Spam & Spammer Profits

CS 161 - Computer Security
Profs. Vern Paxson & David Wagner

TAs: John Bethencourt, Erika Chin, Matthew Finifter, Cynthia Sturton, Joel Weinberger
http://inst.eecs.berkeley.edu/~cs161/

April 21, 2010
Thinking About Economics

• Given the rise of cybercrime-fueled Internet attacks, where should we be investing our limited security resources?
  □ Preventing host compromise?
  □ Policing networks, rolling up botnets?
  □ Other?

• We have structural disadvantages
  □ Defenses public, attacker develops/tests in private
    • Arms race where best case for defender is to “catch up”
  □ Attacker not tied to any particular technology; cheaper for them to change than us
  □ Minimal deterrence
  □ Significant value proposition for attacker
Thinking About Economics, con’t

• Given the rise of cybercrime-fueled Internet attacks, where should we be investing our limited security resources?

  Premise: We’re unlikely to spend efficiently until we understand the economics of the bad guy

• We have structural disadvantages
  - Defenses public, attacker develops/tests in private
    - Arms race where best case for defender is to “catch up”
  - Attacker not tied to any particular technology; cheaper for them to change than us
  - Minimal deterrence
  - Significant value proposition for attacker
Monetizing Spam

• In what ways can spammers make money off of sending spam?
  – And who has incentives to thwart these?
    • (Other than law enforcement)

• Scheme #1: advertise goods or services
  – Examples: fake Rolexes, Viagra, university degrees
  – Profit angle: increased sales
  – Who’ll try to stop: brand holders

• Scheme #2: phishing
  – Profit angle: transfer $$$ out of accounts; sell accounts to others; use accounts for better spamming (e.g. Facebook)
  – Opponents: issuers of accounts
  – Note: targeted phishing (“spear-phishing”) doesn’t actually need much in the way of spam due to low volume
Monetizing Spam, con’t

• Scheme #3: scams
  – Examples: pen pal relationships, 419 (“Nigerian”)
  – Profit angle: con victim into sending money
  – Opponents: scambaiters (419eater.com)

• Scheme #4: recruiting crooks/underlings
  – Examples: money mules, reshippers
  – Profit angle: more efficient cybercrime
  – Opponents: ?

• Scheme #5: recruiting bots
  – Examples: “important security patch!”, “someone sent you a greeting card!”
  – Profit angle: get malware installed on new machines
  – Opponents: ?
Monetizing Spam, con’t

• Scheme #6: pump-and-dump
  – Example: “Falcon Energy (FPK) is about to go through the roof! Don’t miss out on seriou$ Profit$!”
  – Profit angle: penny-stock momentarily goes up, dump pre-bought shares when it does
  – Opponents: Securities and Exchange Commission
  – Note: unlike other monetization techniques, the “back channel” is out-of-band
    • No link in messages back to the scammer
Are Bots & Spam the New Black Gold?

- Spam finance elements:
  - Retail-cost-to-send vs. Profit-per-response
  - Key missing element: spams-needed-per-response, i.e., conversion rate

Storm worm 'making millions a day'

Compromised machines sending out highly profitable spam, says IBM security strategist

Clive Akass, Personal Computer World 11 Feb 2008

The people behind the Storm worm are making millions of pounds a day by using it to generate revenue, according to IBM’s principal web security strategist.

Joshua Corman, of IBM Internet Security Systems, said that in the past it had been assumed that web security attacks were essential ego driven.

How can we measure this? Seemingly only knowable by the spammers themselves.
Welcome to **Storm**!

Would you like to be one of our newest bots? Just read your postcard!

(Or even easier: just wait 5 seconds!)
The Storm botnet

Reachability check

Overnet (UDP)
The Storm botnet

Operated using "bulletproof hosting"
Hosting Plans

We offer a complaint-resistant hosting to host your sites, which are specified in mass mailings.

We decided to bring visitors to your web site through unsolicited mass emails? Wonderful idea! You certainly expect a boom visits. But! As in any ointment and then not pass without a spoon of tar ... Alas, but your wonderful site, shortly after the start of spam mail, will be closed due to flood of complaints from postal services. Is there a way to avoid these problems? Of course! Our complaint-resistant hosting simply ignores any complaints, all postal services, and you can be rest assured about the performance of their sites - they will not be closed. And you get new customers, expand their business and increase their sales and revenue, thanks to spam mailing lists.
**Obuzoustoychivy hosting** is more expensive than usual, but you will have the full guarantee that your site no one ever closes, it will always be available to your customers!

