Internet Crime as a Financial Ecosystem

As the Internet evolved, so did cybercrime...

- Liberal “slide borrowing” from C. Wilson @ NEU
- Most content from recent papers by Savage/Voelker et al.
Drive-by Exploits

• Browsers are extremely complex
  – Millions of lines of source code
  – Rely on equally complex plugins from 3rd party developers
    • e.g. Adobe Flash, Microsoft Silverlight, Java

• Must deal with untrusted, complex inputs
  – Network packets from arbitrary servers
  – HTML/XML, JavaScript, stylesheets, images, video, audio, etc.

• Recipe for disaster
  – Attacker directs victim to website containing malicious content
  – Leverage exploits in browser to attack OS and gain persistence
Browser Architecture circa-2018

- Browsers handle many types of complex input
  - HTML/XML
  - JavaScript
  - Stylesheets
  - Images/video/audio
  - Java and Flash bytecode

- Parsing bugs may be exploitable

- JavaScript gives attackers the ability to stage exploits
Example IE Exploit

$exploit = "<html>" . "\n" . "<div id="msie_xmlbof_vista">
<script>
\n$.innerHTML = xmlrox; \n</script>" . "\n" . "</div>" . "\n" . "</html>";
Executing a Drive-by

- Host exploits on a *bulletproof host*
  - No need to distribute (expensive) exploit code to other websites
  - Resist law enforcement takedowns

- Victim acquisition
  - Spam containing links (email, SMS, messenger)
  - Compromise legitimate websites & add booby-traps (e.g. via XSS)
    - Hidden *iframes* that load exploit website
    - Force a redirect to the exploit website
For all your cloud-based exploitation needs

EXPLOITS-AS-A-SERVICE
Structure of the Underground

- **Bulletproof Hosting**
  - Stolen Account Credentials
  - Carders, Cashiers, and Money Mules
  - Credit Card and Bank Account Theft
  - DDoS and Ransomware Extortion
  - Blackhat SEO
  - Click Fraud and Ad Injection
  - Fake Antivirus
  - Malware Attachments
  - Spammering
  - Pharma
  - Counterfeit Goods

- **Pay-per-Install and Exploit-as-a-Service**
  - Botnets

- **Zero-day Development**
  - Crimeware Development

- **DDoS and Ransomware Extortion**
Decoupling and Specialization

• In old days, compromise and monetization were coupled
  – Criminals develop exploits, use them to launch attacks, then use hacked machines to make money

• Today, these facets of criminal underground decoupled
  – Exploit developers sell exploits kits or packs
  – Other actors leverage the kits to attack hosts
    • Often via spam and/or compromised web servers
  – Compromised hosts are then sold on the black market

• Pay-per-install model of malware
“Manufacturing Compromise: The Emergence of Exploit-as-a-Service”

• Authors identify the exploit-as-a-service malware distribution model
  – Relies on drive-by-download attacks against browsers
  – Blackhole, MPack, and other exploit kits

• Two styles of attacks
  – Miscreant can buy an exploit kit and deploy it themselves
  – A miscreant can rent access to an exploit server that hosts an exploit kit

1. Miscreants responsible for acquiring traffic
  – Direct victims to exploit kits using spam or phishing

2. Traffic-PPI (Pay-per-install) services simplify this process
  – Bundle a traffic acquisition mechanism and an exploit server
  – Attacker supplies a binary (typically a Trojan) and purchases “installs”
Exploit Kits and Traffic-PPI Services

• MPack dates back to 2006

• Many others: Blackhole, Incognito, Eleonore, Phoenix

• Exploit kit pricing:
  – Buying a Phoenix license cost $400 in 2009, $2200 in 2012
  – Renting a Blackhole server costs $50/hour or $500/month

• Traffic-PPI pricing:
  – SellMeYourTraffic – between $0.80-$3.00 per thousand visits
  – Traffbiz – $1.60 per thousand visits
  – Only 9-14% of visits will results in a successful infection
Traffic-PPI Example
- Blackhole malware kit, released in 2010, dominated market in 2012-2013
- Annual license of $1500, or $200/week, targeted Java, Flash, Windows, PDFs
- Suspect arrested in Oct 2013
## Exploits Used by Blackhole

