CYBER SECURITY DIVISION

2013 PRINCIPAL INVESTIGATORS'

UNDERSTANDING AND DISRUPTING THE ECONOMICS OF CYBERCRIME

Carnegie Mellon University (and subcontractors)

Nicolas Christin

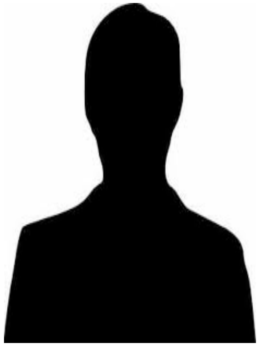
September 19, 2013

Homeland Security

Science and Technology

Team profile

- **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**, National Cyber Forensics Training Alliance (co-PI)
- **Richard Clayton**, Cambridge University (senior personnel)



Why we should look at economics

- Cyber-security attacks cost money
  - Estimates vary and are highly disputed, but:
  - A couple of hundreds of millions of dollars per year in **direct costs** to victims
- **Indirect costs** are killing us!

Criminal revenue	Cost in policing
Large botnet: **1/3 of the spam on the Internet** Made its owners **2.7 million USD** in a year	How much did we invest in email spam reduction over that year? **> 1 Billion USD**

- Can we be smarter? How?
  - Focusing limited law enforcement resources on the points where they matter the most

Approach overview

- Criminals are mostly in it for the money
  - Do cost/benefit analysis too!
- **Very** economically rational
  - **Will** give up if costs become too high
    - "Visa is burning us with napalm" (some illicit Rx seller on the Internet)
    - "Will close shop until Bitcoin value stabilizes" (a drug dealer on the Silk Road anonymous marketplace)

- Need to find and exploit **concentration** points (that can lead to effective financial pressure on criminals
- Need to understand why victims fall for attacks, what are defenses deemed acceptable by the public

Network measurements + economic and behavioral analysis

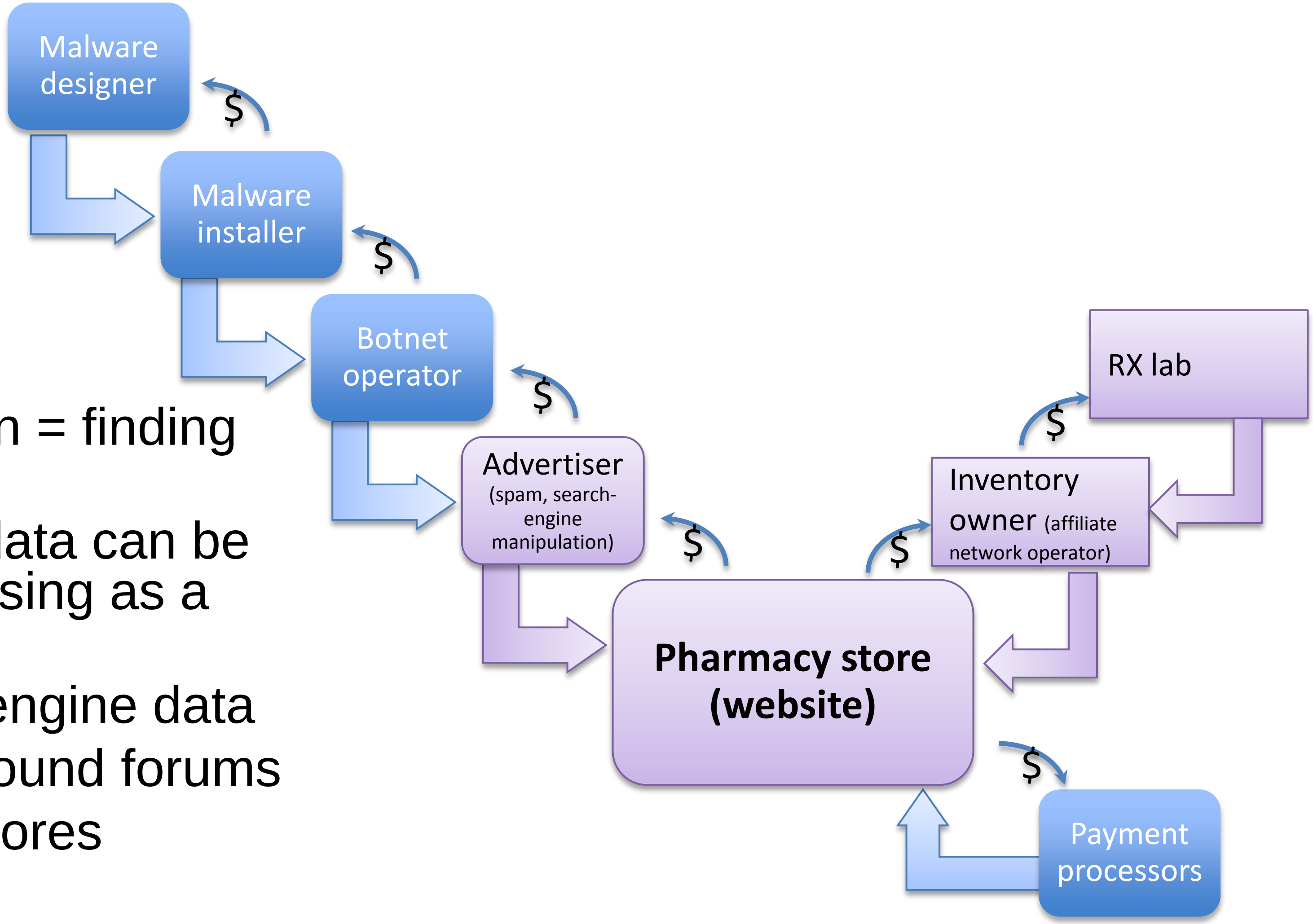
Task 1: Designing cybercrime indicators

- Catalog available data sources for input
  - Survey vantage points of data collection for different cybercrime categories
- Categorize availability of inputs (public vs. private, incentive vs. disincentive to share, …)
- Examples of existing inputs:
  - Known “bad” URLs (e.g., malware databases)
  - Known “bad” IPs
- Design novel indicators
  - E.g., Indicators of certain website platforms known to be vulnerable to compromise (might be measured)
    - “Google dorks”
    - Features of vulnerable CMS
    - …

