Analyzing and Understanding Botnets

Dr. Jose Nazario

Botnets

- **Pressing problem for network operators**
- **ISPs - number 1 pressing issue**
- **Enterprises**
  - Unknown threat scale
  - Big concern to many

- **Estimated 1-5% of the Internet is affected**

ARBOR
NETWORKS

Botnet Basics

- **Network of compromised computers**
- **Connect to one or more servers**
- **Receive commands, act on them**

- **Communication protocols**
  - IRC
  - HTTP
  - Peer to peer
  - DNS

Bots in the Malware Taxonomy

- **Bots exhibit worm characteristics**
  - Use network exploits to propagate
- **Bots exhibit backdoor characteristics**
  - Start up a network listener service, inbound connections
    - FTP server, web server, etc
  - Connect outbound to receive connections
- **Bots utilize rootkits**
  - Rootkits hide their presence
- **Bots have spyware components**
  - Keystroke loggers for information theft

- **Bots are extensible and may download additional software**
- **A botnet herder may load adware and/or spyware on a compromised system**

ARBOR
NETWORKS

How Machines are Infected

- **Operating system vulnerabilities**
  - Classic mechaism
    - RPC-DCOM, MSSQL, Symantec, etc
  - No user required
  - Easily stopped with firewalls and HIPS
- **Client-side attacks**
  - Spammed (email, IM) downloaders
    - When run it installs bot
  - Malicious websites
  - Hacked websites
 - Exploits the human factor

ARBOR
NETWORKS

Botnets in the Internet Underground

- **Bots are distributed computing and resources**
- **Help build a buffer between criminals and victims**
- **Botnets have aggregate storage and bandwidth**
- **Excellent for illicit activities**
  - Spam (increasingly pump and dump)
  - DDoS
  - Warez, stolen media
  - Access to confidential information

ARBOR
NETWORKS

Botnet Uptimes

Botnet Uptime Durations
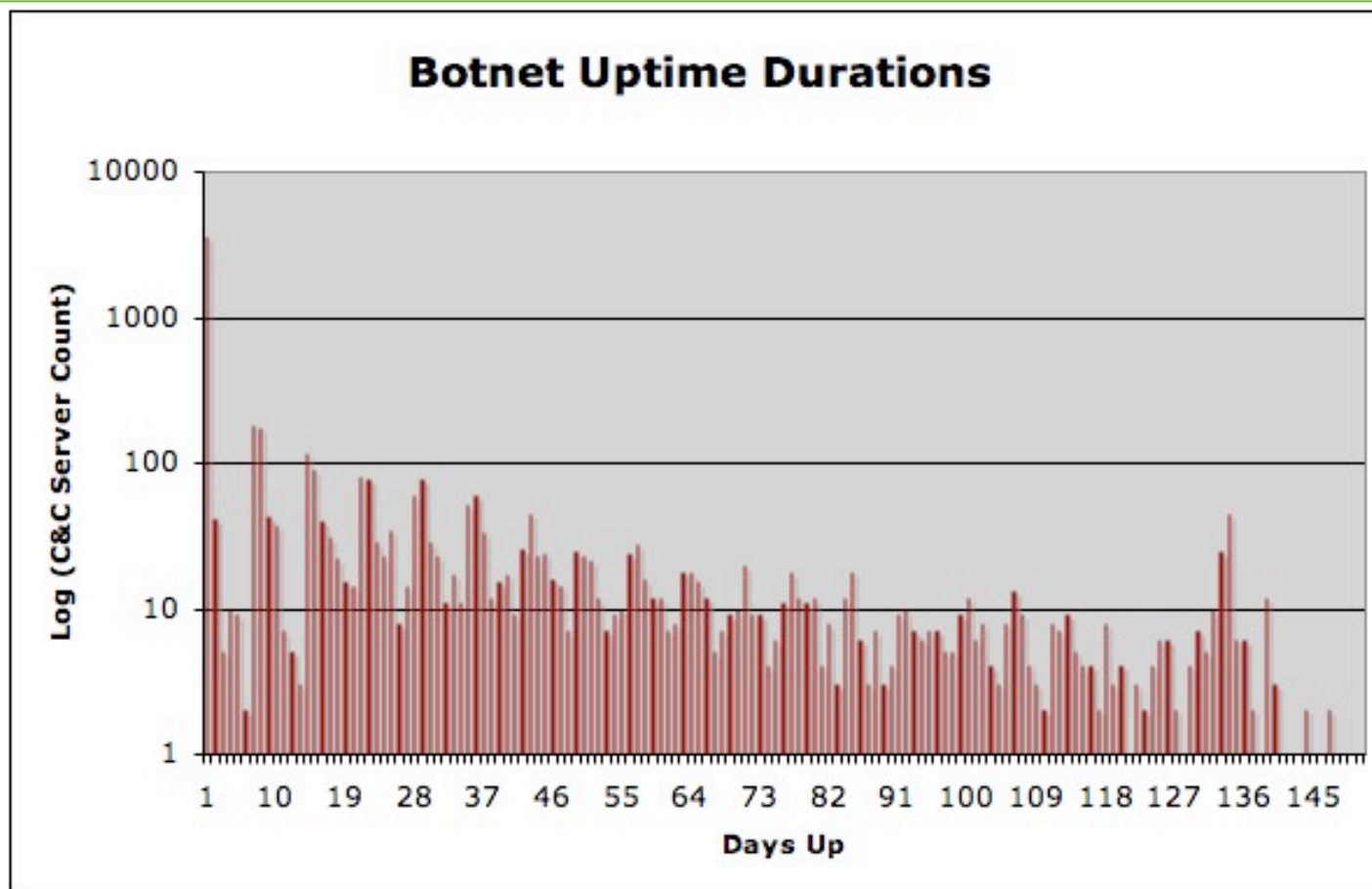
2/3 of botnets are up for 1 day or less
About 1000 botnet servers are active a day

