Cybersecurity via Signaling Games

William Casey, Jose Morales, Evan Wright, Rhiannon Weaver at CMU SEI with Bud Mishra at Courant Institute.

Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213

Cybersecurity via Signaling Games

William Casey, Jose Morales, Evan Wright, Rhiannon Weaver at CMU SEI, with Bud Mishra at Courant Institute.

Software Engineering Institute | Carnegie Mellon University

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the United States Department of Defense.

NO WARRANTY. THIS CARNEGIE MELLON UNIVERSITY AND SOFTWARE ENGINEERING INSTITUTE MATERIAL IS FURNISHED ON AN “AS-IS” BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

DM-0002797

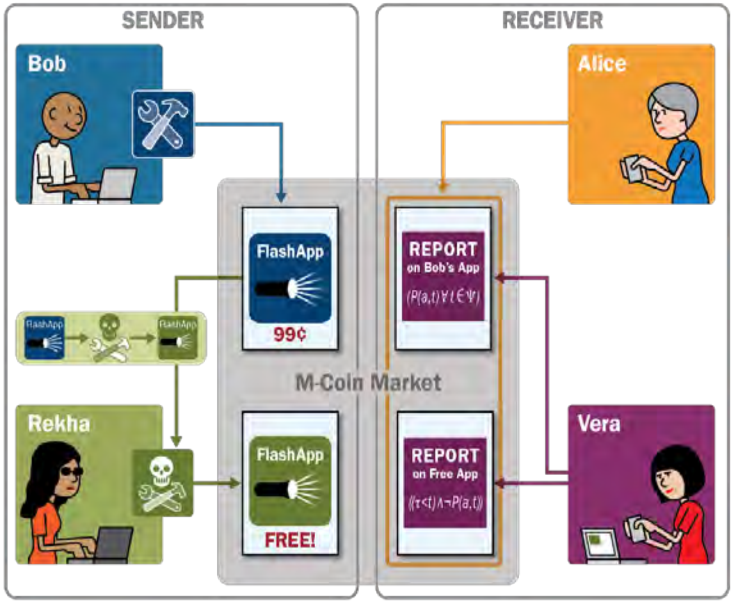
Problem: Cybersecurity

How can we establish trust, manage risk, and mitigate deceptive cyber attacks when our decision-making is constrained to partial information concerning unknowns vulnerabilities, system properties, and threats?

Our approach: A fundamental model of humans actions and the safety properties they affect.

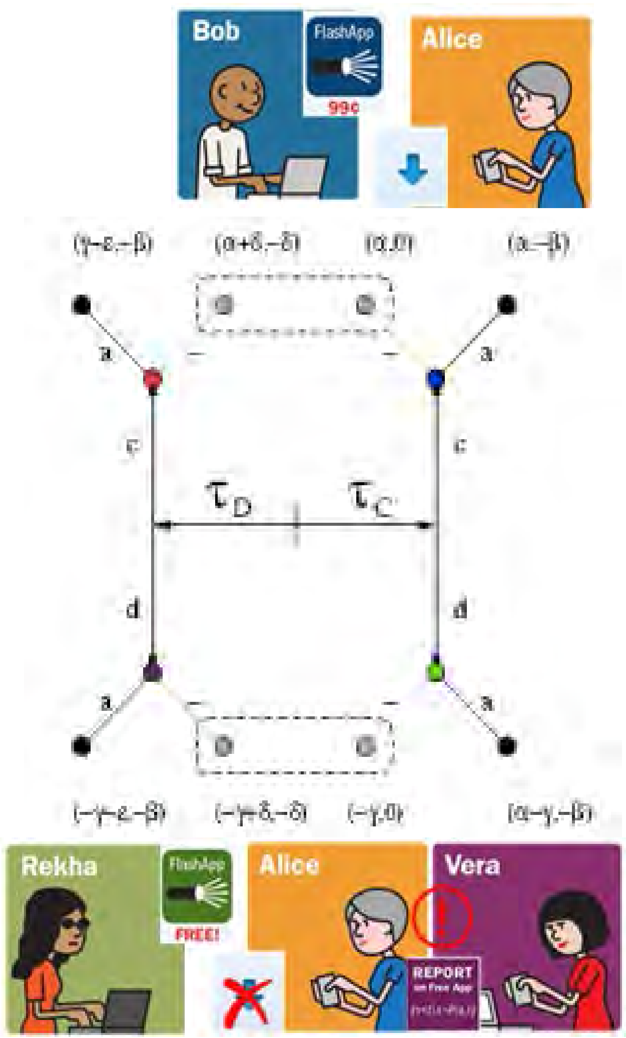
- **Game-theoretical** model to simultaneously study human and system properties within a social technological systems:
  - Deceptions are definable, allowing risk estimation and policy optimization.
  - Mathematical (and virtual) means to create, explore and design a wide range of mechanisms, including agent based models, simulation, evolutionary games, and analytic calculation of equilibria.

