Getting your hands dirty: How to Analyze the Behavior of Malware Traffic and Web Connections

Workshop in Botconf 2016, Lyon, France

Date: 29/11/2016
Length: 3.5hs
Attendants: 29
Instructors: Sebastian Garcia, Veronica Valeros

Abstract

Nowadays there are a lot of tools to analyze traffic, but the most important thing to have is the experience and knowledge of a malware analyst. The goal of the workshop is to give a hands-on experience on analyzing the behavior of malware and botnet traffic in the network by studying their **web patterns** and their **traffic behavior.** The workshop will use both pcap files of real malware captures and real normal captures. Participants will learn a proven approach on how to do their traffic analysis, how to recognize malicious connections, how to separate normal behaviors from malicious behaviors, how to recognize anomalous patterns and how to deal with large amounts of traffic. Analyzing *only* malware traffic may not be so complicated for some people, but accurately **separating it from normal traffic** is harder.

The most important lesson of the workshop is **not** about how to use wireshark or tcpdump. The goal is to transmit the **experience** of recognizing the malicious actions of malware in the network. Specifically how malware hides, how to recognize the encryptions, how to analyze the web patterns and how to discard false connections. The participants should leave with a good knowledge about how to do an overall analysis picture of the traffic to recognize if there are malicious behaviors on it.

Attendants Requirements to Bring ready to the Workshop

- Laptop + Power cord (can be Linux, Mac or Windows)
- Minimal tools installed: wireshark, tcpdump
- Download in **advance** all the pcap files that we are going to use.
- Optional: attendees may use directly a booted Kali Linux image that has all the tools we are going to use.

Outline of the Workshop

1. Introduction (40')

- About the teacher and attendants
- Introduction to what is this workshop about.
    - It is not about tools
    - It Is about learning to analyze malware traffic and to separate it from normal traffic.
- Why to analyze network traffic?
- What can we do with this knowledge? Explain the potential of this information.
- Start of notebooks with Kali, connection to Internet.
- What is an attack? What is the difference with normal?
- What is Malware? What is a botnet?

2. How network protocols work. A baseline reminder (5')

- Current knowledge about networking?
- Network protocols, TCP/IP layers, how do they work?
- Horizontal and vertical communication
- Basic protocols. What are they for? Which ports do they use?
    - Ethernet, ARP, ICMP, IP, TCP, UDP, HTTP, DNS, SSH, SSL/TLS

3. The experience of analysing malware and normal (130')

- **Basic Tools**
    - **Wireshark**
        - Start wireshark and capture some of your traffic.
        - Identify the hosts, ports and protocols used.
        - See the different layers of protocols and encapsulations.
        - Follow a TCP stream
    - **Tcpdump**
        - Use tcpdump to see information from your network
            1. tcpdump -n -s0 -i eth0 | less
        - Use filters for tcpdump
            1. host, port

        - Use -A to see the ASCII text inside packets.
        - Read packets
            1. -r output.pcap
        - Search inside less (/)
            1. Web connections:
                - GET|POST|Host:
- **Real life traffic practice**
    - **1st Example: (30') Analysis of capture3.pcap.bz2**
        - Download from: https://mega.nz/#!MkpgjTIR!_IIOQ4ra2CGh9JkZYfhkhwCCDJWy3IPIenkrlV5AWqA
        - Uncompress it and load it in wireshark:
            1. bzip2 -d capture3.pcap.bz2
        - What can you say about it? What is going on?
        - Analysis of the behavior of the connections.
        - Malware or normal?
        - Introducing Indicators of Compromise.
        - **IP & Hostname reputation**
            1. https://www.virustotal.com/
                - Search for IPs, domains or URLs
                - See if you can infer something about the reputation of:
                    - 89.108.101.61
                    - 95.163.121.33
                    - 93.184.220.29
                    - 13.107.4.50
            2. https://sitereview.bluecoat.com/sitereview.jsp
            3. https://www.senderbase.org/
            4. https://www.passivetotal.org
        - **Web traffic analysis**
            1. What are web access logs?
            2. How to generate them?
                - urlsnarf -p capture3.pcap > capture3.weblogs
            3. Download the logs from (so you dont generate them): https://mega.nz/#!fMwWmSrJ!XvtkjBRI-7My_uAnO7ACLgW-IVvB67iFCTYnKw6mcTk
            4. What information do we see here?
            5. Look for patterns, common things, characteristics.
            6. What to do with patterns?
                - Identify 'behaviors' of the botnet.
                - Strong 'pivot' for hunting: search your network
                - Create your own rules
                - Emerging Threats pattern matching example: http://doc.emergingthreats.net/bin/view/Main/2018340
    - **2nd Example: (15') Analysis of capture2.pcap.bz2**

