NCSU Lablet Summary

Laurie Williams
Munindar Singh
Lablet Co-directors

Science of Security Lablet

NC STATE UNIVERSITY
Department of Computer Science

Faculty (20, 6 sub schools)

- Laurie Williams
- Munindar Singh
- Jon Doyle
- Rada Chirkova
- Emily Berglund
- Chris Mayhorn, Psychol
- Ninghui Li, Perdue
- Robert Proctor, Perdue
- Andy Meneely, RIT
- Jeff Carver, Alabama
- Dave Roberts
- Rob St. Amant
- Helen Gu
- Will Enck
- Emily Berglund, CE
- Mladen Vouk
- Huaiyu Dai, ECE
- Ehab El Shaer, UNC-Char
- Michael Reiter, UNC-CH
- Kevin Sullivan, UVa

Agenda: Missions

- Solve hard problems
- Build science of security community
- Develop and use scientifically rigorous methodology

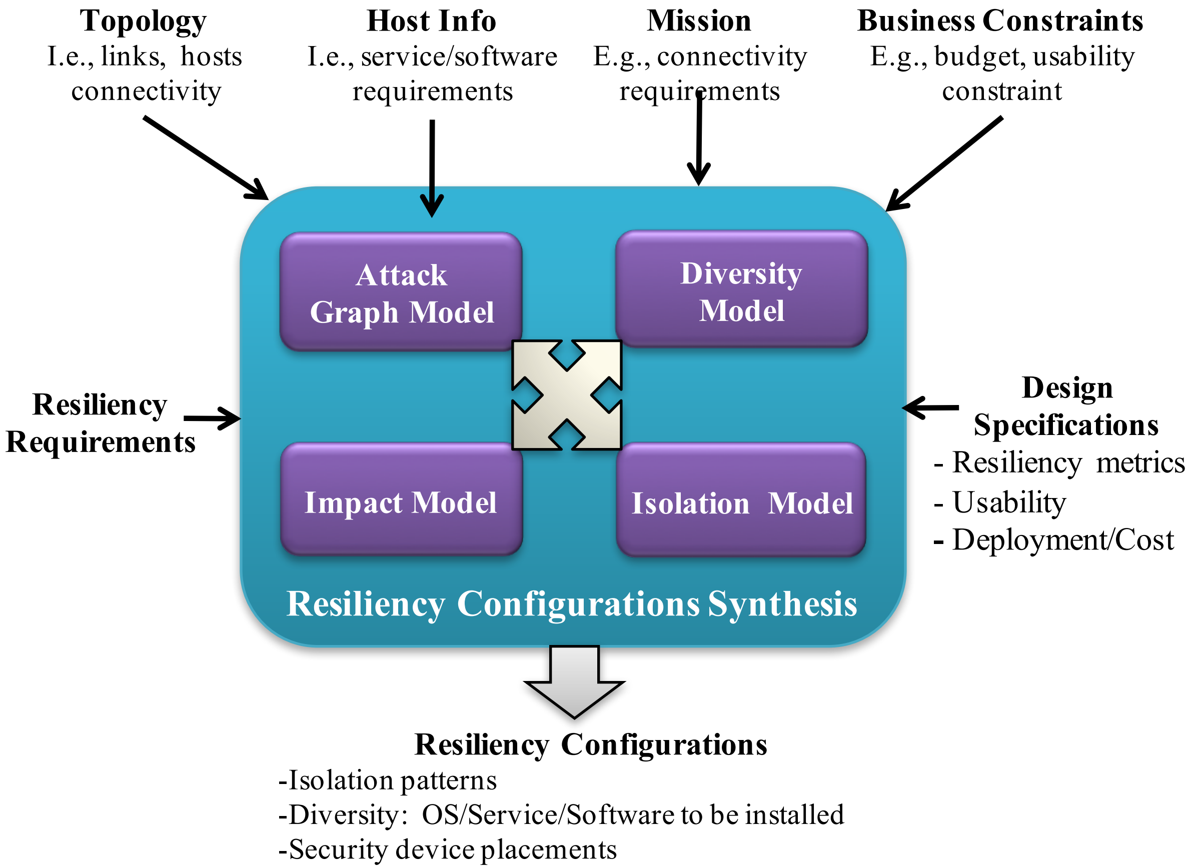
Hard Problems

- Resilience (5 projects, 6 PI)
  - Automated Synthesis of Resilient Architectures
  - Redundancy for Network Intrusion Prevention Systems (NIPS)
  - Smart Isolation in Large-Scale Production Computing Infrastructures
- Policy
- Humans
- Metrics

Automated Synthesis of Resilient Architectures (El Shaer, UNC-Char)

Before our work, we lacked adequate development of metrics and models for static and dynamic assessment of resilience of software. We have defined two resiliency metrics: (1) the isolation metric to quantify the counter-measure resistance on any path; and (2) the diversity metric to quantify the required attack vector by adversary based on the different disjoint attack surface due to OS and application diversity. We have also developed a formal framework for synthesizing configurations.

Automated Synthesis of Resiliency Configurations



Topology
I.e., links, hosts connectivity

Host Info
I.e., service/software requirements

Mission
E.g., connectivity requirements

Business Constraints
E.g., budget, usability constraint

Resiliency Requirements

Attack Graph Model

Diversity Model

Impact Model

Isolation Model

Resiliency Configurations Synthesis

Design Specifications
- Resiliency metrics
- Usability
- Deployment/Cost

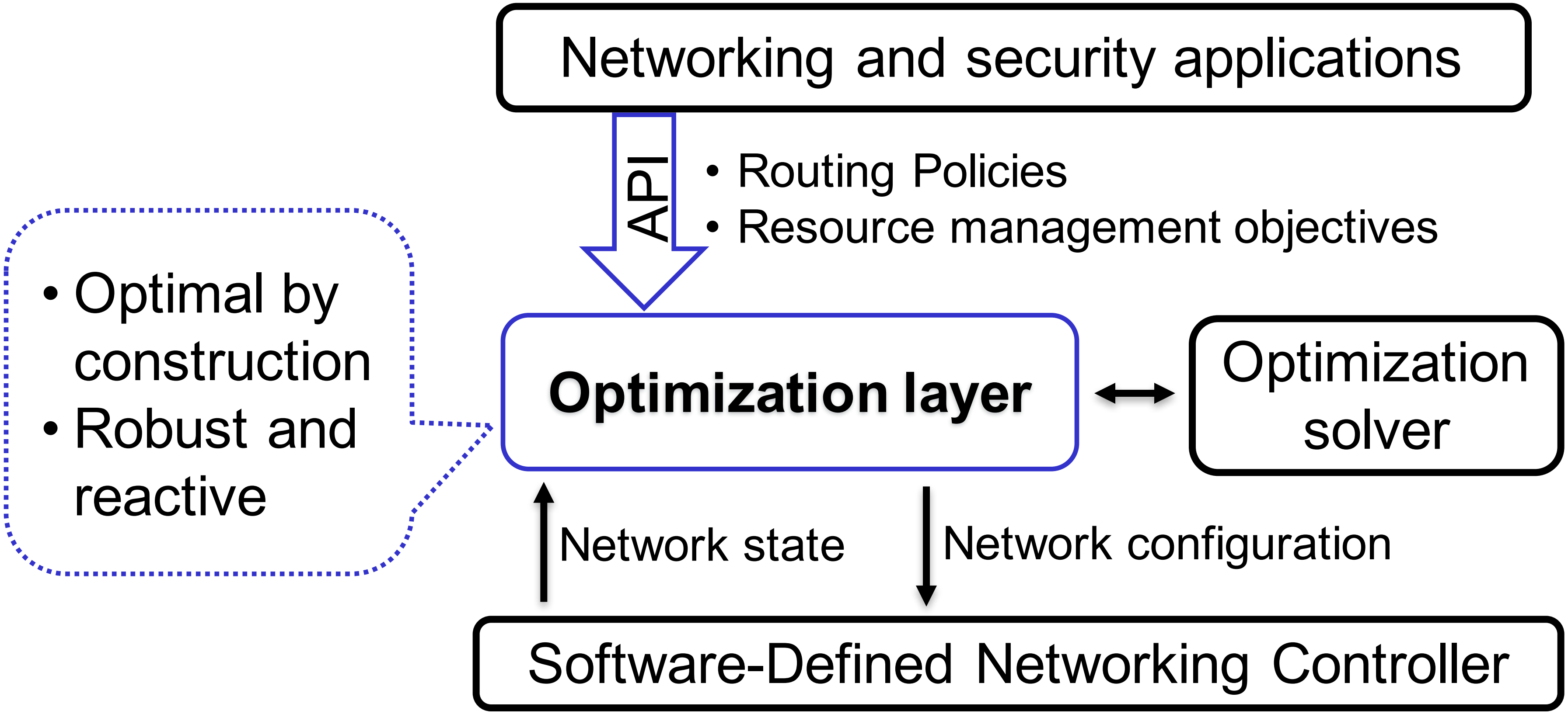
Resiliency Configurations
-Isolation patterns
-Diversity: OS/Service/Software to be installed
-Security device placements

Redundancy for Network Intrusion Prevention Systems (NIPS) (Reiter, UNC)

Prior to this research, SDN optimization applications were generated manually, requiring considerable expertise in algorithm design and networking to develop. Our research has now made it possible to generate such applications with far less expertise and effort, bringing new classes of security applications (such as our SNIPS scalable network intrusion-prevention architecture) within reach for network managers.

Network Optimization Layer

Goal: Improve resiliency of network security systems (e.g., intrusion prevention) by leveraging recent advances in networking

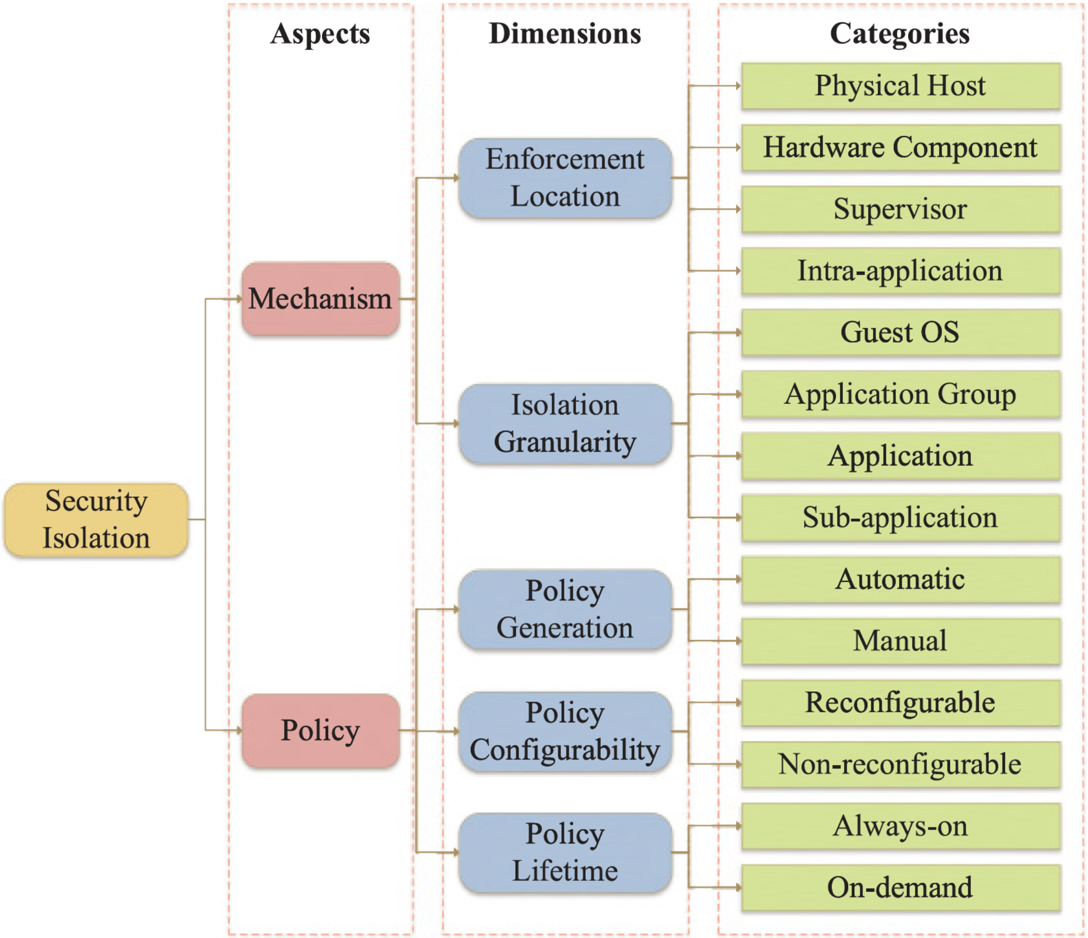

Networking and security applications
API
• Routing Policies
• Resource management objectives
• Optimal by construction
• Robust and reactive
Optimization layer
Optimization solver
Network state
Network configuration
Software-Defined Networking Controller