<table>
<thead>
<tr>
<th>MINI PLAN</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume disc</td>
<td>400 MB</td>
</tr>
<tr>
<td>Domains</td>
<td>1</td>
</tr>
<tr>
<td>Traffic *</td>
<td>Unlimited</td>
</tr>
<tr>
<td>FTP-access</td>
<td>there is</td>
</tr>
<tr>
<td>MySQL database</td>
<td>there is</td>
</tr>
<tr>
<td>Control panel</td>
<td>there is</td>
</tr>
<tr>
<td><strong>COST</strong></td>
<td><strong>4 000 rub. / 1 month.</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>STARTER PLAN</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume disc</td>
<td>500 mb</td>
</tr>
<tr>
<td>Domains</td>
<td>3</td>
</tr>
<tr>
<td>Traffic *</td>
<td>Unlimited</td>
</tr>
<tr>
<td>FTP-access</td>
<td>there is</td>
</tr>
<tr>
<td>MySQL database</td>
<td>there is</td>
</tr>
<tr>
<td>Control panel</td>
<td>there is</td>
</tr>
<tr>
<td><strong>COST</strong></td>
<td><strong>5 000 rub. / 1 month.</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BUSINESS PLAN</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume disc</td>
<td>1000 mb</td>
</tr>
<tr>
<td>Domains</td>
<td>7</td>
</tr>
<tr>
<td>Traffic *</td>
<td>Unlimited</td>
</tr>
<tr>
<td>FTP-access</td>
<td>there is</td>
</tr>
<tr>
<td>MySQL database</td>
<td>there is</td>
</tr>
<tr>
<td>Control panel</td>
<td>there is</td>
</tr>
<tr>
<td><strong>COST</strong></td>
<td><strong>7 000 rub. / 1 month.</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PREMIUM PLAN</th>
<th></th>
</tr>
</thead>
</table>
Spam campaign mechanics

Botmaster

HTTP proxies

HTTP

Proxy bots

TCP

Workers
Campaign mechanics: harvest

Diagram showing the relationship between Botmaster, HTTP proxies, HTTP, Proxy bots, TCP, and Workers.
Campaign mechanics: spamming
<table>
<thead>
<tr>
<th>MACRO</th>
<th>SEEN LIVE</th>
<th>FUNCTIONALITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0)</td>
<td>✓</td>
<td>Spam target email address.</td>
</tr>
<tr>
<td>(A)</td>
<td>✓</td>
<td>FQDN of sending bot, as reported to the bot as part of the preceding C&amp;C exchange.</td>
</tr>
<tr>
<td>(B)</td>
<td>✓</td>
<td>Creates content-boundary strings for multi-part messages.</td>
</tr>
<tr>
<td>(Cnum)</td>
<td>✓</td>
<td>Labels a field’s resulting content, so it can be used elsewhere through (V); see below.</td>
</tr>
<tr>
<td>(D)</td>
<td>✓</td>
<td>Date and time, formatted per RFC 2822.</td>
</tr>
<tr>
<td>(E)</td>
<td>✓</td>
<td>ROT-3–encodes the target email address.</td>
</tr>
<tr>
<td>(Fstring)</td>
<td>✓</td>
<td>Random value from the dictionary named string. ²</td>
</tr>
<tr>
<td>(Gstring)</td>
<td>✓</td>
<td>Line-wrap string into 72 characters per line.</td>
</tr>
<tr>
<td>(Hstring)</td>
<td>✓</td>
<td>Defines hidden text snippets with substitutions, for use in HTML- and plain-text parts.</td>
</tr>
<tr>
<td>(I)</td>
<td>✓</td>
<td>Random number between 1 and 255, used to generate fake IP addresses.</td>
</tr>
<tr>
<td>(Istring)</td>
<td>✓</td>
<td>Produces quoted-printable “=20” linewrapping.</td>
</tr>
<tr>
<td>(K)</td>
<td></td>
<td>IP address of SMTP client.</td>
</tr>
<tr>
<td>(M)</td>
<td>✓</td>
<td>6-character string compatible with Exim’s message identifiers (keyed on time).</td>
</tr>
<tr>
<td>(N)</td>
<td></td>
<td>16-bit prefix of SMTP client’s IP address.</td>
</tr>
<tr>
<td>(Ostring;num)</td>
<td>✓</td>
<td>Randomized message identifier element compatible with Microsoft SMTPSVC.</td>
</tr>
<tr>
<td>(Pnum₁[-num₂];string)</td>
<td>✓</td>
<td>Random string of num₁ (up to num₂, if provided) characters taken from string.</td>
</tr>
<tr>
<td>(Qstring)</td>
<td>✓</td>
<td>Quoted-printable “=&quot; linewrapping.</td>
</tr>
<tr>
<td>(Rnum₁-num₂)</td>
<td>✓</td>
<td>Random number between num₁ and num₂. Note, special-cased when used with (D).</td>
</tr>
<tr>
<td>(Ustring)</td>
<td>✓</td>
<td>Randomized percent-encoding of string.</td>
</tr>
<tr>
<td>(Vnum)</td>
<td>✓</td>
<td>Inserts the value of the field identified by (Cnum).</td>
</tr>
<tr>
<td>(W)</td>
<td></td>
<td>Time and date as plain numbers, e.g. “20080225190434”.</td>
</tr>
<tr>
<td>(X)</td>
<td></td>
<td>Previously selected member of the “names” dictionary.</td>
</tr>
<tr>
<td>(Ynum)</td>
<td>✓</td>
<td>8-character alphanumeric string, compatible with Sendmail message identifiers.</td>
</tr>
<tr>
<td>(Z)</td>
<td>✓</td>
<td>Another Sendmail-compatible generator for message identifiers.</td>
</tr>
</tbody>
</table>