        - Download form: https://mega.nz/#!p4xViQ7J!wenCMFUOPGLlfk5rKNcqCNan1rojY5myHjoc0cR3KV8
        - What can you say about it? malware or normal?
    - **3rd Example: (15') Analysis of file1.small.pcap (of advanced malware)**
        - Download from https://mega.nz/#!J4oyyYTB!_L5I5IAti-d3YQ0ZT0MBnbKanB2qw3ZMh_t1qGYiL5Q
        - What can you say about it? is it malware or normal?
    - **4th Example: (15') Analysis of capture1.pcap.bz2**
        - Download from: https://mega.nz/#!xswWyYbS!HZ_6t7iGeS4yF75gwNXkfVNobTploC1Kd6A2K2S-CNU
    - **5th Example: (20') Analysis of file7.pcap.bz2**
        - Download from: https://mega.nz/#!1lwymCyQ!Dv7CvDds_mSFrnTKD7hEJTGRTBO9d_VGh7yyl7U_XN0
    - **6th Example: (30') Analysis of capture-windowst723-6.pcap.tar.gz**
        - Warning: Large file! >50Mb.
        - Download from: Link not publicly available. Contact us for it.
        - This is a large example. Your mission, if you accept it, is to discover if it is an attack or not, and what happened. You have 30mins. We expect your report.

5. Working with Large Files: Flows and Behaviors (20')

- Goal:
    - Getting an overview of the packet capture
    - Determining what is it important to see.
    - Establishing a general activity timeline
- Break down the packet capture into small pieces to analyze. (e. g. 1 day)
    - editcap -i 86400 capture-windowst723-6.pcap capture-time.pcap
- Use tshark with less filters to quickly find things.
- Use bro IDS logs.
- Use flows
    - What are flows?
- Machine Learning
    - Why machine learning?
- **Stratosphere IPS Project (https://stratosphereips.org/) (5')**
    - What is Stratosphere?
        - Behavioral-based IDS/IPS, free software.
    - Behaviors instead of IoC.
    - Flows instead of packets.
    - Using stratosphere for detection

6. Take aways

- Experience is the most important asset.
- See a lot of traffic, know what to expect. Train your eyes.
- What to look at is more important that the tools you use. Focus on the traffic.
- What is normal for you? What is normal in your network? Learn.
- Know what to expect: normal, attack, malware, misuse traffic.
- When you don't have time or too much data, flows and behaviors.
- Getting the big picture first allows you to react fast.

About Veronica Valeros

Veronica Valeros is a security researcher from Argentina. Since 2013, she has worked as a malware analyst at Cognitive Threat Analytics (CTA, a part of Cisco Systems), Prague, Czech Republic, where she specializes in malware network traffic analysis, network behavioral patterns, and threat categorization. Prior to CTA, Veronica worked independently on various projects involving data analysis, machine learning, and malware sandboxing. Veronica is also the co-founder of MatesLab hackerspace, Buenos Aires, Argentina.

Email: vvaleros@cisco.com
Twitter: @verovaleros
https://researchgate.net/profile/Valeros_Veronica

About Sebastian Garcia

"Sebastian is a malware researcher and security teacher. He did his PhD on the detection of botnets/malware by analyzing their network traffic and creating behavioral models of their actions. He likes to analyze network patterns with machine learning tools, specially on malware and botnet traffic. He is a researcher in the ATG group of Czech Technical University in Prague. He believes that free software and machine learning tools can help better protect users from abuse of their digital rights. He has been teaching in several countries and Universities and working on penetration testing for both corporations and governments. As a co-founder of the MatesLab hackspace he is a free software advocate that worked on honeypots, malware detection, distributed scanning (dnmap) keystroke dynamics, bluetooth analysis, privacy protection, intruder detection, robotics and biohacking.

In the CTU University he is managing the Stratosphere IPS project, where they are developing a free-software behavioral-based IPS.

Email: sebastian.garcia@agents.fel.cvut.cz
Twitter: @eldracote
https://www.researchgate.net/profile/Sebastian_Garcia6
http://stratosphereips.org