Science of Security Lablet

Smart Isolation in Large-scale Computing Infrastructures (Enck, Gu)

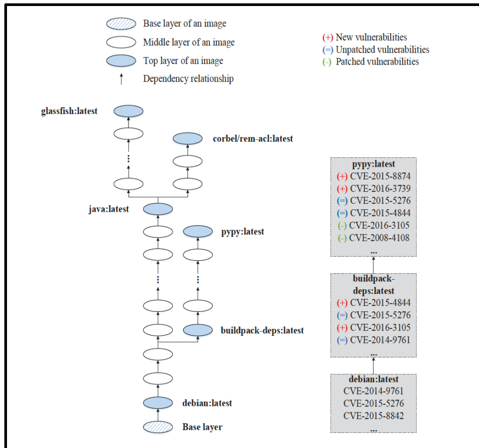
Prior to our work, security isolation was viewed as a static operation for achieving resilient architectures. Through a systematic survey, we characterized different facets of security isolation and identified the need for dynamic smart isolations. We further studied how smart isolation can be practically incorporated into systems, subsequently discovering novel mechanisms (e.g., lazy polyinstantiation) and properties (e.g., vulnerability inheritance) that enable better design of resilient architectures.

Isolation Taxonomy



Smart Isolation

- Goal:
  - Identify and explore primitives that enable dynamic and adaptive security isolation within systems
- Using information flow control as an adaptive policy for smart isolation
  - "Lazy polyinstantiation" creates a new instance of a resource only if needed, i.e., if there is no existing instance whose secrecy context matches the caller's.
- Vulnerability inheritance: isolation actually facilitates vulnerability spreading when people build new container images based on existing container images



(a) Image dependency graph example

(b) Vulnerability propagation results

Science of Security Lablet

Hard Problems

- Resilience
- Policy (5 projects, 7 PI)
  - Formal Specification and Analysis of Security-Critical Norms and Policies
  - Scientific Understanding of Policy Complexity
  - How Good is a Security Policy Against Real Breaches?
- Humans
- Metrics

Formal Specification and Analysis of Security-Critical Norms and Policies (Singh, Doyle, Berglund, Chirkova)

Before our work, policy approaches did not adequately characterize correctness requirements for secure collaboration and did not provide guidance for building privacy-aware software tools. We formalize norms as standards of correct collaborative behavior. The approaches we have developed tackle aspects of secure collaboration ignored in previous research.

Goal and Motivation

- Goal: To aid security analysts and administrators in verifying security properties of sociotechnical systems comprising users and computers through formally modeling and reasoning about such systems.
- Motivation
  - Security is inherently sociotechnical: need social **and** technical tier
  - Traditional computer science focuses on the technical tier
  - Traditional social science focuses on the social tier
  - How can we formalize elements of the social architecture on par with the technical architecture so as to computationally reason about them cohesively?
  - Key challenge: people are autonomous; attempts at regimenting user actions frequently backfire and lead to ad hoc workarounds
- Unifying construct: norms as standards of correct (collaborative) behavior
- Specific questions concern representations for norms; consistency of norms; how norms delimit actions; how we may elicit and extract norms; how are norms adopted

Runtime Reasoning about Norm Conflicts

- How can we detect norm conflicts at runtime to determine what action is normatively appropriate, so a norm-aware agent can act appropriately?
  - Disclose patient data (norm to save life)
  - Do not disclose patient's data without consent
- Contributions:
  - Representation for norms based on non-monotonic logics that captures conditional dominance
  - Reasoning algorithm to determine whether unique consistent non-dominated set of norms can be satisfied

Computing Norm States over Information Stores

- How can we determine states (such as satisfaction and violation) defined in a norm lifecycle model based on an underlying event store?
- Contributions
  - A new language to express norms that supports specifying norms that refer to other norms; over low-level information stores
  - A formal semantics for the language
  - A mapping from norm schemas to relational (SQL) queries to compute the lifecycle states of norm instances from underlying relational information stores

Scientific Understanding of Policy Complexity (Li, Proctor, Perdue)

Before our work, some fundamental reasons why certain policies are complex and error-prone were not well understood. Now we identified lack of adequate abstractions in policy languages as a main reason for policy complexity and introduced ways to address the problem in some policy languages.

Firewall Policy Language

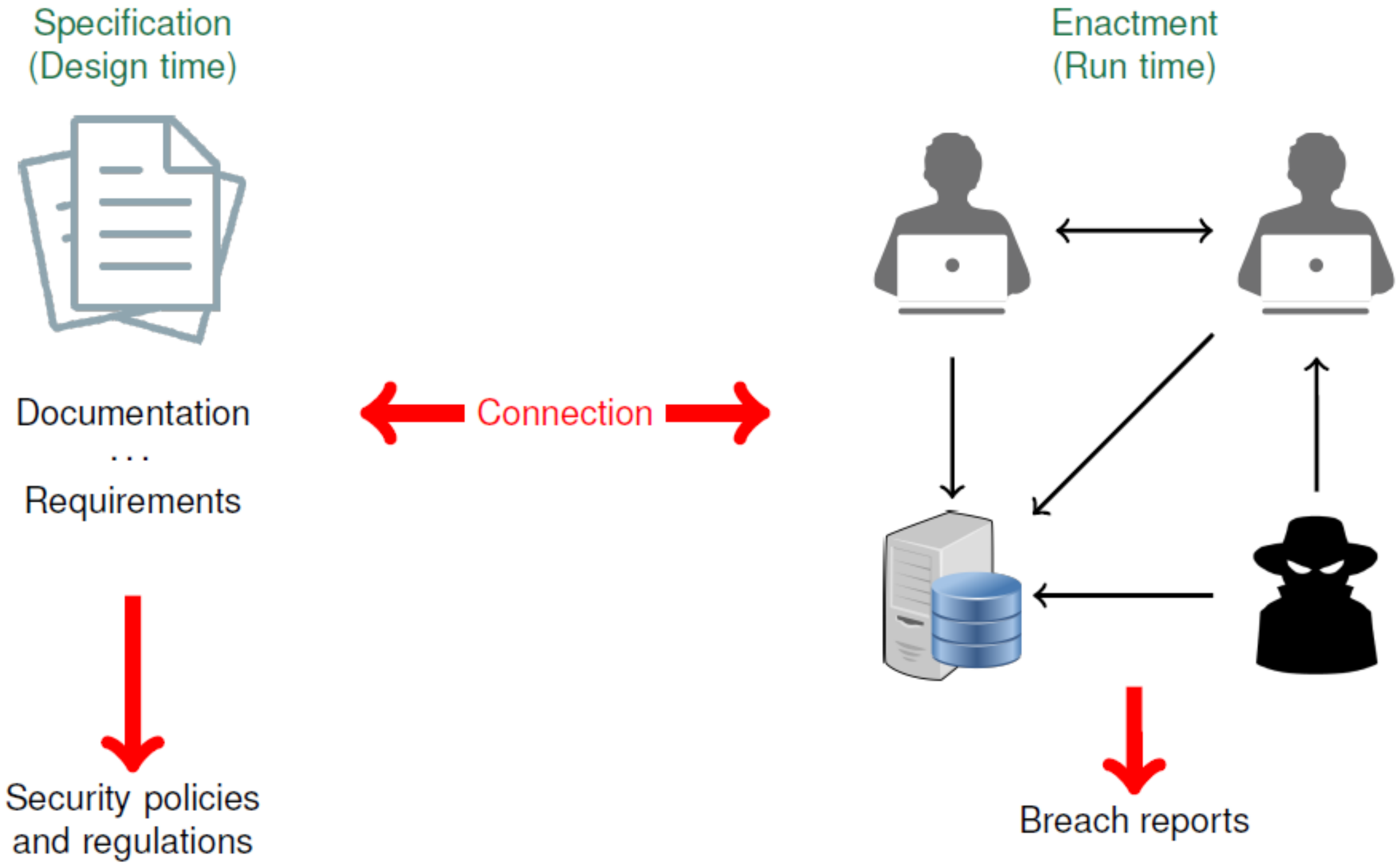
- Firewall policies are complex and error-prone
  - Current Firewall policy language is too low-level and lacks suitable abstraction/organizing principle for expressing complex policies
- We developed a firewall policy language
  - that enables modular policies
  - and that provides abstractions that help write/understand policies

TMFL: A Tri-Modular Firewall Language

- Concept of primary address:
  - Choose one of source and destination IP as the primary address (destination IP is often the better choice)
- A policy is organized into three kinds of modules
  - Primary modules, Auxiliary modules, Template modules
- Expressing real-world policies in TMFL naturally reveals previous unknown configuration errors

How Good is a Security Policy Against Real Breaches? (Singh, Williams)

Before our work, there was no formal way of connecting security policies and regulations with real breach data. We provided a systematic and repeatable methodology to represent policies as norms, breaches as norm violations, and identify gaps/holes in policies accordingly.



Specification
(Design time)

Documentation
. . .
Requirements

Connection

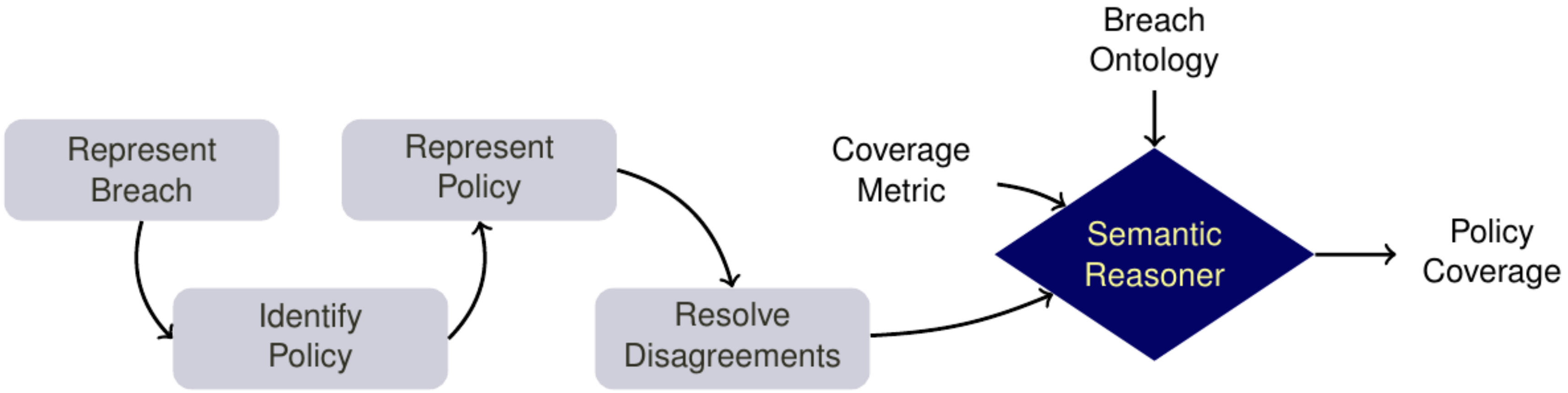
Security policies
and regulations

Enactment
(Run time)

