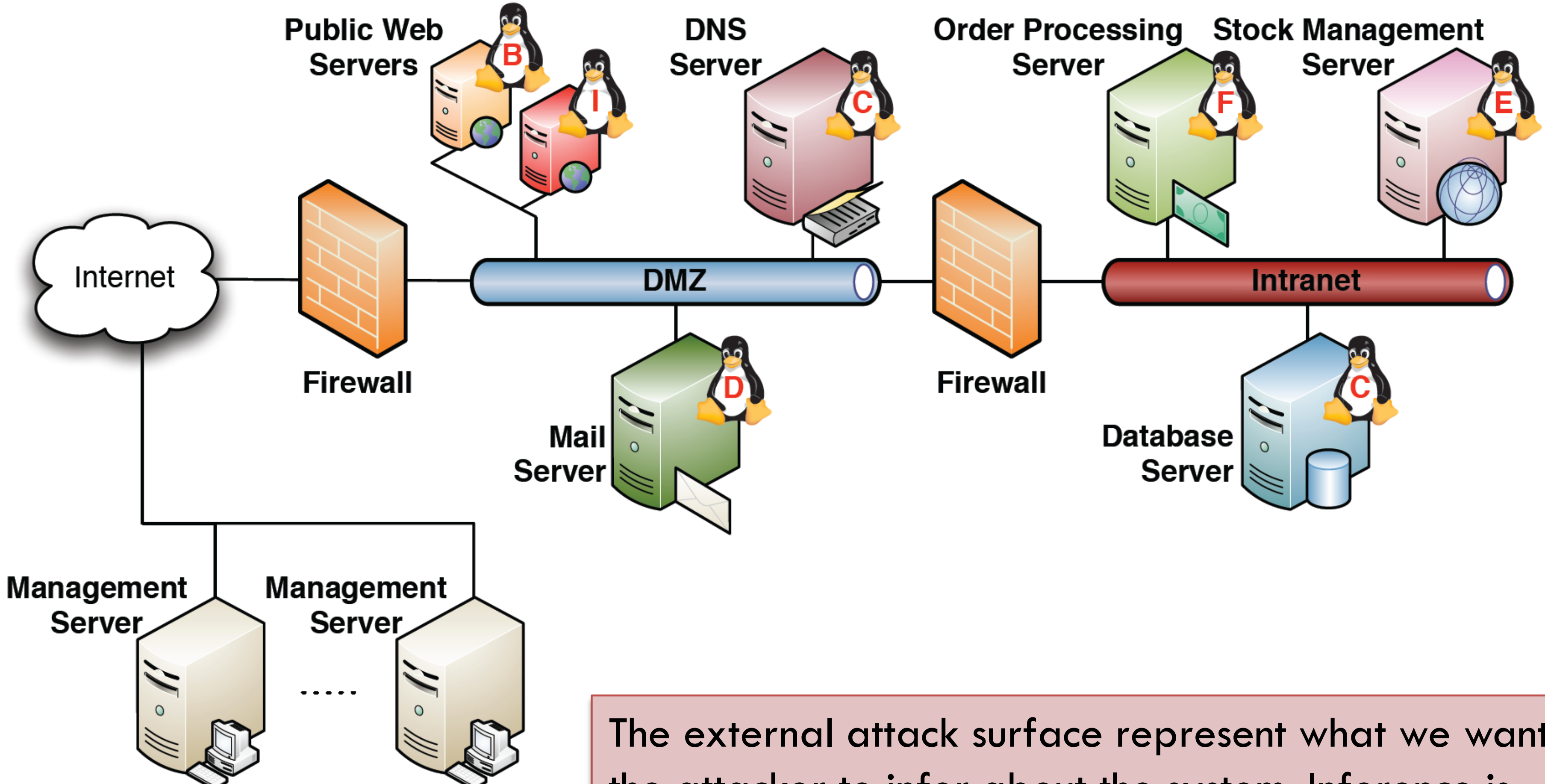
- Manipulating responses to an attacker's probes
  - **Goal:** altering the attacker's perception of a system's attack surface
- Creating distraction clusters
  - **Goal:** controlling the probability that an intruder may reach a certain goal within a specified amount of time
- Increasing diversity
  - **Goal:** increasing the complexity and cost for attackers by increasing the diversity of resources along certain attack paths
    - Different metrics are proposed to measure diversity

Example: Internal Attack Surface



The internal attack surface represent insider knowledge about the system, and can use topology graphs, attack graphs, dependency graphs, or a combination of them. For the sake of presentation, this example only shows topology information.

