Discussion Groups Report Out

NSF Secure and Trustworthy Cyberspace
Principal Investigators Meeting
Nov 27-29 2012

GROUP 1 – TEACHING SAFE PROGRAMMING

Discussion Group 1

Question:

How can we teach, and encourage and evaluate the teaching of, safe programming practices to reduce the vulnerability of future software systems?

Co-leads: Bill Pugh, Matt Bishop

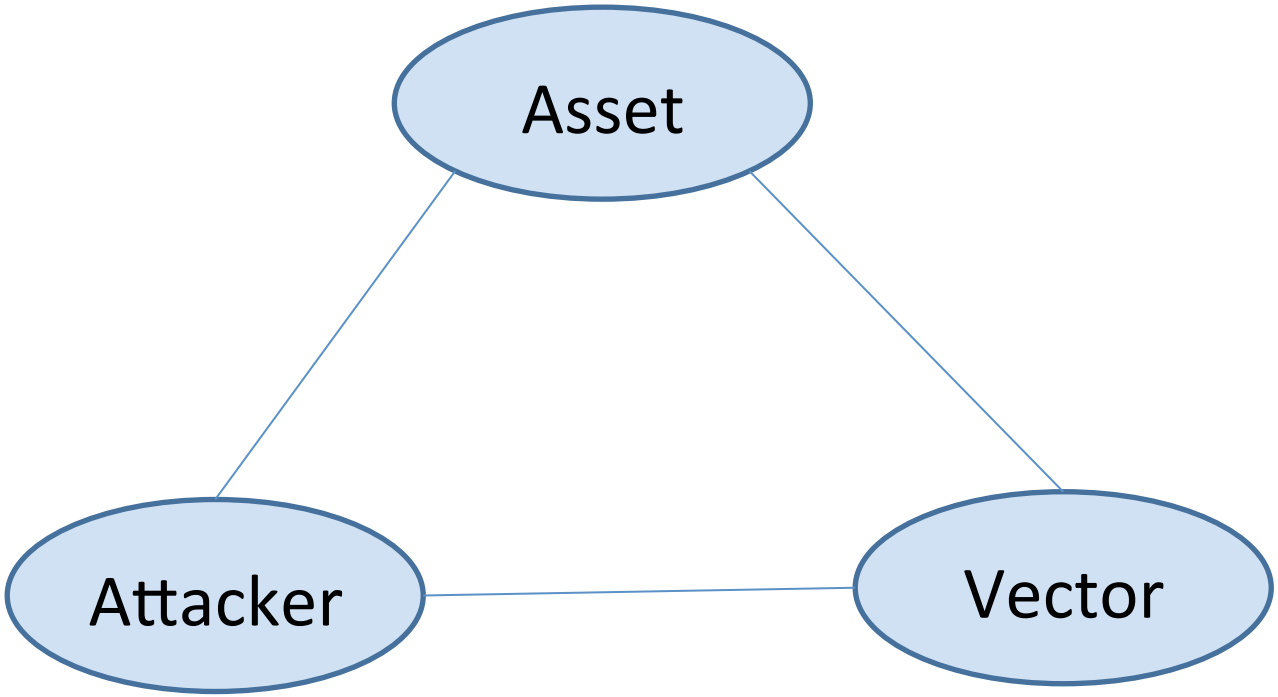
- Better way to organize, incentivize sharing of resources
  - Current systems don't work well
- Develop units appropriate for high schools, introductory, advanced programming courses
  - Assignments, slides, MOOC components and other materials
  - Units/components, rather than entire courses/curriculums
- How do you evaluate that what you are doing is meeting your goals?
  - And get meaningful feedback on it

GROUP 2 – THREAT MODELS

DG2
What threat models should guide future SaTC research?

Carl A. Gunter (Chair)
Lina Zhou (Scribe)
William Enck, Marco Gruteser, Sang Kim, Jung-Min Park, Stolfo Salvatore, Ravinder Shankesi, K. Subramani, XiaoFeng Wang

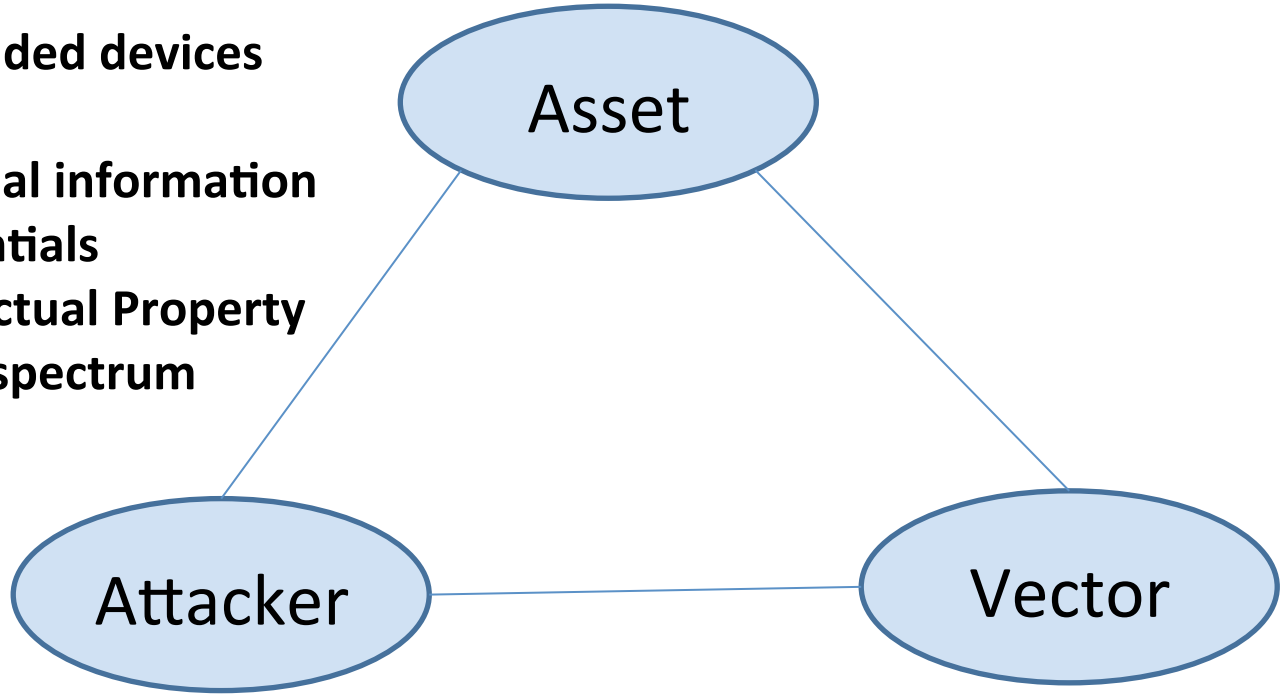
Traditional Threat Model



Guiding Future Research

- New threats derive from changes in the nodes of this graph.
- The overall model is still good but there are changes in the nodes.
- Ideally research plans should cover as much of the model as possible, but
- Sometimes strong work can be done without all elements being present.

Threats



Embedded devices
Clouds
Personal information
Credentials
Intellectual Property
Radio spectrum

Attackers for hire
States
“Hacktivists”
Black market
Insiders

Exploit:
New side channels
Social networks
App stores
Cell phone sensors
Parallel eval

GROUP 3 – CHARACTERISTICS OF TRANSITION TO PRACTICE

Discussion Group 3

What are the characteristics of SaTC ideas/technologies that are ready for transition to practice?