Breach reports

Goal & Methodology

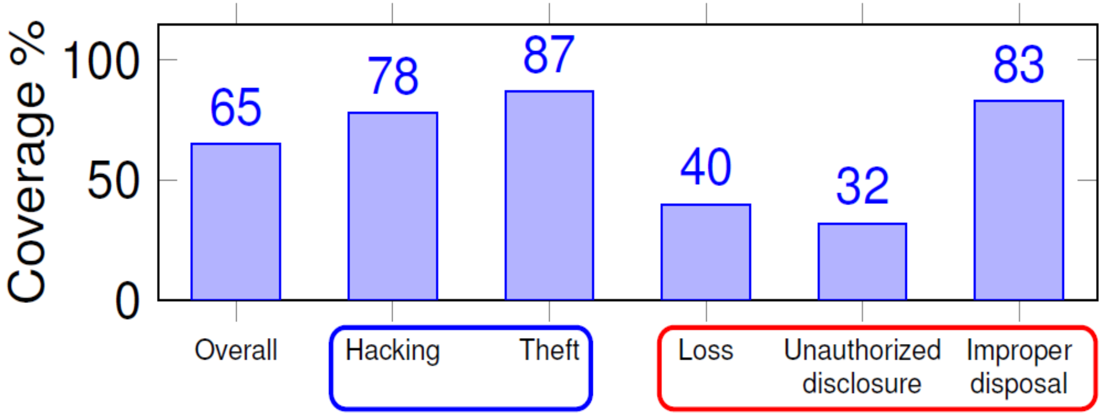
Goal: To help analysts identify the gaps/holes in security policies by developing a systematic process for formal investigation of reported breaches



Data & Results

- Investigated 1,577 breaches reported by HHS [1]
- 44% accidental misuses and 56% malicious misuses
- Our results corroborate findings in recent cybersecurity reports [2, 3]

- Computed coverage of various breach categories by HIPAA
- Better coverage for malicious misuses than accidental misuses



[1] Notice to the Secretary of HHS breach of unsecured protected health information affecting 500 or more individuals: https://ocrportal.hhs.gov/ocr/breach/
[2] The United States Department of Defense (DoD). Cybersecurity culture and compliance initiative. 2015
[3] The healthcare information and management systems society (HIMSS) cybersecurity study. 2016

Hard Problems

- Resilience
- Policy
- Humans (3 projects, 4 PI)
  - A Human Information Processing Analysis of Online Deception Detection
  - Leveraging the Effects of Cognitive Function on Input Device Analysis to Improve Security
- Metrics

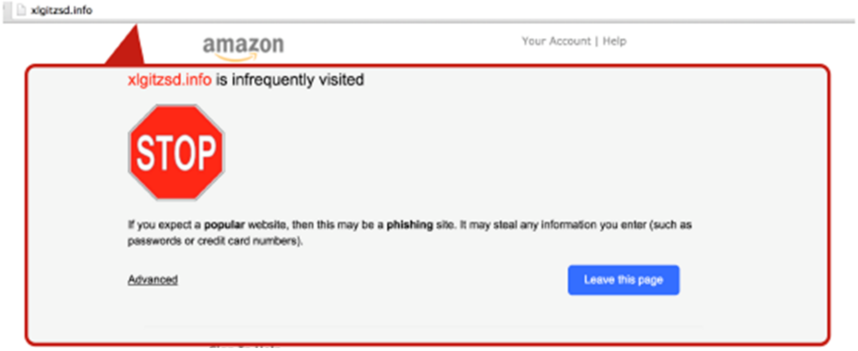
A Human Information Processing Analysis of Online Deception Detection (Proctor, Li)

A significant percentage of users still fall for phishing attacks when a warning is presented. Through our work, we established that training increases warning compliance rate and enables more accurate identification of phishing webpages in the absence of a warning.

Science of Security
Lablet

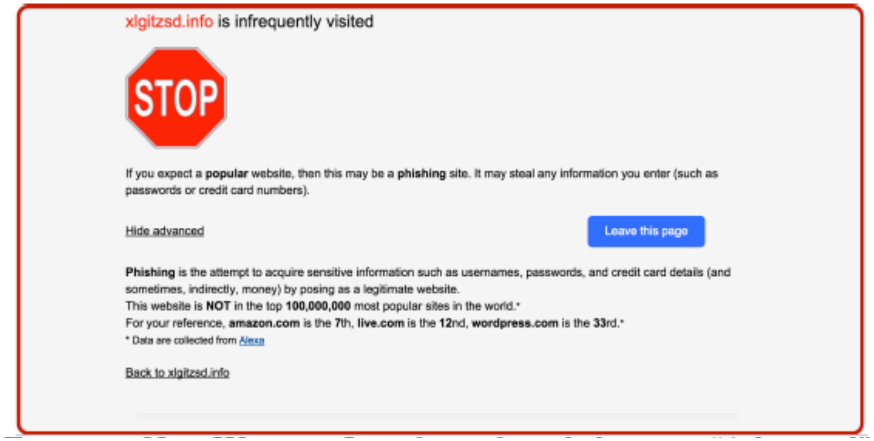
NC STATE UNIVERSITY
Department of Computer Science

Use of phishing training to improve security warning compliance



xlgitzsd.info is infrequently visited

If you expect a **popular** website, then this may be a **phishing** site. It may steal any information you enter (such as passwords or credit card numbers).



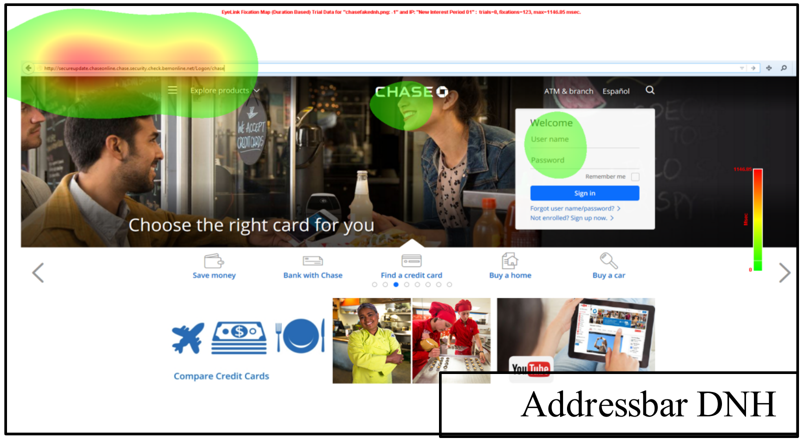
xlgitzsd.info is infrequently visited

If you expect a **popular** website, then this may be a **phishing** site. It may steal any information you enter (such as passwords or credit card numbers).

Phishing is the attempt to acquire sensitive information such as usernames, passwords, and credit card details (and sometimes, indirectly, money) by posing as a legitimate website.
This website is **NOT** in the top **100,000,000** most popular sites in the world.*
For your reference, **amazon.com** is the **7**th, **live.com** is the **12**nd, **wordpress.com** is the **33**rd.*
* Data are collected from Alexa

Back to xlgitzsd.info

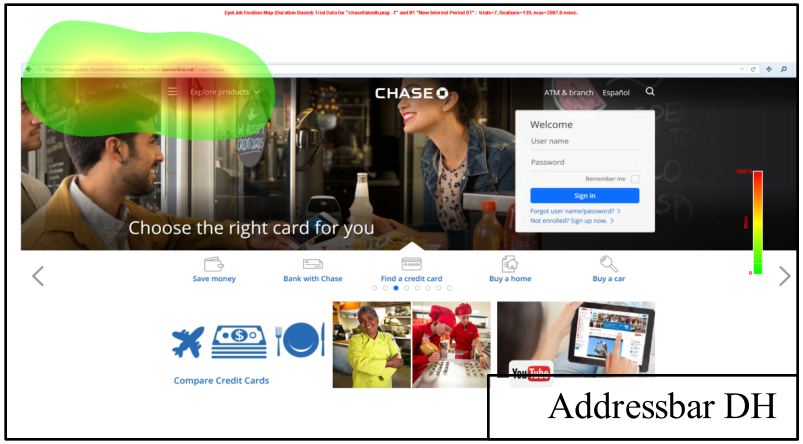
Domain highlighting



EyeLink Fixation Map (Duration Based) Trial Data for "chasefakednh.png: -1" and IP: "New Interest Period 01" : trials=8, fixations=123, max=1146.85 msec.

http://secureupdate.chaseonline.chase.security.check.bemonline.net/Logon/chase

Addressbar DNH

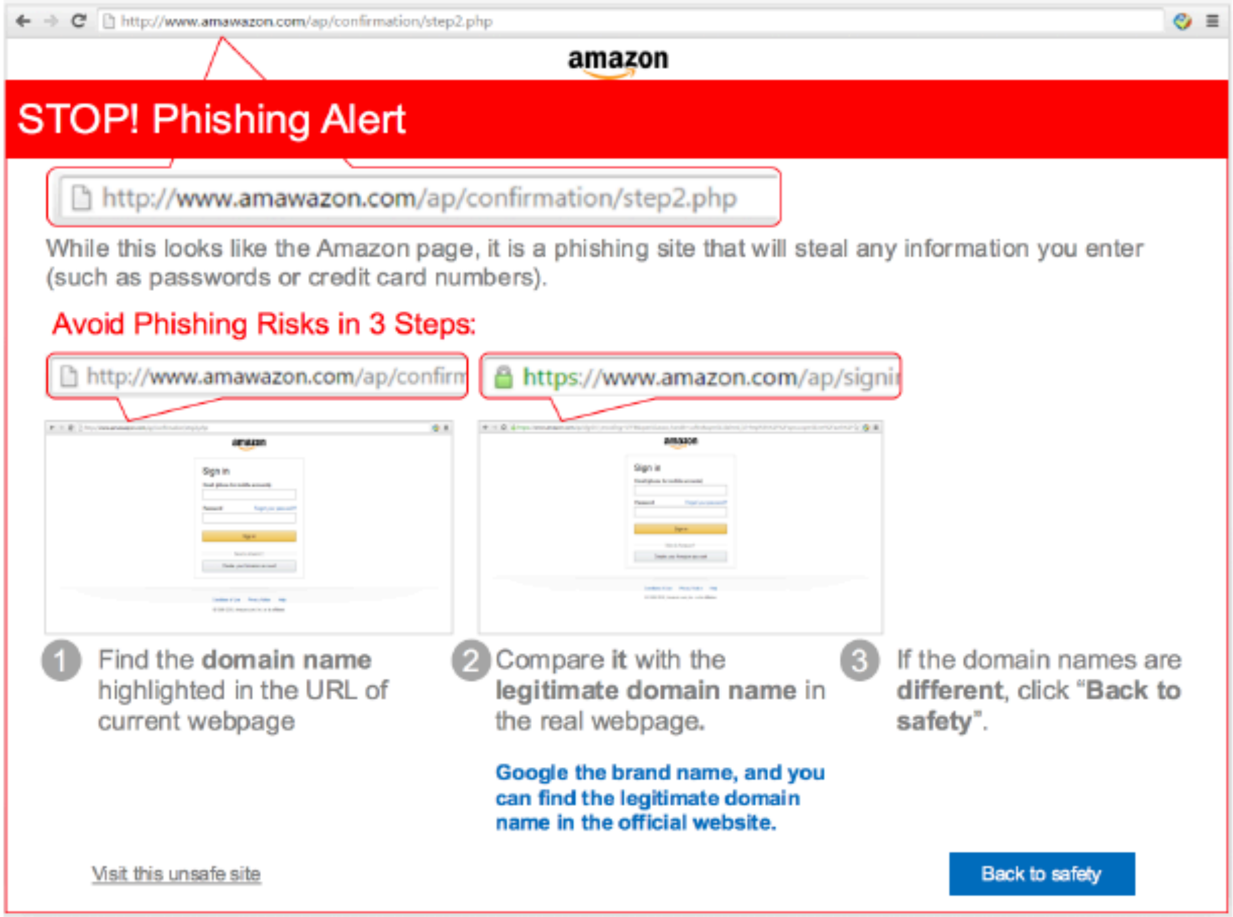


EyeLink Fixation Map (Duration Based) Trial Data for "chasefakedh.png: -1" and IP: "New Interest Period 01" : trials=7, fixations=139, max=2087.8 msec.

http://secureupdate.chaseonline.chase.security.check.bemonline.net/Logon/chase

Addressbar DH

Embedding anti-phishing training within cybersecurity warnings



http://www.amawazon.com/ap/confirmation/step2.php

amazon

STOP! Phishing Alert

http://www.amawazon.com/ap/confirmation/step2.php

While this looks like the Amazon page, it is a phishing site that will steal any information you enter (such as passwords or credit card numbers).

