CSE 490K Lecture 14

Botnets and Spam

Tadayoshi Kohno

Some slides based on Vitaly Shmatikov’s
Botnets

- Botnet = network of autonomous programs capable of acting on instructions
  - Typically a large (up to several hundred thousand) group of remotely controlled “zombie” systems
    - Machine owners are not aware they have been compromised
  - Controlled and upgraded via IRC or P2P

- Used as the platform for various attacks
  - Distributed denial of service
  - Spam and click fraud
  - Launching pad for new exploits/worms
Building a Botnet

- Attacker
  - Compromise attempt
  - Install bot software
    - Win XP compromised
    - FreeBSD
    - Mac OS X
    - Win XP compromised
Typical Infection Path

- Exploit a vulnerability to execute a short program (shellcode) on victim’s machine
  - Buffer overflows, email viruses, etc.
- Shellcode downloads and installs actual bot
- Bot disables firewall and antivirus software
- Bot locates IRC server, connects, joins channel
  - Typically need DNS to find out server’s IP address
    - Especially if server’s original IP address has been blacklisted
  - Authentication password often stored in bot binary
- Botmaster issues authenticated commands
Joining the IRC Channel

Win XP
/connect jade.va.us.dal.net
/join #hacker
...

Win XP
/connect jade.va.us.dal.net
/join #hacker
...

Win XP
/connect jade.va.us.dal.net
/join #hacker
...

jade.va.dal.net
(12:59:27pm) -- A9-pcgbdv (A9-pcgbdv@140.134.36.124) has joined (#owned) Users : 1646

(12:59:27pm) (@Attacker) .ddos.synflood 216.209.82.62

(12:59:27pm) -- A6-bpxufrd (A6-bpxufrd@wp95-81.introweb.nl) has joined (#owned) Users : 1647

(12:59:27pm) -- A9-nzmpah (A9-nzmpah@140.122.200.221) has left IRC (Connection reset by peer)

(12:59:28pm) (@Attacker) .scan.enable DCOM

(12:59:28pm) -- A9-tzrkeasv (A9-tzrkeasv@220.89.66.93) has joined (#owned) Users : 1650
Botnet Propagation

- Each bot can scan IP space for new victims
  - Automatically
    - Each bot contains hard-coded list of IRC servers’ DNS names
    - As infection is spreading, IRC servers and channels that the new bots are looking for are often no longer reachable
  - On-command: target specific /8 or /16 prefixes
    - Botmasters share information about prefixes to avoid

- Evidence of botnet-on-botnet warfare
  - DoS server by multiple IRC connections ("cloning")
  - Patch victim against other bots

- Active botnet management
  - Detect non-responding bots, identify “superbots”
Denial of Service (DoS) Redux

- **Goal**: overwhelm victim machine and deny service to its legitimate clients
- **DoS** often exploits networking protocols
  - Smurf: ICMP echo request to broadcast address with spoofed victim’s address as source
  - Ping of death: ICMP packets with payloads greater than 64K crash older versions of Windows
  - SYN flood: “open TCP connection” request from a spoofed address
  - UDP flood: exhaust bandwidth by sending thousands of bogus UDP packets
Distributed Denial of Service (DDoS)

- Build a botnet of zombies
  - Multi-layer architecture: use some of the zombies as “masters” to control other zombies

- Command zombies to stage a coordinated attack on the victim
  - Even in case of SYN flood, SYN cookies don’t help (why?)

- Overwhelm victim with traffic arriving from thousands of different sources
DDoS Architecture

- Attacker
- Master machines
- Zombie machines
- Victim
DDoS Tools: Trin00

- Scan for known buffer overflows in Linux & Solaris
  - Unpatched versions of wu-ftpd, statd, amd, ...
  - Root shell on compromised host returns confirmation
- Install attack daemon using remote shell access
- Send commands (victim IP, attack parameters), using plaintext passwords for authentication
  - Attacker to master: TCP, master to zombie: UDP
  - To avoid detection, daemon issues warning if someone connects when master is already authenticated
- August of 1999: a network of 227 Trin00 zombies took U. of Minnesota offline for 3 days
DDoS Tools: Tribal Flood Network

- Supports multiple DoS attack types
  - Smurf; ICMP, SYN, UDP floods

- Attacker runs masters directly via root backdoor; masters talk to zombies using ICMP echo reply
  - No authentication of master’s commands, but commands are encoded as 16-bit binary numbers inside ICMP packets to prevent accidental triggering
  - Vulnerable to connection hijacking and RST sniping

