CIS 700/003: Distributed Systems meet Social Networks

Phishing

March 30, 2010
Where are we?

- Measurement and analysis of MAD systems
- Building MAD systems
- Faults and Misbehavior
- Internet crime
  - Spoofing
  - Sybil attack
  - Denial of Service
  - Phishing
- Privacy and confidentiality
- Novel opportunities
- Experience
An Inquiry into the Nature and Causes of the Wealth of Internet Miscreants

Jason Franklin  Vern Paxson  Adrian Perrig  Stefan Savage

CCS 2007
Discussion points

- **Characteristics of the market**
  - Size, diversity, professionalism, types of services offered
  - Exploitation chain: Cashiers, confirmers, ...
  - Verification procedures (to get the 'voice' attribute)

- **Challenges in doing such a study**
  - Why not start your own underground market?
  - Why not hack the server to get at the private messages?
  - Why not participate in the market a bit to see what actually happens?

- **Purpose of the automated bots**

- **Phishing the phishers?**

- **Idea of establishing a 'lemon market'**
Some of your discussion points

- Privacy concerns
- Most details unknown or unverifiable
- How comprehensive is this study?
  - How many other markets are there?
  - Other types of fraud?
- Undersampling
- What about the private messages?
- Quality of the analysis
A Profitless Endeavor: Phishing as Tragedy of the Commons

Cormac Herley and Dinei Florêncio

NSPW 2008
Motivation

- Conventional wisdom: Phishers are making huge amounts of money
  - FTC Identity Theft Survey: $47 billion (!) lost annually
  - No great skills required; phishing money is 'easy money'
  - 18-year old claims to be making $3-4k a day

- But: No hard data
  - Phishers don't file tax returns and don't report to the SEC
  - Most stories are anecdotal or based on surveys

- Question: Can all this really be true?
Observation: Phishing is a low-skill job and requires no significant investments

- Basic computer skills are enough
- No barrier to entry
- All that is needed is ~$60 worth of software

Shouldn't low-skill jobs pay like low-skill jobs?

- Suppose number of dollars to be phished is approx. constant
- If phishers really made as much as surgeons, wouldn't new entrants increase competition and drive returns down?
- So, are phishers living by a different set of economic laws than the rest of us?
Idea: Apply basic economics

- Let's apply some basic economic theories
  - Sustainable harvest as a function of sustained effort
  - The same theory that is also applied to real 'fishing'

- $X := \text{Total pool of phishable dollars}$
  - Exact amount irrelevant; what matters is that it is finite

- $E := \text{Total effort of all phishers}$
  - Perhaps the number of hours they put in

- $H(X, E) := \text{Total dollar harvest per unit time}$
How does $X$ develop over time?

- Without phishing, $X$ grows quickly
- Growth dep. on $X$

$$\frac{dX}{dt} = f(X)$$

- Exponential growth pattern, but bounded resource pool
- Exact shape does not matter, only that it goes down at the end
What if phishing occurs?

- Phished dollars are removed from the pool:
  \[ \frac{dX}{dt} = f(X) - m \cdot H(X, E) \]

- Each dollar stolen may cause more than one dollar to leave the pool (factor m)
  - Victim may be especially careful with any remaining money
  - People learn about danger and become more careful
  - Actual value of m does not matter; can assume m=1

- How does X develop over time?
  - \( f(X) < m \cdot H(X, E) \): Pool of phishable dollars shrinks
  - \( f(X) > m \cdot H(X, E) \): Pool grows, but as \( f(X) \) falls, growth limited
  - In equilibrium, \( f(X) = m \cdot H(X, E) \) ('sustainable' harvest!)
Calculating the sustainable harvest

- At sustained level, pool is stable

\[ H(X, E) = \frac{1}{m} \cdot f(X) \]

- But the stable X is related to the effort E that is used to phishing it
  - No effort: X is maximal; lots of effort: X is zero
  - In between: Inverse relationship

- So we can express the sustainable harvest as a function of the effort alone
  - \( H(X, E) = H(X(E), E) = H(E) \)
Summary so far

- Sustainable harvest depends on phishing effort
  - E=0: No phishing, therefore no harvest
  - Large E: Lots of phishing causes 'overfishing'; no harvest!
The tragedy of the commons

What are the economics?
- Cost (time, opportunity cost) is proportional to $E$: $C(E) = a \cdot E$
- Average return: $R_{avg}(E) = H(E)/E$