Avoid Phishing Risks in 3 Steps:

http://www.amawazon.com/ap/confirm

https://www.amazon.com/ap/signi

1 Find the **domain name** highlighted in the URL of current webpage

2 Compare **it** with the **legitimate domain name** in the real webpage.

Google the brand name, and you can find the legitimate domain name in the official website.

3 If the domain names are **different**, click "**Back to safety**".

Leveraging the Effects of Cognitive Function on Input Device Analysis to Improve Security (Roberts, St. Amant)

Before our work, approaches to distinguish authorized users of computer systems from automated bots were less mature. This project has demonstrated the initial feasibility of Human Subtlety Proofs (HSP), which probe human cognition through subtle task modification to provide the increased security of interactive proofs with the lowered cognitive burden of observational proofs.

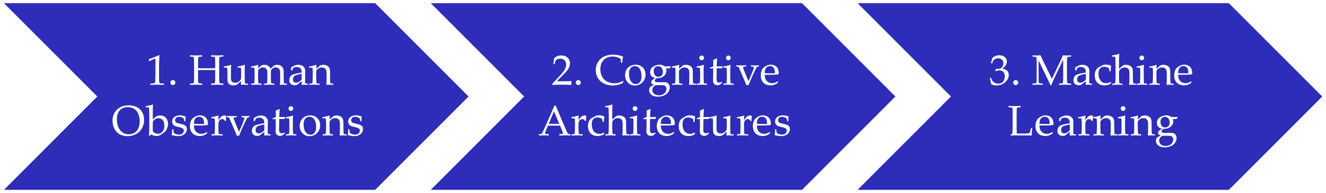
Distinguishing human users from bots

Human Subtlety Proofs are…

- able to differentiate humans from bots.
- as secure as Human Interactive Proofs.
- as unobtrusive as Human Observational Proofs.
- able to use cognitive models to explain behavior.
- applicable in almost any interactive system.

Is characteristically human behavior detectable in input device usage?

Technical approach:



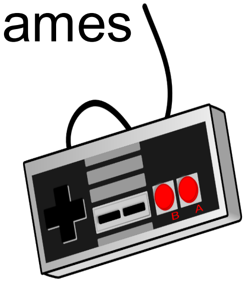
1. Human Observations

2. Cognitive Architectures

3. Machine Learning

1. Human observations as ground truth.
2. Cognitive models for explanation.
3. Machine learning for classification.

Testbed: Casual games

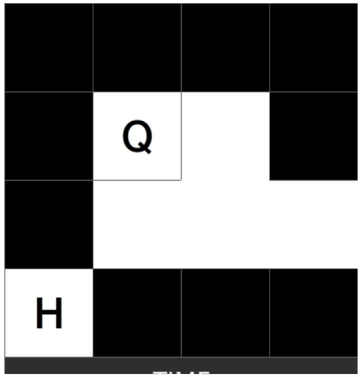


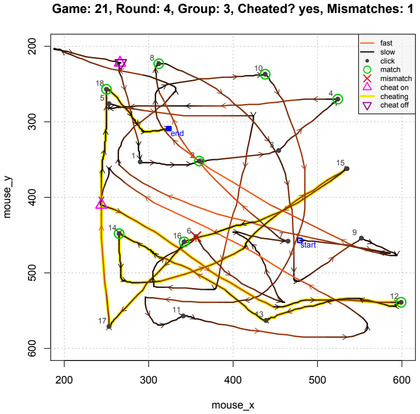
Science of Security Lablet

NC STATE UNIVERSITY
Department of Computer Science

Casual games as testbeds

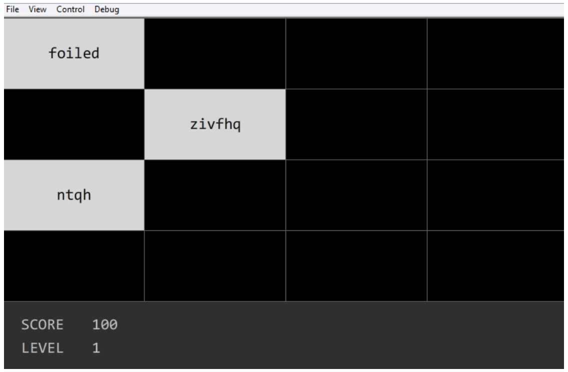
Concentration game





- Cognitive modeling: Player speed vs accuracy.
- Statistical modeling: 70% accuracy in identifying cheating/deception.

Typing game



foiled

zivfhq

ntqh

SCORE 100

LEVEL 1

- Cognitive modeling: Transcription typing of words, non-words.
- Findings: Speed improves with practice and familiarity—typing accuracy does not.

Science of Security Lablet

NC STATE UNIVERSITY

Department of Computer Science

Hard Problems

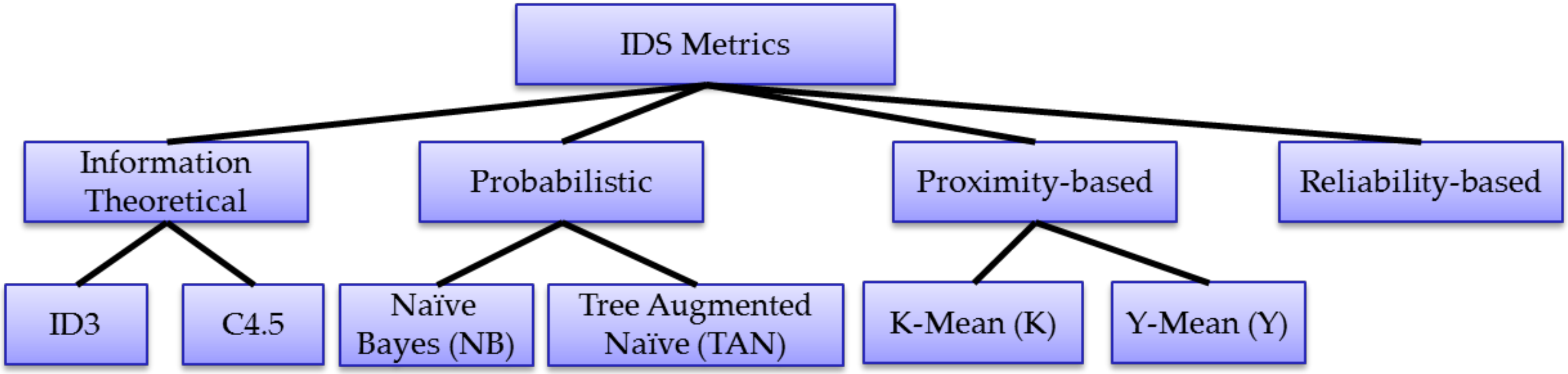
- Resilience
- Policy
- Humans
- Metrics (3 projects, 4 PI)
  - – Systemization of Knowledge of from Intrusion Detection Models
  - – Attack Surface and Defense in Depth Metrics

Systemization of Knowledge of from Intrusion Detection Models

(Dai, Meneely)

- Before our work, there were hundreds of disparate publications about intrusion-detection systems, each with varying methods and evaluation approaches. Our work has led to a taxonomy to compare those studies and to systematize that knowledge.

Systematization of Metrics in Intrusion Detection Systems



IDS Metrics

Information Theoretical

Probabilistic

Proximity-based

Reliability-based

ID3

C4.5

Naïve Bayes (NB)

Tree Augmented Naïve (TAN)

K-Mean (K)

Y-Mean (Y)

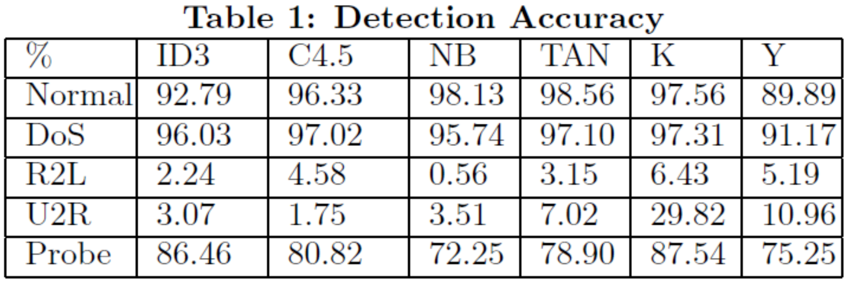


Table 1: Detection Accuracy

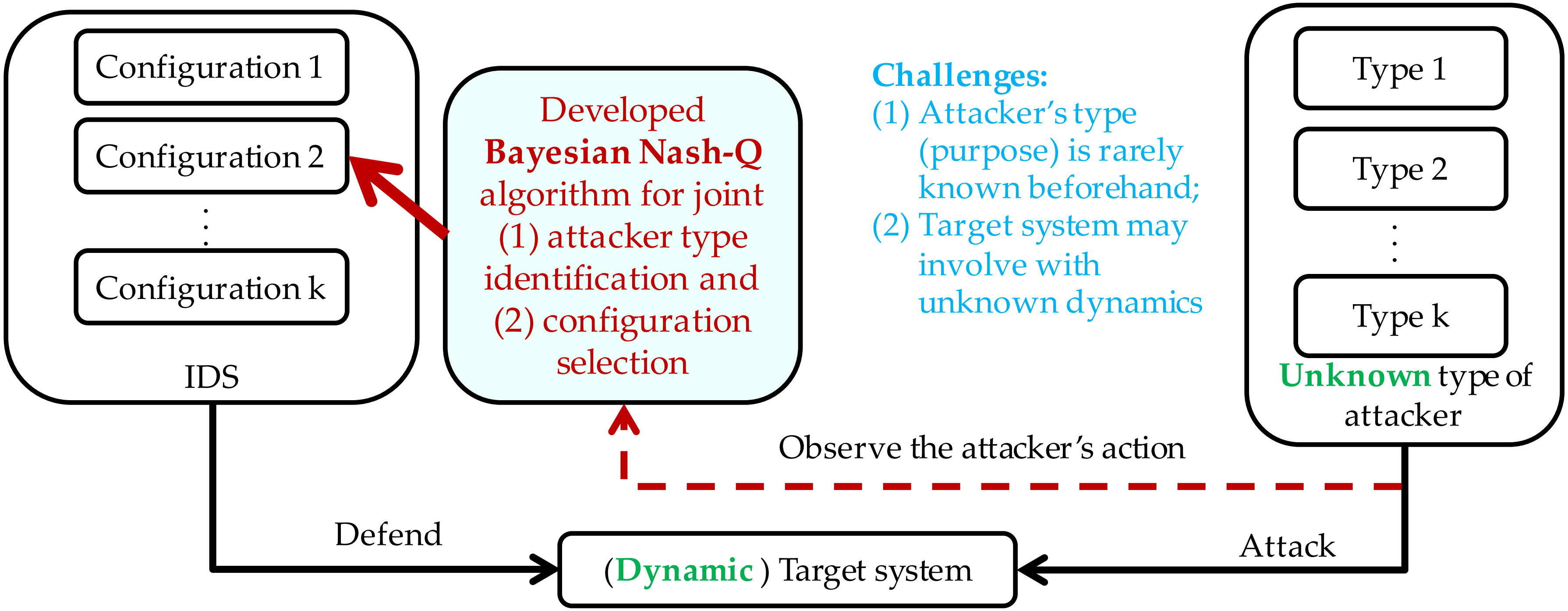
%	ID3	C4.5	NB	TAN	K	Y
Normal	92.79	96.33	98.13	98.56	97.56	89.89
DoS	96.03	97.02	95.74	97.10	97.31	91.17
R2L	2.24	4.58	0.56	3.15	6.43	5.19
U2R	3.07	1.75	3.51	7.02	29.82	10.96
Probe	86.46	80.82	72.25	78.90	87.54	75.25

Science of Security Lablet

NC STATE UNIVERSITY
Department of Computer Science

Dynamic IDS Configuration in the Presence of Intruder Type Uncertainty

IDS configuration: Due to limited resource, IDS needs to be properly configured to achieve the best performance


