FORENSICS CYBER-SECURITY

MEIC, METI

2019/2020 1st Semester

2nd Exam February 3, 2020 Duration: 2h00

- Use a pen only; no extra material is allowed, such as calculator, scratch paper, etc.
- Write your answers in the free space after each question.
- The exam can be answered in Portuguese or in English.
- Identify all sheets; **unidentified pages will not be graded!**

I. (0.5 + 0.5 + 1 + 0.5 + 0.5 + 0.5 + 0.5 = 4 points)

1. A fundamental goal of digital forensics is to produce admissible evidence.

   a. Indicate the four stages of the Kruse model that help to attain admissible evidence.

   b. Tell the name of the three possible forms of offense reconstruction typically used.

2. One of the desirable properties of a high-quality forensic tool is determinism. Explain what it means for a forensic tool to be deterministic.

3. Consider the following investigative scenario:

   > TurstAngol is a company that was discovered to be involved in a major scam. After obtaining a search warrant, the police headed toward their facilities in order to gather all possible evidence. The forensic team identified four important computers operating in different conditions. A *mail server* that contained all the email of the organization. It was up and running, connected to the Internet, but you found that a logical bomb had been programmed to delete all its content in just 5min. A *backup server* that was used to store backups of the company. You don't know what data it actually contains nor how frequently the backups have been performed. The backup server contained four 1TB hard disks and it was powered off. A *file server* that was dedicated for storing the working documents of the company. This was a Linux server equipped with a single 500GB hard disk and it was running. You were told the root password, then you logged in, and discovered that that disk had been formatted in the previous day. A *Windows workstation* that belonged to the company's CEO. You moved the mouse and saw that it was password protected. Based on the LED blinkers, you could see that some disk and network activity was taking place. The system administrator can tell you the login password to that computer.

   Explain how you would proceed to deal with each of these four computers. Justify your responses:

   a. Mail server.

   b. Backup server.

   c. File server.

   d. Windows workstation.

II. (1 + 1 + 0.5 + 1 + 0.5 + 0.5 + 0.5 + 0.5 + 1 + 0.5 + 1 = 8 points)

1. Mike wants to hide a secret message inside a cover photo using a steganographic tool. The cover photo consists of a 32-bit RGB image with the resolution of 1600×1200 pixels. The tool implements a LSB-2 encoding scheme applied to each color channel. What is the maximum size of the secret message that can be hidden inside that cover photo?

2. Consider the following unallocated disk space containing the blocks of six deleted files. The empty blocks filled with zeros are indicated with the symbol "–":

–	A_0	A_1	A_2	F_0	F_1	A_3	B_1	B_0	D_1	D_0	E_0	E_1	–	E_2	C_0

The files (and respective blocks) are listed below according to their file formats:

- GIF: file A (A_0, A_1, A_2, and A_3), and file F (F_0, and F_1)
- JPEG: file B (B_0, and B_1), file C (C_0), file D (D_0, and D_1), and file E (E_0, E_1, and E_2)

The formats of these files are as follows:

- GIF: "0x47 0x49 0x46 0x38 0x37 0x61" header, "0x00 0x3B" footer
- JPEG: "0xFF 0xD8" header and "0xFF 0xD9" footer

Your job is to recover the deleted files using two file carving tools. Each tool implements a different file carving technique: single-pass structure-based and bifragment gap carving.

a. Write the blocks produced by each tool for each file (use "×" if no output is generated):

File	Single-pass structure-based carving tool	Bifragment gap carving tool
A		
B		
C		
D		

b. "File F is likely to be older than file A." Based on the disposition of blocks in the unallocated disk space, do you agree with this statement? Justify. (No justification: 0 points)

3. Provide two examples of relevant artifacts that can be obtained using the *volatility* tool.

4. Consider a volume image taken from a 8GB thumb drive (1GB = 1024×1024×1024 bytes). The hex dump shown below depicts the last bytes of the MBR retrieved from that image. Analyze the partition layout of the volume by answering the following questions.

```
0x01A0: .. .. .. .. .. .. .. .. .. .. .. .. .. .. .. ..
0x01B0: 00 00 00 00 00 00 00 00 A8 E1 A8 E1 00 00 80 00
0x01C0: 00 00 83 00 00 00 01 00 00 00 FF FF 3F 00 00 00
0x01D0: 00 00 07 00 00 00 00 00 80 00 00 00 A0 00 00 00
0x01E0: 00 00 83 00 00 00 00 00 40 00 00 00 20 00 00 00
0x01F0: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 55 AA
```

Master Boot Record Format

A	*S*	*Description*
0x000	446	Bootstrap code area
0x1BE	16	Partition table entry #1
0x1CE	16	Partition table entry #2
0x1DE	16	Partition table entry #3
0x1EE	16	Partition table entry #4
0x1FE	2	Signature (0xAA55)

Partition Table Entry Format

A	*S*	*Description*
0x0	1	Bootable flag (0x80 active, else 0x0)
0x1	3	Starting CHS address *(obsolete)*
0x4	1	Partition type (e.g., 0x0: empty, 0x07: NTFS, 0x83: Linux)
0x5	3	Ending CHS address *(obsolete)*
0x8	4	Starting LBA address
0xC	4	Size (in sectors)

A: address (hex), *S*: field size (bytes)

Note: Consider the MBR format as specified in the reference tables above. The multi-byte fields are little-endian, the LBA addressing starts in 0, and the sector size is 512 bytes.