ARBOR
NETWORKS

Botnet Server Geography

Country	Number of servers	Percentage
US (United States)	311	40.6%
KR (South Korea)	78	10.2%
CN (China)	49	6.4%
DE (Germany)	48	6.3%
CA (Canada)	41	5.4%
TW (Taiwan)	26	3.4%
GB (Great Britain)	24	3.1%
IT (Italy)	15	2.0%
JP (Japan)	13	1.7%
NL (Netherlands)	12	1.6%
Other	149	19.5%

Data taken from ATLAS, March 1, 2007

ARBOR
NETWORKS

Botnet Server by OS

Botnet Command Server OS

- Linux 2.2 or later
- Multiple guesses
- Apple OS X
- FreeBSD 3.x or later
- Other

6%
10%
12%
21%
51%
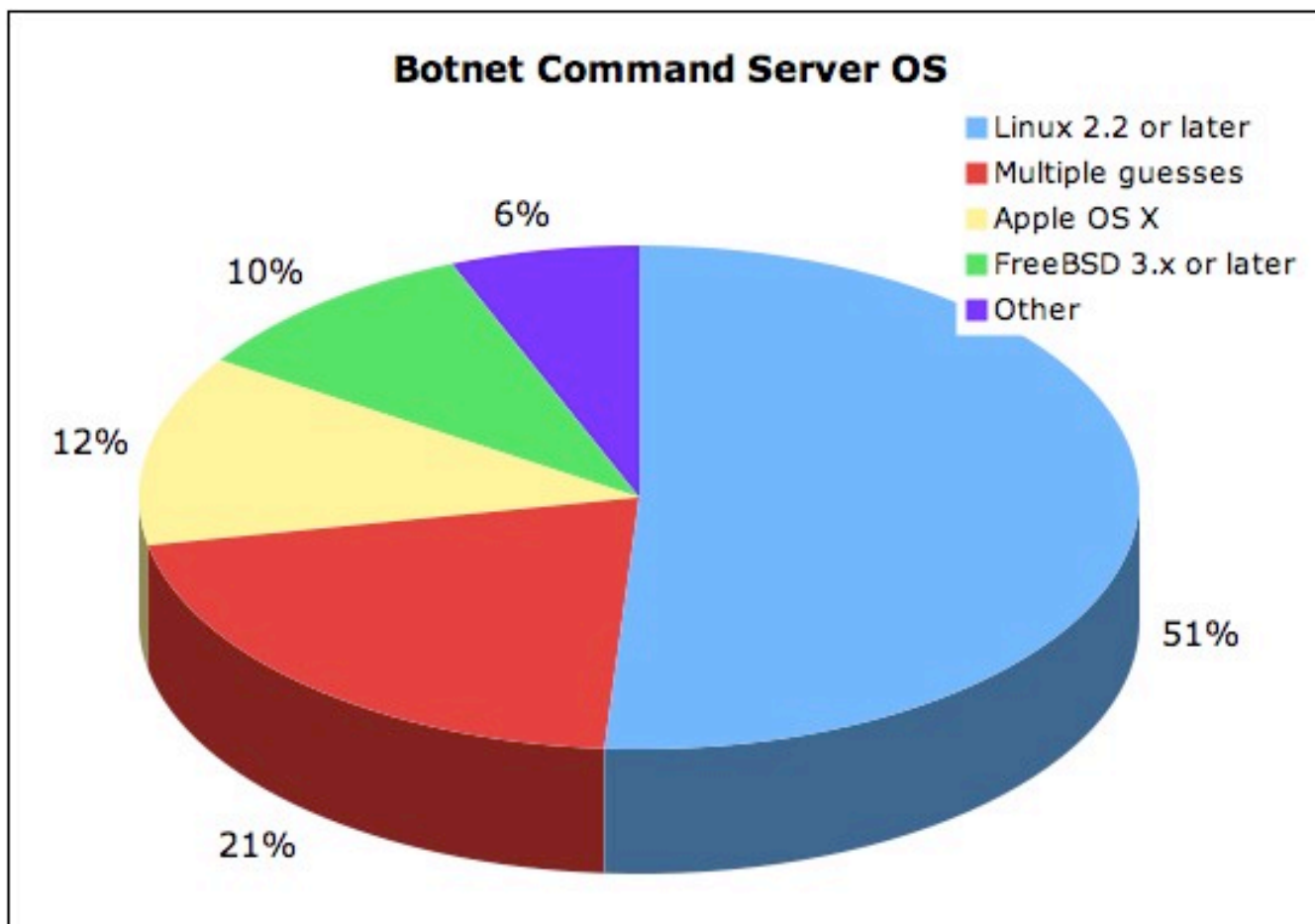
Based on 500 Botnet C&C servers, NMAP 4.0 fingerprinting
Data from Arbor Networks' botnet detection policies, 1 day snapshot