Configuration 1
Configuration 2
Configuration k
IDS
Developed **Bayesian Nash-Q** algorithm for joint (1) attacker type identification and (2) configuration selection
Type 1
Type 2
Type k
Unknown type of attacker
Observe the attacker's action
Defend
Attack
(**Dynamic**) Target system

Challenges:

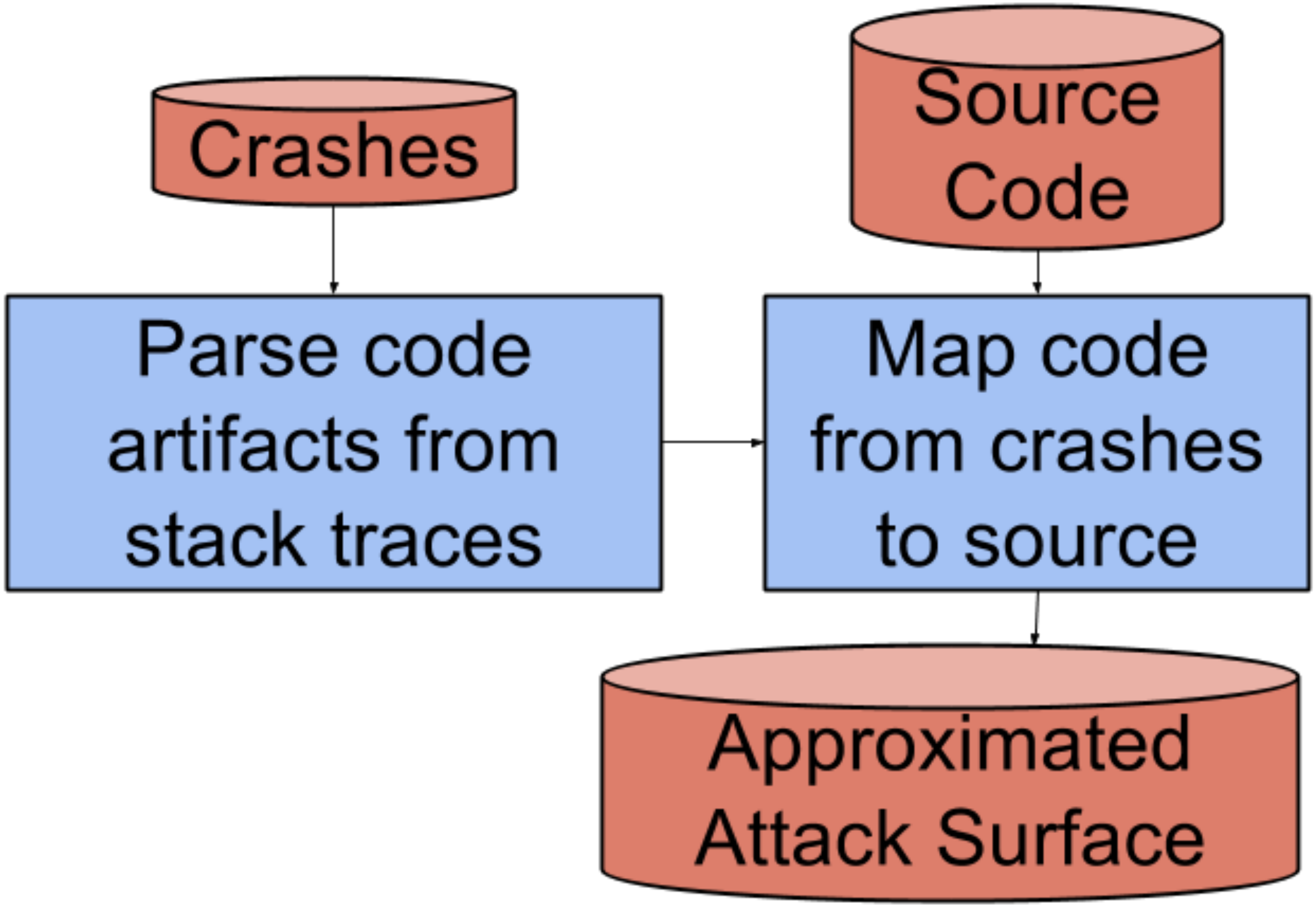
(1) Attacker's type (purpose) is rarely known beforehand;

(2) Target system may involve with unknown dynamics

Attack Surface and Defense in Depth Metrics (Williams, Meneely)

Before our work, we did not know if vulnerabilities could be predicted more effectively by incorporating attack-surface data. Today, we know that we can improve the prediction quality of models when attacker behavior is simulated via random walks on the call graph starting from the attack surface and via approximating the attack surface by recording all the files that appear on stack traces from crash dumps.

Methodology - RASA


Crashes
Source Code
Parse code artifacts from stack traces
Map code from crashes to source
Approximated Attack Surface

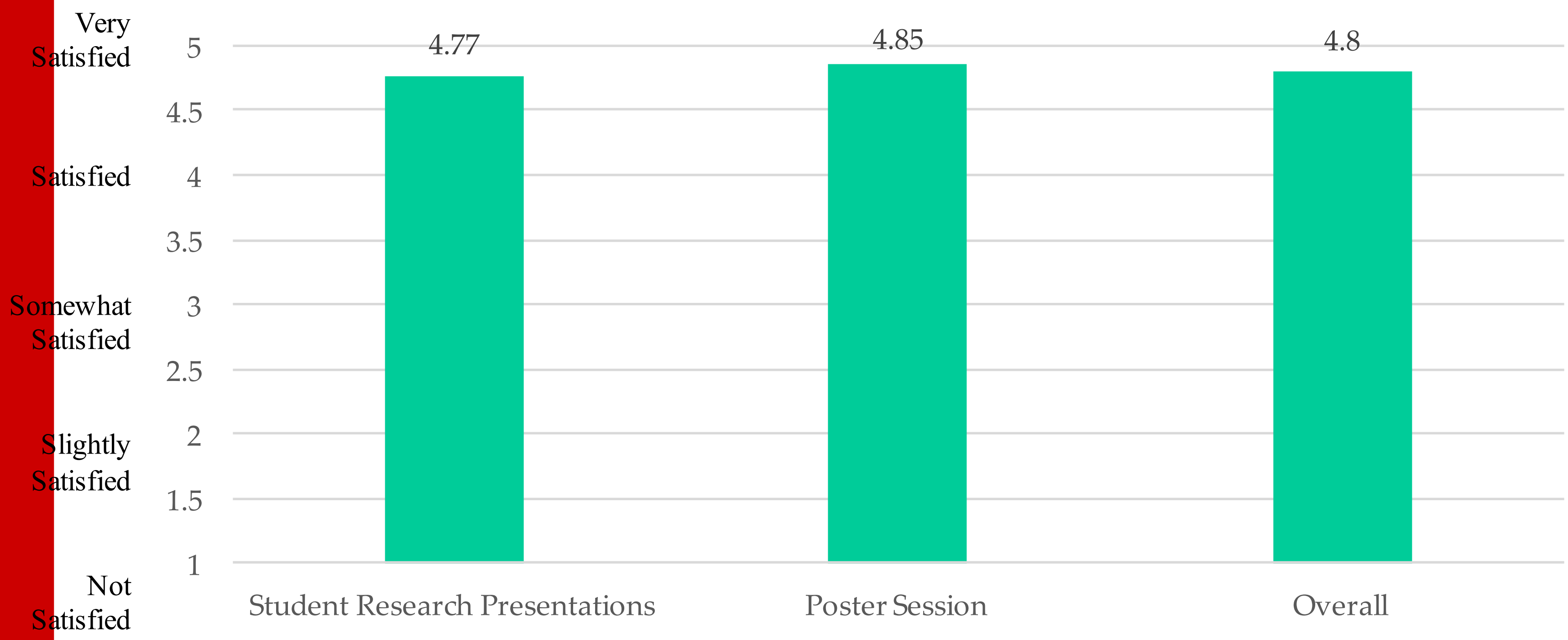
Agenda: Missions

- Solve hard problems
- Build science of security community
- Develop and use scientifically rigorous methodology

Community Building

- Annual community day with industry and government organization
- Annual summer workshop
- HotSoS 2014
- HotSoS 2018: April 10-11