What are the success paths and pitfalls for different approaches to transition?

Recommendations for NSF - 1

- Do NOT forget your mission – SCIENCE!!
  - Do not **sacrifice** "R" in order to **grow** "T"
  - Must consider "cultural" impact on NSF community
- DECOUPLE "Transition" from Research Proposals
  - At least for SMALLS – **Sacred** for Basic Research
    - NSF needs to consider the impact on students
  - Maybe for MEDIUMS
- Reconsider proposal evaluation for Transition
  - Current review panels are not experienced in transition and are unable to adequately evaluate
- Improve "Marketing of Successes" as you move forward with "Transitions"

Recommendations for NSF - 2

- Provide "Clarity" in terms and definitions
  - What does "**transition"** mean (especially for an organization that is funding basic research)?
  - What does "**practice**" mean? Does it have to be a "widget" that gets deployed somewhere?
  - What's the difference between "transition phase" and "transition perspective" and the "option for transition to practice"? Very confusing
  - Still some confusion between requirements for "Broader Impact" and "Transition to Practice"
- NSF Organizational Changes (OCI now in CISE)
  - Consider "Transition" only in OCI programs?

Technology characteristics for TTP

- Maturity (DOD TRLs 1-9)
  - How is NSF going to evaluate maturity for "transition purposes"? (DEF.)
- A "Need" identified (and can be articulated)
  - User/Partner/Customer
- Well documented and reproducible
  - Academic SW does not equal production SW
- Usable by others (in the future) (DEF.)

Approaches to Transition – Success/Pitfalls

Open source	• Wide adoption – possibly commercial • Community that contributes back • Existence of Evangelists (champions)	• Continually maintain • Lack of Evangelists (champions) • QA versus security (is open source more secure?)
Commercialization	• People and network • Plan – Value proposition • Uptake, market, users	• Stifling creativity • Team, poor execution, etc. • Innovation understanding – is I-Corps the right approach? • [See Paul's slides for more]
(Broader) Impact	• Publications, papers, if you can measure value • Advancing the field (SaTC, inter-disciplinary) • Training students	• Missing significant technology component • Riskiness of research – too risky vs. not risky enough
Influencing or community building	• Internet measurement conference (as an example) – people centric • Develop curricula and training (course books, videos)	• No champion • Timing – not enough interested parties
Licensed technologies (Idea Transfer)	• Shortcut to commercialization • Ease piloting • Market scale – suited for smaller	• Lack of support • University knowledge lacking • Wrong kind of IP • Loss of control

GROUP 4 – USABILITY BARRIERS

What are the barriers to creating systems with security and privacy properties that users can understand and use?

Discussion Group 4
Alessandro Acquisti
Angela Sasse
Provocateur: Chris Clifton

NSF Secure and Trustworthy Cyberspace
Inaugural Principal Investigator Meeting
Nov. 27 -29th 2012, National Harbor, MD

Top Barriers

1. We don't understand the users
2. Misaligned incentives for system owners
3. Security depends …
4. Complexity of Systems

1 We Don't Know Enough About Users

- Lack of understanding of user capabilities
  - Particularly when viewed in organizational context
- User requirements research leading to quantifiable data
- Limited studies, often do not generalize
  - Insufficient to inform designers
    Action item: Research leading to design recommendations
- Action item: Testbeds
  - *PlanetLab / DETER for usability studies*
    - *Common IRB?*
  - We have methodologies for usability,
    Need research on methodologies for security

2 Misaligned Incentives for System Owners

- Privacy policies ≠ Privacy Practices
- Lack of user choice – accept T & C or get lost
- Unclear or non-existent auditing/feedback
- Solutions:
  - *Shakespeare: Henry The Sixth, Part 2 Act 4, scene 2, 71–78*
  - Digital wallets / real control
  - Unionise users, campaign against coercive systems

3 Security Depends On Others

- On others' expertise, behavior, motivation, diligence
- Example Developers think only about "their" system
  - *User deals with multiple systems*
- Example: leak in one place can compromise many accounts
- Key challenges:
  - Robust, usable authentication without need for backup
  - Systems that don't disclose information about people without consent - e*ven if data are published by others*
  - *Negotiating (or learning) expectations*

4 Complexity of Systems

- Security mechanisms are hard to get right – even by technical people *(e.g., TPM)*
- Need better abstractions, training
  - Design patterns for security mechanisms
  - *Side benefit – single learning curve for end user*
- Paradigm shift: Rewards and incentives for reuse
  - *Need Open source spirit and practices in the Security Community*

GROUP 5 – BUILDING CODE FOR CRITICAL INFRASTRUCTURE

Discussion Group 5

What might a **building code** for critical infrastructure software/hardware look like?

NIAP/CC/ISO-15408
DO 178b->c
DO 133
FDA CDRH 514(k)
DoD OT&E guidance
NASA 8739.8
NRC 1.172
ISO-9000
CMMI
DHS best practices
NIST guidance
IEEE stds, practices
OMG stds, practices

Bill Scherlis (CMU) with Sol Greenspan (NSF) and Dan Massey (Colo St)
and
David Naumann, Dighao Wu, Zhong Shao, Ron Perez, Alvaro Cardenas,
Joseph Kielman, Patrick Schaumont, Xenopfon Koutsoukos,
Ken Mai, Elaine Shi, Jim Pasquale, Manimaran Govindrasu, Mladen Vouk,

Building codes, idealized

Five salient features

(1) Engineering constraints
(2) Predicted quality outcomes
(3) Visible evidence of quality
(4) Explicit support for response
(5) Continuous evolution

They exist.

They work.

Consensus and compromise

(1) Enable innovation
(2) Protect IP
(3) Limit impacts on cost, performance, schedule, quality
(4) Fairly allocate risk and responsibility
(5) Afford measurement and visibility of risk and cost

Accommodations and Possibilities

(1) Fast pace of technology and ecosystem advancement
- More goals (what); less mechanism (how)
- Require a positive case with concrete evidence

(2) Scale, interconnection, customization unlike physical systems
- Composition is key

(3) Diversity and inter-relatedness of quality attributes
- Build models, analyses, metrics, composition for each
- Combine quality and security attributes – breakage and threats

(4) Hardware special needs and opportunities
- Rethink trusted hardware

(5) Economics and measurement as fundamental drivers
- Address incentives in building code – from EVM to IDE
- Fairly allocate risk mitigation benefit

GROUP 6 – MODELS IN CYBERSECURITY RESEARCH

Models in Cybersecurity Research

What's used, what's needed

Joshua Guttman

S Appleby, M Burmester, D Canas, Q Gu, R Herklotz,
Y Liang, H Nissenbaum, A Pollington,
A Scedrov, B Sunar

November 28, 2012

What's a model?