Cybersecurity and the Actions of People



- Information is incomplete.
  - Decisions rely on signals (info) available at the time actions are needed.
  - Deceptive strategies which leverage information asymmetries arise naturally.
- Many ways to minimize information asymmetries with respect to the desired properties of a system.
  - We organizing these into agent types:
    - Recommenders for liveness.
    - Verifiers for safety.

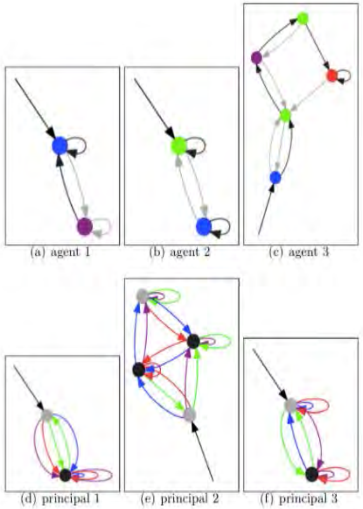
Information Asymmetric Signaling Games



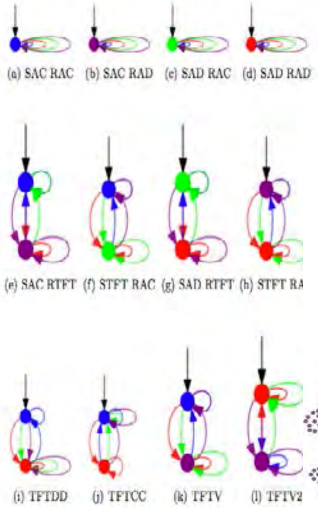
- A signaling game describes a scenario with two players:
  - A ‘sender’ has a type determined by nature and transmits a signal (information) to a ‘receiver.’
  - The ‘receiver’ having interpreted the signal selects an action with various equity outcomes – result will depend on the sender’s type (unknown), their signal, and the receiver action.
- We have specialized signaling game models to:
  - App malware.
  - Multiple vulnerabilities and deceptive exploits.
  - Managing and estimating risks from non-compliance.

