

## Botnets: The Web Killer

Chris Lee  
Computer Network Security  
March 7th, 2008

---

---

---

---

---

---

---

## Online Crime

- Motives
  - Money
  - Thuggery
  - Espionage
  - Money
- Tactics
  - Spam
  - DDoS
  - Targeted attacks
  - Proxies
  - Phishing
  - Spyware
  - Spam

---

---

---

---

---

---

---

## Tools of Online Crime

- Phishing Kits
  - Still needs spam to lure people
- Malware
  - Botnets
    - DNS fast flux networks
    - Spamming
    - Web servers for phishing and spam
    - DDoS
  - Spyware
    - Passwords
    - Software keys
    - Fraudulent banking transactions

---

---

---

---

---

---

---

## Botnets are used for Cybercrime

- DDoS (\$500~\$1500 per attack)
- Phishing (~\$2B/year)
- Keylogging/Spying (Sharma \$150K)
- Software license key stealing
- Spamming (2 Men, ~\$2B in 5 years)
- Attack Evasion
- Click fraud
- Adware (DollarRevenue \$430/day)

---

---

---

---

---

---

---

## Why Use Botnets?

- Power. Lots of nodes = Lots of bandwidth, storage, processing power, and IPs
- Anonymity. Botmaster uses compromised hosts, effectively hiding herself
- Hard to block. Since the botmaster has many IPs, blocking spam and attacks become difficult
- Availability. Lots of source code in the underground and lots of victims waiting for the taking.

---

---

---

---

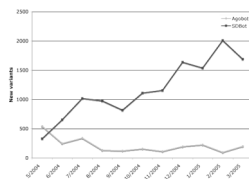
---

---

---

## Botnet History and Codebases

- Started as useful agents for managing systems and IRC channels
- 1993 The IRC Wars
- 1999/March Pretty Park
- 1999/May SubSeven
- 2000 GTBot
- 2002 SDBot, AgoBot
- 2003 MyDoom, Sobig
- 2003 Sinit
- 2004 PhatBot + WASTE
- 2006 Nugache.A
- 2007 Storm Bot



---

---

---

---

---

---

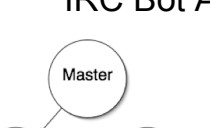
---

# Botnet Architectures

- C&C Discovery
  - IP Address
  - DNS Name
  - Bootstrap Peers
  - Random Scanning
- Services
  - Spam
  - DNS
  - Proxy
  - Web Server
  - Scanning
- Spreading
  - Emails
  - Remote exploit
  - Trojan
- Command Channel
  - IRC
  - HTTP
  - P2P Network
  - DNS
  - ICMP

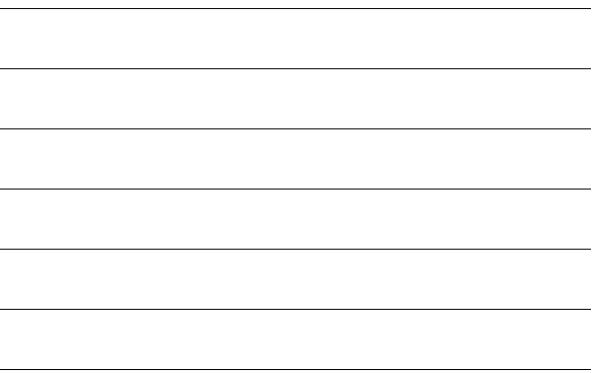
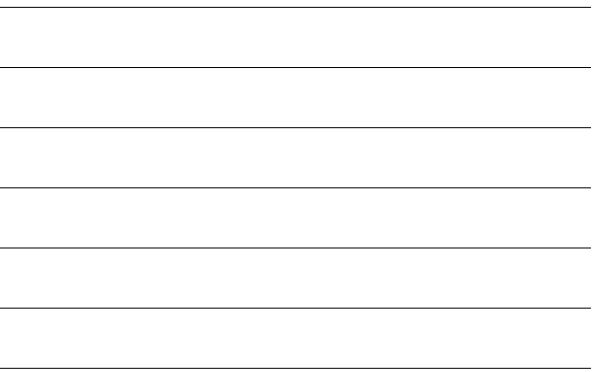
- 
- 
- 
- 
- 
- 

# IRC Bot Architecture

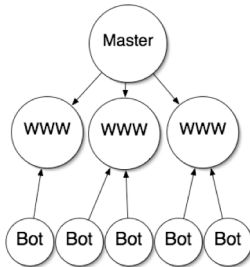


```
graph TD; Master((Master)) --> IRC1((IRC)); Master --> IRC2((IRC)); IRC1 <--> IRC2; IRC1 <--> IRC3((IRC)); IRC2 <--> IRC3; Bot1((Bot)) --> IRC1; Bot2((Bot)) --> IRC1; Bot3((Bot)) --> IRC2; Bot4((Bot)) --> IRC3; Bot5((Bot)) --> IRC3;
```