- A model carves out sets of
  - entities and properties to study
  - explanatory principles

  allowing us to
  - acquire data
  - make predictions
  - test hypotheses

Models are extremely various

- Physical models, often systems of differential equations
- Tractable summary of data and experience
- Set of guidelines for design — Eg “Risk exposure”
- Threat model — Eg Capabilities of attacker
- Game theoretic model — Goals, payoffs
- Norms predicting behavior expectations, basis of laws and ethics

Some models are big

- Noninterference, information flow
- Access control — Eg RBAC
- Computational model of crypto — PPT
- Dolev-Yao (symbolic) model of crypto

Other models are single-purpose

- Electric power self-stabilization
- BGP routing properties
- Design choices of expert developers
- Empirical modeling for decision making
- Reduce mass of English to corpus of formal rules
- Attack patterns in an intrusion

Simplify configurations

Other models are single-purpose

- Electric power self-stabilization
- BGP routing properties
- Design choices of expert developers
- Empirical modeling for decision making
- Reduce mass of English to corpus of formal rules
- Attack patterns in an intrusion

Simplify configurations

Single purpose models have high payoff

Security analysis: A tissue of models

- Systems have layers, and need layered models
- Models at different layers often very different
- Composability of components similar, horizontally
- One big model leaves out too much reality
- But: Attackers seek model join points

Security analysis: A tissue of models

- Systems have layers, and need layered models
- Models at different layers often very different
- Composability of components similar, horizontally
- One big model leaves out too much reality
- But: Attackers seek model join points

Need: "smooth weld" analysis methods

GROUP 7 – DECONSTRUCTIVE SECURITY IN THE RESEARCH PORTFOLIO

Group 7: The Role of Deconstructive Security in the Research Portfolio

Fred B. Schneider

Cornell University

The Big Picture

Deconstructive research: Activity where the primary focus is on *real attacks* to *real systems*.

Rationale(s):

- *Research*: Insights that could be useful for other research.
- *Social Good*: Call attention to important societal risks.
  - Analogy with investigative journalism.

Meta-questions for such research:

- What venue is most effective for disseminating such work?
  - Oakland, CCS, … **vs** DEF CON, Black Hat, … **vs** NY Times, Wash Post…
- What vehicle efficiently incentivizes or funds such effort?
  - Gov funding agencies **vs** industry **vs** individuals/groups on speculation.

Attacks as "Research"

Axiom: *Research* means the work should have a broad audience and long life.

Attacks as research if ….

… shows need for new **kinds** of defenses

- Ideally, work proposes those new defenses.

… illustrates new **classes** of vulnerabilities

- Perhaps due to new requirements or properties.
- People are part of the system; they can be vulnerabilities.

… extends our understanding of applicability for class of attacks and/or defenses.

… the retargeting of existing attack is itself novel and valuable as research.

Attacks as "Research"

… as a means to an end? Answers a question.

- How important is that question (audience / durability)?
- How novel is the answer?

- Good **deconstructive research** will:
  - Threat model is articulated.
  - Vulnerability and system described in enough detail for work to be reproducible
- Good **social good** work will:
  - Threat model is articulated.
  - Give risk assessment that is defensble (*Avoids hyperbole*)

Terms of Reference

Where does deconstructive security research belong in the research portfolio?

Discussion/additional questions: NSF and others have funded research that exposes flaws in current systems. Often the types of flaws found are not novel, although their context is.

- What does this type of research do to advance science or provide broad impact to society?
- What economic, political, and social impacts should be taken to account in determining whether to fund such research?
- Should NSF seek partner agencies in funding this kind of research?

GROUP 8 – POLICIES & NORMS IN THE ERA OF CYBERWAR

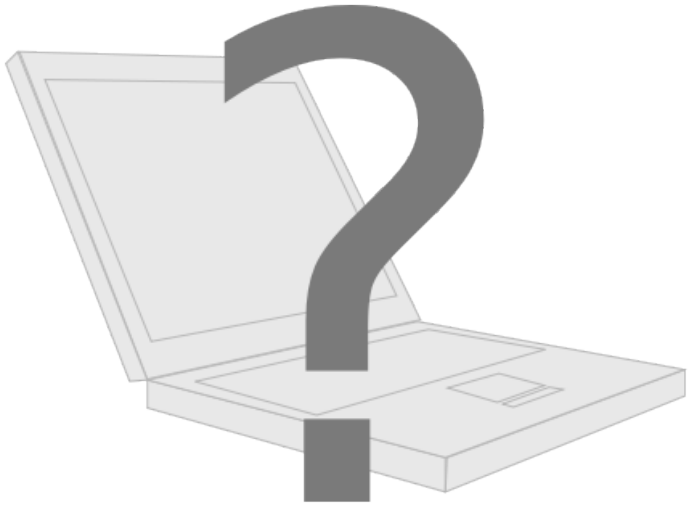
Policies and Norms in the Era of **Cyberwar**

The Challenge of “Commons”

- Cyberspace has some properties of Commons
- Security issues don’t fit well with the Commons framework paradigm
  - A single actor can ruin things for everyone

Attribution Remains Important

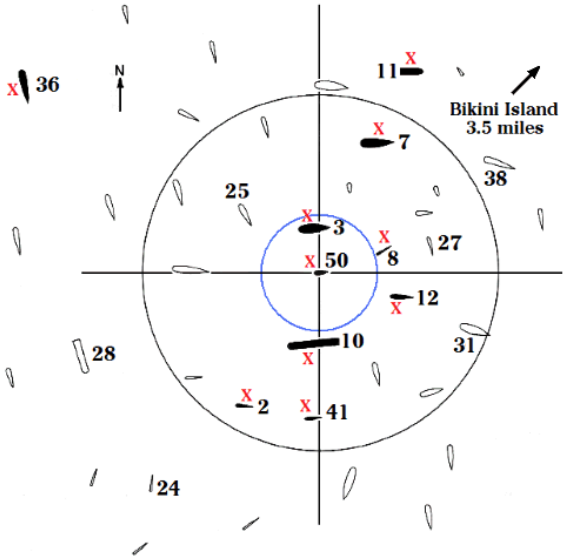
- Politically critical
- Does not apply to all threats



- Context-specific attribution questions
- Time/utility curves for forensic research
- Non-technical attribution data and analysis

Damage assessment – for us and them

- Us: Understanding offensive use
  - What is the 'blast radius' of a cyber operation?
- Them: defining the laws of cyberwar
  - How should we measure the damage of an attack to shape policy?
- What can technologists offer?



Computer Scientists as Organizational Innovators

- New types of institutions
  - Ad hoc
  - International
  - Cross-domain
  - Focused
  - Informal
- Example: Conficker working group

GROUP 9 – RESEARCH METHODS

Group 9 – Research Methods

Gail-Joon Ahn (ASU), Don Goff (CyberPack),
Mina Guirguis (TSU), Adele Howe (CSU),
Zbigniew Kalbarczyk (UIUC), Apu Kapadia (IU),
Stuart Krohn (NSA), Cristina Nita-Rotaru (Purdue),
Adam Smith (PSU)

Roy Maxion, Chair (CMU)

Computer Science Department
Carnegie Mellon University
Email: maxion@cs.cmu.edu

29 November 2012

NSF SaTC PI Meeting, Discussion Brief-Out
National Harbor, Maryland

Question

How do research methods vary across the disciplines involved in cyber security?

Method

- Ground rules – chair represents group
- Define "method"
- Enumerate range of methods
- Consider formal & experimental methods
- Enumerate disciplines in cyber security
- Discuss
- Reconsider question and viewpoints
- Make recommendations to NSF & PIs

What is a research method?

- A procedure for investigating a phenomenon in pursuit of a valid, credible and reproducible result.

- A procedure for establishing the highest quality evidence possible for supporting a claim.
  - Exposing the procedure ... supports claims ...
  - Avoids problems of confounded variables, internal/external invalidities, etc.
  - Allows judgments of validity and reproducibility of the experiment.