ARBOR
NETWORKS

One network over time
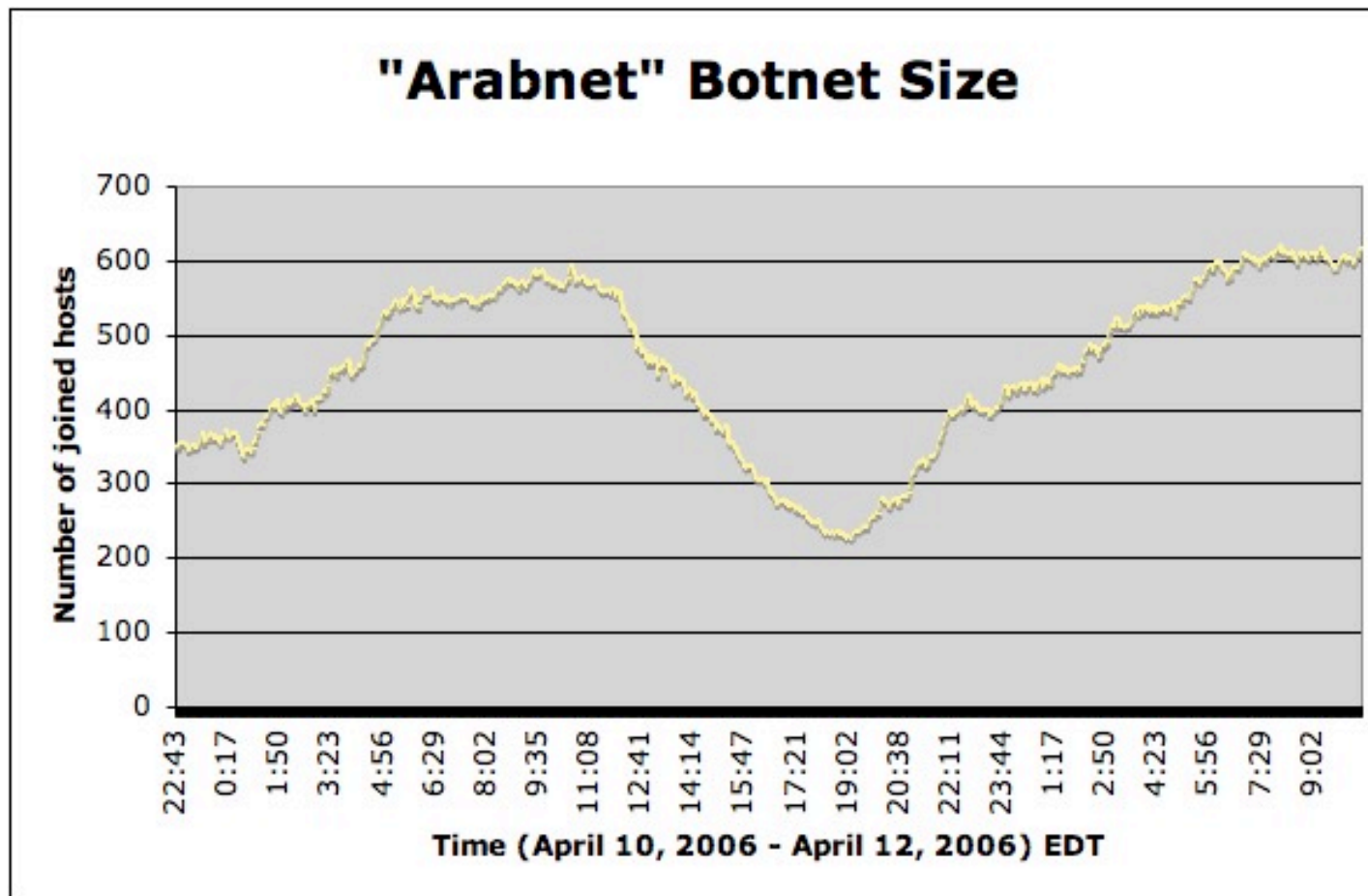
Based on active botnet monitoring by Arbor Networks

