

# Understanding & Disrupting the Economics of Cybercrime

## Cyber Security Division 2012 Principal Investigators' Meeting

October 10, 2012

**Nicolas Christin, PI**  
**Senior Systems Scientist**  
**Carnegie Mellon University, CyLab/INI/ECE/EPP**  
[nicolasc@cmu.edu](mailto:nicolasc@cmu.edu)  
**+1 412-268-4432**

# Introduction

- Project addresses **TTA #9 “Cyber Economics”**
  - **“Develop new theories and models of cyber economics** and scientific understanding of the social dimensions of cyber economics.”
  - **“Develop scientific frameworks to incentivize vendors of cyberspace-related technologies** (e.g., encourage use of secure software engineering and analysis practices, software vulnerability detection, security incident forensics) through acquisition, regulation, and standards.”
  - **“Promote an environment** where (1) **users are well informed** about cyber security; and, (2) **individuals have “ownership” of their personal data**, are aware of its provenance, and control its authenticated and authorized distribution, use, destruction with improved understanding of the economic value of such data.”
  - **“Empower cyberspace service providers** e.g., Internet Service Providers, Application Service Providers, registrars, registries, banks, countries, nation-states, etc., to reduce abusive or criminal behavior.”

# Team

- **Nicolas Christin**, Carnegie Mellon University (PI)
- **Alessandro Acquisti**, Carnegie Mellon University (co-PI)
- **Ross Anderson**, Cambridge University (co-PI)
- **Tyler Moore**, Southern Methodist University (co-PI)
- **Ryan Williams**, NCFTA (co-PI)
- **Richard Clayton**, Cambridge University (senior personnel)



# Technical approach

- **Key objective:**
  - Obtain a holistic view of cybercrime economics to better understand which intervention policies work
- Four interdependent technical tasks
  - **Task 1: Designing cybercrime indicators** (*Moore, Acquisti, Anderson, Christin, Clayton, Williams*)
  - **Task 2: Designing data interchange formats and standards** (*Anderson, Christin, Clayton, Moore*)
  - **Task 3: Modeling cybercrime supply chains and the role of Internet intermediaries in reducing crime** (*Christin, Acquisti, Clayton, Moore, Williams*)
  - **Task 4: Modeling attacker/defender behavioral psychology** (*Acquisti, Anderson, Christin*)

# Task 1:

# Designing cybercrime indicators

- Traditional economic indicators (e.g., GDP, unemployment rates, stock market indices) help guide business decisions
- Our goal: design cybercrime indicators to help consumers and firms understand how to best defend themselves
- Designing *feasible* indicators is hard
  - Reporting can be biased
  - Victims may underreport losses
  - Security vendors may overhype threats
  - Selection and sample biases must be avoided
  - Tension between detail and relevancy over time

# Designing cybercrime indicators: our approach

- Step 1: Catalog available data sources for inputs
  - Survey vantage points of data collection for different cybercrime categories
  - Example: how to measure incidence of malware
    1. Study vulnerabilities that are exposed
    2. Track web servers that distribute infections
    3. Count prevalence of infected servers in web search
    4. Track botnets that organize infected computers
    5. Count infected client machines at ISPs
    6. Count criminals perpetrating attacks
    7. Count number of victims
    8. Estimate financial loss caused by malware

# Step 1: Catalog available data sources for inputs

- Examine for different classes of cybercrime the most suitable vantage point for indicators
- Categorize availability of inputs
  1. Privately held, positive incentive to share
  2. Privately held, negative incentive to share
  3. Publicly available, already being collected
  4. Publicly available, not yet collected
  5. Not currently collected, unable to publicly collect

# Step 2-3: devise indicators & match available data

- Step 2: devise indicators with different units of measurement
  - Number of victims
  - Number of criminals/attackers
  - Financial losses
  - Timeliness and comprehensiveness of response
- Step 3: Collaborate with partners to match data to indicators
  - Apply indicators to data already collected in our own cybercrime research
  - Partners who collaborate data will be encouraged to do so via PREDICT

# Task 2: Data interchange format & standards

- **Even if we had good indicators (task 1), how do we share data?**
- Lots of logs record cybercrime activity
- How can we share information about activity
  - without infringing the privacy of innocent individuals ?
  - without compromising commercial confidentiality ?
- How can disparate log data be integrated?
  - logs must stay where they generated, and queries run upon them, but how do ensure that queries are proportionate?
- Much study of these issues for fixed datasets (e.g., census)
- What do appropriate solutions look like for the dynamic world of the Internet ?