Task 2:
Sharing indicator data

- **Even when we have 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), less so for dynamic data (Internet)
- Which data can be made public?
  - Easy answer: data that is already public in the first place (fortunately there is lots of it, see next slide)
  - What about non-public data?
    - Necessary: Anonymization
    - Necessary: Non-interference with measurement itself (cf. Heisenberg principle)
    - No “sufficient” condition – case by case evaluation?

Task 3:
Uncovering cybercrime supply chains



- Monetization = finding customers
- So a lot of data can be found by posing as a customer
  - Search engine data
  - Underground forums
  - Actual stores
  - …

Task 4:
Modeling attacker and victim behavior

Conduct user experiments to:

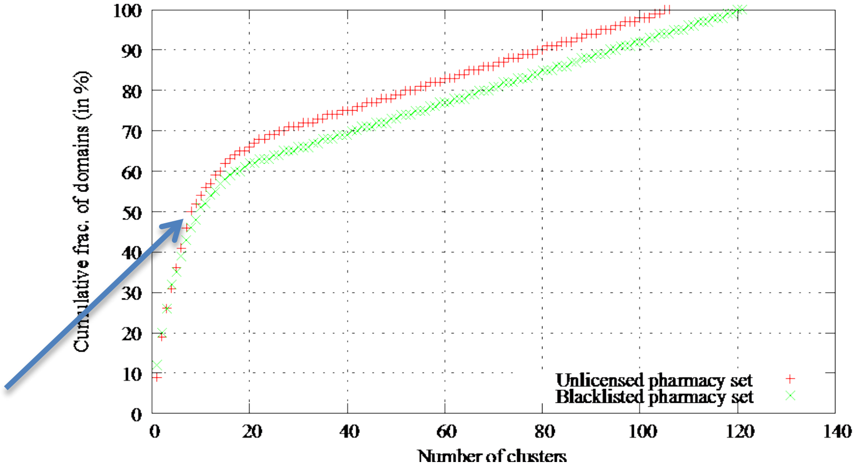
1. **Understand the impact of framing**
   - E.g., how do individuals' judgment and condemnation of cybercrime vary as function of the characteristics of the crime?

2. **Understand 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. **Improve 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.

Benefits of the approach

- **Tangible impact on society**
  - Impact adversary's behavior
    - Some evidence from pharmaceutical affiliates after payment processor crackdown
- **Reduce cost of law enforcement and policing**
  - Taking down ~8-10 pharmaceutical labs vs. ~4,000 online pharmaceutical shops



- Help us **determine what can be addressed by social norms vs. economics vs. technological means**
  - Evidence from pharmaceutical research: people are interested in buying from these shops; why?
- **Help us come up with appropriate defenses by understanding attackers**
  - Syrian Electronic Army ≠ "Canadian Pharmacy" ≠ Nation-state adversary

Alternatives

- **Formal economic models**
  - Lots of assumptions that do not necessarily hold in practice
    - Perfect information
    - Perfect strategy execution…

- **Traditional computer security research**

1. Find an attack (or invent a new attack)
2. Build a defense
3. Repeat

- **Other cybercrime measurement research**
  - Stefan Savage, Vern Paxson, and their collaborators
  - Less focus on building economic models; no behavioral work

  - *Not so much competition as much as complement to our work*
    - The more data we get, the better picture we have

Current status

- **Major milestones so far: academic contributions**

  Identifying Risk Factors for Web Server Compromises
  M. Vasek and T. Moore. Working paper (in submission).
  Empirical Analysis of Factors Affecting Malware URL Detection
  M. Vasek and T. Moore. Proc. *E-Crime'13.*
  Pick Your Poison: Pricing and Inventories at Unlicensed Online Pharmacies
  N. Leontiadis, T. Moore and N. Christin. Proc. *ACM EC'13.*
  (more to come in Y2)

- **Deliverables** (besides academic contributions)
  - Monthly reports delivered as needed
  - Software & data: see transition activities
- **Schedule**
  - Behavioral task started a bit late; catching up right now
  - Data interchange standards task slightly more complex than thought initially (adverse incentives for industrial actors)
    - Work on indicators (task 1) very helpful
  - Rest of the project on schedule

Next steps

- **Plans for remainder of the effort**
  - Continue on our four tasks
  - Significant work on indicators (task 1), behavioral analysis (task 4) to take place in Y2
  - Connection with related efforts we are starting
    - E.g., analysis of zero-day markets
      - As part of cybercrime supply chains research (task 3)
- **Technology Transition Activities**
  - Peer-reviewed publications: knowledge product
    - Models, methodologies, description
  - Discussion/transition of knowledge with relevant agencies

  - Working on making (part of) our datasets public (part of task 2)
    - Harmless for data that was publicly available in the first place
    - Conservative approach with non-public data
  - Working on making measurement software (as well as software helpers) public/open-source as well

Contact Information

Nicolas Christin

Assistant Research Professor
Carnegie Mellon University
Electrical and Computer Engineering, and CyLab
CIC Room 2108
4720 Forbes Ave
Pittsburgh, PA 15213, USA

Email: nicolasc@cmu.edu
Web: https://www.andrew.cmu.edu/user/nicolasc
Twitter: @nc2y
Phone: 412-268-4432 (rarely used)