No One is Immune

```
BRA|042537931!~BRA042537@SaguenayNet-
1B7DA413.sa02.rprbbe.pae.procergs.rs.gov.br #DIA#
['[SCAN]: Random Port Scan started on 200.198.x.x:445 with
a delay of 5 seconds for 0 minutes using 150 threads.']
```

- **Access to government or military sites occurs**
- **Such access is usually accidental**

Know Your Goals

- **Malware Collection**
  - Popular with AV, security companies

- **Attack Traceback**
  - Our primary goal

- **Attacker Profiling and Assessment**
  - Small, specialized field

Botnet Tracking Requirements

- **Origins**
  - Can't do this from your desktop!

- **Targets**
  - Botnet server, passwords, bot characteristics, etc

- **Malware**
  - Have to know what a bot would do

- **Client**
  - Have to have a botnet client to participate

ARBOR
NETWORKS

Secondary Requirements

- **Distant origins**
  - Don't want it to tie back to you

- **Multiple origins**
  - Don't want to be too obvious

- **Familiarity with attacker underground**
  - Exploits, vulnerabilities, underground economy

- **Language skills**
  - Be able to read and write foreign languages

How to Actively Monitor Botnets

Sacrificial Lambs

- **One binary at a time**
  - Repeat for every new bot

- **High risk of participating in an attack**

- **Lower risk of looking "out of place"**

Custom Clients

- **Multiple nets at once**

- **Easy to customize**

- **May look "different" (and hence suspicious)**

This is what we'll use

ARBOR
NETWORKS

Botnet Tracking Client Requirements

- **Secure**

- **Scalable**

- **Flexible**

- **Easy to retarget**

- **Records everything it sees**

- **Stealthy**

ARBOR
NETWORKS

Project Bladerunner

- **Botnet infiltration**
  - Active monitoring
  - Multiple networks at once
- **Uses Python and irclib module**
- **Also wrote a Kaiten tracking tool**
  - Kaiten affects Linux systems

- **Focused only on IRC-based botnets**

ARBOR
NETWORKS

About Bladerunner

- **Mimics a basic bot**
- **Understands "login", "join"**
- **Chooses to be quiet rather than misspeak**

- **Logs everything**

ARBOR
NETWORKS

Why a Custom Bot?