Simple examples of unsupported claims

- "According to statistics supplied to the commission by San Francisco-based service provider CloudFlare Inc., attacks account for about 15 percent of global Internet traffic on any given day.“

- Attacks do $50 million damage per year.

- **Really?**
  .... I wonder how they measure that ...
  ... in a reproducible way.

Recognize broad types of procedures

Empirical vs formal

- Empirical research
  - Range of methodologies, usually adapted to the situation
  - Not often described in detail in papers
- Formal methods research
  - The "method" is the proof

Three general methods

- Observational
  - Mostly descriptive
- Inductive
    - Classic scientific method
    - Hypothesis testing
- Deductive
    - Mathematical
    - Formal proofs

- Each method breaks down into several specialized methods, each providing a particular level or quality of evidence.

What disciplines are in cybersecurity?

- Tried to cluster topics into groups
  - Intrusion detection, Malware detection, Trustworthy hardware / software, usability, privacy, crypto, data collection, criminology, forensics, economics .....

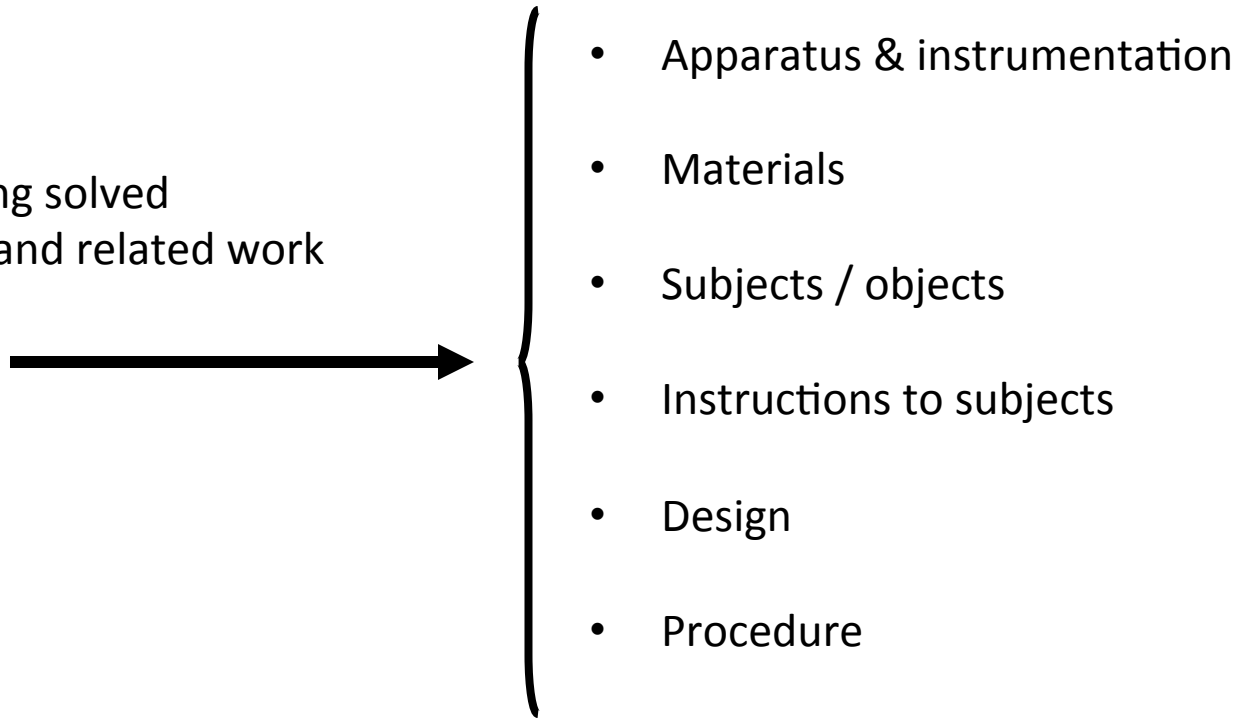
- Topics didn't coalesce neatly into disciplines.
- Some topics are inherently interdisciplinary.
  - Behavioral sciences, physical sciences
- Recognized that other sciences have their own methodologies, so borrow from them.

Recommendations

- Lay out a straw methodology that can be adapted to various experimental situations; encourage use.
- Write a book. Teach it. Inculcate students.
- Make research methods a required course ...
  (as many disciplines do).
- Change publication traditions to require method section ... (as many disciplines do).
- Provide evidence for claims.

Straw -- Parts of an experimental paper

- Title
- Author(s)
- Abstract
- Introduction
- Problem being solved
- Background and related work
- Approach
- **Method**
- Data
- Analysis
- Results
- Discussion
- Limitations
- Conclusion
- Future
- Acknowledgements
- References
- Appendices
- Endnotes and footnotes



- Apparatus & instrumentation
- Materials
- Subjects / objects
- Instructions to subjects
- Design
- Procedure

Thank you ...

- Contact information
  - Roy.Maxion@cs.cmu.edu
  - Carnegie Mellon University
  - Computer Science and Machine Learning
  - 412-268-7556

GROUP 10 – MODELING FOR HUMANS IN COMPLEX CYBER SYSTEMS

Discussion Group 10

What modeling techniques should we use to account for the role of humans in complex cyber systems?

Brent Rowe, RTI International

Rich Wash, Michigan State University

First… who do we mean by *humans*?

- Users (home) – lots of research
- Users (in companies) – some research
- Attackers – some research
- SW developers – little research
- Security professionals – little research

Defining *models* that include humans

- Qualitative (e.g., cognitive or mental models) versus quantitative (e.g., mathematical or computational models)
- Micro (focused on actors/agents) versus macro (systems)
- Static versus dynamic.
- Status quo (current behavior) focus versus focus on predicting factors involved in behavior change

- Also, combinations of these types of models are used

How are such models *developed*?

- Based on theory (psychology, economics, network theory, etc.)
- Based on observational data(observe actions or estimate with surveys or interviews)
- Based on experimental data (test impact of a "stimulus" in lab, real world, or through surveys or interviews)

- Note–Models developed based on theory can be tested by collecting exp or obs data, and models developed based on exp or obs data can be used to develop theories…

What models have been *used*?

- Cognitive models of users, attackers, developers, etc. (including based on psychology theories)
- Models of individual agents/actors (micro)
  - Rational utility models
  - Behavioral economic models – assume irrationality; seek to identify and add to model
- Models of a system of agents/actors (macro)
- Models of threats – types, frequency, etc.
- Models of network traffic aggregated to the individual user level

Recommendations

- More models of SW developers & security producers (use experimental & observational studies for guidance)
- New models of sophisticated attackers who conduct their own research (e.g., what data do they collect before attacking?)
- New models of the value of victims' characteristics to attackers – financial information, computer resources, etc.
- More models of status quo / base case (for all actors/ systems/threats)
- New models of users' ideal online activities (to be used to design security around users)

Recommendations

- More models acknowledging the evolutionary nature of threats and actors
- More modelers should be encouraged to reuse their models, build on them, share them, and combine them with other researchers models
- Leveraging of modeling work in other disciplines – e.g., climate change models and epidemiological models involve complex

GROUP 11 – METHODS FOR EVALUATING STABILITY/ TRUSTWORTHINESS OF COMPLEX DIGITAL INFRASTRUCTURE SYSTEMS

Predicting the next "flash crash" or blackout: What **methods** are available for evaluating the stability/ trustworthiness of complex digital infrastructure systems?