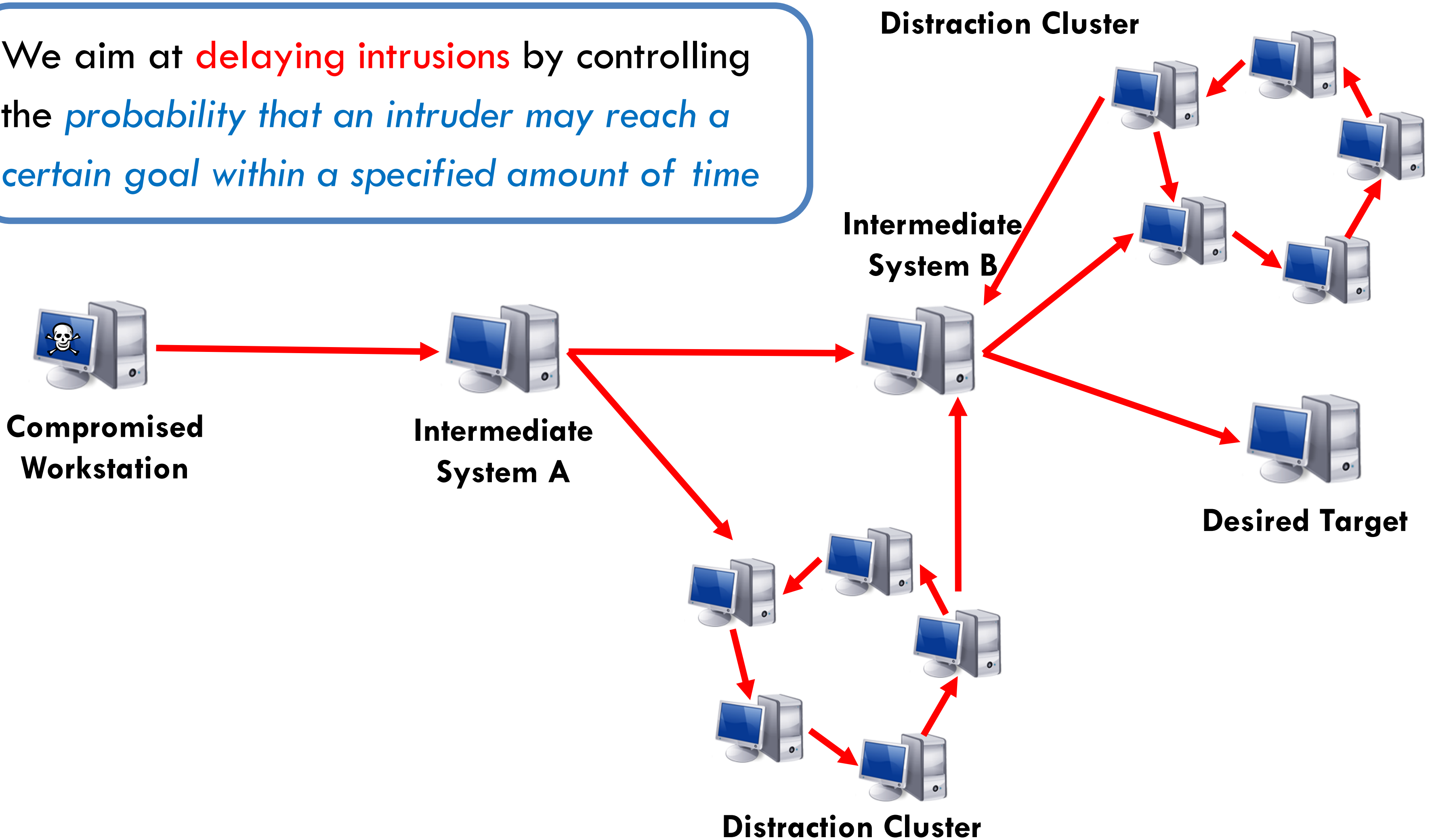
Example: External Attack Surface



The external attack surface represent what we want the attacker to infer about the system. Inference is based on probing and sniffing.

Distraction Clusters

We aim at delaying intrusions by controlling the *probability that an intruder may reach a certain goal within a specified amount of time*