- **Time consuming to defang a bot**
- **Only needed very basic functionality**
- **Knew code very well**
- **Little risks (DDoS, installations, etc)**

- **Bladerunner was about 300 LoC**

ARBOR
NETWORKS

Botnets and DDoS

- **About half of all botnets we tracked performed DDoS attacks**
  - Most attacks are not against a significant target
  - Most attacks are not crippling to the endpoint

- **Did observe a set of high profile attacks in the spring of 2006**
  - Against a series of anti-spam and anti-DDoS companies

- **DDoS nets use different bots than spyware or adware bots**
  - Not all bots have DDoS capabilities
  - Type of bot used can often indicate intent of herder

ARBOR
NETWORKS

Botnet Tracking as DDoS Traceback

- **Looked at DosTracker archive**
  - Arbor project to analyze global DDoS prevalence
  - Over 20,000 DDoS attacks measured between Sept 2006 and January 2007

- **Looked at Shadowserver botnet tracking logs of DDoS attacks**
  - Over 21,000 attacks in this timeframe
  - Over 400 unique IRC servers

- **Attack intersection results**
  - **2%** of all DDoS attacks measured by Arbor had clear botnet cause
  - **13%** of all DDoS attacks recorded by botnet tracking showed up in Arbor monitors

ARBOR
NETWORKS

Our Current Position in Botnet Response

- **(Community position)**

- **Collection**
  - Nepenthes or other honeypots

- **Communication**
  - Whitestar list, DA, NSP-SEC, Shadowserver, etc

- **Analysis**
  - Sandboxing (Norman dominates)

- **Tracking**
  - Shadowserver, some private tracking

ARBOR
NETWORKS

Where the Botherders Are

- **Source code is widely available**
  - GPL licensed, using CVS!
  - GUI-based configuration, no coding skills needed
  - Bug fixing
    - Compare SpyBot in 2004 and 2006
    - Lots of little bugs fixed: string bounds checks, etc

- **Multiple types of bots**
  - SpyBot, SDBot, Reptile, Agobot, Rbot, RxBot, Kaiten, etc …
  - Lots of overlapping capabilities, not all support DDoS
  - Which codebase you use depends on your intentions

- **Proliferation of spyware, adware provides money**

ARBOR
NETWORKS

Low Skilled Attackers

```
Nads!hacker@208.189.38.32 #asdf ['do a whois']
Nads!hacker@208.189.38.32 #asdf ['and paste it']
Nads!hacker@208.189.38.32 #asdf ['btw']
Nads!hacker@208.189.38.32 #asdf ['have you tried logging into it in PM?']
jowww!N0002@Netadmin.net #asdf ['Oo']
jowww!N0002@Netadmin.net #asdf ['na a :D']
Nads!hacker@208.189.38.32 #asdf ['k']
Nads!hacker@208.189.38.32 #asdf ['what you do is']
Nads!hacker@208.189.38.32 #asdf ['log into it in a PM']
Nads!hacker@208.189.38.32 #asdf ['k']
Nads!hacker@208.189.38.32 #asdf ['then remove it']
Nads!hacker@208.189.38.32 #asdf [':\\']
jowww!N0002@Netadmin.net #asdf ['what is PM? :D']
Nads!hacker@208.189.38.32 #asdf ['like this']
Nads!hacker@208.189.38.32 #asdf ['./query']
Nads!hacker@208.189.38.32 #asdf ['and .rm the bot']
Nads!hacker@208.189.38.32 #asdf [':\\']
Nads!hacker@208.189.38.32 #asdf ['Only way for it to go away']
jowww!N0002@Netadmin.net #asdf ['dont wana let me in :P']
Nads!hacker@208.189.38.32 #asdf ['hm...']
Nads!hacker@208.189.38.32 #asdf ['ok']
Nads!hacker@208.189.38.32 #asdf ['have you tried glining it?']
jowww!N0002@Netadmin.net #asdf ['glining?']
Nads!hacker@208.189.38.32 #asdf ['./gline 198F1F9E.37ACCE37.162073EA.IP']
Nads!hacker@208.189.38.32 #asdf ['will get rid of bot']
Nads!hacker@208.189.38.32 #asdf ['on the server']
Nads!hacker@208.189.38.32 #asdf ['from connecting']
```

High Firepower DDoS

- **Part of the BlueSecurity, Prolexic, Rackspace, 6 Apart DoS**

- **Continued for more than a week**

```
Tue Apr 25 17:58:37 2006 pubmsg sadf!tsinternetuser@room
#usa# ['.tusa ack 72.52.6.3 80 40']
Tue Apr 25 17:58:57 2006 pubmsg sadf!tsinternetuser@room
#usa# ['.x x']
Tue Apr 25 17:58:57 2006 pubmsg sadf!tsinternetuser@room
#usa# ['.tusa ack 72.52.6.3 80 40']
Tue Apr 25 17:59:18 2006 pubmsg sadf!tsinternetuser@room
#usa# ['.x x']
Tue Apr 25 17:59:18 2006 pubmsg sadf!tsinternetuser@room
#usa# ['.tusa ack 72.52.6.3 80 40']
```