Modeling Cyber and Social Technological Systems

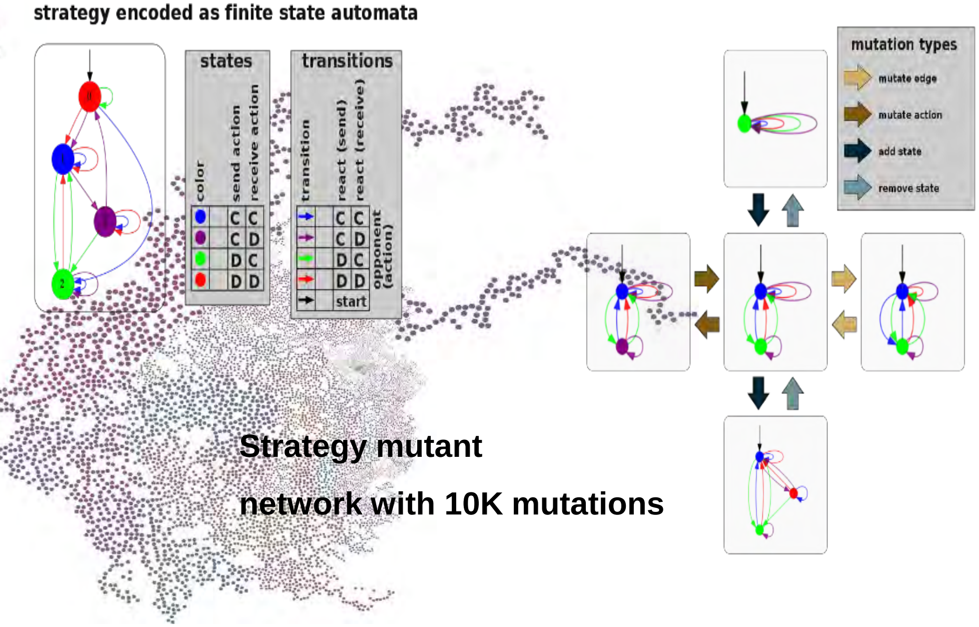
Signaling games are played repetitively in social technological systems. The **deception** or *gap between what ones says and does* is an increasingly important risk factor. We build upon evolutionary game theory (EGT) to describe systems of non-cooperative agents which explore and exploit utilities in cyber.



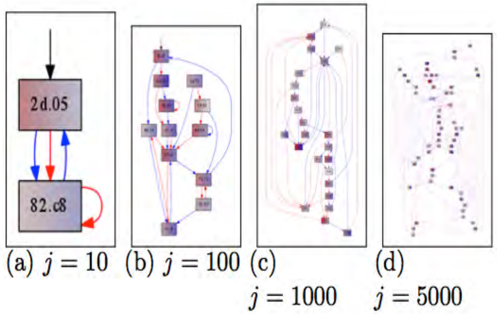
Asymmetric compliance strategies



signaling game strategies



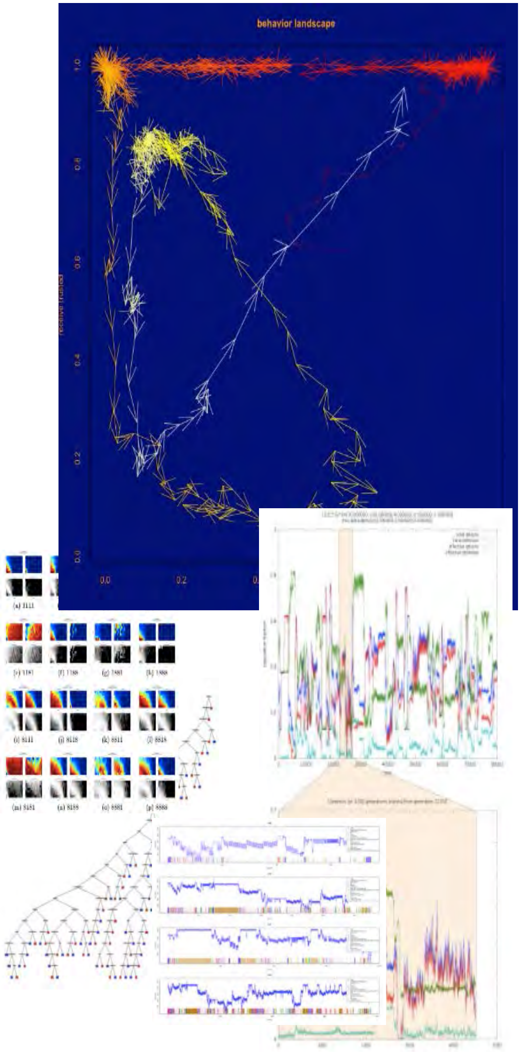
Strategy mutant network with 10K mutations