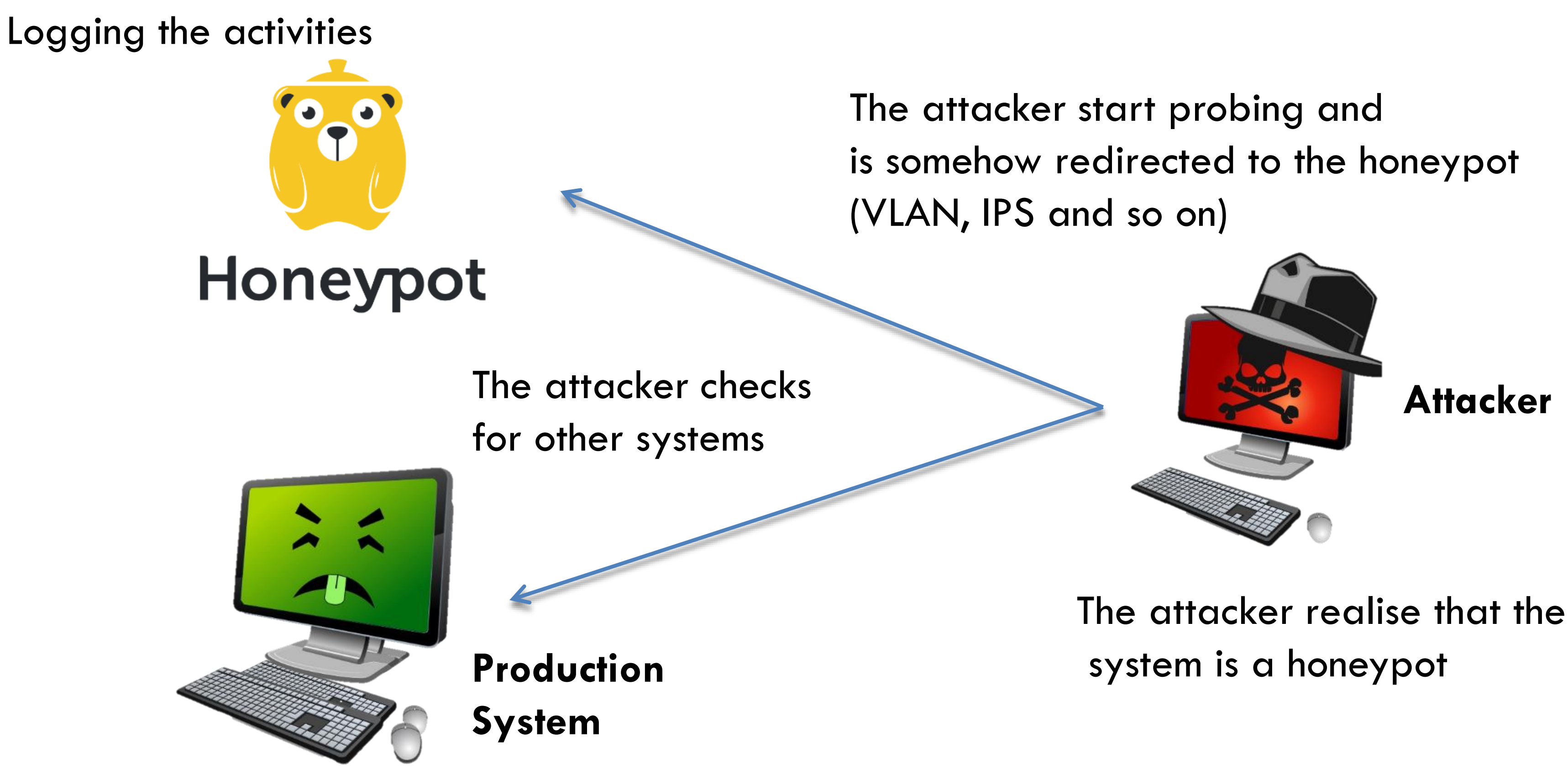
Network diversity

- We take the first step towards formally *modeling network diversity as a security metric*
  - We propose a network diversity function based on well known mathematical models of biodiversity in ecology
  - We design a network diversity metric based on the least attacking effort
  - We design a probabilistic network diversity metric to reflect the average attacking effort
  - We evaluate the metrics and algorithms through simulation
- The modeling effort helps understand diversity and enables quantitative hardening approaches

Solving Real-world Problems