How do phishers behave?
- A single rational phisher would operate at $E_x$
- Multiple phishers: Incentive to increase the effort, plus incentive for additional phishers to enter the market
  - An instance of the well-known tragedy of the commons!
- Result: Stabilization where $R_{avg}(E) = a$
Implications for phishing

- In economic equilibrium, total revenue equals total cost incurred by all the actors
  - No 'easy money'
- Each participant makes only as much as he would have made in opportunities elsewhere
  - If phishing does not require special qualifications, the revenue will therefore be also low
- Increasing phishing effort causes the total phishing revenue to decline!
  - The harder the phishers try, the worse their situation gets!
  - 1954 economics paper: "This is why fishermen are not wealthy, despite the fact that the fishery resources of the sea are the richest and most indestructible available to man. By and large, the only fisherman who makes it rich is the one who makes a lucky catch."
  - Increasing effort is a sign of failure and not of success!
New and better ways of phishing

- What if phishing becomes easier?
  - More automation, better tools $\rightarrow$ New factor $b < a$
  - Consequence: New equilibrium at $R_{avg}(E) = b$
  - Phishers with lower cost structure drive out the old ones
  - Revenue determined by the least skilled people able to accomplish the task!
  - Thus, total amount of losses due to phishing are probably decreasing over time!
Why are people still phishing?

- **Expectation:** Phishers stop phishing when it becomes unprofitable for them
  - When revenue falls below their opportunity cost, \( R_{\text{avg}}(E) < a \)
  - But: Several counterexamples in economics, e.g., English privateers pillaging Spanish ships in 1625-1630. Why?

- **Several potential reasons:**
  - Emotional ties
    - Example: Fisherman in the third generation, inherits ship from father
  - Gambling
    - Example: Gold rush. Most prospectors get little, but some get lucky!
  - Poor information
    - Example: English privateers from above

Probably apply to phishing scenario: Constant reports of "easy money" in the media
Objections

- What about all that data showing that phishing losses are huge?
  - Almost certainly enormous amount of exaggeration
  - Example: $47bn figure in FTC report
    - Total reported 2003 profits of the five top U.S. banks was $59bn
    - $50bn would give $500k income to 100k identity thieves, which is four times the number of cardiologists in the U.S.

- Paper criticizes methodology in several studies
  - Example: Selection bias, self-reporting of losses, ...
  - In most victim surveys, margin of error for 95% confidence interval is larger than the estimated phishing rate!
Summary

- Phishers probably make less than you think
  - Paper estimates annual U.S. losses at $61 million
  - Average phisher probably makes 100s of $, not thousands

- Phishing: A tragedy of the commons
  - Phishers cannot prevent others from entering the market
  - Result: Competition and 'overphishing' push the return towards the opportunity cost of the phishers
    - In other words, they cannot make more money through phishing than they could in other occupations for which they are qualified

- Many phishing studies are problematic
  - Reliance on self-reporting, email surveys with low return rates, etc.
Take-away points

- A brief look at Internet crime
  - Well organized, with markets, specialization, exploitation chains, mechanisms for establishing credibility...
  - Large (113,000 unique nicks!) and diverse (CC fraud, identity theft, spam, extortion/DoS, botnets, logins, ...)

- An example of how economics can be applied in this context
  - Use concepts like sustainable harvest and tragedy of the commons to study phishing
Where are we?

- Measurement and analysis of MAD systems ✔
- Building MAD systems ✔
- Faults and Misbehavior ✔
- Internet crime
  - Spoofing ✔
  - Sybil attack ✔
  - Denial of Service ✔
  - Phishing ✔
- Privacy and confidentiality
- Novel opportunities
- Experience
Recap of this section

- **Spoofing**
  - Piatek et al.: Triggering DMCA notices against arbitrary IPs
  - Handling of complaints; examples of court cases involving IPs

- **Sybil attacks**
  - SybilGuard: Using social networks to mitigate Sybil attacks
  - SybilLimit: Improvement of SybilGuard, provable guarantees

- **Denial of service**
  - Defense by offense: Using bandwidth as a currency
  - Capabilities: Explicitly authorizing all network flows

- **Phishing**
  - Wealth of miscreants: Study of an underground market
  - Herley/Florêncio: Phishing as a tragedy of the commons
Where are we?

- Measurement and analysis of MAD systems ✔
- Building MAD systems ✔
- Faults and Misbehavior ✔
- Internet crime
  - Spoofing ✔
  - Sybil attack ✔
  - Denial of Service ✔
  - Phishing ✔
- Privacy and confidentiality
- Novel opportunities
- Experience

Next time
Next week you will learn:

Why an 'anonymized' data set doesn't necessarily protect your privacy