CSE543 - Introduction to Computer and Network Security

Module: Botnets

Professor Patrick McDaniel
Fall 2008
Story
Botnets

• A **botnet** is a network of software robots (bots) run on **zombie machines** which run are controlled by **command and control** networks
  ‣ **IRCbots** - command and control over IRC
  ‣ **Bot herder** - owner/controller of network
  ‣ "**scrumping**" - stealing resources from a computer

• Surprising Factoid: the IRC server is exposed.
Statistics (controversial)

- The actual number of bots, the size of the botnets and the activity is highly controversial.
  - As of 2005/6: hundreds of thousands of bots
  - 1/4 of hosts are now part of bot-nets
  - Growing fast (many more bots)

- **Assertion:** botnets are getting smaller(?!?)
What are botnets being used for?

Activities we have seen

**Piracy**

Stealing CD Keys:

```
ying!ying@ying.2.tha.yang PRIVMSG #atta :BGR|0981901486 $getcdkeys
BGR|0981901486!nmavmkmyam@212.91.170.57 PRIVMSG #atta :Microsoft Windows
Product ID CD Key: (55274-648-5295662-23992).
BGR|0981901486!nmavmkmyam@212.91.170.57 PRIVMSG #atta :[CDKEYS]: Search completed.
```

**Mining**

Reading a user's clipboard:

```
B][!Guardian@globalop.xxx.xxx PRIVMSG ##chem## :~getclip
Ch3m|784318!~zbhibvn@xxx-7CCCB7AA.click-network.com PRIVMSG ##chem## :-
[Clipboard Data]- Ch3m|784318!~zbhibvn@xxx-7CCCB7AA.click-network.com PRIVMSG
##chem## :If You think the refs screwed the seahawks over put your name down!!!
```

**Attacks**

DDoS someone:

```
devil!evil@admin.of.hell.network.us PRIVMSG #t3rr0r0Fc1a :!pflood 82.147.217.39
443 1500 s7n|2K503827!s7s@221.216.120.120 PRIVMSG #t3rr0r0Fc1a :\002Packets\002
\002D\002one \002;\002>\n s7n|2K503827!s7s@221.216.120.120 PRIVMSG #t3rr0r0Fc1a flooding....\n```

**Hosting**

Set up a web-server (presumably for phishing):