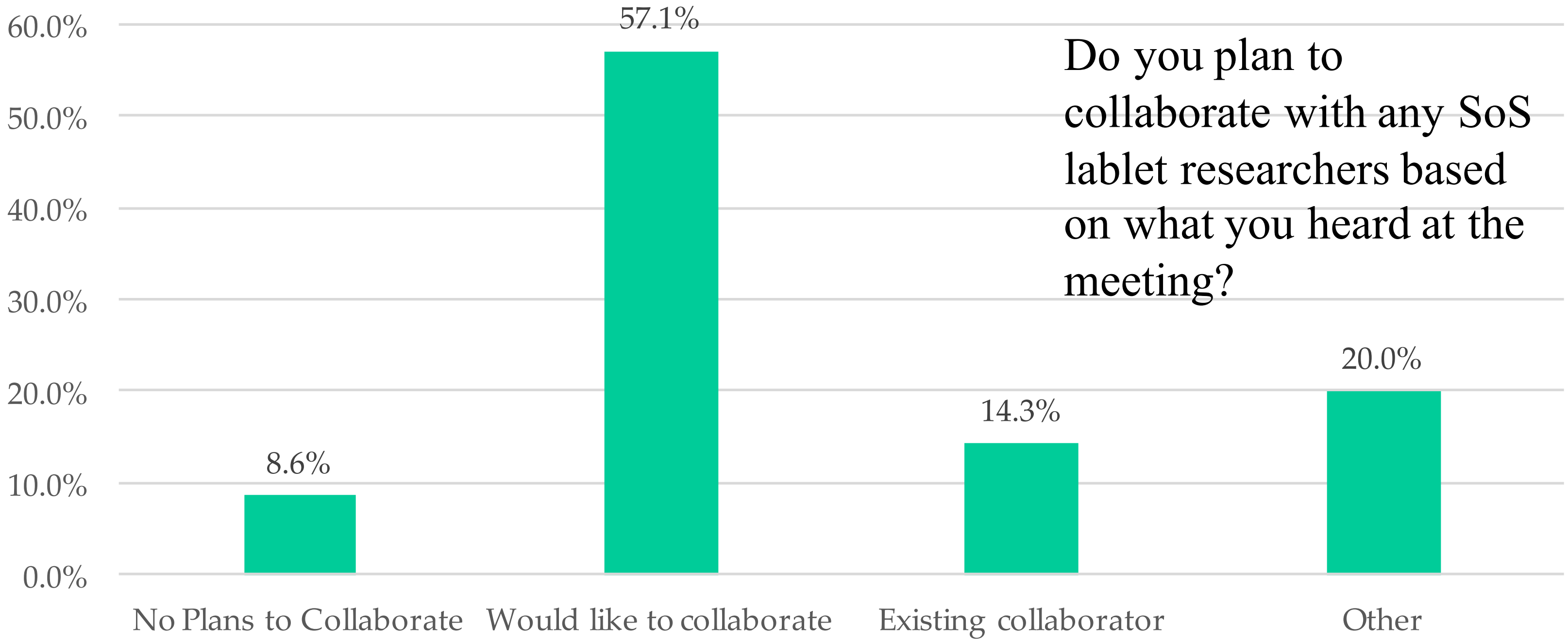
Satisfaction with Community Meetings



Science of Security Lablet

NC STATE UNIVERSITY
Department of Computer Science

Community Day Attendee Intent to Collaborate

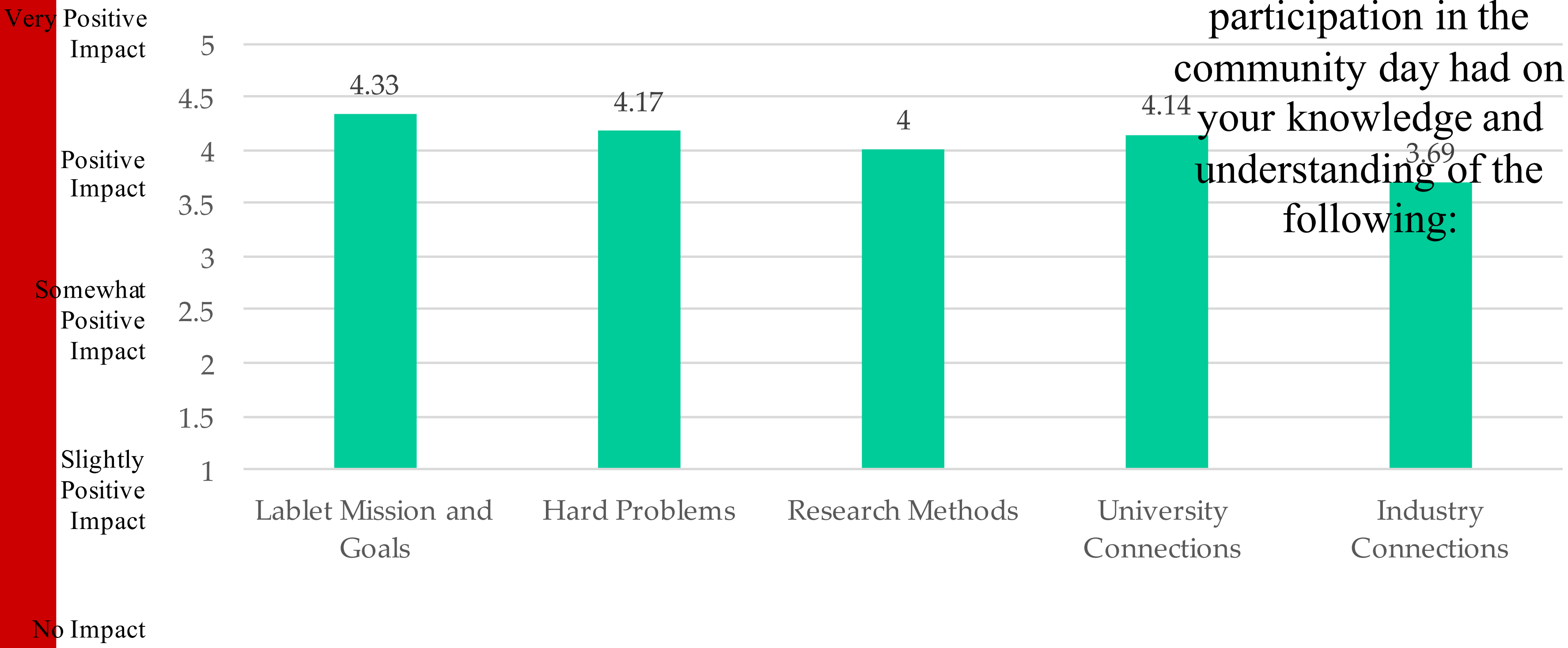


Do you plan to collaborate with any SoS lablet researchers based on what you heard at the meeting?

Science of Security Lablet

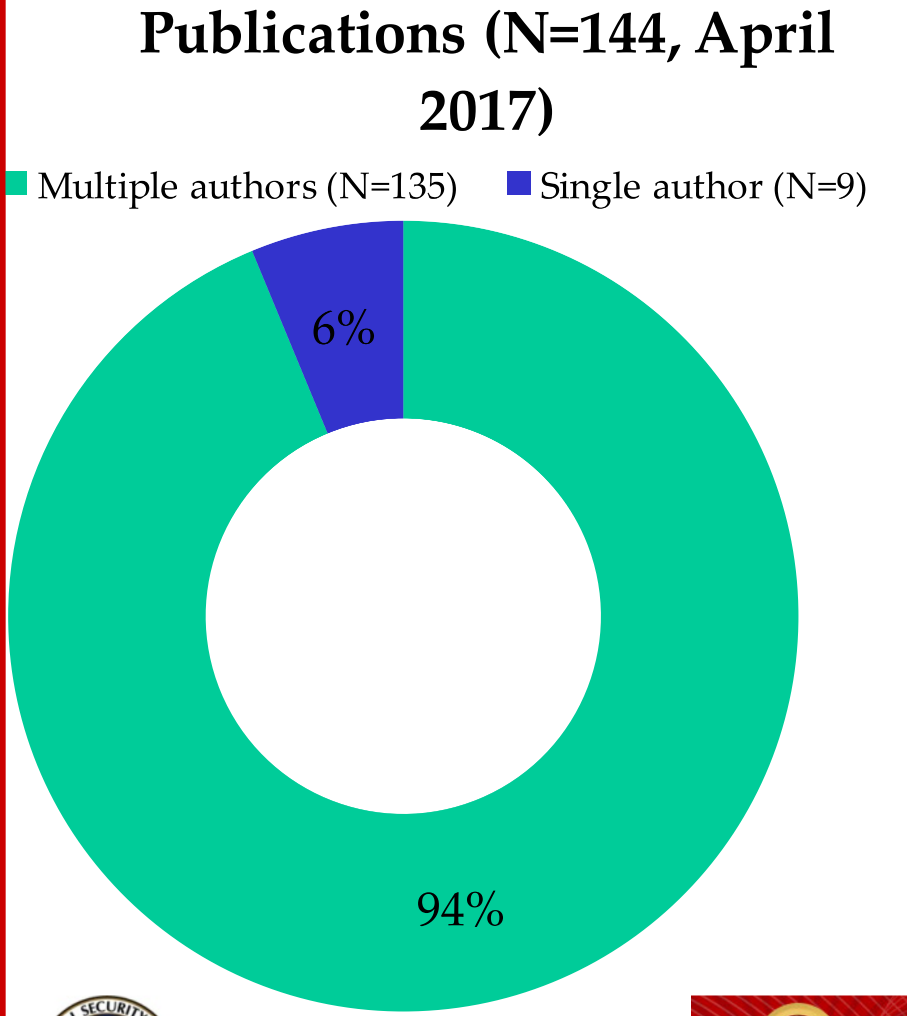
NC STATE UNIVERSITY
Department of Computer Science

Community Day Impact

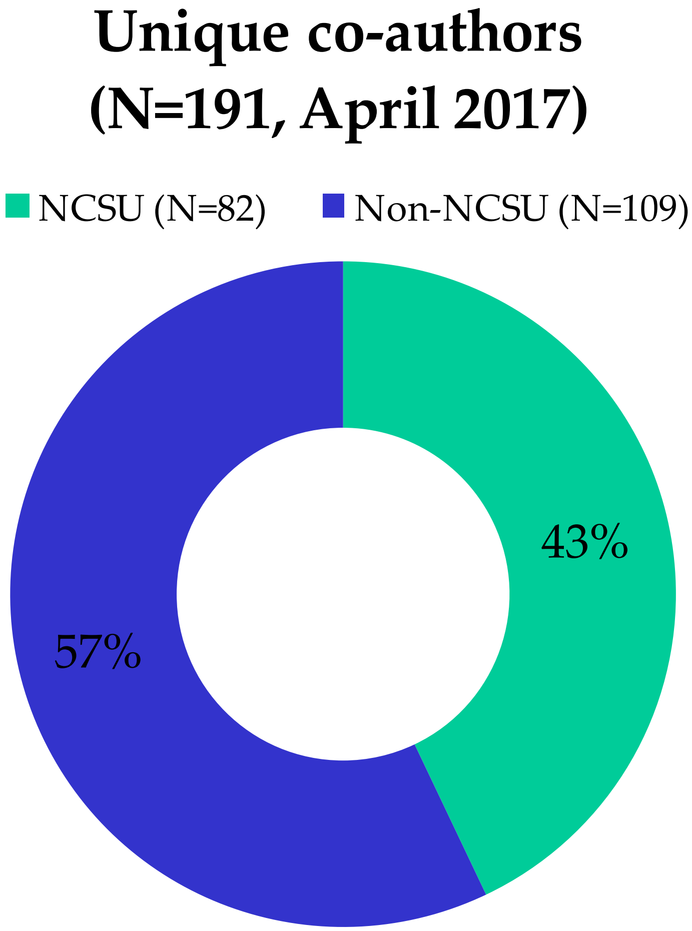


What impact has participation in the community day had on your knowledge and understanding of the following:

Co-authorship Analysis: Results

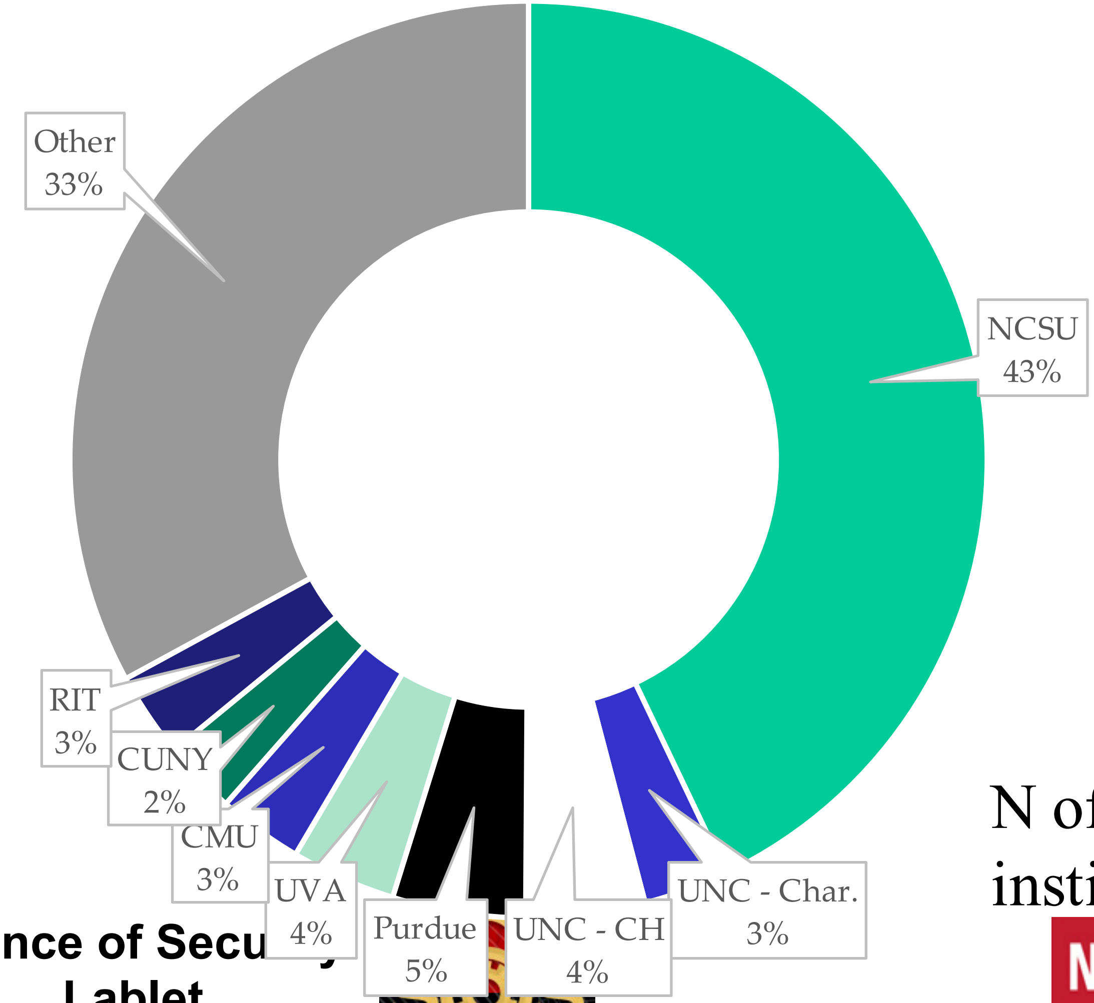


Publications (N=144, April 2017)



Unique co-authors (N=191, April 2017)