ARBOR
NETWORKS

Where the Botherders Aren't

- **IRC**
  - Too many snoops on IRC
  - Too easy to break into
  - Lots its "elite" factor some time ago
  - Growing number of HTTP, IM, and other bots

- **Web Forums (eg Ryan 1918)**
  - They know these are monitored

ARBOR
NETWORKS

We've Peaked!

- **This combination reached its peak in early 2006**
- **Good guys**
  - Lots of basic RE analysts
  - Armed with tools like sandboxes
  - Lots of collection networks (ie Nepenthes)
  - Rapidly caught, analyzed, and tracked botnets
- **Bad guys**
  - Explosion in bots and botnets launched
  - Only a few botnet groups were actively thwarting attacks
  - HTTP and P2P bots were not very popular yet (still IRC heavy)
  - Lots of botnets were very visible

- **This confluence meant we peaked**

ARBOR
NETWORKS

The Revolt by Botnet Operators

- **More and more bots are defeating the basic techniques**

- **Sandboxes are being defeated**
  - Increased use of debugger checks
  - Delays in revealing useful information
  - Poisoning data
  - Inject fake bots to detect people who mine Norman for data

- **Honeypots and honeynets**
  - Detected or ignored

- **IRC tools**
  - Fingerprinted and blocked, or simply ignored

- **It's all downhill from here!**

ARBOR
NETWORKS

The Botnet Herder Ability Curve

Can barely use IRC

DDoS as a pissing match

Lured by adware dollars
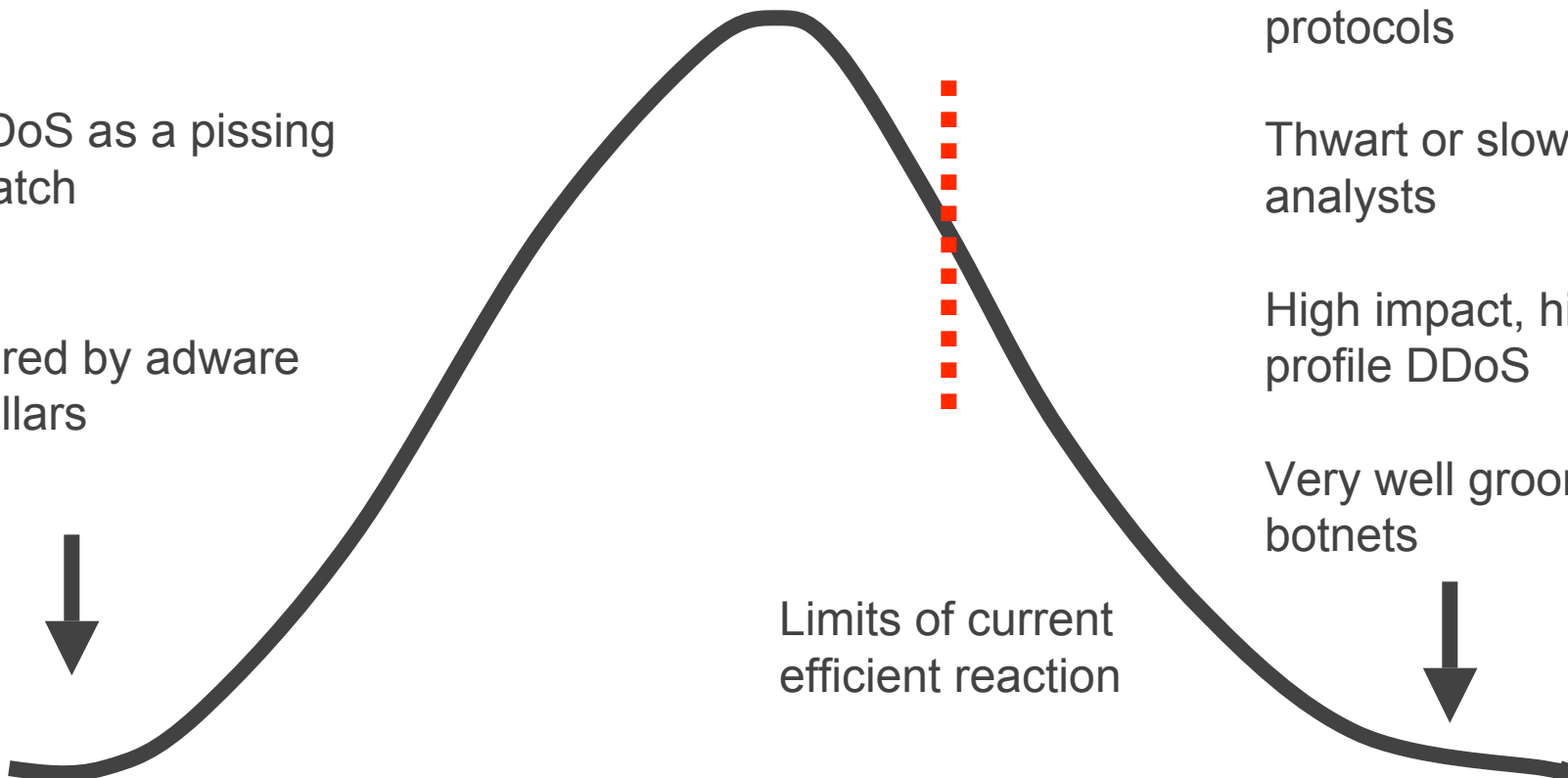
Write their own communication protocols