SaTC PI Meeting – DG 11
Presented by
Bill Sanders
University of Illinois

Group Participants (16)

- Lance Joneckis, IDA
- Wayne Burleson, Univ. Mass. Amherst
- Yan Chen, Northwestern Univ.
- Li-Chiou Chen, Pace Univ.
- Chunxiao Chigan, Univ. Mass. Lowell
- Thomas Eisenbarth, Worcester Polytechnic Institute
- George Kesidis, Penn. State (co-lead)
- Lorie Liebrock, New Mexico Tech
- Jeff Rowe, UC Davis
- Bill Sanders (lead)
- Simha Sethumadhavan, Columbia Univ. (co-lead)
- Haiying Shen, Clemson Univ.
- Ankur Srivastava, Univ. Maryland
- K. Subramani, WVU
- Weichao Wang, UNC Charlotte
- Wei Yu, Towson Univ.

Attributes of Complex Digital Infrastructure Systems (CDIS)

- Systems of systems (networks of networks), each with independently designed system so exhibit interoperation challenges
- Heterogeneity
- Potential for highly dynamic, stochastic, and chaotic behavior
- Distributed control
- Big Data
- Human-in-the-loop

Example CDIS

Mostly autonomous, leverage Internet or private Intranet, e.g.,

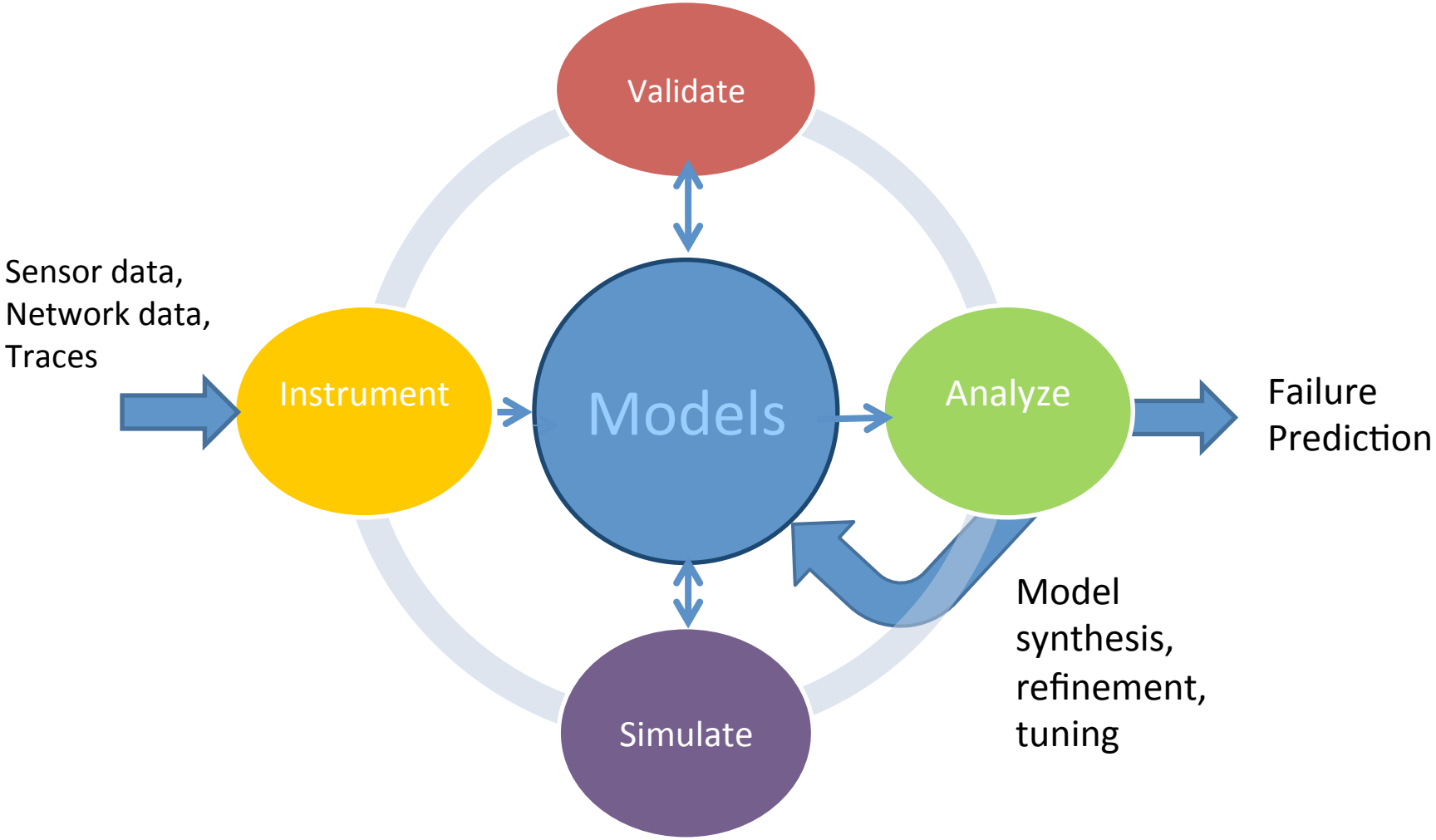
- Smart Grid spanning physical power system, marketplace, communication/control infrastructure (SCADA)
- Financial Systems
- Smart transportation systems, e.g., smart car, airline systems

Answer to Question?

Predicting the next "flash crash" or blackout: What **methods** are available for evaluating the stability/ trustworthiness of complex digital infrastructure systems?

- Simple Answer: Empty Set
- Reality: Practitioners are doing this all the time
  - E.g., active monitoring of transmission portion of power grid
- Truth: Somewhere in-between

Methods – The Big Picture



Methods Details

Instrumentation

- Placement strategy for lightweight sensors
- Use of mobile sensor agents
- Privacy preserving collection and storage
- In-line processing

Model Synthesis

- Hybrid state-space (cont. and/or discrete-time) with discrete-events
- Known “normal” operating range and well-defined decision functions for anomalies, attacks and faults (known vulnerabilities).
- Incorporate automated feedback control (including containment actions), human factors and economic considerations (e.g., through game theoretic framework)
- Hierarchical, decomposed structure

Analytical Model Solution

- Solvers of formal models including both hybrid state-spaces and discrete-events
- Differential equation solvers
- Graph-theoretic models, link analysis of interconnectedness (suitably weighted to account for trust/reliability, priority), etc.

Methods Details, cont.

Simulation Model Solution	• "Brute" force simulation at scale for, e.g., Chinese power system and financial markets – cf. combinatorially big computational challenges • Test how systems would handle hypothetically abnormal (e.g., attack/defense) situations • Trace-driven experiments for stress testing on-line/off-line at system level
Analysis (Producing Predictions)	• Decision makers act on mixture of experts: model, simulation, human operators
Validation of Predictions	• Cross-check simulation and analytical modeling result in terms of behavioral details • Expert prediction consensus, comparisons against "real world" historical data

Limitations of Existing Methods / Research Challenges

- **Scaling techniques applied to isolated systems to complex systems of complex system**
  - Scalability of predictions
  - Scalability of the computational infrastructure
  - Integrating federated models from different disciplines (CS, Econ, Finance, Social)
- **CDIS model creation**
  - Very large space of available techniques. Which is most applicable to the specific complex system?
  - Identifying critical variables from the very large available space.
- **Model validation**
  - Hypothesis driven experimentation is difficult without a science of complex systems
  - Validation on full-scale CDIs is not possible
  - Can catastrophic events be recorded and used for validation?