- Adversarial defense of enterprise systems
  - Pareto-optimal solutions that allow defenders to simultaneously maximize productivity and minimize the cost of patching
- Optimal scheduling of cyber analysts
  - Given limited resources, the analyst workforce must be optimally managed for minimizing risk

Classical Approach



A Different Approach

Logging the activities

The attacker sees directly
the Production System

Option 2. The attacker keep
interacting with the system

Attacker

The attacker interact
with the system

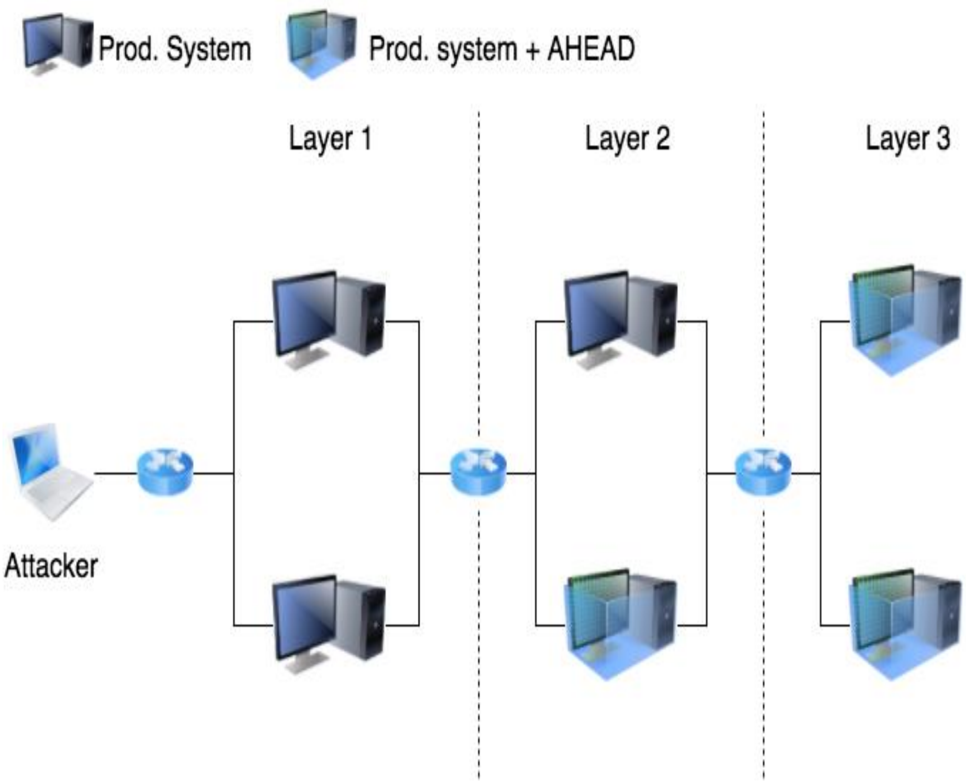
Option 1
He thinks that the
system is a Honeypot,
look for other systems