<table>
<thead>
<tr>
<th>CVE</th>
<th>Target</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVE-2011-2110</td>
<td>Flash</td>
<td>Adobe Flash Player unspecified code execution</td>
</tr>
<tr>
<td>CVE-2011-0611</td>
<td>Flash</td>
<td>Adobe Flash Player unspecified code execution</td>
</tr>
<tr>
<td>CVE-2010-3552</td>
<td>Java</td>
<td>Skyline</td>
</tr>
<tr>
<td>CVE-2010-1885</td>
<td>Windows</td>
<td>Microsoft Windows Help and Support Center</td>
</tr>
<tr>
<td>CVE-2010-1423</td>
<td>Java</td>
<td>Java Development Toolkit insufficient argument validation</td>
</tr>
<tr>
<td>CVE-2010-0886</td>
<td>Java</td>
<td>Unspecified vulnerability</td>
</tr>
<tr>
<td>CVE-2010-0842</td>
<td>Java</td>
<td>JRE MixerSequencer invalid array index</td>
</tr>
<tr>
<td>CVE-2010-0840</td>
<td>Java</td>
<td>Java trusted methods chaining</td>
</tr>
<tr>
<td>CVE-2010-0188</td>
<td>Adobe Acrobat</td>
<td>LibTIFF integer overflow</td>
</tr>
<tr>
<td>CVE-2010-4324</td>
<td>Adobe Acrobat</td>
<td>Use after free vulnerability in doc.media.newPlayer</td>
</tr>
</tbody>
</table>
End of Blackhole

• 2013: Dmitry “Paunch” Fedotov arrested along with his dev team
  – Author and maintainers of Blackhole
  – Over 1,000 customers, $50k/month in revenue
  – Roughly $2.3M in total revenue

• 2016: sentenced to 7 years in a penal colony
The backbone of the underground

BOTNETS
Structure of the Underground

- Bulletproof Hosting
- Stolen Account Credentials
- Credit Card and Bank Account Theft
- DDOS and Ransomware Extortion
- Blackhat SEO
- Spam
- Click Fraud and Ad Injection
- Bitcoin Mining
- Carders, Cashiers, and Money Mules
- Phishing
- Pharma
- Counterfeit Goods
- Fake Anti-virus
- Malware Attachments
- Botnets
- Pay-per-Install and Exploit-as-a-Service
- Zero-day Development
- Crimeware Development

- Click Fraud and Ad Injection
- Bitcoin Mining
- Malware Attachments
- Fake Anti-virus
- Counterfeit Goods
- Pharma
- Phishing
- Carders, Cashiers, and Money Mules
- Stolen Account Credentials
- DDOS and Ransomware Extortion
- Blackhat SEO
- Spam
- Pay-per-Install and Exploit-as-a-Service
- Bulletproof Hosting
- Crimeware Development
- Zero-day Development
From Crimeware to Botnets

• Infected machines are a fundamentally valuable resource
  – Unique IP addresses for spamming
  – Bandwidth for DDoS
  – CPU cycles for bitcoin mining
  – Credentials

• Early malware monetized these resources directly
  – Infection and monetization were tightly coupled

• Botnets allow criminals to rent access to infected hosts
  – Infrastructure as a service, i.e. the cloud for criminals
  – Command and Control (C&C) infrastructure for controlling bots
  – Enables huge-scale criminal campaigns
Old-School C&C: IRC Channels

- snd spam: <subject> <msg>
- snd spam: <subject> <msg>
- snd spam: <subject> <msg>

- Problem: single point of failure
- Easy to locate and take down
P2P Botnets

Botmaster

Master Servers

Structured P2P Distributed Hash Table (DHT)

Insert commands into the DHT

Get commands from the DHT
Fast Flux DNS

But: ISPs can blacklist the rendezvous domain

HTTP Servers

www.my-botnet.com

Botmaster

Change DNS → IP mapping every 10 seconds
Domain Name Generation (DGA)

Bots generate many possible domains each day.

...But the Botmaster only needs to register a few.

Can be combined with fast flux.
“Your Botnet is My Botnet”

• Takeover of the Torpig botnet
  – Random domain generation + fast flux
  – Team reverse engineered domain generation algorithm
  – Registered 30 days of domains before the botmaster!
  – Full control of the botnet for 10 days

• Goal of botnet: credential theft and phishing spam
  – Steals credit card numbers, bank accounts, etc.
  – Researchers gathered all this data

• Other novel point: accurate estimation of botnet size
Attacker places a redirect on the vulnerable server

Rootkit installation

Collect stolen data

Capture banking passwords

Researchers Infiltrated Here
Man-in-the-Browser Attack

Injected DLL steals information entered into forms
Torpig Rendezvous Algorithm

1. Try to connect to a computed a weekly domain
   - Append a list of TLDs, in order
   - Example: adlfn.com → adlfn.net → adlfn.biz