.

a. In order to properly interpret the information contained in the MBR, is it important to consider the characteristics of the underlying storage technology, e.g., whether it is based on hard disk or solid state drive technology? Justify your answer.

b. How many partitions are there in this volume and what are their respective types?

c. A forensic analyst executed the command below. What it the purpose of this command? Be specific in your answer taking into account the parameters provided as input.

```
# dd if=disk1.dd of=part2.dd bs=512 skip=0x1 count=0x3FFFFF
```

5. Give two examples of cases where inspecting the Windows Event Log may be useful.

6. A network administrator detected a sequence of three SYN packets sent from a remote client to a local server. The server IP address is 146.193.41.153. Each SYN packet had a different destination port number, respectively: 77, 71, and 75. After this sequence, a fourth SYN packet was sent to the destination port 72, initiating a TCP/IP connection with the server on port 72. To investigate this, the network administrator executed *nmap* as follows:

```
$ sudo nmap -sS 146.193.41.153
Starting Nmap 7.70 ( https://nmap.org ) at 2020-01-29 13:48 WET
Host is up (0.00042s latency).
Not shown: 996 closed ports
PORT     STATE SERVICE
22/tcp   open  ssh
25/tcp   open  smtp
80/tcp   open  http
443/tcp  open  https
MAC Address: 0C:C4:7A:4C:EC:E0 (Super Micro Computer)
```

   a. The flag `-sS` instructs *nmap* to perform a SYN scan. What is a SYN scan?

   b. Considering the *nmap* output, suggest an explanation to the observed suspicious traffic.

7. An organization was hit by a spear phishing attack. Explain what this attack consists of and what would be the primary source of evidence you would investigate in this case.

III. (2 + 0.5 + 0.5 + 1 + 1 + 0.5 + 0.5 + 1 + 1 = 8 points)

1. For each of the following statements, indicate whether it is true (T) or false (F). Each correct answer is awarded 0.25 points; each wrong answer is penalized by subtracting 0.10 points.

   a. ____: In conventional search engines, some content cannot be found through URL traversal because some sites are protected by captchas.

   b. ____: In centralized botnet architectures, the IRC protocol is primarily used for communication between the zombies.

   c. ____: In order to hide their presence, many rootkits deploy their own versions of userland commands, replacing *ps* in order to hide the rootkits' malicious processes.

   d. ____: In the Bitcoin system, the transaction records preserved by the blockchain contain precious information about the Bitcoin account address of the transaction issuers.

   e. ____: The IMEI of an Android device can be modified by rooting the device.

   f. ____: Browser fingerprinting allows to determine the IP address of clients hidden behind a NAT by analyzing the packet size distribution of the associated HTTPS traffic.

   g. ____: In the NTFS file system, the $DATA attribute is used inside MFT entries to refer to the file's data.

   h. ____: Virtual Machine Introspection is not adequate for detecting the presence of malware on guest virtual machines.

2. DeathStar is a botnet that uses HTTP for the communication between zombies and the C&C server. It uses primarily drive-by-download methods for infecting its victims.

   a. What traffic patterns would you expect to see in the local area network of your organization if the collaborators' workstations had been infected by the DeathStar botnet?

   b. "Ensuring that each collaborator uses strong passwords helps prevent infection by this botnet." Do you agree with this statement? Justify. (No justification: 0 points)

3. Explain why kernel-level rootkits are particularly difficult to detect. Provide an example of a stealth technique implemented by this brand of rootkits.

4. Describe one strategy for establishing the relationship between the Bitcoin account address and the IP address of a specific user based on the analysis of the Bitcoin network traffic.

5. To analyze a suspect malicious binary (*log.exe*) the following command was executed:

```
$ strings -a log.exe
!This program cannot be run in DOS mode.
Rich
.text
`.rdata
@.data
L$"%
h4z@
128.91.34.188
%04d-%02d-%02d %02d:%02d:%02d %s
```

   a. This method can be classified as a static or as a dynamic analysis technique? Justify.

   b. Based on the output listed above, suggest a hypothesis about the behavior of this malware. Justify your answer.

6. A cybercrime investigator managed to control one relay node of the Tor anonymity network. Describe one possible attack that the investigator may attempt to launch in order to learn information about the Tor users that have chosen that node as the exit node.

7. Describe how a mobile device can be located based on the information collected exclusively by the cellular network infrastructure.