# Data processing complexity

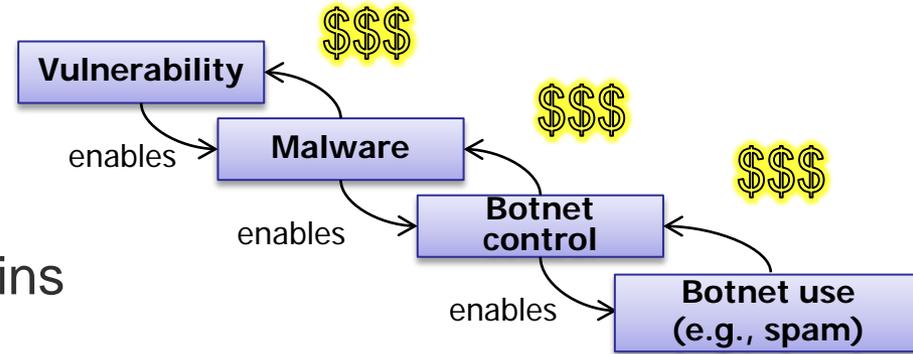
- Email log data tells us about spam senders
- If all email is spam, then sender likely a botnet member
- Would like to map the botnet and mitigate
- Email receiver doesn't want to divulge the levels of traffic they see (or how good their spam detector is)
- An intervening agency might wish to prioritize the worst cases, but need to collate with other data to avoid double counting when IPs are dynamic or carrier grade NAT
- Clear evidence of spamming involves sight of email headers, and false positives can be privacy intrusive

# Task 3: Cybercrime supply chains & intermediaries

- A few cybercrime case studies exist, but rarely connected to the “big picture”
- **Key contributions:**
  - Measurements collected over long time periods from numerous online sources (tight connection w/ Tasks 1 & 2)
  - Unbiased scientific analysis of measurements
  - Methodology to continue measurements and analysis on an sustainable basis (longitudinal work)
- The task two key objectives are to:
  - 1. Uncover concentration points in cybercrime**
  - 2. Build economic models of criminal revenues**
- Pilot study (Moore/Christin) on illicit online pharmaceutical commerce showed approach promising

# Uncovering concentration points

- Measure, overlay and connect different networks to algorithmically identify bottlenecks in criminal supply chains



- Initial approach**

- Build upon our existing measurement infrastructure to collect data on different abuses (malware, search poisoning, phishing, scams, ...)
- Identify characteristics of the infrastructure miscreants rely on
  - E.g., common registrars, hosting platforms, financial mechanisms, common software development techniques...
- Observe changes to tactics used by miscreants after defender intervention and evaluate effectiveness of intervention

# Building economic models

- **Analysis**
  - Uncover collusion across offending domains using network analysis
  - Regression analysis to derive most viable attacker tactics
- Upon identifying sources of concentration
  - Build economic models of costs and revenue sources
  - Identify the most cost-effective intervention points
- **Outcomes**
  - Estimate how much legitimate intermediaries may profit from crimes
  - Find set of incentives needed to change behavior

# Task 4:

## Modeling attacker (and victim) behavior

- **Two objectives**
  - Understand behavioral principles that explain why cyber attacks work even on informed/sophisticated agents
  - Develop behavioral counter-measures to those attacks
- **Approach:** Build upon the nascent interdisciplinary research on cyber-security behavior, informed by studies from behavioral economics, human-computer interaction, and cognitive psychology.

# Modeling attacker and victim behavior

- 1. Understanding the impact of adversarial framing**
  - E.g., how do individuals' judgment and condemnation of cybercrime vary as function of the characteristics of the crime?
- 2. Understanding user biases when dealing with computer risks**
  - Explore behavioral traits and mechanisms that make cybercrime work and security fail
    - E.g., deception (online attackers cheat victims by exploiting similar psychological and behavioral mechanisms as their offline counterparts).
- 3. Improving risk management through better interventions such as messaging and re-personalization**
  - Design soft paternalistic solutions to counter or anticipate those biases.
    - Design technical systems and public policies in manners that take into account the possible or likely biases in individuals' behavior.