Summary

- What methods?
- There is no unified or perfect method
- Practitioners are using methods to predict these events today on mission critical CDIS, albeit imperfectly
- We believe it is important to work to alleviate the limitations of existing methods in emergent CDIS interoperating at greater scale

GROUP 12 – ANONYMITY AND ACCOUNTABILITY TRADEOFFS

Anonymity and Accountability: How Do We Enable Tradeoffs?

Jeannette Wing and Rebecca Wright, co-chairs

Mihir Bellare, Dan Boneh, Rohit Chadha, Nicolas Christin, Anupam Datta, Roger Dingledine, Yingfei Dong, Zhenhai Duan, Nelly Fazio, Yong Guan, Andreas Haeberlen, Aaron Jaggard, Ping Ji, Aggelos Kiayias, April Kontostathis, Anna Lysyanskya, Mohammad Mahmoody, Steve Myers, Paul Reber, abhi shelat, Micah Sherr, Stephen Tate, Michael Taylor, Nicholas Weaver, Emmett Witchel, Matthew Wright, Li Xiong, Grace Hui Yang, Yong Zhao, Ye Zhu

NSF Secure and Trustworthy Computing PI meeting
28 November 2012

A Motivating Scenario

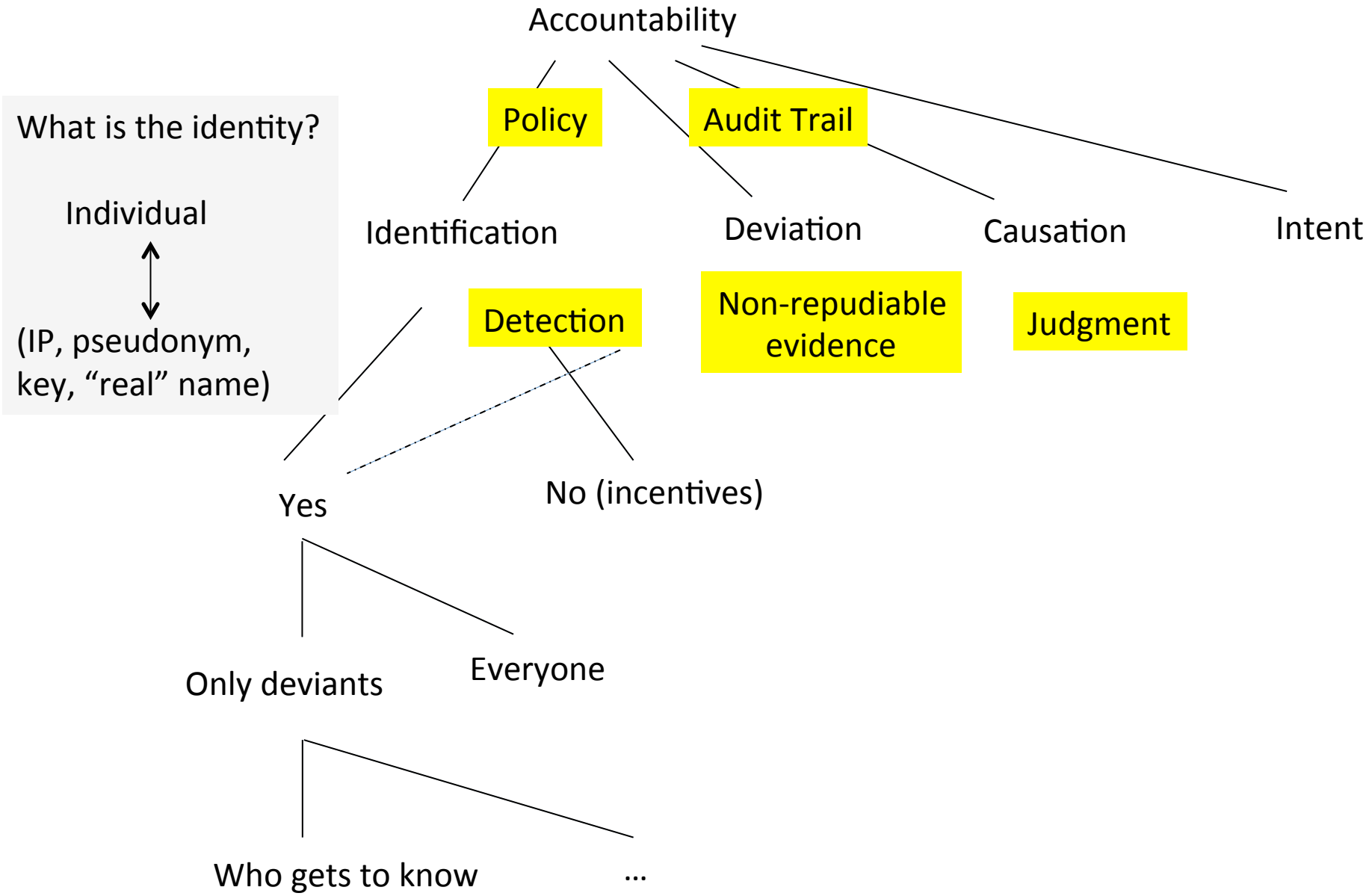
- Healthcare chat room:
  - people should be able to participate anonymously to safely/comfortably discuss their health concerns. Sometimes posting photos can help convey information.
  - however, this can attract inappropriate users – e.g., people posting child porn photos.

- Should have accountability to deter such misbehavior, without compromising anonymity for appropriate use.

Anonymity and accountability: Can we have both?

- Yes, we can!
  - In some settings, under certain well-defined conditions, for certain functions, for well-defined notions of anonymity and accountability, using certain known and practical cryptographic primitives, we can have both.

Accountability and Related Issues



Primitives, Tools, Systems, and Concepts for Anonymity [partial list]

- Tor (System)
- Crypto primitives:
  - group signatures [systems: Nymble, Jack]
  - (revocably) anonymous credentials [system: BLACR]
  - blind signatures
- reputation-based systems
- pseudonym-based systems
- Helios (system), and other e-voting systems
- Bitcoin (system), and other anonymous e-cash systems
- Accountable Internet Protocol
- PeerReview (and related systems)

Open Research Questions

- What is anonymity? accountability? metrics for quantifying them?
- What is a meaningful definition of accountable anonymity? anonymous accountability?
- How do we balance accountability and privacy when different kinds of participants have different constraints/policies (e.g., voting systems, online privacy policies)?

Open Research Questions

- How much anonymity can you get as a function of the power of the adversary to control the network? (and possibly of efficiency of solution)
- efficient cryptographic solutions for a larger class of policies, particularly more flexible policies, dynamic policies, and policies that may depend on private data.
- Retaining anonymity or unlinkability at all in today's world (where identity can be leaked or partially leaked by devices, applications, network, users…).

Open Research Questions

- How strong are the accountability and anonymity properties that can be achieved, relative to the cost of obtaining an identity (or a credential)?
- Can we change the costs to effectively dissuade bad behavior?
- Design anonymity system that enables data mixing for utility goals (e.g., deviation detection, pricing, targeted advertising while preserving privacy)

Open Research Questions

- Beyond computer science research:
  - How to enable users to make informed decisions about anonymity-accountability tradeoffs? Can the informed consent model be useful here?
  - Where must regulatory frameworks supplement technical approaches?
  - Do users care, and in what contexts? (Knowledge, time-dependence?)
  - How much do people value their identity?

