Lecture notes for botnets

Def: Bot
- Software that
  - executes without intent of system owner
  - accepts commands from remote attacker via the network
  - executes network-based attacks
  - [draw picture]
- Bots are resources to attackers
  - install to remain persistent across reboot
  - try to minimize impact to system
  - patch software flaws to keep other bots out

Def: Botnet
- A network of bots under control of same master
- Essentially a malicious distributed system created without consent of node owners
- [draw picture]

Def: Botmaster
- The attacker who controls an entire botnet

Evolution
- evolved from viruses & worms
- worms: automated propagation for disruption
- bots: automated propagation to create distributed system that makes an attacker money
- essentially worms with profit capability

Bot network use
1. Propagation for bot installation
2. Rallying messages sent by bots to botmaster
3. Controlling messages sent by botmaster to bots
4. Attack traffic sent by bots in response to command

Def: Command & Control

- The network communications channel used by a botnet for rallying and controlling messages
  - Rallying: bots advertising presence
  - Control: botmaster issuing commands to bots
- Common C&C: IRC, HTTP, P2P

Uses for botnets

- DDoS
  - 4% of global internet traffic is DDoS
- Spam generation
  - 95-100% of global spam comes from botnets
- Host phishing websites
- These uses make money: extortion, junk mail, phishing

Defenses

- Software security (prevent propagation)
- Detect propagation (unusual/many network connections)
- Detect rallying and C&C
  - DNS-based detection
  - Bots do not use hardcoded IP addresses: reveals attacker, provides straightforward shutdown of C&C
  - Bots use DNS to allow botmaster to change IP address of C&C server
  - Detect unusual DNS use & take over account
- Withstand attacks (iron, spam prevention)

Spam prevalence

- Extremely low cost to send
- Even minuscule response rates can generate income

Spam prevention

- Analyze message content
  - Bayes filtering
  - Identify distinguishing words
  - Requires per-user training period
  - Random messages, random blocks of text in spam used to confuse

training
- Analyze message properties
  - SpamAssassin
  - Identify distinguishing characteristics
- Analyze traffic
  - Identify unusual message sending patterns / rates
- Collaborative identification
  - Only few recipients need mark as spam for all recipients to receive filtering
  - Gmail, any centralized service
- Increase cost of sending
  - Computation, micro-payments
- Legal repercussions