Co-authorship Analysis: Multi-Institutional Results

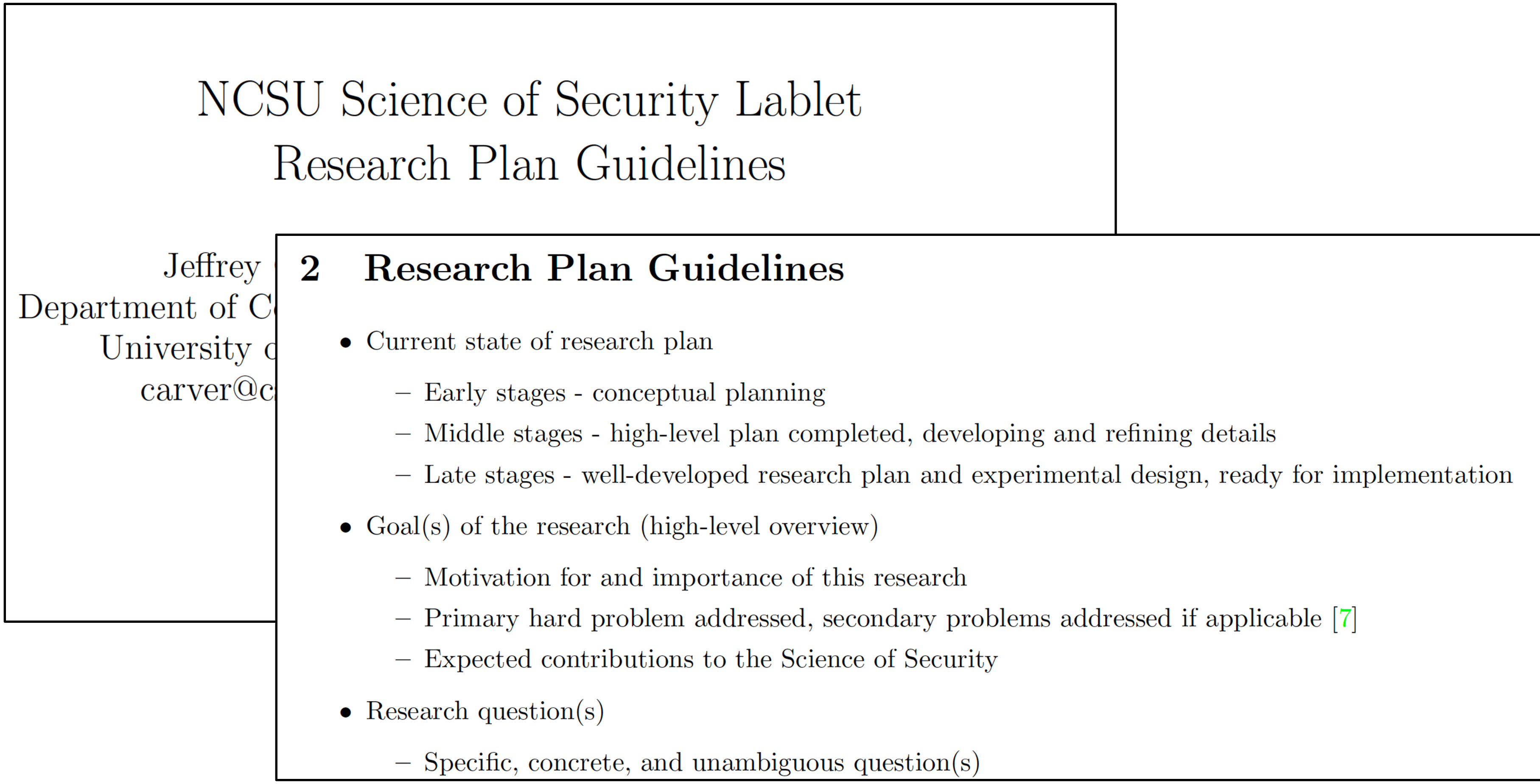


N of co-author institutions = 45

Agenda: Missions

- Solve hard problems
- Build science of security community
- Develop and use scientifically rigorous methodology

Weekly seminar: Research Plan



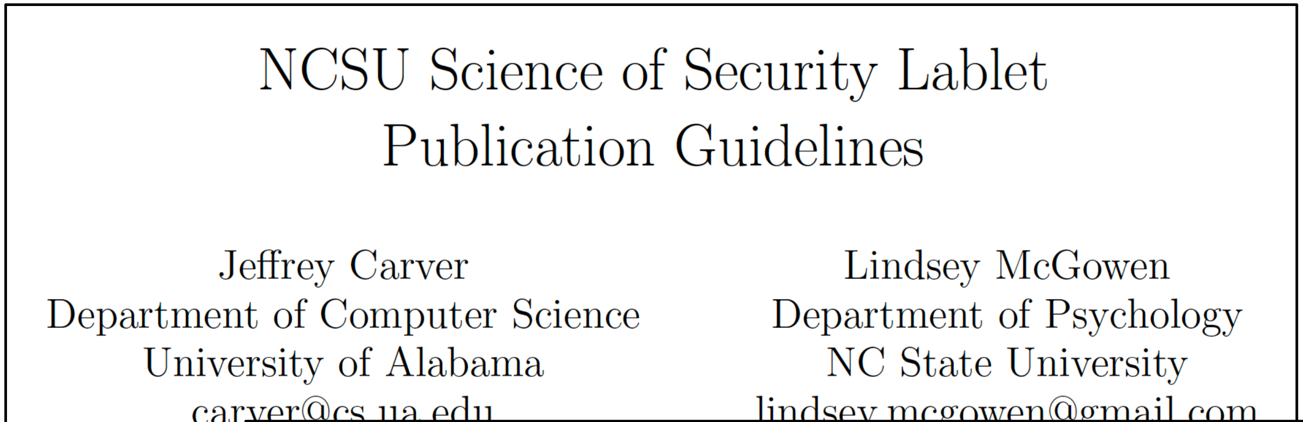
NCSU Science of Security Lablet
Research Plan Guidelines

Jeffrey
Department of C
University c
carver@c

2 Research Plan Guidelines

- Current state of research plan
  - Early stages - conceptual planning
  - Middle stages - high-level plan completed, developing and refining details
  - Late stages - well-developed research plan and experimental design, ready for implementation
- Goal(s) of the research (high-level overview)
  - Motivation for and importance of this research
  - Primary hard problem addressed, secondary problems addressed if applicable [7]
  - Expected contributions to the Science of Security
- Research question(s)
  - Specific, concrete, and unambiguous question(s)

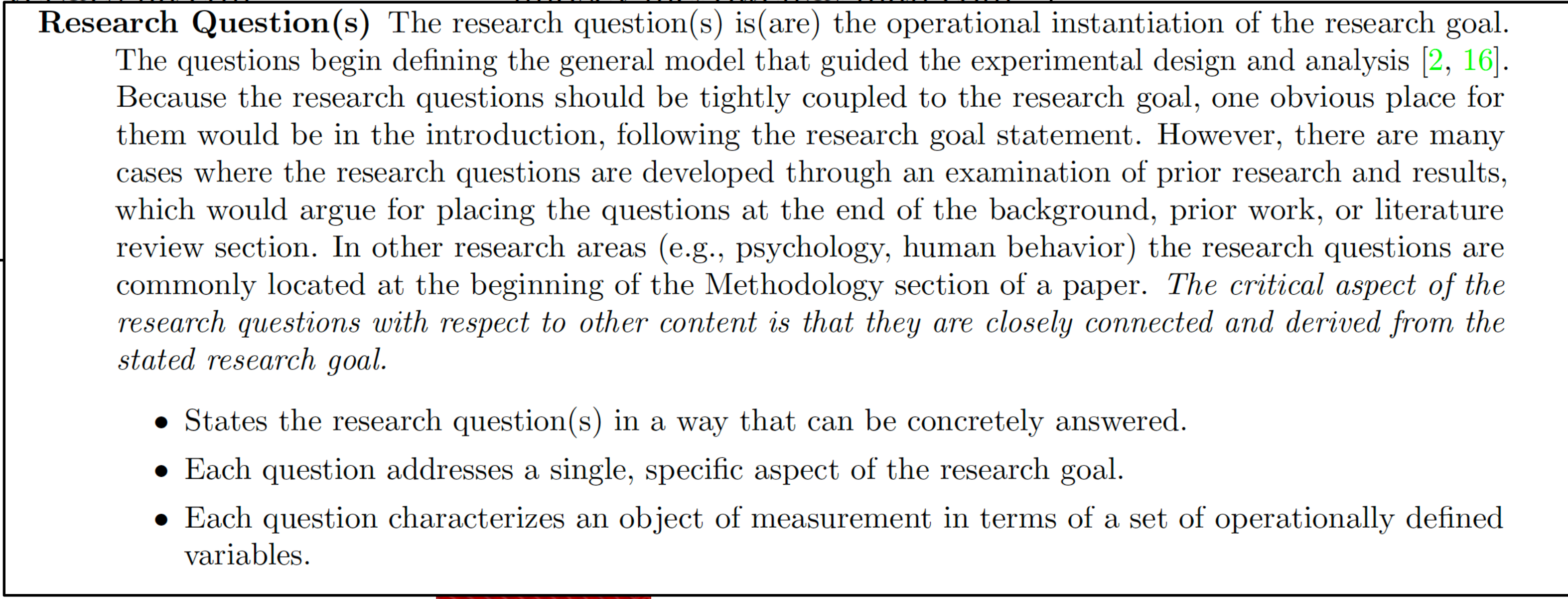
Weekly seminar: Paper



NCSU Science of Security Lablet
Publication Guidelines

Jeffrey Carver
Department of Computer Science
University of Alabama
carver@cs.ua.edu

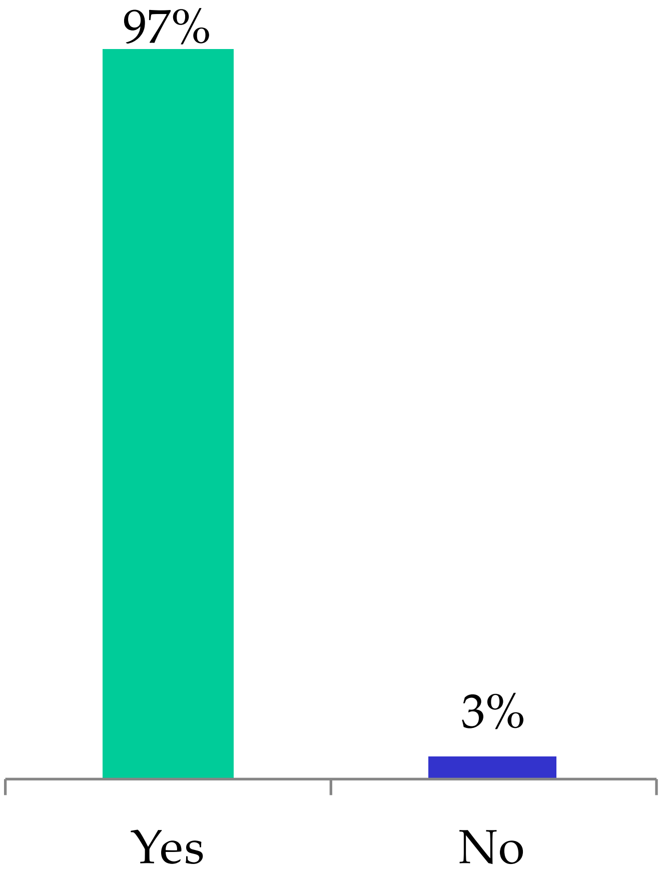
Lindsey McGowen
Department of Psychology
NC State University
lindsey.mcgowen@gmail.com



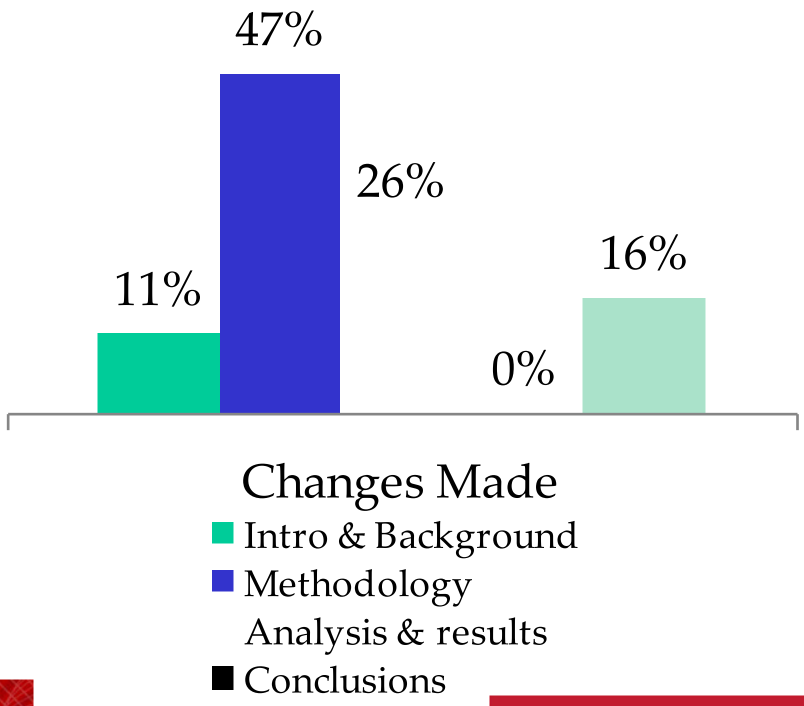
Research Question(s) The research question(s) is(are) the operational instantiation of the research goal. The questions begin defining the general model that guided the experimental design and analysis [2, 16]. Because the research questions should be tightly coupled to the research goal, one obvious place for them would be in the introduction, following the research goal statement. However, there are many cases where the research questions are developed through an examination of prior research and results, which would argue for placing the questions at the end of the background, prior work, or literature review section. In other research areas (e.g., psychology, human behavior) the research questions are commonly located at the beginning of the Methodology section of a paper. *The critical aspect of the research questions with respect to other content is that they are closely connected and derived from the stated research goal.*

- States the research question(s) in a way that can be concretely answered.
- Each question addresses a single, specific aspect of the research goal.
- Each question characterizes an object of measurement in terms of a set of operationally defined variables.