- List of zombie daemons’ IP addresses is encrypted in later versions of TFN master scripts
  - Protects identities of zombies if master is discovered
DDoS Tools: Stacheldraht

- Combines “best” features of Trin00 and TFN
  - Multiple attack types (like TFN)
- Symmetric encryption for attacker-master connections
- Master daemons can be upgraded on demand
- February 2000: crippled Yahoo, eBay, Amazon, Schwab, E*Trade, CNN, Buy.com, ZDNet
  - Smurf-like attack on Yahoo consumed more than a Gigabit/sec of bandwidth
  - Sources of attack still unknown
Spam
Email in the Early 1980s

- Mail relay: forwards mail to next hop
- Sender path includes path through relays
Email Spoofing

- Mail is sent via SMTP protocol
  - No built-in authentication
- MAIL FROM field is set by the sender
  - Classic example of improper input validation
- Recipient’s mail server only sees IP address of the direct peer from whom it received the msg
Open Relays

- SMTP relay forwards mail to destination
  1. Bulk email tool connects via SMTP (port 25)
  2. Sends list of recipients via RCPT TO command
  3. Sends email body (once for all recipients!)
  4. Relay delivers message

- Honest relay adds correct Received: header revealing source IP
- Hacked relay does not
A Closer Look at Spam

Received: by 10.78.68.6 with SMTP id q6cs394373hua;
   Mon, 12 Feb 2007 06:43:30 -0800 (PST)
Received: by 10.90.113.18 with SMTP id l18mr17307116agc.1171291410432;
   Mon, 12 Feb 2007 06:43:30 -0800 (PST)
Return-Path: <wvnlwee@aviva.ro>
Received: from onelinkpr.net ([203.169.49.172])
   by mx.google.com with ESMTP id 30si11317474agc.2007.02.12.06.43.18;
   Mon, 12 Feb 2007 06:43:30 -0800 (PST)
Received-SPF: neutral (google.com: 203.169.49.172 is neither permitted nor
   denied by best guess record for domain of wvnlwee@aviva.ro)
Message-ID: <20050057765.stank.203.169.49.172@ASAFTU>
From: "Barclay Morales" <wvnlwee@aviva.ro>
To: <raykwatts@gmail.com>
Subject: You can order both Viagra and Cialis.
A Closer Look at Spam

Received: by 10.78.68.6 with SMTP id q6cs394373hua;
    Mon, 12 Feb 2007 06:43:30 -0800 (PST)
Received: by 10.90.113.18 with SMTP id l18mr17307116agc.1171291410432;
    Mon, 12 Feb 2007 06:43:30 -0800 (PST)
Return-Path: <wvnlwee@aviva.ro>
Received: from onelinkpr.net ([203.169.49.172])
    by mx.google.com with ESMTP id 30si11317474agc.2007.02.12.06.43.18;
    Mon, 12 Feb 2007 06:43:30 -0800 (PST)
Received-SPF: neutral (google.com: 203.169.49.172 is neither permitted nor
denied by best guess record for domain of wvnlwee@aviva.ro)
Message-ID: <20050057765.stank.203.169.49.172@ASAFTU>
From: "Barclay Morales" <wvnlwee@aviva.ro>
To: <raykwatts@gmail.com>
Subject: You can order both Viagra and Cialis.
A Closer Look at Spam

Received: by 10.78.68.6 with SMTP id q6cs394373hua;
    Mon, 12 Feb 2007 06:43:30 -0800 (PST)
Received: by 10.90.113.18 with SMTP id l18mr17307116agc.1171291410432;
    Mon, 12 Feb 2007 06:43:30 -0800 (PST)
Return-Path: <wvnlwee@aviva.ro>
Received: from onelinkpr.net ([203.169.49.172])
    by mx.google.com with ESMTP id 30si11317474agc.2007.02.12.06.43.18;
    Mon, 12 Feb 2007 06:43:30 -0800 (PST)
Received: from onelinkpr.net ([203.169.49.172] is neither permitted nor
    trusted by best guess record for domain of wvnlwee@aviva.ro)
Message-ID: <20050057765.stank.203.169.49.172@ASAFTU>
From: "Barclay Morales" <wvnlwee@aviva.ro>
To: <raykwatts@gmail.com>
Subject: You can order both Viagra and Cialis.
A Closer Look at Spam