2. Try to connect to a computed a daily domain
   - Same list of TLDs, in order

3. Try to connect to a hardcoded list of fallback domains
   - rikora.com, pinakola.com, and flippibi.com

• First successful connection wins
   - If the whitehat owns the weekly .com domain, they win
def generate_daily_domain():
    return generate_domain(GetLocalTime(), 8)

def scramble_date(t, p):
    return ((t.month ^ t.day) + t.day) * p) + t.day + t.year

def generate_domain(t, p):
    if t.year < 2007: t.year = 2007
    s = scramble_date(t, p)
    c1 = (((t.year >> 2) & 0x3fc0) + s) % 25 + 'a'
    c2 = (t.month + s) % 10 + 'a'
    c3 = (((t.year & 0xff) + s) % 25 + 'a'
    if t.day * 2 < '0' or t.day * 2 > '9': c4 = (t.day * 2) % 25 + 'a'
    else: c4 = t.day % 10 + '1'
    return c1 + 'h' + c2 + c3 + 'x' + c4 + suffix[t.month - 1]
Stolen Information

• Data gathered from Jan 25-Feb 4 2009

<table>
<thead>
<tr>
<th>User Accounts</th>
<th>Data Type</th>
<th>Data Items (#)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mailbox account</td>
<td>54,090</td>
</tr>
<tr>
<td></td>
<td>Email</td>
<td>1,258,862</td>
</tr>
<tr>
<td></td>
<td>Form data</td>
<td>11,966,532</td>
</tr>
<tr>
<td></td>
<td>HTTP account</td>
<td>411,039</td>
</tr>
<tr>
<td></td>
<td>FTP account</td>
<td>12,307</td>
</tr>
<tr>
<td></td>
<td>POP account</td>
<td>415,206</td>
</tr>
<tr>
<td></td>
<td>SMTP account</td>
<td>100,472</td>
</tr>
<tr>
<td></td>
<td>Windows password</td>
<td>1,235,122</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bank Accounts</th>
<th>Country</th>
<th>Institutions (#)</th>
<th>Accounts (#)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>US</td>
<td>60</td>
<td>4,287</td>
</tr>
<tr>
<td></td>
<td>IT</td>
<td>34</td>
<td>1,459</td>
</tr>
<tr>
<td></td>
<td>DE</td>
<td>122</td>
<td>641</td>
</tr>
<tr>
<td></td>
<td>ES</td>
<td>18</td>
<td>228</td>
</tr>
<tr>
<td></td>
<td>PL</td>
<td>14</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>162</td>
<td>1,593</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>410</td>
<td>8,310</td>
</tr>
</tbody>
</table>

• How much is this data worth?
  • Credit cards: $0.10-$25 each, banks accounts: $10-$1000 each
  • Estimated total: $83K-$8.3M
How to Estimate Botnet Size?

- Passive data collection methodologies
  - Honeypots
    - Infect your own machines with Trojans
    - Observe network traffic
  - Look at DNS traffic
    - Domains linked to fast flux C&C
  - Networks flows
    - Analyze all packets from a large ISP and use heuristics to identify botnet traffic
- None of these methods give a complete picture
Size of the Torpig Botnet

- Why the disconnect between IPs and bots?
  - Dynamic IPs, short DHCP leases
- Casts doubt on prior studies, enables more realistic estimates of botnet size
“A Botmaster’s Perspective of Coordinating Large-Scale Spam Campaigns”

• Takeover of the Pushdo/Cutwail botnet
  – First appeared in 2007
  – Almost exclusively used for spam

• Failed past takeovers
  – McColo in 2008
  – 3FN in 2009
  – FireEye in 2010

• Used dynamic analysis to identify the IPs of C&C servers
  – Shut down 20, took over 16
  – Covers ½ to 2/3 of all Cutwail C&C servers
Size