- Internet Relay Chat
  - Servers relay messages
- Bots can connect to different servers
- Botmaster can use any server on that IRC network to control bots
- Botmasters often uses [undernet.org](http://undernet.org) due to their lousing policing or their own custom IRC

[illegible]

## HTTP Bot Architecture



- Botmaster updates webpage with instructions
- Bots periodically check website using standard HTTP
- Hard to detect and block

---

---

---

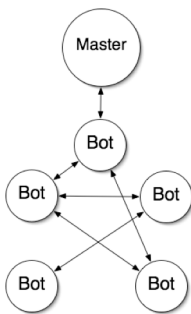
---

---

---

---

## P2P Botnet Architectures



- Uses a peer to peer protocol
- Highly resilient to take downs
- Master can use any peer on the network as a controller
  - But typically doesn't
- Encryption is common
- Usually very professionally built

---

---

---

---

---

---

---

## Theme #1: Evolution

- Botnets evolve making research reactive and difficult
- Researchers have difficult barriers to perform botnet research
- There are a lot of areas for researchers to help (more on that later)

---

---

---

---

---

---

---

## Botnet Networking Evolution

- Bad Guys
  - Simple IRC Botnets
  - Dynamic Nick/Channels
  - Setting up IRC servers
  - DNS redirection
  - Polymorphism and alternate channels of communication (e.g., P2P)
- Good Guys
  - Closing channels, Locking out IPS/Nicks
  - Behavioral blocking
  - Take down
  - Multiple take down, DNS “poisoning”
  - Intrusion Detection Systems & Antivirus

---

---

---

---

---

---

---

## Botnet Binary Evolution

- Bad Guys
  - More bot code bases
  - More bots
  - Packers and obfuscation
  - More botheaders
  - Leaving IRC
  - Encryption
  - Debugger/VM detection
- Good Guys
  - Behavioral analysis
  - Sandboxes
  - Process dump tools
  - More analysts
  - Honeypots
  - Reverse engineering
  - ??

---

---

---

---

---

---

---

## Theme #2: Different Strokes

- Botnets come in different shapes and sizes for different purposes.

---

---

---

---

---

---

---

### Tale of Two Botnets

- How: SSH Brute force
- What: IRC proxybot
- Who: Romanian kids
- Why: DDoS other kids off of IRC
- How: eCard social engineering spam with web-drive by downloads
- What: P2P bot with differentiated services
- Who: Russian mafia
- Why: Spam \$\$\$\$

---

---

---

---

---

---

---

### Storm Botnet

- Infects machines using browser exploits on webpages for games or ecards.
- P2P Botnet using Overnet for communicating updates
- Mainly used for spam and protective auto-DDoS
- Used RSA encryption to encrypt updates, now uses XOR encryption on all messages.
- Has tiered services, spammers, DNS, web proxies, and web servers
- Auto-DDoS triggered by probing web proxies

---

---

---

---

---

---

---

### Theme #3 Counting is hard

- 1, 2, 3, 7, 4, 1... where was I?
- In a P2P botnet, enumerating peers is hard
  - Constant DHCP churn
  - Nodes joining and leaving
  - Peers giving only partial peer lists
  - Liars

---

---

---

---

---

---

---

## Enumeration

- Goal: to find the size and topology of the Storm botnet
- Method: join hundreds of nodes to the botnet and "ride"
- Technologies:
  - Ultra, ultra lightweight virtualization with 3~10MB footprint per host, (IP bound, not CPU or RAM)
  - Threaded overnet crawler in PERL
  - Spamtraps
  - Honeypots
  - Instrumented Virtual Machine (QEMU)

---

---

---

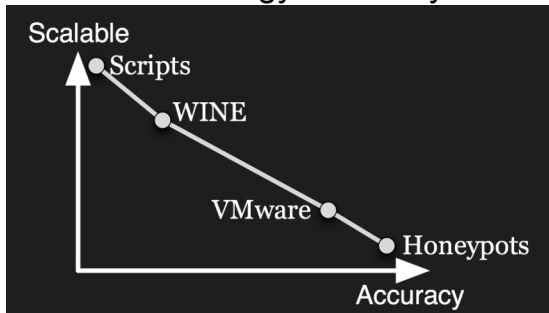
---

---

---

---

## Technology Accuracy



---

---

---

---

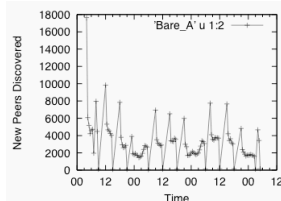
---

---

---

## Bare Metal

- Real machine, real os, real infection
- Bad traffic blocked, all else passed and recorded
- Connects to peers allowing us to naturally enumerate them