Epistatic Strategies with k-vulnerabilities

Yielding Understanding of System Dynamical Modes, Emergent Properties, Risk, and Controls

- **"Cyber Security via Signaling Games: Toward a Science of Cyber Security,"** 2014 International Conference on Distributed Computing and Internet Technology. Simulations studies reveals 'talk is cheep and costly signaling via checking market m-coin' is an effective pathway for systems recovery.
- **"Agent-Based Trace Learning in a Recommendation-Verification System for Cybersecurity,"** 2014 IEEE International Conference on Malicious and Unwanted Software. An outline of a ML defenses with a Recommendation/Verifier System that creates agent types for desired properties.
- **"Cyber Security via Minority Games with Epistatic Signaling,"** 2014 International Conference on Bio-inspired Information and Communications Technologies. Preferential early mover advantages have similar effects to maintaining strong global effectiveness measures but will be easier to do.
- <u>**Awarded Best Paper**</u>**: "Compliance Control: Managed Vulnerability Surface in Social-Technological Systems via Signaling Games,"** 2015 ACM CCS International Workshop on Managing Insider Security Threats. Consider a risk-sensitive control to actuate behavior.



Signaling Game Systems and Behavior Dynamics

Rekha
FlashApp
FREE!
Alice

Bob
FlashApp
99¢
Alice

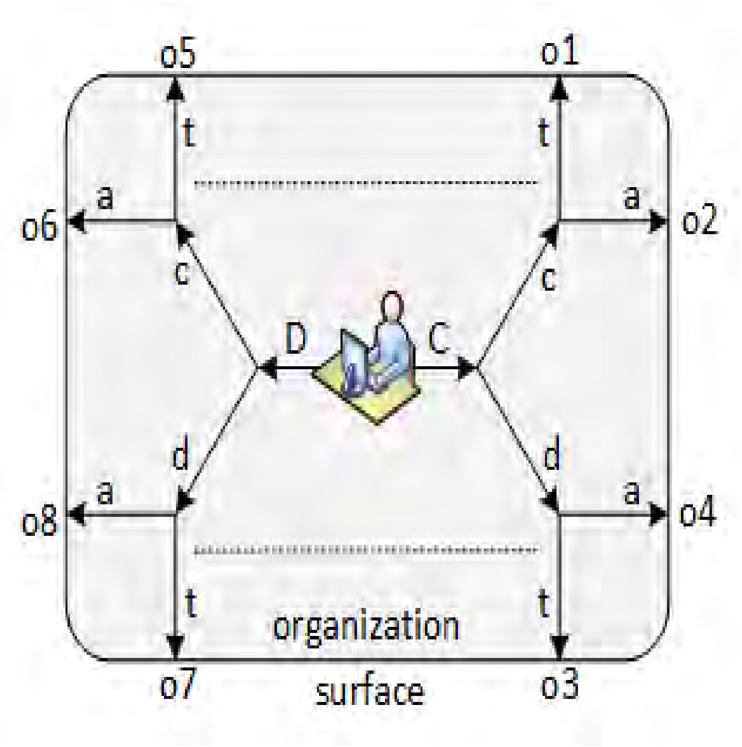
▶

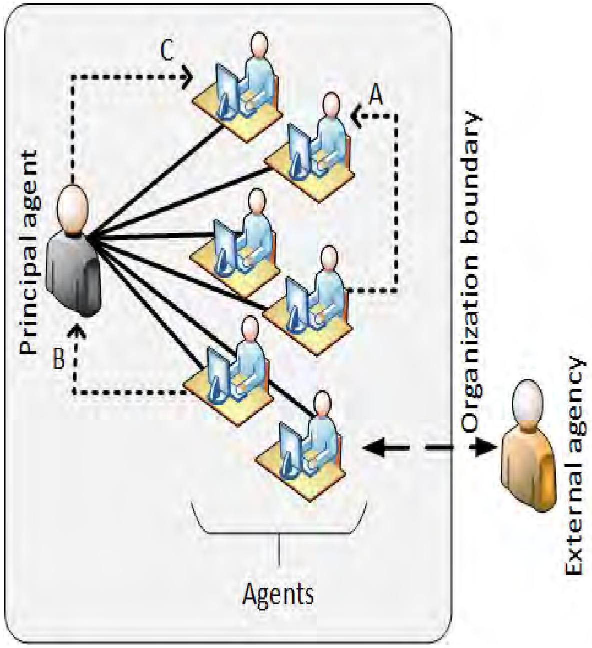
Rekha
FlashApp
FREE!
Alice
Vera
REPORT
on Free App

Bob
FlashApp
99¢
Alice
Vera
REPORT
on Bob's App

Compliance Control and Managing a Vulnerability Surface

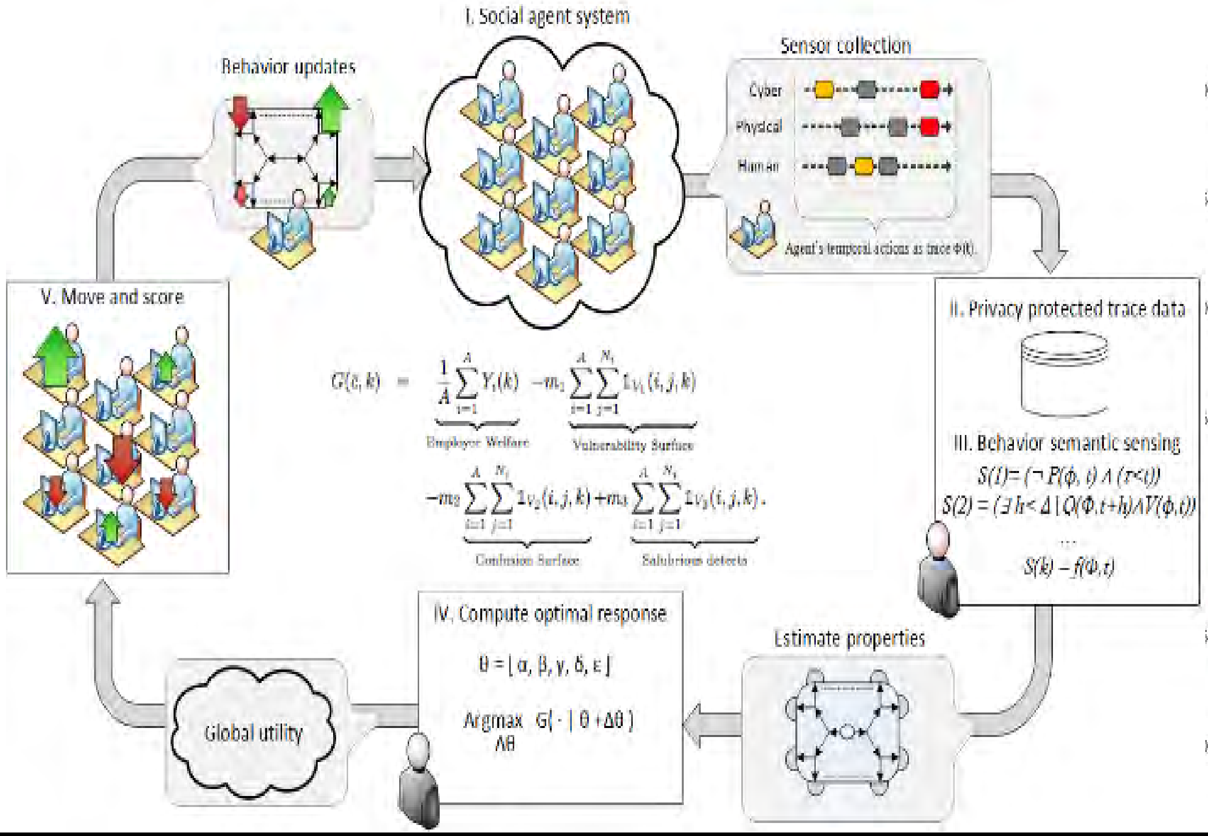
- Game focuses on organizational policies, compliance, and behavioral patterns arising from atomic actions within an organization.
- A deceptive type may optimize a local utility with a non-compliant action. Non-compliance creates vulnerability and confounds risk estimation for a principal.