Open Research Questions

- Many Tor-related questions:
  - scalability and efficiency
  - censorship prevention
  - cryptographic questions (see Roger Dingledine for specifics)
  - incentives: how do we get more people to run Tor relays? how do you manage those who use too much resources?

Open Research Questions

- Many identity infrastructure questions:
  - Credentials: weak link is identity infrastructure. How do we get the identity infrastructure to be strong enough to not be the weak link?
  - How to build Internet-scale identity infrastructure that will enable use of anonymous credentials and group signatures?
  - How to compare two identity infrastructures? what properties (technical, human factors)? how to measure them?

Broader Impacts Need

- Educate policymakers and system designers about what is possible.

GROUP 13 – DATA DELETION

WG 13: Data Deletion

What policies and technologies should be required to enforce the expiration of data?

Participants:

Kevin Bauer, Nikita Borisov, Hao Chen, Johannes Gehrke, Yong Guan, Apu Kapadia, Chris Kanich, Tadayoshi Kohno (Chair), Corin Pitcher, Thomas Ristenpart (Recorder), Roberto Tamassia, Jan Whittington, Tse-Chuan Yang

High Level Results

- In many cases we desire the ability to enforce the expiration of data
- But area is very complicated: No solutions without controversy

- But two main "contributions"
  - 1: Identified situations (data use cases) for which we may for wish data expiration
  - 2: Identified "axes" for the problem
- Claim: Any progress for one of these use cases and specific points on these axes could be valuable

WG 13 on Data Deletion: What policies and technologies should be required to enforce the expiration of data?

Data Use Cases

- Corporate email (internal)
- Corporate email (between companies)
- Gmail
- Laptop and phone data (e.g., lost or stolen device)
- Social network data
- Captured public data (drones, ATM cameras)
- Health records
- Financial records
- Childhood records (go away at age 18)
- Death (may want records to go away or come back)
- Deletion as a solution for account compromise
- Sexting
- Digital media for our own photos/movies
- *Digital media such as movies*
- *Criminal uses*

WG 13 on Data Deletion: What policies and technologies should be required to enforce the expiration of data?

Axes of the problem

- Consumer versus corporate data
- Clean versus comingled data
- Structured versus unstructured data
- Trust in second party (e.g., Google), third party (e.g., ad networks that buy data from Google), other first parties (e.g., Alice and Bob)
- Prevention vs. auditing (for a solution)
- Desired lifetime of data (disappear immediately, days, years, forever)
- Type of data:
  Data about you (tracking data, photo of you that's not shared with you),
  vs. Data authored by you
  vs. Data shared with you
- Who implements mechanism: Client only, client+cloud, cloud
- (Dis)Incentives for adoption: economic, government, ethical
- Role of policy in the solution: none, law
- People's changing preferences: retroactive change in preference?

WG 13 on Data Deletion: What policies and technologies should be required to enforce the expiration of data?

Some Interesting Ideas that Arose

- Solution idea: The law could say that "if a user clicks on 'delete,' data is legally deleted" even if data persists
- Solution idea: Audit companies to verify that they've deleted data (third-party auditors)
- Solution idea: Decrease signal-to-noise (so much noisy data, doesn't matter if data persists)
- Solution idea: Pay to delete (to incentivize companies) (instead of right to delete)
- Solution idea: Service-level agreement for deletion
- Solution idea: Watchdog timer for deletion (data deleted every week unless user requests otherwise)
- Challenge: Inferences about deleted data from non-deleted data
- Challenge: Cognitive overhead of data expiry (hard for users if some data disappears and other persist)

High Level Results

- In many cases we desire the ability to enforce the expiration of data
- But area is very complicated: No specific solutions that weren't without controversy

- But two main "contributions"
  - 1: Identified situations (data use cases) for which we may for wish data expiration
  - 2: Identified "axes" for the problem
- Claim: Any progress for one of these use cases and specific points on these axes could be valuable

GROUP 14 – PROVENANCE, INTEGRITY, LONGEVITY OF SCIENTIFIC RECORDS

First NSF SATC PI's Meeting
Discussion Session 14

Ravi Sandhu
Executive Director and Endowed Professor
Nov. 29, 2012

ravi.sandhu@utsa.edu
www.profsandhu.com
www.ics.utsa.edu

Question

14. How can we assure provenance, integrity, longevity of scientific records?

Discussion/additional questions: Scientific results depend critically on understanding data captured from experiments and from observations of the natural world. Very little of today's scientific data is captured or stored without the involvement of a computing system. Further there is increasing demand for data from publicly funded science to be made available to the public as soon as possible. This discussion should explore the policies, requirements, and mechanisms available for assuring provenance, integrity, and preservation of scientific data in the face of potentially malicious behavior.

Participants

- Leader: Ravi Sandhu
- Co-Leaders: Elisa Bertino, Sharad Mehrotra
- Participants: Genevieve Bartlett, Kevin Butler, Keith Frikken, Gabriel Ghinita, Jeff Hoffstein, Wei Jiang, Murat Kantarrcouglu, Sorin Lerner, Lee Osterweil, Don Porter

Provenance Solutions

- Impossible: Time machine
- Impractical: Record all context relevant or irrelevant

- Practical
  - Capture what is relevant for the purpose we want to use provenance

Provenance Threats

- Bad data leading to bad science without bad intent
- Deliberate scientific fraud by insiders
- Deliberate mischief by malicious outsiders

Provenance Challenges

- Scientific data manipulation processes are complex
- Provenance data is big

- Usability and adoption by scientists
- Automated capture including human-in-the-loop

- Annonymized data
  - Medical data
  - Sociological data
  - Cyber security data

Provenance Opportunity

- Great opportunity for NSF SaTC
  - Inherently interdisciplinary
  - Its all about enabling good data-based science
  - Center scale funding

GROUP 15 – IDENTITY MANAGEMENT – WHY SO SLOW?

Identity Management: Why So Slow?

Susan Landauchair

(Danny Weitzner standing in)

Participants: Jim Basney, Sasha Boldyreva, Laura Dabbish, Minaxi Gupta, Jeff Hancock, Ken Klingenstein, Adam Lee, Tien Nguyen, Mischa Rabinovich, Zhijie Shi, Steve Weber, Dan Wolf

Identity Federation

- Single sign-on: single authentication that enables access to multiple resources.

my_email_address@gmail.com

Use Cases

- Start with use cases and understand them thoroughly.
- CAC:DoD implementation with 17 M cards

  highly constrained environment.
- InCommon: SAML based,

  many federations.
- Orthogonal use cases: fake reviews in Amazon, fake Twitter postings, posting on Craiglist, etc.?

A Real-World Problem: DOE Research Labs

- DoE Labs want to share data archives, wikis, supercomputer with users around the world.
- Solution: Identity Management.
- **But** how do you perform a risk assessment to determine your needed level of assurance? How do you **really** figure that out?

Accountability and Anonymity

- Accountability v. anonymity is a policy issue.

- Anonymity is also a layered issue: identity anonymity doesn't provide full anonymity (think Petraeus).

Research Questions I

- Where do federated systems work (that is, please all stakeholders)?
- Economics of anonymous credentials: we know cost of privacy spill; what do I get from anonymous credentials?
- Is adoption of federated systems a necessary set of tradeoffs (economic, privacy, and political/policy)? If so, how do you provide incentives?

Research Questions II

- UI is **really** hard. What does the user need to know about attribute release?
- How does the system trust the user (what the user does to get trust)?
- Dynamic and increasingly rich world for metadata: how does that change federation?