---

---

---

---

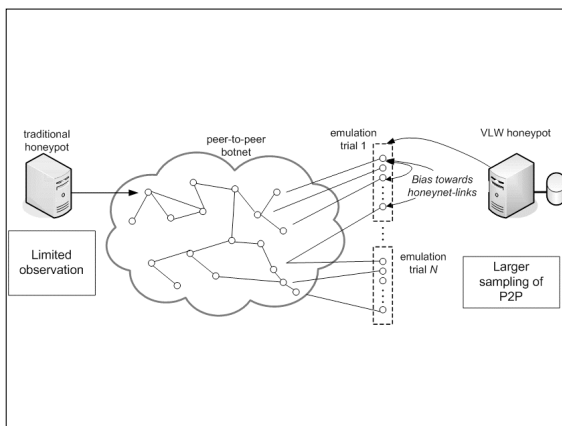
---

---

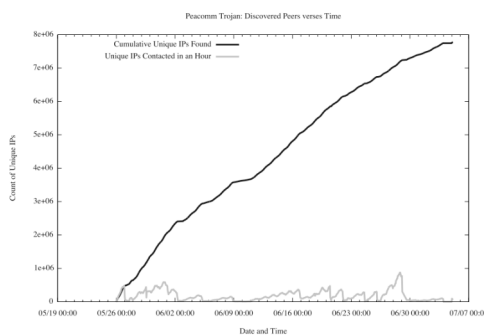
---

## WineBots

- WINE converts WIN32 API calls to Linux system calls
- Modified to separate each instance's mutex, registry, and network stack
- Launched hundreds of instances of Storm bot
- IP allocation limited, not IO, not RAM, not CPU
- Fragile and requires lots of hacking to run different versions of malware

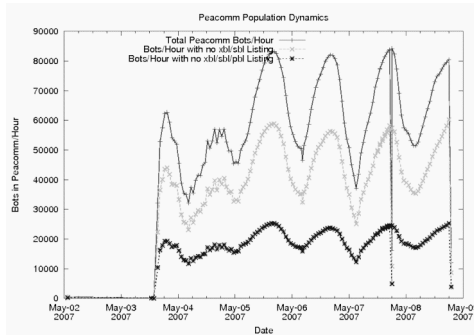


## Storm Bot Discovery Rate





## Population Dynamics



August 2007

---

---

---

---

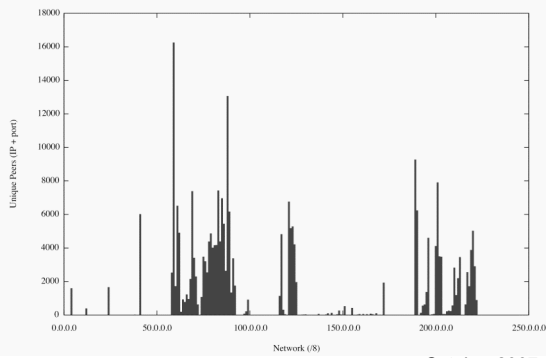
---

---

---

---

Distribution of Storm Peers in the IP Address Space



October 2007

---

---

---

---

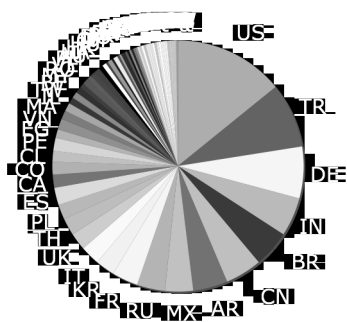
---

---

---

---

## Storm Bot Infection Breakdowns



October 2007

---

---

---

---

---

---

---

---

## Data

- Winebots (Oct. 2007)
  - 1/2 TB of compressed pcap files
  - 27 million unique IPs seen
  - 180~230 K unique hosts daily
- Threaded crawler
  - 20~50 K new ips daily
- SpamTrap
  - 3 GB of spam daily
  - Used to collect Storm spammer IPs and proxy addresses for "ground truth"
- Honeypots
  - 4 bare-metal Storm infected nodes for protocol "ground truth"