Presenter Feedback Utilization



Feedback Used to Make Changes to Research



Science of Security Lablet

NC STATE UNIVERSITY
Department of Computer Science

Research Methods: Analysis of Published Papers

- Goal: Analyze papers from top security conferences to characterize the reporting from the perspective of a science of security

- We developed and refined a series of rubrics, based on information published in the general research literature, to analyze whether papers contain information important for understand, replication, meta-analysis, and theory building

- We analyzed papers from IEEE Security & Privacy (2015 & 2016) and from ACM CCS (2016)

Research Methods: Analysis of Published Papers

- Observations
  - Mixed reception from larger security community
  - Terminology issues are critical (e.g. definition of *case study*)

- Key results:
  - *Processes* and *Protocols* were the most common research subjects
  - Empirical data most commonly gathered about systems (rather than humans) and via automated means
  - Most studies were *observational* rather than *interventional*
  - Papers generally omitted a discussion of the *Threats to validity* of the research methods
  - More than ¾ of papers proposed and evaluated a solution in the same paper (i.e. little evidence of published replications)

- Deliverables:
  - Two HotSoS papers
  - In process: IEEE S&P paper summarizing the two HotSoS papers
  - In process: CACM “Good examples” paper

So much done, so much more to do …

Science of Security Lablet

https://www.nitrd.gov/images/cybersecurity.jpg

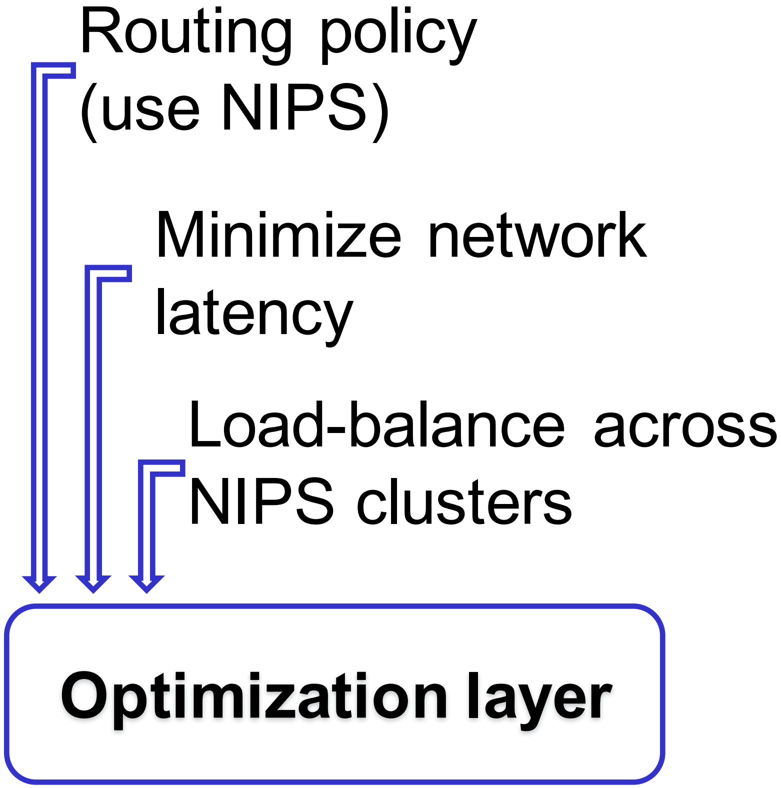
NC STATE UNIVERSITY
Department of Computer Science

Resiliency by Isolation and Diversity

- **Contribution #1 [Isolation & Diversity Measurement and Optimization]:**
  - Developing a new approach based on combining hypothesis testing and automated synthesis to automatically determine the optimal fine-grain isolation (least privilege configuration) between hosts/services in the network with minimal user interaction.
  - Our approach enables the users to automatically generate fine-grain network access control to minimize the global residual risk considering the end-hosts' security weaknesses based on compliance scanning reports (XCCDF), threat exposure, potential damage, impact on usability, and budgetary constraints.
  - We developed and integrate isolation and diversity metrics to optimize resiliency against multi-stage APT attackers.

Resilient Intrusion Prevention

- NIPS overload = false negatives
- Model NIPS traffic steering as an optimization problem
- Leverage the optimization layer
- Up to 5x load reduction
- Robust to network variations



Routing policy (use NIPS)

Minimize network latency

Load-balance across NIPS clusters

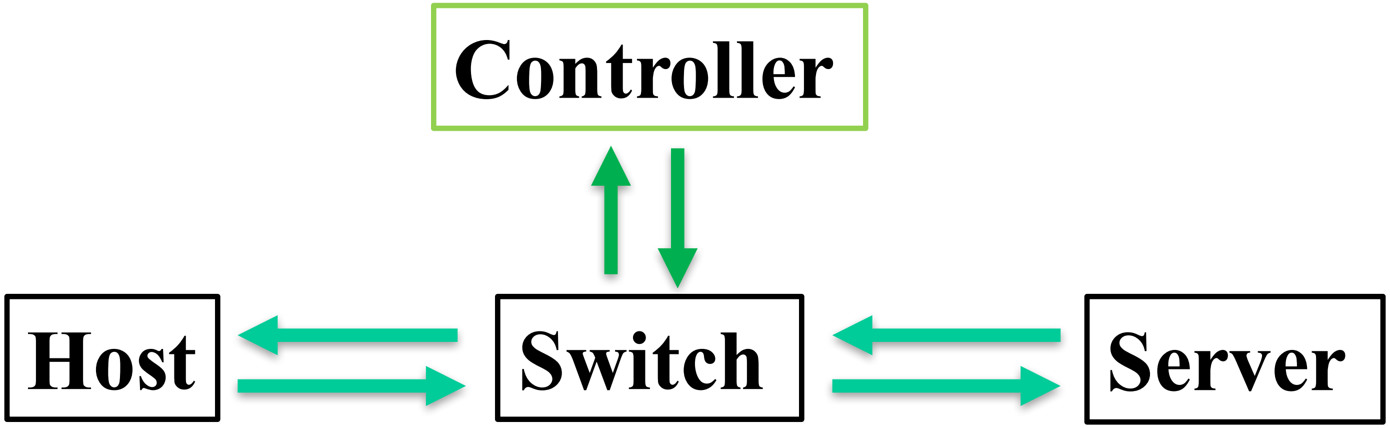
Optimization layer

Science of Security Lablet

NC STATE UNIVERSITY
Department of Computer Science

SDN Flow Reconnaissance Attack

- Attacker determines whether two parties communicated recently by probing network
  - Attacker flows that are routed quickly indicate recent flows that caused switch to retrieve covering rules from controller



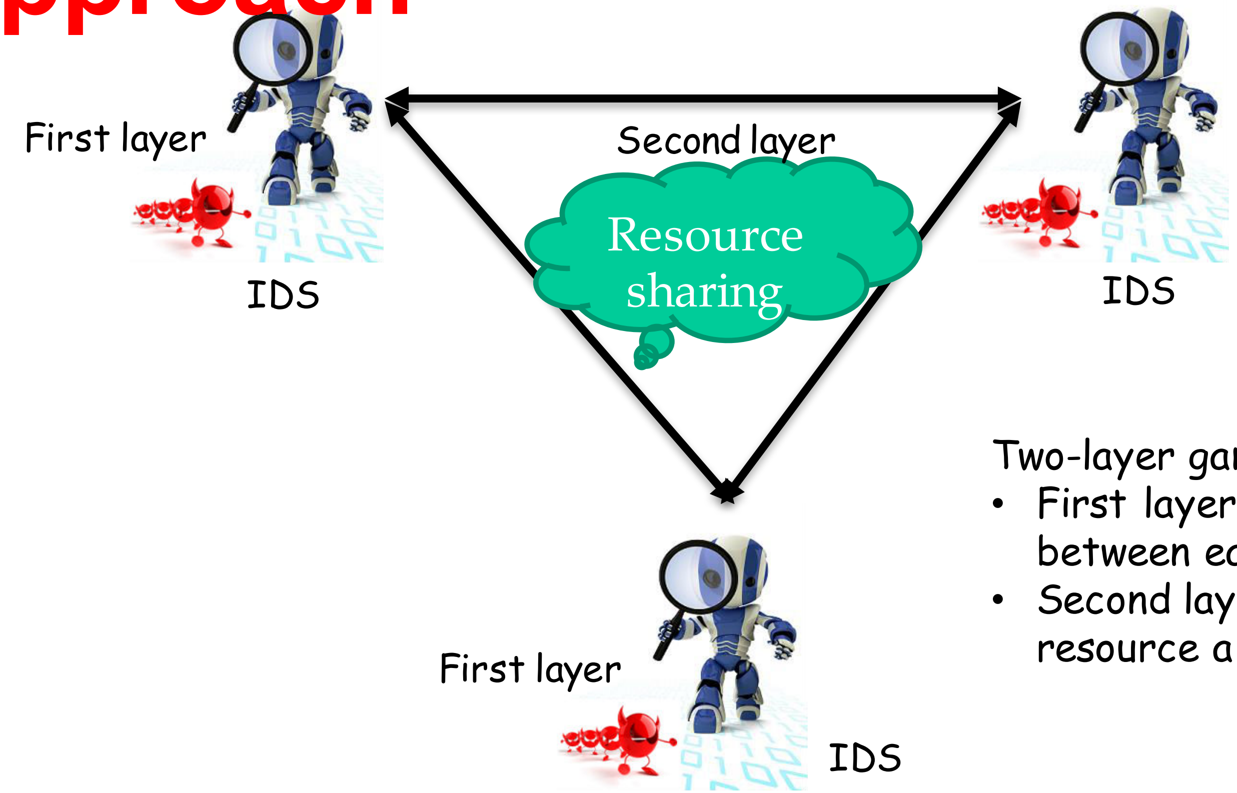
Android Policy Complexity

- Android has multiple layers of access control: Unix/Linux, SEAndroid, and so on
- Policies are large and complex
- We want to identify potential policy misconfigurations in SEAndroid policies in part by comparing against underlying Unix/Linux policies

Android Policy Analysis Findings

- Three kinds of problems
  - Composition privilege escalation: combined effects of multiple rules grant more access than sum of parts.
  - Critical object types accessed by low-privileged domains
  - Some object types are too coarse grained to offer effective protection
- Problems exist across multiple versions, but are improving

Collaborative IDS Configuration: A Two-layer Game-Theoretical Approach



First layer

- Estimate the attacker's attacking strategy

Two-layer game approach

- First layer: a stochastic game between each IDS and its attacker
- Second layer: auction based resource allocation scheme