```
[DeXTeR]!alexo@185-130-136-193.broadband.actcom.net.il PRIVMSG [Del]29466
:.http 7564 c:\ [Del]38628!zaazbob@born113.athome233.wau.nl PRIVMSG _[DeXTeR] :
[HTTPD]: Server listening on IP: 10.0.2.100:7564, Directory: c:\\.
```
Other goals of a botnet ...

- SPAM relays
- Click fraud
- Spamdexing
- Adware
IRC botnets

• An army of compromised hosts ("bots") coordinated via a command and control center (C&C). The perpetrator is usually called a "botmaster".

"A botnet is comparable to compulsory military service for windows boxes"

-- Bjorn Stromberg
Bots usually require some form of authentication from their botmaster.
Infection

- Worms, Trojan horses, backdoors

**Note:** the software on these systems is updated

**Bot theft:** bot controllers penetrate/"steal" bots.
IRC

- 1988 - one-to-many or many-to-many chat (for BBS)
- Client/server -- TCP Port 6667
- Used to report on 1991 Soviet coup attempt
- Channels (sometimes password protected) are used to communicate between parties.
  - Invisible mode (no list, not known)
  - Invite only (must be invited to participate)
Not only for launching attacks ...

- Some botmasters pay very close attention to their bots
  - hence *covert* infiltration is important
- In many cases, Botmasters “inspect” their bots fairly regularly, and isolate certain bots (“*cherry picking*”)

That’s a lot of movies served! (~ 300)
Lots of bots out there

• Level of botnet threat is supported by the conjecture that large numbers of bots are available to inflict damage

• Press Quotes
  ▸ “Three suspects in a Dutch crime ring hacked 1.5 million computers worldwide, setting up a “zombie network””, Associated Press
  ▸ “The bot networks that Symantec discovers run anywhere from 40 systems to 400,000”, Symantec
Measuring botnet size

- Two main categories
  - **Indirect** methods: inferring botnet size by exploiting the side-effects of botnet activity (e.g., DNS requests)
  - **Direct** methods: exploiting internal information from monitoring botnet activity
Indirect Methods

• Mechanism
  ‣ DNS blacklists
  ‣ DNS snooping

• What does it provide?
  ‣ DNS footprint

• Caveats
  ‣ DNS footprint is only a lower bound of the actual infection footprint of the botnet
  ‣ DNS records with small TTLs
  ‣ DNS servers blocking external requests (~50%)
DNS Blacklist

- The value of a bot is related to its status on the **DNS blacklists**
  - Compromised hosts often used as SMTP servers for sending spam.
  - DNS blacklists are lists maintained by providers that indicate that SPAM has been received by them.
  - Organizations review blacklists before allowing mail from a host.
- A "**clean**" bot (not listed) is worth a lot
- A listed bot is largely blocked from sending SPAM
**Observation**: bot controllers/users need to query for BL status of hosts to determine value.

**Idea**: if you watch who is querying (and you can tell the difference from legitimate queries), then you know something is a bot

**Understanding the in/out ratio**:

\[ \lambda_n = \frac{d_{n,\text{out}}}{d_{n,\text{in}}} \]

**Q**: what does a high ration mean? Low?
Results
Direct Methods

• Mechanisms
  ‣ Infiltrate botnets and directly count online bots
  ‣ DNS redirection (by Dagon et al.)

• What do they provide?
  ‣ Infection footprint & effective size (infiltration)
  ‣ Infection footprint (DNS redirection)

• Caveats
  ‣ Cloning (infiltration)
  ‣ Counting IDs vs. counting IPs (infiltration)
  ‣ Measuring membership in DNS sinkhole (DNS redirection)
  ‣ Botmasters block broadcasts on C&C channel (infiltration)
Estimating size [Monrose et. al]

- DNS redirection “sinkhole”
  - Identify, then self poison DNS entries
- DNS cache hits
  - Idea: query for IRC server to see if in cache
  - If yes, at least one bot in the network within the TTL (see [14])
  - Limitations: TTL, not all servers answer, lower bound on bots
How many bots?

• Approach: infiltration templates based on collected *honeynet* data, e.g., observing compromised hosts that are identified within the channel

• How many?
  ‣ 1.1 million distinct user IDs used
  ‣ 425 thousand distinct IP addresses

• Issues:
  ‣ NAT/DHCP?
  ‣ “Cloaked” IP address (SOCKS proxies?)
  ‣ Botnet membership overlap
Botnet size, what does it mean?

- **Infection Footprint**: the total number of infected bots throughout a botnet’s lifetime
  - Relevance: how widespread the botnet infection

- **Effective Botnet Size**: the number of bots simultaneously connected to the command and control channel
  - Relevance: the botnet capacity to execute botmaster commands (e.g., flood attacks)

- An Example:
  - While a botnet appeared to have a footprint of 45,000 bots, the number of online bots (i.e., its effective size) was < 3,000
Botnet footprint estimates

- Redirection results:
  - Botnets with up to 350,000 infected hosts [Dagon et al.]
Large botnets may not be so big!

Footprints

Effective size
Are we counting unique infections?

- Cloning activity observed in 20% of the botnets tracked (moving between bot channels)
- 130,000 bots created more than 2 million clones during our tracking period
Summary

• Size estimation is harder than it seems
• Botnet size should be a qualified term
  ‣ Different size definitions lead to radically different estimates
• Current estimation techniques are laden with a number of caveats
  ‣ Cloning, counting method, migration, botnet structures, DHCP, NAT, etc.
• A prudent study of the problem requires persistent multifaceted tracking of botnet activity