---

---

---

---

---

---

---

## Overnet Crawlers

- Built a variety of crawlers
- Different methods yield very different results
- Some methods are more covert than others
  - UCSD keeps scanning every minute even after all other nodes stop talking to us
  - Another kept giving different hash values for the same IP/port combination
  - Others will use every port and IP in a small allocation
- Active crawling will not be able to contact NATted victims

---

---

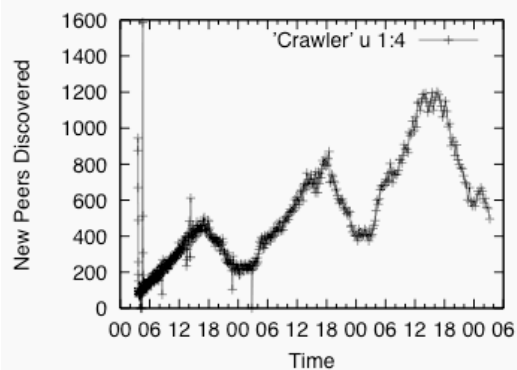
---

---

---

---

---



---

---

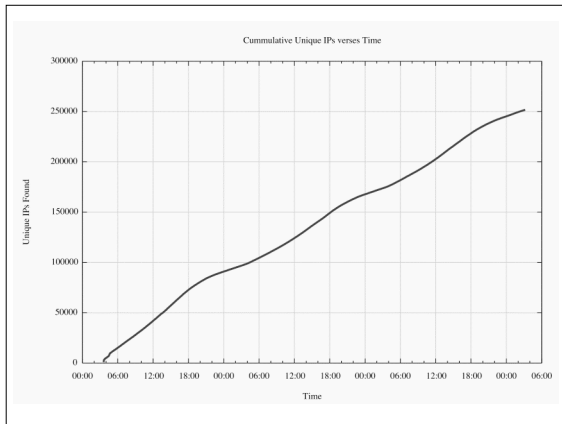
---

---

---

---

---




---

---

---

---

---

---

---

---

## Passive Protocol Monitoring

- Since Storm uses Overnet, it uses the distributed hash table (DHT) to search for updates
- The DHT uses peer hashes to “address” peers in routing tables
- We can chose our own peer hash, so we choose peer hashes across the entire hash space
- Many nodes will come to us because we're well situated in routing tables

---

---

---

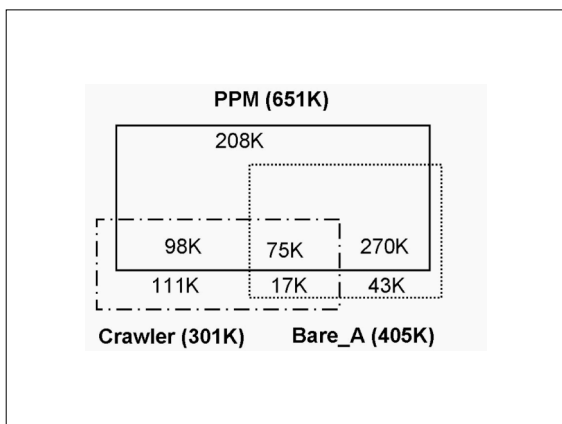
---

---

---

---

---




---

---

---

---

---

---

---

---

## Storm Conclusion

- Measuring storm is hard
  - Constant flux of IPs and online population
  - Various measurement techniques
  - NATs and Firewalls
- Using multiple techniques yields good results
  - Bare metal gives ground truth
  - WINE gives a large number of peers
  - Crawler operates quickly
  - PPM uses the searching protocol to its advantage

---

---

---

---

---

---

---

## Research Directions

- Research Topics
  - Botnet detection at the network level
  - Binary reverse engineering
  - Botnet and covert malware modeling
  - Victim enumeration
  - Worst-of-breed botnet architectures
    - E.g., Wireless only botnets
  - Tracking/attacking the criminal economy
    - E.g., Forging billions of fake CC accounts
  - Anti-spam techniques

---

---

---

---

---

---

---

## Places to Start

- Barford - An Inside Look at Botnets
- Dagon - A Taxonomy of Botnets
- Honeynet - KYE: Tracking Botnets
  
- Cymru - The Underground Economy
- FTC - Spam Summit - Uncovering the Malware Economy
  
- SRI International - A Multi-perspective Analysis of the Storm Worm

---

---

---

---

---

---

---

## Interested?

- If you think you'd like to research botnets and/or online crime, let's talk.
  - chrislee at gatech
    - PGP ID: 5AED 522C 4A53 BC89 7494 6A17 F69C 9528 14E4 4DBF
    - Available from subkeys.pgp.net or <http://chrislee.dhs.org/files/chrislee.pub>
- Thank you.

---

---

---

---

---

---

---