Production System

Client/ Server/ Honeypot/
Network Component

Joint work with Prof Luigi Mancini, U of Rome

Evaluation of our Approach



31 last year MSc students

3-layer experiment:

- L1 - No AHEAD deployed
- L2 - AHEAD on one machine
- L3 - AHEAD on both machines

Goal: root privilege in L3 machine

L3 machines and L1 machines had same vulnerable service

Results

Layer	Machine	Success %	Time to Success	Traffic (GB)	Avg. Individual Traffic
L1		90.32%	1h 9m 36s	21.23	0.68
	Prod. System 1	5.34%		7.4305	0.24
	Prod. System 2	84.98%		13.7995	0.44
L2		61%	14h 37m 26s	78.88	2.82
	Prod. System 3	61%	14h 37m 26s	52.0608	1.86
	Prod. System + AHEAD	0%	∞	26.82	0.96
L3		6%	48h 25m 42s	54.89	2.89
	Prod. System1 + AHEAD	0%	∞	23.6027	1.24
	Prod. System2 + AHEAD	6%	48h 25m 42s	31.29	1.65

Optimal Scheduling of Cyber Analysts for Minimizing Risk*

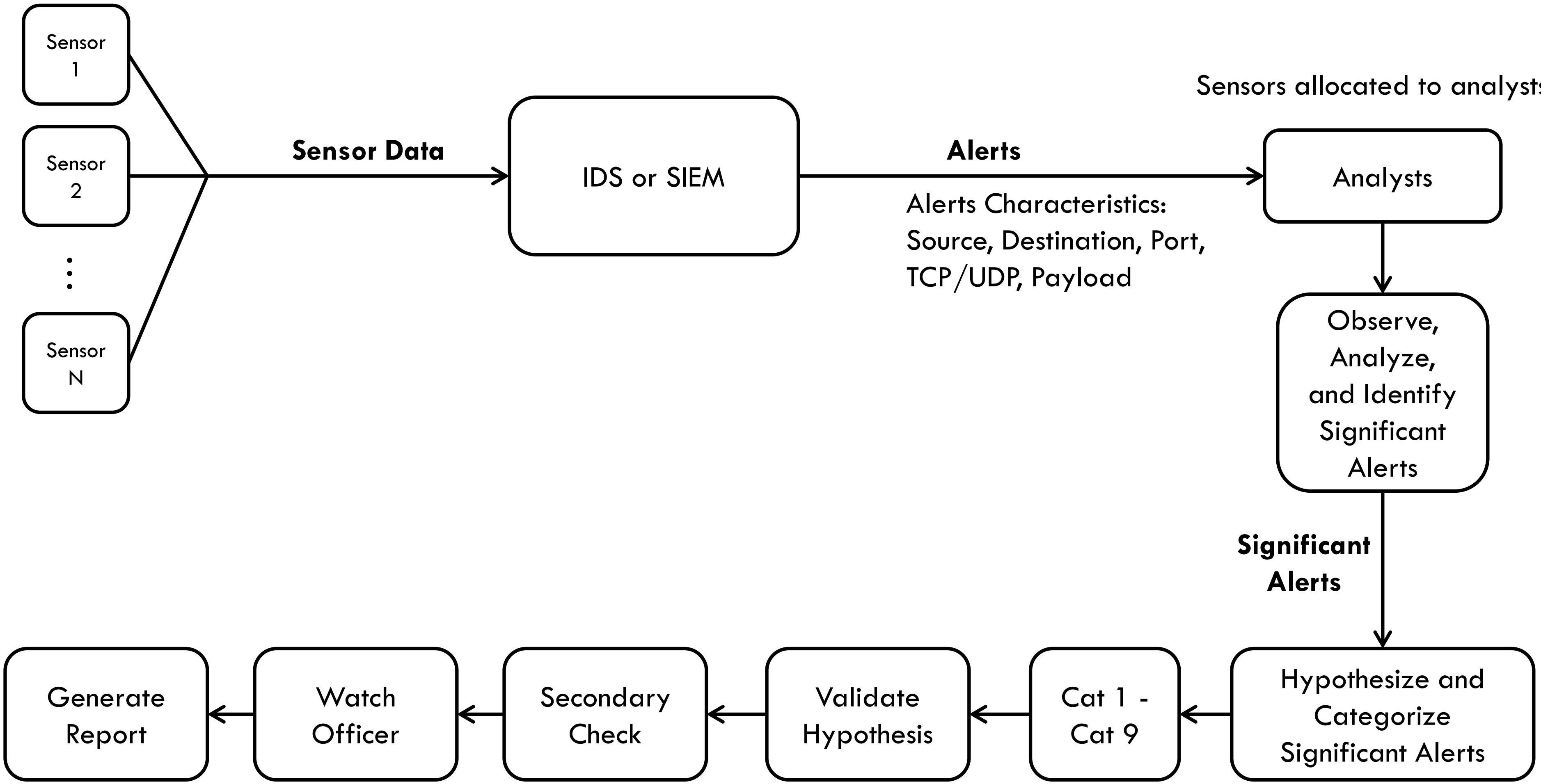
*Joint work with Rajesh Ganesan (GMU), Ankit Shah (GMU), Hasan Cam (ARL)