Table 2: Storm’s spam-generation templating language.
Figure 2: Snippet of a spam template, showing the transformation of an email header from template (top) to resulting content (bottom). The $>$-symbol indicates line continuations. Bold text corresponds to the formatting macros and their evaluation.
Campaign mechanics: spamming

Botmaster

HTTP proxies

HTTP

Proxy bots

TCP

SPAM cans
Campaign mechanics: reporting
Welcome to Storm! What can we sell you?
Anatomy of a modern Pharma spam campaign

Spammer
- Paid commission
- Bulk email sender
- Open proxies/botnet

Spam emails
- Email with link
- Other recipients

Email recipient
- Impotent? Curious?
- Clicks on link

Affiliate program
- email target page (disposable hosting)

Redirects
- Commission (up to 40%)

Viagra/ pharmacy site

Sale made
- Typical value $50-100?

Generic pills
- Sildenafil citrate
- Non-US supply (India, Mexico etc.)
- Patent expired
- Mere cents per pill = massive markup

Shipped to customer

Diagram by Stuart Brown
modernlifefisrubbish.co.uk
These folks seem trustworthy …
... how about these?
If we control these... we can monitor & influence these...
Botnet infiltration

• **Key idea:** distributed C&C is a **vulnerability**
  - Botnet authors like de-centralized communications for scalability and resilience, but…
  - … to do so, they trust their bots to be good actors
  - If you can *modify* the right bots you can *observe* and *influence* actions of the botnet

• Thanks to *E-Card* spam, we can easily acquire Storm bot binaries …
  - … and run them within controlled *GQ* honeyfarm environment

• With a lot of elbow grease, we reverse-engineered the C&C protocol …

• … so we can *record* all C&C sent through us …
Template points to spammer’s server

Modified template points to our server

Under our control

Bot Controller

Proxying Bot

Re-writing Proxy

WWW server

real email accounts

our webmail accounts

our domain

n00b
Spam conversion experiment

- Experimented with Storm March 21 – April 15, 2008
- Instrumented roughly 1.5% of Storm’s total output

<table>
<thead>
<tr>
<th></th>
<th>Pharmacy Campaign</th>
<th>E-card Campaigns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Postcard</td>
</tr>
<tr>
<td>Worker bots</td>
<td>31,348</td>
<td>17,639</td>
</tr>
<tr>
<td>Emails</td>
<td>347,590,389</td>
<td>83,665,479</td>
</tr>
<tr>
<td>Duration</td>
<td>19 days</td>
<td>7 days</td>
</tr>
</tbody>
</table>
## Spam pipeline