Received: by 10.78.68.6 with SMTP id q6cs394373hua; Mon, 12 Feb 2007 06:43:30 -0800 (PST)
Received: by 10.90.113.18 with SMTP id l18mr17307116agc.1171291410432; Mon, 12 Feb 2007 06:43:30 -0800 (PST)
Return-Path: <wvnlwee@aviva.ro>
Received: from onelinkpr.net ([203.169.49.172]) by mx.google.com with ESMTP id 30si11317474agc.2007.02.12.06.43.18; Mon, 12 Feb 2007 06:43:30 -0800 (PST)
Received-SPF: neutral (google.com: 203.169.49.172 is neither permitted nor
rejected by best guess record for domain of wvnlwee@aviva.ro)
Message-ID: <20050057765.stank.203.169.49.172@ASAFTU>
From: "Barclay Morales" <wvnlwee@aviva.ro>
To: <raykwatts@gmail.com>
Subject: You can order both Viagra and Cialis.
A Closer Look at Spam

Received: by 10.78.68.6 with SMTP id q6cs394373hua;
Mon, 12 Feb 2007 06:43:30 -0800 (PST)
Received: by 10.90.113.18 with SMTP id l18mr17307116agc.1171291410432;
Mon, 12 Feb 2007 06:43:30 -0800 (PST)
Return-Path: <wvnlwee@aviva.ro>
Received: from onelinkpr.net ([203.169.49.172])
by mx.google.com with ESMTP id 30si11317474agc.2007.02.12.06.43.18;
Mon, 12 Feb 2007 06:43:30 -0800 (PST)
Received-SPF: neutral (google.com: 203.169.49.172 is neither permitted nor
denied by best guess record for domain of wvnlwee@aviva.ro)
Message-ID: <20050057765.stank.203.169.49.172@ASAFTU>
From: "Barclay Morales" <wvnlwee@aviva.ro>
To: <raykwatts@gmail.com>
Subject: You can order both Viagra and Cialis.

A Closer Look at Spam

Inserted by relays

Bogus!

Puerto Rico

Mongolia
Why Hide Sources of Spam?

◆ Many email providers blacklist servers and ISPs that generate a lot of spam
  - Use info from spamhaus.org, spamcop.net

◆ Real-time blackhole lists stop 15-25% of spam at SMTP connection time
  - Over 90% after message body URI checks

◆ Spammers’ objective: evade blacklists
  - Botnets come very handy!
Thin Pipe / Thick Pipe

- Spam source has high-speed broadband machine (HSB) and controls a low-speed zombie (LSZ)

- Hides IP address of HSB; LSZ is blacklisted
Send-Safe Mailer

Send-Safe is a bulk email software that allows you to send email from your own computer, or a remote computer with or without the use of proxies. In those countries where it is legal to use proxies, the Send-Safe program can make it impossible for anyone to trace the email back to your ISP and thus keeps your connection to the internet safe. This gives you a safe haven in which to send your mail.

Download demo | Purchase for Euro 50 | Members Area

Send-Safe Standalone

Send-Safe Standalone is a standalone version of one of the most successful and efficient bulk email software in the industry. It's designed specially for those top mailers who prefer to pay once for the software once and then don't spend extra money for the huge volume of their mailings. Purchasing Send-Safe Standalone, you get all the power of our famous Send-Safe mailer but pay only one time fee, with no troubles of credits overrun, account expiration, lost server connections and so on.

Download demo | Purchase for Euro 599

Send-Safe Enterprise
Send-Safe Spam Tool

- Deliverability: 87%
- Avg speed: 960244 mails/hour

Total good proxies: 527, Using 317 fastest proxies. Reply time: min=0.4534s, max=2.9521s
Open Relays vs. Open Proxies

- Open proxy
  - Spammer must send message to each recipient through the proxy

- Open relay
  - Takes a list of addresses and sends to all
  - Can host an open relay on a zombie

- Listing services for open proxies and relays
  - http://www.multiproxy.org/
  - http://www.stayinvisible.com/
  - http://www.openproxies.com/ ($20/month)
Bobax Worm

- Infects machines with high bandwidth
  - Exploits MS LSASS.exe buffer overflow vulnerability
- Slow spreading (and thus hard to detect)
  - On manual command from operator, randomly scans for vulnerable machines
- Installs hacked open relay on infected zombie
  - Once spam zombie added to blacklist, spread to another machine
  - Interesting detection technique: look for botmaster’s DNS queries (trying to determine who is blacklisted)
Distribution of Spam Sources