# Milestones and deliverables

- **Deliverables**
  - **Monthly reports:** describe activity, technical progress achieved against goals, difficulties encountered, any recovery plans (if needed), plans for the next calendar month, and financial expenditures.
  - **Final technical report:** comprehensive, stand-alone document that will describe the entire research carried out (measurements, models, designs), as well as a list of all financial expenditures incurred through the course of the project.
  - **Project conferences, meetings, and reviews:** In person or by video-conferencing
  - **Peer-reviewed manuscripts** related to the project
  - (If applicable) Standard drafts submitted for public review and commentary
  - (Potentially) Subset of acquired or measured online crime data that could be shared through the PREDICT infrastructure
  - (Potentially) Software prototypes of online crime detection algorithms.

# Schedule

	Year 1	Year 2	Year 3
<b>Cybercrime indicators</b>			
Catalog available data sources	***	**	*
Devise indicators	**	***	**
Match w/ available data	*	*	***
<b>Data interchange standards</b>			
Enumerate privacy barriers	***	**	*
Enumerate aggregated data types	*	**	***
<b>Cybercrime supply chains</b>			
Uncover concentration points	***	**	*
Build economic models	*	**	***
<b>Attacker behavior</b>			
Understand user biases	***	**	*
Understand adversarial framing	**	**	**
Improve risk management	*	**	***

\*\*\* high effort, \*\* medium effort, \* light effort

# Technology transition plan

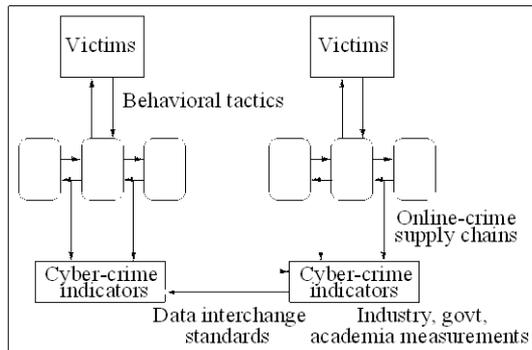
- **Main outcomes of this work:** methods, algorithms, and data exchange standards to aid industry and law enforcement in combating online crime
- Development of a commercial technology (e.g., software) not a primary objective of this project but work may lead to open source products
- **Plan to make publicly available gathered measurement data** (whenever possible)
  - Allows experimental reproducibility
  - Possibly using PREDICT repository
- More generally, concrete transition of our research results to the general public is possible, and yields tangible outcomes
  - Prior impact on search-engine operators, government agencies

# Quad chart

BAA Number: Cyber Security BAA-11-02  
 Title: Understanding and Disrupting the Economics of Cybercrime

Offeror Name: Carnegie Mellon University  
 Date: October 10, 2012

Photograph or artist's concept:



Holistic view of cyber-criminal economics

The figure represents the different areas of investigation and their connections with each other

**Operational capability:**

**Performance targets:** achieve operational understanding of how cyber-crime supply chains work, taxonomy of behavioral tactics used by malfesants to compromise their targets, data interchange standards for sharing cyber-crime data, design of a set of cyber-crime indicators.

Performance of key parameters will be evaluated by their usefulness to law enforcement and industry; as well as peer-reviewed publication output.

No cost of ownership: knowledge and standards will be publicly disseminated.

Project directly addresses all four main topics of TTA #9.

**Proposed Technical Approach:**

Directly addresses all main topics (g(1), g(2), g(3) and g(4)) of TTA#9, "Cyber Economics."

**Tasks:** (1) Designing cyber-crime indicators, (2) Designing data interchange formats and standards, (3) Modeling online-crime supply chains, and (4) Modeling attackers' behavioral psychology.

**Current status:** Fundamental research; design phase.

**Actions done to date:** considerable expertise in acquiring cyber-crime data; preliminary published research in behavioral economics applied to online crime; industry partnerships under way.

This research inscribes itself into the research agendas of all five PIs.

**Schedule, Cost, Deliverables, & Contact Info:**

Three years, Type I project (New Technologies). Yearly retreat planned to refine objectives and assess progress.

**Deliverables:** Peer-reviewed publications related to all four tasks describing recommended algorithms and methodologies; data interchange standard drafts; subset of online crime data that could be shared through PREDICT; (if applicable) software prototypes of online crime detection algorithms;

**Corporate Information:** Offeror: Carnegie Mellon University; **Administrative P.O.C.:** Kristen Jackson; Office of Sponsored Programs; 5000 Forbes Ave, Warner Hall, 4<sup>th</sup> Floor; Pittsburgh, PA 15213; **Technical P.O.C.:** Nicolas Christin; CIC Room 2108; 4720 Forbes Ave; Pittsburgh, PA 15213