Per Hour

Per Day
## Spam Campaigns

<table>
<thead>
<tr>
<th>Client (ID)</th>
<th>Instances (#)</th>
<th>Unique Bot IPs (#)</th>
<th>Avg. Lifespan (Days)</th>
<th>Mails Sent (#)</th>
<th>Average Mails/Active Bot (Per Day)</th>
<th>Campaign Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
<td>2,251,156</td>
<td>17</td>
<td>98,401,907,545</td>
<td>2,571</td>
<td>Phishing, Malware</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>40,924</td>
<td>168</td>
<td>45,555,535,375</td>
<td>6,626</td>
<td>Phishing</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>56,733</td>
<td>54</td>
<td>155,098,090,946</td>
<td>50,626</td>
<td>Diplomas</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>34,742</td>
<td>22</td>
<td>17,941,545,204</td>
<td>23,473</td>
<td>Phishing, Pharm.</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>21,993</td>
<td>8</td>
<td>60,169,427,197</td>
<td>341,980</td>
<td>Money Mule</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>29,471</td>
<td>13</td>
<td>4,309,066,448</td>
<td>11,247</td>
<td>Pharmaceuticals</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>27,658</td>
<td>55</td>
<td>9,408,910,232</td>
<td>6,185</td>
<td>Phishing</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>30,503</td>
<td>135</td>
<td>12,485,832,067</td>
<td>3,032</td>
<td>Phishing</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>29,415</td>
<td>18</td>
<td>2,365,652,828</td>
<td>4,467</td>
<td>Real Estate</td>
</tr>
</tbody>
</table>
Blacklisting

• Blacklisting common technique to filter spam
  – IPs of machines sending spam are recorded and distributed
  – Email providers filter emails from these IPs
  – E.g. Spamhaus

• Cutwail bots queried their own blacklist status periodically!
  – SORBS, SpamCop, DNSBL
  – Reported their status to the C&C
  – C&C would divert spam to other “clean” bots
Time to Blacklist

<table>
<thead>
<tr>
<th>Time Since Spam Campaign Started</th>
<th>Fraction Blacklisted</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 hours</td>
<td>29.6%</td>
</tr>
<tr>
<td>3 hours</td>
<td>46.4%</td>
</tr>
<tr>
<td>6 hours</td>
<td>75.3%</td>
</tr>
<tr>
<td>18 hours</td>
<td>90%</td>
</tr>
</tbody>
</table>

![Graph showing the number of IP addresses online over time](image)
Stopping Botnets

• Individual perspective: ridding your network of bots
  – Anti-virus and anti-malware
  – Intrusion and anomaly detection to identify infections, block traffic

• Global perspective: takedowns and arrests
  – Create a sinkhole (fake C&C server)
  – Track down and arrest the perpetrators
Classic Detection of Bots

- Defeated by using standard ports
  - HTTP(S) ports 80/443
- Defeated by encryption

- Unusual ports or protocols
  - IRC port 6667
- Message signatures
  - “cmd=spam; target=…”
Detection of DGA and Fast Flux

- For DGA: many failed DNS lookups
- For fast flux: multiple DNS lookups for one name, response has short TTL
  - 10 seconds – 10 minutes
  - Most DNS names have TTL of hours or days
Detection of P2P

- Many connections to seemingly random hosts
  - Bursty traffic patterns
  - Unexpected geographic patterns (connections to hosts in other countries)
## Infamous Takedowns

<table>
<thead>
<tr>
<th>Botnet Name</th>
<th>Timeframe</th>
<th>Estimated Size</th>
<th>Taken Down by…</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNS Changer</td>
<td>2006-2011</td>
<td>4M</td>
<td>FBI, Trend Micro</td>
</tr>
<tr>
<td>Rustock</td>
<td>2006-2011</td>
<td>150K-2.4M</td>
<td>FBI, Microsoft, Fireeye, Univ. of Washington</td>
</tr>
<tr>
<td>Grum</td>
<td>2008-2012</td>
<td>560K-840K</td>
<td>Fireeye, Spamhaus</td>
</tr>
<tr>
<td>Conficker</td>
<td>2008-2009</td>
<td>4M-13M</td>
<td>FBI, Microsoft, Symantec, ICANN</td>
</tr>
<tr>
<td>Citadel</td>
<td>2011-2013</td>
<td></td>
<td>FBI, Microsoft</td>
</tr>
<tr>
<td>Gameover Zeus/Cryptolocker</td>
<td>2012-2014</td>
<td></td>
<td>DoJ, FBI, Europol, Dell, Microsoft, Level3, McAfee, Symantec, Sophos, Trend Micro, Carnegie Mellon, Georgia Tech, etc.</td>
</tr>
<tr>
<td>SIMDA</td>
<td>2011-2015</td>
<td>770K</td>
<td>INTERPOL, Trend Micro, Microsoft, Kaspersky Lab</td>
</tr>
<tr>
<td>DRIDEX</td>
<td>2014-2015</td>
<td></td>
<td>FBI, Trend Micro</td>
</tr>
<tr>
<td>Avalanche</td>
<td>2009-2016</td>
<td>500K</td>
<td>FBI, Symantec, Fraunhofer</td>
</tr>
</tbody>
</table>