Thwart or slow RE analysts

High impact, high profile DDoS

Very well groomed botnets

Limits of current efficient reaction

ARBOR
N E T W O R K S

Getting Botnets Taken Down

- **Getting the information in the right hands**
  - Thousands of botnets a week, only so much operators can do
  - Cannot blindly block
- **Focus is on active, high profile DDoS networks**
- **Coordination is a pain in the neck**
  - DNS registrar
  - DNS server network(s)
  - C&C host network(s)

- **Botnet operators can easily stay a few steps ahead**

- **Complement is egress filtering for victims**

Botnet Takedown Success Rates

- **Based on Arbor Networks' aggregated data of botnet controller addresses**

ASN8376	GO-JOR	9/9
ASN3320	DTAG	11/11
ASN3225	Gulfnet	15/15
ASN4814	CHINA169-BB	26/26
ASN22927	Telefonica	35/35

- **Most targeted networks**

ASN30058	FDCSERVERS	466 C&C servers	72% takedown
ASN25761	STAMINUS-COMM	400 C&C servers	53% takedown
ASN19318	NJIIX-AS-1	220 C&C servers	60% takedown

ARBOR
NETWORKS

Technical Challenges

- **Encrypted communications channels**

- **Defeating rapid analysis techniques**

- **New or custom command languages**
  - HTTP, peer to peer

ARBOR
N E T W O R K S

Encrypted Channels

- ## Encryption
  - Windows "Somelender" bots - homegrown Caesar cipher

```
(66.186.35.22:8080) :ckodg!j@tyrant PRIVMSG
## :=GoU6jyt7xCuvfRamp+NOAeNFFF/q/h9EHT/H6DV5fxcD7RoX9Pt5a/o2AST9N+j4Y4jf
(66.186.35.22:8080) :ckodg!j@tyrant PRIVMSG ## :=rvyJWDmfvujXJ4XDKp5
(66.186.35.22:8080) :ckodg!j@tyrant PRIVMSG ## :=+rhlS+/trmwFfUNtERLa
```

Decrypts to:

```
(66.186.35.22:8080) :ckodg!j@tyrant PRIVMSG ## :40% ddos tcp
65.77.140.140 6667 900 -s -f -i -2
(66.186.35.22:8080) :ckodg!j@tyrant PRIVMSG ## :* kill dos
(66.186.35.22:8080) :ckodg!j@tyrant PRIVMSG ## :* kill ddos
```

Fallout from Encrypted Commands

- **Very time consuming**

- **Two options**
  - Mimic bot
    - Must reverse encryption algorithm
    - Must implement
  - Honeypot the bot and monitor it
    - Doesn't scale well

- **This dramatically slows down botnet tracking**

ARBOR
N E T W O R K S

Defeating AV Detection

- **Polymorphism is rare**
  - Achieve polymorphism by simply repackaging bots
  - New or modified packer
  - Fresh compile
  - Bingo, AV fails to detect