GROUP 16 – LEVERAGING R&D FOR EDUCATION

Discussion Group 16: How can we leverage R&D work done to improve cybersecurity education?

David Balenson, SRI International (Lead)
Justin Cappos, NYU Poly
Art Conclin, University of Houston
Wenliang (Kevin) Du, Syracuse University
Haibo He, University of Rhode Island
Manish Karir, DHS S&T Cyber Security Division
Di Ma, University of Michigan-Dearborn
Jelena Mirkovic, USC/ISI
Gookwon Suh, Cornell University
Jaideep Vaidya, Rutgers University
Venkat Venkatakrishnan, University of Illinois at Chicago
Xiaohong Yuan, North Carolina A&T State University

General Approach

- Establish a wiki/repository through which instructional materials and learning activities derived from funded research can be disseminated
- Incentivize researchers to produce instructional materials from their research
- Incentivize educators to consume and evaluate materials in their classes
- Motivated, in part, by existing educational facilities:
  - DETERlab Education Portal, Jelena Mirkovic, USC/ISI
  - SEATTLE Open Peer-to-peer Computing, Justin Cappos, NYU Poly
  - SEED (Security Education) Lab, Kevin Du, Syracuse U.

How identify key concepts during the research phase?

- Be open and inclusive
  - Narrow and broad concepts
  - Fundamental and applied concepts
- Not all research projects are applicable
- Leverage ACM / IEEE Computer Science Curriculum as a taxonomy for identifying concepts
- Look for concepts that are repeatable
- Observe security trends, industry practice
- “Coolness” factor

How make integral part of R&D?

- Establish an “open source” community
  - Materials undergo peer review, cross testing, ongoing maintenance
- Incentivize researchers and educators by offering additional funding (after the fact) for
  - Contributing material that is adopted by others
  - Adopting and evaluating material provided by others
- Facilitate via
  - (Semi-)standard format
  - Unified set of platforms for lab exercises

Types of materials?

- Lecture materials
  - Reading materials, lecture notes, slide decks, videos, animations, etc.
- Hands-on labs and exercises (with instructor manuals)
- Tests and quizzes
- Data sets
- Source code
- Case studies

How do we disseminate materials to educational enterprise?

- Types of consumers
  - Graduate, upper-level undergraduate
  - Lower-level undergraduate
  - K-12
  - Games and competitions
  - Workforce development
  - Online learning
- Target materials to different types of consumers
- Community wiki/repository (supported by NSF)
  - Catalog w/ metadata and actual instructional materials
  - Or links to material on other sites
- Promote throughout the community

How to evaluate materials and community concept?

- Incentivize educators to evaluate and highlight effective material
- Crowd-sourced ratings
  - Reputation-based scoring (1-5 stars)
  - +1/-1 from users like Amazon reviews
  - Online reviews
- Collect metrics regarding contributions and adoptions
- Test students before and after to evaluate learning
- Integrate evaluation into materials (feedback-based learning)
- Survey students/teachers
- Survey 1-2 years after the class to gauge impact

GROUP 17 – AGENDA FOR A CYBERSECURITY WORKFORCE

An Agenda for Cyber Security Workforce

Discussion Group 17

SaTC PI Meeting
Nov 28, 2012

Goal

- Prepare students for workforce
  - Breadth vs. depth
  - Train for roles, non specialist + specialist
  - Approaches to scale
  - Common body of Knowledge

Need for Skills/Background Thinking

- Broader than STEM
  - Critical thinkers
  - Risk Analysis
  - Economic models
  - Communication with domain experts
  - Fixing skills vs. Breaking skills
  - Practical vs. formal/theoretical

Raise Cyber Security Awareness

- K12
- Add security literacy courses to basic undergraduate curriculum
- Work across disciplines
- Motivate new generation of students
- Delivery options for scale
- Cost issues

Diversity

- Pipeline/graduation/hiring
  - Effects of competition on populations
  - Identify and invest in parallel programs
  - Team membership
  - Education system (industry, government, and academia)

Recommendations

- Attract
  - Reward, fellowships, post docs, internships, forgivable education loans
- Invest
  - Support for faculty in teaching, hiring
- Create
  - Security Guru

GROUP 18 - CANCELLED

GROUP 19 – INCENTIVES AND NORMS

Social, Cognitive, and Economic Perspectives Applied to Cybersecurity and Cyberprivacy

Contributors:

Becky Bace, Sandra Carpenter, Yingying Chen, Jenne Lindqvist, Kevin McCabe, Nasir Memon, Lisa PytlikZillig, Laura Razzolini, Alan Tomkins

Discussion Question 19

What incentives, norms, attitudes, habits, cognitive limits, or other mechanisms present the most important obstacles to cybersecurity, and how might such factors be utilized to benefit cybersecurity?

We addressed the question primarily from the users' perspectives

Incentives and Attitudes

- Incentives can be monetary, norm-based (e.g., shame), fame; economic theories apply, broadly, in terms of costs/benefits
- Security and privacy attitudes are related to trust and perceived risk (but costs related to risks may be unknown)
- Social influence can impact attitudes and behavior (e.g., "green" behaviors), such that new norms might develop; potentially target children

Users' Cognitive Processes & Human Factors

- Users' mental models of security and privacy issues may not match threat models
- Users may have poor or incomplete knowledge of types of risk they may face, their likelihood, their severity
- Users may lack the ability and/or willingness to attend to, comprehend, or remember security or privacy mechanisms (e.g., long passwords)
- Users' cognitive load and/or stress impacts accuracy

Challenges

- Norms, attitudes, incentives, and cognitive capabilities are variable
  - Age, experience, knowledge, culture
- Many theories and research findings in social and cognitive psychology show context-specific and domain-specific patterns
- The SBE literature, as it relates to cybersecurity and cyberprivacy, is not well-organized or accessible
  - Relevant publications in social science may appear in "marginal" journals; visibility may be low

Suggestions

- Identify theories from economics, cognitive and social psychology that may apply to the cyber-domain and test them through empirical research
- Identify intersections between threat models, users' mental models, and models of economic incentives
- Provide an infrastructure to warehouse, update, and disseminate economic and social science research conducted in the cyber-domain (e.g., CyLab at CMU; Anderson) and make its existence more salient to social scientists
- The Federal Cybersecurity R&D Strategic Plan could explicitly call for social and cognitive psychology research

GROUP 20 - GROUP, ORGANIZATIONAL, INSTITUTIONAL, AND POLICY OBSTACLES TO CYBERSECURITY

What are the group, organizational, institutional, and policy obstacles to cybersecurity?

Obstacle	NSF Research & Education Directions
Cyber-risk has not been quantified	• Cyber-insurance • Risk modeling • Standards and certification • Policy, regulation, and liability
A gap in understanding of privacy and security user/ behavioral models exists	• Development of behavioral models • Understanding of usage patterns • Characterization of individual and group differences
Insufficient formal cybersecurity education	• Optimal educational delivery mechanisms, e.g. general education requirement, interdisciplinary education. • Online education, e.g. Standford model • Assessment, evaluation, and effectiveness of the educational design, content, and delivery
Misinformation about cybersecurity and lack of understanding about the consequences of inaction.	• Communicating with policy makers • Awareness for a wider audience, e.g. YouTube

THANK YOU!

Discussion Groups Report Out

NSF Secure and Trustworthy Cyberspace
Principal Investigators Meeting
Nov 27-29 2012