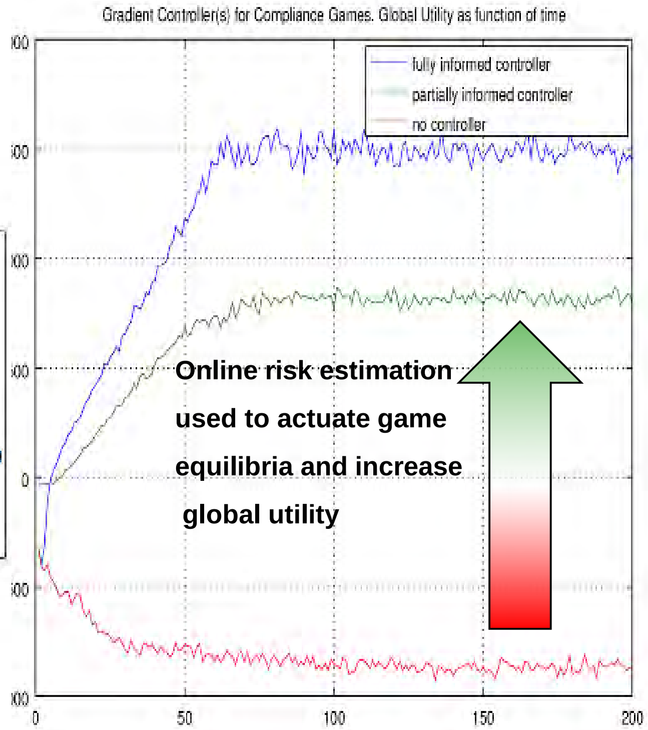




Compliance Control: Managed Vulnerability Surface

- We suggest a counter-strategy: observable risk measures with a ‘honey surface.’
- And create a closed control loop to optimize utility by forming risk estimators from observables and show that in principle: deception is a controllable.


I. Social agent system
Behavior updates
Sensor collection
Cyber
Physical
Human
Agent's temporal actions as trace Φ(t).
II. Privacy protected trace data
III. Behavior semantic sensing
$S(1)= (\neg P(\phi, t) \wedge (\tau<t))$
$S(2) = (\exists h< \Delta \mid Q(\Phi,t+h) \wedge V(\phi,t))$
...
$S(k) - f(\Phi,t)$
V. Move and score
$G(t,k) = \underbrace{\frac{1}{A}\sum_{i=1}^{A} Y_i(k)}_{\text{Employee Welfare}} - m_1 \underbrace{\sum_{i=1}^{A}\sum_{j=1}^{N_1} 1_{V_1}(i,j,k)}_{\text{Vulnerability Surface}} - m_2 \underbrace{\sum_{i=1}^{A}\sum_{j=1}^{N_2} 1_{V_2}(i,j,k)}_{\text{Confusion Surface}} + m_3 \underbrace{\sum_{i=1}^{A}\sum_{j=1}^{N_3} 1_{V_3}(i,j,k)}_{\text{Salubrious defects}}.$
IV. Compute optimal response
θ = [α, β, γ, δ, ε]
Argmax G[· | θ + Δθ]
Δθ
Estimate properties
Global utility



Online risk estimation used to actuate game equilibria and increase global utility

The end, thank you for your interest.

Question?

Contact: wcasey@sei.cmu.edu