[Ramachandran, Feamster]
IP Blacklisting Not Enough

[Ramachandran, Feamster]

More than half of client IPs appear less than twice
## Distribution Across Domains

<table>
<thead>
<tr>
<th>AS Number</th>
<th># Spam</th>
<th>AS Name</th>
<th>Primary Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>766</td>
<td>580559</td>
<td>Korean Internet Exchange</td>
<td>Korea</td>
</tr>
<tr>
<td>4134</td>
<td>560765</td>
<td>China Telecom</td>
<td>China</td>
</tr>
<tr>
<td>1239</td>
<td>437660</td>
<td>Sprint</td>
<td>United States</td>
</tr>
<tr>
<td>4837</td>
<td>236434</td>
<td>China Network Communications</td>
<td>China</td>
</tr>
<tr>
<td>9318</td>
<td>225830</td>
<td>Hanaro Telecom</td>
<td>Japan</td>
</tr>
<tr>
<td>32311</td>
<td>198185</td>
<td>JKS Media, LLC</td>
<td>United States</td>
</tr>
<tr>
<td>5617</td>
<td>181270</td>
<td>Polish Telecom</td>
<td>Poland</td>
</tr>
<tr>
<td>6478</td>
<td>152671</td>
<td>AT&amp;T WorldNet Services</td>
<td>United States</td>
</tr>
<tr>
<td>19262</td>
<td>142237</td>
<td>Verizon Global Networks</td>
<td>United States</td>
</tr>
<tr>
<td>8075</td>
<td>107056</td>
<td>Microsoft</td>
<td>United States</td>
</tr>
<tr>
<td>7132</td>
<td>99585</td>
<td>SBC Internet Services</td>
<td>United States</td>
</tr>
<tr>
<td>6517</td>
<td>94600</td>
<td>Yipes Communications, Inc.</td>
<td>United States</td>
</tr>
<tr>
<td>31797</td>
<td>89698</td>
<td>GalaxyVisions</td>
<td>United States</td>
</tr>
<tr>
<td>12322</td>
<td>87340</td>
<td>PROXAD AS for Proxad ISP</td>
<td>France</td>
</tr>
<tr>
<td>3356</td>
<td>87042</td>
<td>Level 3 Communications, LLC</td>
<td>United States</td>
</tr>
<tr>
<td>22909</td>
<td>86150</td>
<td>Comcast Cable Corporation</td>
<td>United States</td>
</tr>
<tr>
<td>8151</td>
<td>81721</td>
<td>UniNet S.A. de C.V.</td>
<td>Mexico</td>
</tr>
<tr>
<td>3320</td>
<td>79987</td>
<td>Deutsche Telekom AG</td>
<td>Germany</td>
</tr>
<tr>
<td>7018</td>
<td>74320</td>
<td>AT&amp;T WorldNet Services</td>
<td>United States</td>
</tr>
<tr>
<td>4814</td>
<td>74266</td>
<td>China Telecom</td>
<td>China</td>
</tr>
</tbody>
</table>
Most Bots Send Little Spam

Most bot IP addresses send very little spam, regardless of how long they have been spamming...
Where Does Spam Come From?

[Ramachandran, Feamster]
Where Does Spam Come From?

- IP addresses of spam sources are widely distributed across the Internet

[Ramachandran, Feamster]
Where Does Spam Come From?

[Ramachandran, Feamster]

- IP addresses of spam sources are widely distributed across the Internet
  - In tracking experiments, most IP addresses appear once or twice; 60-80% not reachable by traceroute
Where Does Spam Come From?

[Ramachandran, Feamster]

◆ IP addresses of spam sources are widely distributed across the Internet
  • In tracking experiments, most IP addresses appear once or twice; 60-80% not reachable by traceroute

◆ Vast majority of spam originates from a small fraction of IP address space
Where Does Spam Come From? [Ramachandran, Feamster]

◆ IP addresses of spam sources are widely distributed across the Internet
  • In tracking experiments, most IP addresses appear once or twice; 60-80% not reachable by traceroute

◆ Vast majority of spam originates from a small fraction of IP address space
  • Same fraction that most legitimate email comes from
Where Does Spam Come From? [Ramachandran, Feamster]

- IP addresses of spam sources are widely distributed across the Internet
  - In tracking experiments, most IP addresses appear once or twice; 60-80% not reachable by traceroute
- Vast majority of spam originates from a small fraction of IP address space
  - Same fraction that most legitimate email comes from
- Spammers exploit routing infrastructure
Where Does Spam Come From?