Statement of Need

- Cybersecurity threats are on the rise
- Demand for Cybersecurity analysts outpaces supply [1] [2]
- Given limited resources (personnel), the analyst workforce must be optimally managed
- Given the current/projected number of alerts it is also necessary to know the optimal workforce size

[1] http://www.rand.org/pubs/research_reports/RR430.html
[2] http://www.rand.org/news/press/2014/06/18.html

Process Flow, Definition of Significant Alerts



Significant Alerts = 1% of all Alerts Generated

Categories 1-9

DON CYBER INCIDENT CATEGORY

Cat 1-9	Description
1	**Root Level Intrusion (Incident):** Unauthorized privileged access (administrative or root access) to a DoD system.
2	**User Level Intrusion (Incident):** Unauthorized non-privileged access (user-level permissions) to a DoD system. Automated tools, targeted exploits, or self-propagating malicious logic may also attain these privileges.
3	**Unsuccessful Activity Attempted (Event):** Attempt to gain unauthorized access to the system, which is defeated by normal defensive mechanisms. Attempt fails to gain access to the system (i.e., attacker attempt valid or potentially valid username and password combinations) and the activity cannot be characterized as exploratory scanning. Can include reporting of quarantined malicious code.
4	**Denial of Service (DOS) (Incident):** Activity that impairs, impedes, or halts normal functionality of a system or network.
5	**Non-Compliance Activity (Event):** This category is used for activity that, due to DoD actions (either configuration or usage) makes DoD systems potentially vulnerable (e.g., missing security patches, connections across security domains, installation of vulnerable applications, etc.). In all cases, this category is not used if an actual compromise has occurred. Information that fits this category is the result of non-compliant or improper configuration changes or handling by authorized users.
6	**Reconnaissance (Event):** An activity (scan/probe) that seeks to identify a computer, an open port, an open service, or any combination for later exploit. This activity does not directly result in a compromise.
7	**Malicious Logic (Incident):** Installation of malicious software (e.g., trojan, backdoor, virus, or worm).
8	**Investigating (Event):** Events that are potentially malicious or anomalous activity deemed suspicious and warrants, or is undergoing, further review. No event will be closed out as a Category 8. Category 8 will be re-categorized to appropriate Category 1-7 or 9 prior to closure.
9	**Explained Anomaly (Event):** Events that are initially suspected as being malicious but after investigation are determined not to fit the criteria for any of the other categories (e.g., system malfunction or false positive).