<table>
<thead>
<tr>
<th>Sent</th>
<th>MTA</th>
<th>Inbox</th>
<th>Visits</th>
<th>Conversions</th>
</tr>
</thead>
<tbody>
<tr>
<td>347.5M</td>
<td>82.7M (24%)</td>
<td></td>
<td>10,522 (0.003%)</td>
<td>28 (0.000008%)</td>
</tr>
<tr>
<td>83.6 M</td>
<td>21.1M (25%)</td>
<td></td>
<td>3,827 (0.005%)</td>
<td>316 (0.00037%)</td>
</tr>
<tr>
<td>40.1 M</td>
<td>10.1M (25%)</td>
<td></td>
<td>2,721 (0.005%)</td>
<td>225 (0.00056%)</td>
</tr>
</tbody>
</table>

The diagram shows a flow of targeted addresses through different stages: 
- A: Email not delivered
- B: Blocked by spam filter
- C: Ignored by user
- D: User left site
- E: Crawler/purchaser
Spam pipeline

---

Pharma: 12 M spam emails for one "purchase"

Sent

347.5M

83.6M

40.1M

A

targeted addresses

MTA

Visits

Conversions

40.1 M

10.1M (25%)

2,721 (0.005%)

225 (0.00056%)

E-card: 1 in 10 visitors execute the binary

Effects of Blacklisting (CBL Feed)

Unused

Other filtering

Effective

Delivery Rate Prior to Blacklisting

Delivery Rate Post Blacklisting

Effects of Blacklisting (CBL Feed)
Spam filtering software

- The fraction of spam delivered into user inboxes depends on the spam filtering software used
  - Combination of site filtering (e.g., blacklists) and content filtering (e.g., spamassassin)
- Difficult to generalize, but we can use our test accounts for specific services

<table>
<thead>
<tr>
<th>SPAM FILTER</th>
<th>PHARMACY</th>
<th>POSTCARD</th>
<th>APRIL FOOL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gmail</td>
<td>0.00683%</td>
<td>0.00176%</td>
<td>0.00226%</td>
</tr>
<tr>
<td>Yahoo</td>
<td>0.00173%</td>
<td>0.000542%</td>
<td>none</td>
</tr>
<tr>
<td>Hotmail</td>
<td>none</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>Barracuda</td>
<td>0.131%</td>
<td>N/A</td>
<td>0.00826%</td>
</tr>
</tbody>
</table>

Fraction of spam sent that was delivered to inboxes
Pharma: 12 M spam emails for one "purchase"

- Sent
- MTA
- Visits
- Conversions
- Inbox

Response rates by country

- Response rates by country
- Two orders of magnitude
- No large aberrations based on email topic

Response rates for Pharmacy Email

Response rates for Self-prop Email
Site needs to be up hours to days to reap real users rather than just crawlers.
Spam pipeline

<table>
<thead>
<tr>
<th></th>
<th>Sent</th>
<th>MTA</th>
<th>Inbox</th>
<th>Visits</th>
<th>Conversions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pharma</td>
<td>347.5M</td>
<td>82.7M (24%)</td>
<td></td>
<td>10,522 (0.003%)</td>
<td>28 (0.000008%)</td>
</tr>
<tr>
<td></td>
<td>83.6 M</td>
<td>21.1M (25%)</td>
<td>3,827 (0.005%)</td>
<td>316 (0.00037%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>40.1 M</td>
<td>10.1M (25%)</td>
<td>2,721 (0.005%)</td>
<td>225 (0.00056%)</td>
<td></td>
</tr>
</tbody>
</table>

Pharma: 12 M spam emails for one “purchase”

E-card: 1 in 10 visitors execute the binary
The Spammer’s Bottom Line

- 28 purchases in 26 days, avg. “sale” ~$100
  - Total: $2,731.88, $140/day
- But: we interposed on only ~1.5% of workers:
  - $9,500/day (8,500 new bots per day)
  - $3.5M/year (back of envelope - be very careful!)
    - Though if selling Viagra via Glavmed affiliation, cut is 40%
- Storm: service provider or integrated operation?
  - Retail price of spam ~$80 per million
    - Pharmacy spam would have cost 10x the profit!
  - Strongly suggests Storm operates as an integrated operation rather than a reseller