- **The bot is just a tool to load the real payload on the box**
  - Spyware, adware, spam tools, etc …
  - The bot code itself can be thrown away once it's gotten the second stage payload on board

ARBOR
NETWORKS

Anti Analysis Techniques

- **Increased use of rapid analysis thwarting tools**
  - eg Debugger detection
  - Poisoned "wells" (honeypots)

- **Detection and disabling of sandbox tools**
  - Detect VMWare
  - Detect Norman
  - Result: no results

- **Solution: put a human in the loop**

ARBOR
NETWORKS

Defeating Sandboxes and Honeypots

```
/* Check if running inside VMWare */
int IsVMWare()
{
        int version=VMGetVersion();
        if(version)
                return TRUE;
        else
                return FALSE;
}

/* Fool ProcDump with increasing size */
void FoolProcDump()
{
        __asm {
                mov eax, fs:[0x30]
                mov eax, [eax+0xC]
                mov eax, [eax+0xC]
                add dword ptr [eax+0x20], 0x2000        // increase size variable
        }
}

/* Screw with Connectix VirtualPC */
#define vpcscreen __asm _emit 0x0F __asm _emit 0x3F __asm _emit 0x01 __asm _emit 0x0C
#define vpcadditions __asm _emit 0x0F __asm _emit 0x3F __asm _emit 0x05 __asm _emit 0x01
```

ARBOR
NETWORKS

HTTP Bots

- **Two main mechanisms**
  - Phone home (register, poll for commands: beaconing)
  - Register, await an inbound connection
- **Communication is over HTTP, using URLs**

- **Korgo, Padobot, Bzub, Nuclear Grabber, MachBot**

- **Example registration URL**
  - ```
    http://XXXXXXX/index.php?
    id=jqkooamqechepsegsa
    &scn=0
    &inf=0
    &ver=19
    &cnt=GBR
    ```

ARBOR
NETWORKS

HTTP Bot Implications

- **Harder to spot**
  - No long lived connection

- **Have to know what to look for in URL logs**
  - Hiding in the maelstrom

- **Still uses a central command point**
  - Easy to block

- **Not too hard to lurk**
  - Poll server, understand replies

ARBOR
NETWORKS

Peer to Peer Bots

- **Storm Worm (CME-711, January 2007)**
  - UDP-based eDonkey protocol
  - Used to send spam
  - About 50,000-100,000 hosts active

- **Nugache (Spring, 2006)**
  - Encrypted TCP, custom command protocol
  - No clear use for this network yet
  - Network is being matured
  - About 20,000-75,000 hosts sustained

ARBOR
NETWORKS

Peer to Peer Bot Implications

- **Attempted for years, now in production**
- **Resilient network**
  - No central point to shut down
  - No central point to block
- **Difficult traceback**
  - Network manager can enter network from anywhere

- **Anyone can join network**
- **Reverse protocol, join and lurk**

DNS Signalling

- **Change DNS records to point to DDoS target**
- **DNS server hosted within another botnet or bot-friendly ISP**
- **Most DNS traffic is ignored and permitted to pass**

- **Used in recent Spamhaus DDoS events**

ARBOR
N E T W O R K S

Changes in Botnet Handlers' Intents

- **Previously**
  - Getting the bot on there was the end goal
  - Keeping the bot on there was important

- **Now**
  - The bot is just to bootstrap new code on there
  - The bigger that window of opportunity is, the better
  - Evade AV detection by staying ahead
  - First seen on a wide scale with Zotob

ARBOR
NETWORKS

Success on Their End

- Increased spam volumes
- All attributable to deployed botnets

- High impact DDoS events against high profile crimefighters, antispam groups

- Inter-spam gang fighting

- With success like this, don't expect a slowdown

The Botnet Arms Race

Bad Guys

Then

- More bot families
- More bots
- Packers and obfuscators
- More botherders

- Leaving IRC behind

Now

- Encryption

Good Guys

Scalable

- Behavioral analysis
- Sandboxes
- Process dump tools

- More analysts

- Sacrificial lambs

Not

- Reversing

ARBOR
NETWORKS

Conclusions

- **Botnets have been a sustained growth industry**

- **Botnet herders have increasingly ditched their "minders" (the good guys)**

- **Botnets are increasingly used for high profile problems and crime**

- **We must work hard to adapt to these new realities and increase our monitoring**
  - Collaboration will be crucial

An Untenable Position

Reactive | *Proactive*

How do we get from here …..   To here? We must.