Source: Dept of Navy, Cybersecurity Handbook, page 20

Statement of Need

- Cybersecurity threats are on the rise
- Demand for Cybersecurity analysts outpaces supply [1] [2]
- Given limited resources (personnel), the analyst workforce must be optimally managed for minimizing today's risk
- Given the current/projected number of alerts it is also necessary to know the optimal workforce size to keep risk under a certain threshold

[1] http://www.rand.org/pubs/research_reports/RR430.html
[2] http://www.rand.org/news/press/2014/06/18.html

Definition of Risk

- Alert Coverage is defined as the % of the significant alerts (1% of the total alerts) that are thoroughly investigated in a work-shift by analysts and the remainder (forms the Risk) is not properly analyzed or unanalyzed because of
  - Sub-optimal shift scheduling
  - Not enough personnel in the organization
  - Lack of time (excessive analyst workload)
  - Not having the right mix of expertise in the shift in which the alert occurs
- Risk % = 100 – Alert Coverage %

Note: From this slide onward, the term alert refers to significant alerts only

Requirements

- The cybersecurity analyst scheduling system
  - Shall ensure that an optimal number of staff is available to meet the demand to analyze alerts
  - Shall ensure that a right mix of analysts are staffed at any given point in time
  - Shall ensure that risks due to threats are maintained below a pre-determined threshold
  - Shall ensure that weekday, weekend, and holiday schedules are drawn such that it conforms to the working hours/leave policy

Problem Description

Risk is proportional to Analyst Characteristics

1. Alert generation rate
2. the number of analysts,
3. their expertise mix,
4. analyst's shift and days-off scheduling,
5. their sensor assignment,
6. Category of alert – analyst workload – time to analyze (input)

Two types of problems to solve:

Simulation: Given all of the above, what level of risk is the organization operating at?
Optimization: Given an upper bound on risk, what are the optimal settings for 1-5?

Algorithm Contributions

Optimization Algorithm

- Mixed Integer Programming solved using Genetic Algorithm
- Outputs
  - the number of analysts,
  - their expertise mix,
  - their sensor-to-analyst assignment

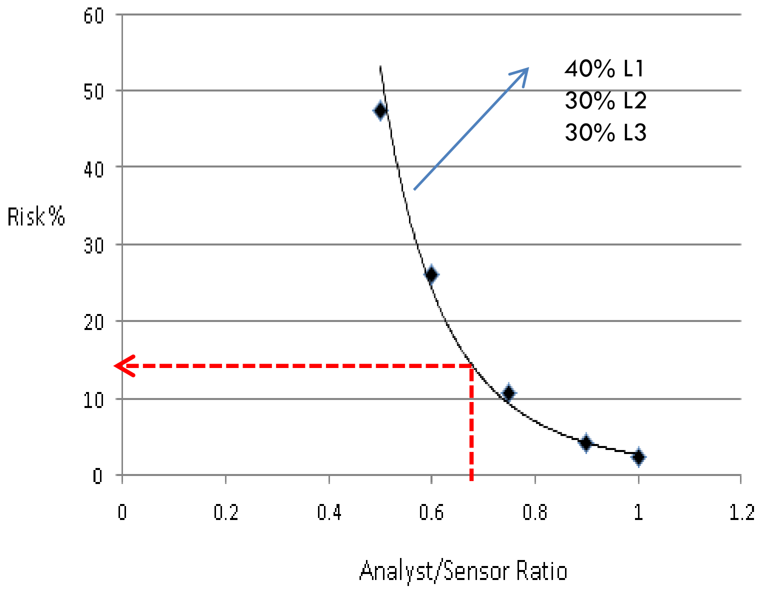
Scheduling Algorithm

- Integer programming and a heuristic approach
- Output
  - Analyst shift and days-off scheduling