[Ramachandran, Feamster]

- IP addresses of spam sources are widely distributed across the Internet
  - In tracking experiments, most IP addresses appear once or twice; 60-80% not reachable by traceroute

- Vast majority of spam originates from a small fraction of IP address space
  - Same fraction that most legitimate email comes from

- Spammers exploit routing infrastructure
  - Create short-lived connection to mail relay, then disappear
Where Does Spam Come From?

[Ramachandran, Feamster]

- IP addresses of spam sources are widely distributed across the Internet
  - In tracking experiments, most IP addresses appear once or twice; 60-80% not reachable by traceroute

- Vast majority of spam originates from a small fraction of IP address space
  - Same fraction that most legitimate email comes from

- Spammers exploit routing infrastructure
  - Create short-lived connection to mail relay, then disappear
  - Hijack a large chunk of unallocated “dark” space
Spambot Behavior

[Ramachandran, Feamster]
Spambot Behavior

- Strong correlation with Bobax infections

[Ramachandran, Feamster]
Spambot Behavior

- Strong correlation with Bobax infections
- Most are active for a very short time
Spambot Behavior

- Strong correlation with Bobax infections
- Most are active for a very short time
  - 65% of Bobax victims send spam once; 3 out of 4 are active for less than 2 minutes

[Ramachandran, Feamster]
Spambot Behavior

- Strong correlation with Bobax infections
- Most are active for a very short time
  - 65% of Bobax victims send spam once; 3 out of 4 are active for less than 2 minutes
- 99% of bots send fewer than 100 messages regardless of their lifetime

[Ramachandran, Feamster]
Spambot Behavior

- Strong correlation with Bobax infections
- Most are active for a very short time
  - 65% of Bobax victims send spam once; 3 out of 4 are active for less than 2 minutes
- 99% of bots send fewer than 100 messages regardless of their lifetime
- 95% of bots already in one or more blacklists

[Ramachandran, Feamster]
Spambot Behavior

- Strong correlation with Bobax infections
- Most are active for a very short time
  - 65% of Bobax victims send spam once; 3 out of 4 are active for less than 2 minutes
- 99% of bots send fewer than 100 messages regardless of their lifetime
- 95% of bots already in one or more blacklists
  - Cooperative detection works, but ...
Spambot Behavior

Strong correlation with Bobax infections

Most are active for a very short time
  - 65% of Bobax victims send spam once; 3 out of 4 are active for less than 2 minutes

99% of bots send fewer than 100 messages regardless of their lifetime

95% of bots already in one or more blacklists
  - Cooperative detection works, but ...
  - Problem: false positives!

[Ramachandran, Feamster]
Spambot Behavior [Ramachandran, Feamster]

- Strong correlation with Bobax infections
- Most are active for a very short time
  - 65% of Bobax victims send spam once; 3 out of 4 are active for less than 2 minutes
- 99% of bots send fewer than 100 messages regardless of their lifetime
- 95% of bots already in one or more blacklists
  - Cooperative detection works, but ...
  - Problem: false positives!
  - Problem: short-lived hijacks of dark address space
Detecting Botnets
Detecting Botnets

Today’s bots are controlled via IRC and DNS

- IRC used to issue commands to zombies
- DNS used by zombies to find the master, and by the master to find if a zombie has been blacklisted
Detecting Botnets

- Today’s bots are controlled via IRC and DNS
  - IRC used to issue commands to zombies
  - DNS used by zombies to find the master, and by the master to find if a zombie has been blacklisted
- IRC/DNS activity is very visible in the network
  - Look for hosts performing scans, and for IRC channels with a high percentage of such hosts
    - Used with success at Portland State University
  - Look for hosts who ask many DNS queries, but receive few queries about themselves
Detecting Botnets

◆ Today’s bots are controlled via IRC and DNS
  • IRC used to issue commands to zombies
  • DNS used by zombies to find the master, and by the master to find if a zombie has been blacklisted

◆ IRC/DNS activity is very visible in the network
  • Look for hosts performing scans, and for IRC channels with a high percentage of such hosts
    – Used with success at Portland State University
  • Look for hosts who ask many DNS queries, but receive few queries about themselves

◆ Easily evaded by using encryption and P2P 😞