Botnet Infiltration
Possibilities & Challenges

Christian Kreibich
International Computer Science Institute
The “Underground” Economy

» Financially motivated, Internet-driven abuse
  » Fraud, identity theft, extortion, money laundering...

» Complex real-world market
  » Vendors, merchants, spammers, malware authors, botmasters, affiliate programs, ...

» Prevention extremely difficult
  » Technical and sociological problem

» What can we do?
n Bot·net
Botnet infiltration

» Botnets: a central *technical* phenomenon
  » Distributed systems: a *weakness* we can attack

» Since around 2007, a *constant presence*
  » Storm, MegaD, Waledac, ...
  » New research experience, lots of interest
  » Even a little scary — DDoS threat? Kneecaps?

» We have infiltrated several botnets longitudinally
  » Storm: 1 year, MegaD: 3 months, others passively
  » *Very fruitful* efforts
  » But fraught with *legal & ethical challenges*
<table>
<thead>
<tr>
<th>Insights</th>
<th>Challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technical</td>
<td>?</td>
</tr>
<tr>
<td>Sociological</td>
<td>?</td>
</tr>
</tbody>
</table>
The Storm Botnet

Hosted infrastructure

HTTP proxies

HTTP

Address harvests

Work requests

Proxy bots

TCP

Spam, DDoS, ...

Botmaster

Workers

Overnet

.work requests
Challenge: Malware Containment

» Must operate malware **safely**
  » Spamming, DDoS, iframe & SQL injections, ...

» Undercover bots are **sensitive** and **scarce**

» Tight containment is **time-consuming**
  » Each botnet unique
  » C&C nature not known ahead of time

» Transparent app-layer **containment proxy**
  » Default-deny, filtering, redirection
  » Iteratively expand understanding of the C&C
Insight: Campaign Awareness

From: Spamcraft: An Inside Look At Spam Campaign Orchestration, LEET'09
Insight: Domain Use & Usability

From: Spamcraft: An Inside Look At Spam Campaign Orchestration, LEET'09
The Storm Botnet

Hosted infrastructure

Botmaster

HTTP proxies

HTTP

Address harvests

Work assignments

(Spam, DDoS,...)

Infected machines

Proxy bots

TCP

Overnet

Workers

Work requests

(Fraud, DDoS,...)
Insight: Rendezvous Infiltration

» Accurate size measurement — carefully!
  » Real bots or former bots' addresses?
  » Machines behind NATs?
  » Bot IDs vs IP addresses?

» More recently: domain generation algorithms

» Colliding experiments!
The Storm Botnet

Infection mechanics

HTTP proxies

HTTP

Address harvests

Work requests

(Spam, DDoS,...)

Botmaster

Work assignments

Infected machines

TCP

Workers

Proxy bots

Overnet
Insight: Spam Conversion

» 1 in 12.5m pharma targets yields sale
» 1 in 265k greeting card targets yields infection
» 1 in 10 visitors of infection site ran offered program
» Revenue: ~$3.5M / year
Challenge: Law Enforcement

» FBI takedown order on our servers was in progress

» Oops!

» Coordination with LE
  » May be prudent
  » May be difficult
  » May be asymmetric
Challenge: Human Subjects

» Privacy concerns
  » We see who is infected
  » We see who is targeted
  » We see what people do
  » **Highly sensitive data!**

» In US: Institutional Review Board (IRB) approval
  » Institutions receiving federal grants must have one
  » 6-8 weeks processing time for basic cases
  » IRBs lack technical depth
The Storm Botnet

- **Botmaster**
- **HTTP proxies**
- **HTTP**
- **TCP**
- **Overnet**
- **Workers**
- **Proxy bots**
- **Address harvests**
- **Work assignments**
- (Spam, DDoS, ...)

Infected machines

Hosted infrastructure
Challenge: Ethical Standards

» “Defense in depth”: C&C filtering, signal jamming

» “White” botnets: takeover / cleanup

» Botnet rental (BBC study)

» Product purchases

» Keylogger data pilfering

» Slippery slope — what to do?
  » Community standards (ethics panels everywhere!)
  » Do results justify the means?
  » Some papers rejected (NDSS'10, LEET'10) ...
  » ... others not (NDSS'10).
<table>
<thead>
<tr>
<th>Insights</th>
<th>Challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Technical</strong></td>
<td><strong>Challenges</strong></td>
</tr>
<tr>
<td>» Understanding</td>
<td>» Arms race advancement</td>
</tr>
<tr>
<td>» Spam awareness</td>
<td>» Invasion resilience</td>
</tr>
<tr>
<td>» Botnet size</td>
<td>» Bot reliability</td>
</tr>
<tr>
<td>» C&amp;C rewriting</td>
<td>» Malware containment</td>
</tr>
<tr>
<td>» Offense in</td>
<td>» Colliding experiments</td>
</tr>
<tr>
<td>depth</td>
<td></td>
</tr>
<tr>
<td><strong>Sociological</strong></td>
<td></td>
</tr>
<tr>
<td>» Victim behavior</td>
<td>» Victim privacy</td>
</tr>
<tr>
<td>» Spammer behavior</td>
<td>» Human subjects</td>
</tr>
<tr>
<td>» Bot herder</td>
<td>» IRB approval</td>
</tr>
<tr>
<td>behavior</td>
<td>» Law enforcement involvement</td>
</tr>
<tr>
<td>» Market analysis</td>
<td>» Ethics</td>
</tr>
<tr>
<td>» Volumes</td>
<td>» Do no harm</td>
</tr>
<tr>
<td>» Profits</td>
<td></td>
</tr>
</tbody>
</table>
C&C inspection and rewriting

From: Spamalytics: An Empirical Analysis of Spam Marketing Conversion, CCS'08