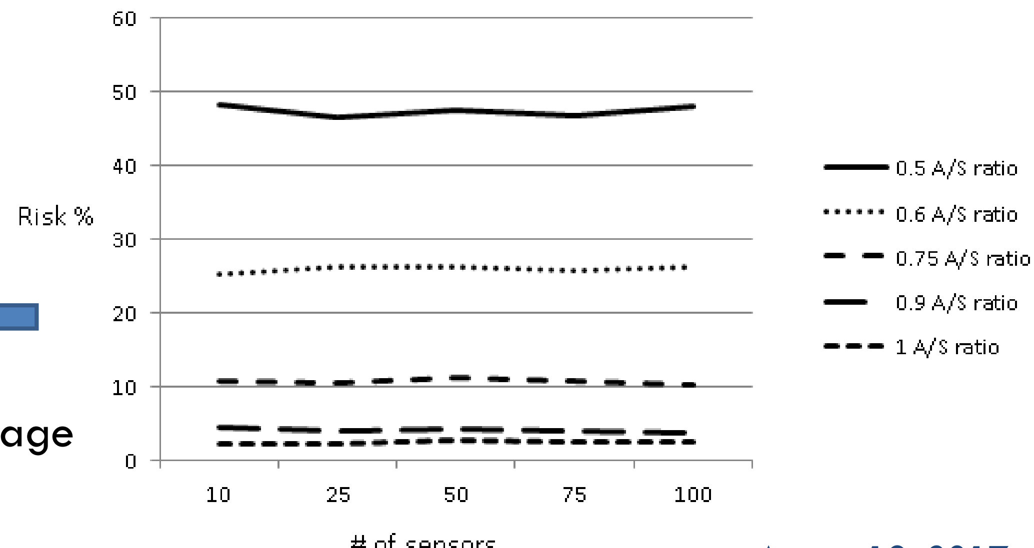
Simulation Algorithm

- Validates optimization
- A tool can be used as a stand-alone algorithm to measure the current risk performance of the organization for a given set of inputs

Main Results



- Risk% varies non-linearly with analyst/sensor (A/S) ratio
- Plot is useful for hiring decisions
- Assumption: All sensors have the same average alert generation rate, and it remains fixed



For a given analyst/sensor ratio risk is independent of the # of sensors, when the average alert arrival and average service rates remain the same

Sample days off Scheduling

- An analyst works 12*6 + 1*8 = 80 hrs in 2 weeks (7 out of every 14 days from Sun to Sat)
- Gets every other weekend off
- Works no more than 5 consecutive days in a 14 day period

Output of the days-off scheduling algorithm or 10 analysts

Day →	S	S	M	T	W	T	F	S	S	M	T	W	T	F	S	S	M	T	W	T	F	S
1	X	X	X	X			X			X	X				X	X	X	X			X	
2	X	X		X	X	X					X	X			X	X		X	X	X		
3	X	X			X	X					X	X	X		X	X			X	X		
4	X	X				X	X			X			X	X	X	X				X	X	
5	X	X	X				X			X		X		X	X	X	X				X	
6			X	X	X			X	X				X	X			X	X	X			X
7				X	X			X	X	X	X			X				X	X			X
8			X		X	X		X	X		X	X					X		X	X		X
9				X		X	X	X	X			X	X					X		X	X	X
10			X				X	X	X	X			X	X			X				X	X

X – off days

Need for Dynamic Scheduling

- Static optimization and scheduling assumes
  - Same average alert generation rates for all sensors, which is drawn from a Uniform distribution.
- What if there are world events or zero-day attacks that could trigger an increase in analyst workload
- What if there are varying alert generation rates per sensor per hour
  - Causes uncertainty in future alert workload to be investigated
    - Workload uncertainty makes it difficult for managing personnel scheduling
      - How many analysts at each level of expertise must report to work?
      - Do we have the flexibility in the schedule to adapt to day to-day changing analyst needs

Research Findings

- Alert estimation is critical for a successful implementation of the dynamic optimization model
- The average alert generation rate must be handled by a static workforce (X matrix)
- Dynamic optimization is capable of adapting to changes in alert generation because the alert estimation model is updated daily and the model learns to bring in adequate on-call personnel by simulating several alert generation rates.
- If estimation accuracy is good then risk is minimized and balanced between the 14-days.

Questions?

Sushil Jajodia

jajodia@gmu.edu

http://csis.gmu.edu/jajodia