FORENSICS CYBER-SECURITY

MEIC, METI

2018/2019 1st Semester

2nd Exam February 4, 2019 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. (1 + 0.5 + 0.5 + 1 + 0.5 + 0.5 = 4 points)

1. A user acquires illegal goods on the Internet by visiting an online web service from his Windows desktop. Give one example of evidence that this operation might leave on the web service backend, and one example of evidence that might be found on the desktop.

2. What is the main advantage of using one-way hash functions for performing integrity checks as opposed to using checksums?

3. What are the main factors that influence the volatility of digital evidence?

4. Consider the following first responder scenario and indicate whether or not all the four guidelines for evidence admissibility have been properly safeguarded. Justify your answer.

   "RentAHouse, a real estate company, was suspect of conducting fraudulent house rental operations. The police authorities obtained a search warrant to confiscate relevant documentation from their systems. Agent Sharp, a certified forensic analyst who was part of the team, was given access to the email server and to the file server, where most critical files were located. To collect the company's emails, he logged into the mail server with root privileges and created a logical copy of all files containing the emails. The copy was saved on a forensically sound external SSD drive. Then, when he logged into the file server, he observed a running script that was deleting files from the hard disk. Immediately, he pulled the power plug and disconnected the machine. Next, he rebooted the server from a USB drive containing a Kali OS distribution and created a bitstream copy of the server's hard disk into another external SDD drive. Afterwards, he placed both external SDD drives inside evidence bags, filled out a chain of custody form containing the ID of the SSDs and his own handwritten signature, and took this material to the forensic lab for future analysis."

5. The Kruse investigation model lays out the forensic workflow according to four stages.

   a. Provide one example in which the scientific method might be useful in the assessment stage.

   b. Indicate one limitation of the Kruse investigation model.

II. (2 + 2 + 1.5 + 1.5 + 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. ____: After performing a low-level formatting of a hard disc, it is not possible to retrieve the metadata of the file systems previously located on that disc.

   b. ____: Data unit viewing is commonly used to extract individual inode data structures from ExtX file systems.

   c. ____: One of the drawbacks of Parallel Unique Path file carving algorithms is that errors tend to propagate in cascade.

   d. ____: UserAssist, Last Visited MRU, and prefetch files provide valuable sources of evidence for investigating the execution of programs in Windows systems.

   e. ____: Export flow is a technique that can be used to analyze the contents of TCP streams from NetFlow logs.

   f. ____: Email spoofing can be detected by determining a mismatch between the domain name in the "From" field and the domain name of the SMTP server indicated in the bottom-most "Received" field of the email header.

   g. ____: Evidence of XSS exploits can sometimes be obtained from the logs of the web server where the application backend runs.

   h. ____: Randomly picking IP addresses and try to connect to them on different ports is done by specialized search engines in order to find insecure NATs.

2. Your goal is to analyze a steganographic tool that allows for embedding a secret message into a 16-bit grayscale bitmap image. We know only three facts about the tool's encoding scheme. First, it uses a variant of LSB. Second, it includes a one-byte prefix before the message which indicates the message size in bytes. For example, encoding a message of 5 bytes would result in a total of 6 bytes embedded into the image; the value of the prefix would be 5 (i.e., 00001001 in binary). Third, the message can be protected by a user-provided password. The algorithm that tells which pixels contain the encoded data (i.e., both prefix and message) is:

   * the first pixel (p_0) to encode bits is given by: $p_0 =$ password *mod* 5;
   * the next pixels are found by: $p_i = p_{i-1} + 1$;
   * pixel numbering begins in 0, i.e., the index of the first pixel of an image is 0.

   To determine the LSB scheme used by the tool, we encoded a test message into a dummy image using password = 8. The test message consisted of the two-character sequence: "ZZ". The ASCII value of character "Z" is 01011010 in binary notation. The hex dump shown below lists the pixel data of the greyscale bitmap image generated by the tool. How many LSB bits per pixel are used by the tool? Justify your answer (pure guesses earn 0 points).

```
00: 2c66 e611 b3c5 c188 2cd8 1694 c515 c3bd
10: cc61 31d3 d87a ffee bed8 bb9b e888 e9ea
20: b97d 0097 a619 fa2a bb87 4c41 448c 9691
30: 8e15 018e ac1e f441 8063 2466 20b6 33b4
40: .... .... .... .... .... .... .... ....
```

3. Indicate three techniques for memory acquisition. Write one strength of each technique.

4. In a company, some illegal content (file X.avi) was found on a backup server. That file belongs to Mr. Smith, one of the employees, and it was backed up from his workstation on January 20th at 20h00. The backup service creates copies of his "My Document" files daily at the same time. While interviewing Mr. Smith, he said he was not aware of such file: on that day, he had only deleted file B.doc and edited file A.doc; by the time he left work, around 19h00, no other files existed inside folder "My Documents", he claimed. Based on the additional findings listed below, indicate whether his story might be plausible or not. Justify your answer by suggesting a timeline of events that may support your hypothesis.
   * On the backup server, the snapshot of January 19th contained file B.doc only, and the snapshot of January 20th contained files: A.doc and X.avi.
   * The file modification timestamps on the workstation's file system were: A.doc – `2019/01/20 14h21m32s765ms`, B.doc – `2019/01/20 13h46m52s232ms`, X.avi – `2019/01/19 19h00m00s000ms`.
   * File B.doc was found to have been deleted and some of its previous clusters overwritten by X.avi.
   * The workstation contained malware which is known to download illegal files from the Internet.

5. A network administrator observed abnormal traffic in the internal network. The IP address range of the network is 192.168.30.0/24. The listing below shows a fragment of the collected tcpdump trace. Suggest a hypothesis for what might be happening. Justify your answer.

```
13:21:45.012014 192.168.30.101.1090 > 192.168.30.27.80: S 92946:92946(0) win 8192
13:21:45.013095 192.168.30.101.1092 > 192.168.30.28.80: S 92932:92932(0) win 8192
13:21:45.014107 192.168.30.101.1093 > 192.168.30.29.80: S 93094:93094(0) win 8192
13:21:45.015865 192.168.30.101.1095 > 192.168.30.30.80: S 93016:93016(0) win 8192
13:21:45.016763 192.168.30.101.1096 > 192.168.30.31.80: S 93106:93106(0) win 8192
13:21:45.018001 192.168.30.101.1097 > 192.168.30.32.80: S 93076:93076(0) win 8192
13:21:45.018456 192.168.30.101.1100 > 192.168.30.33.80: S 93154:93154(0) win 8192
13:21:45.018997 192.168.30.101.1102 > 192.168.30.34.80: S 93280:93280(0) win 8192
```

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

1. A sent message M to B using onion routing. M was relayed by three nodes, R_1, R_2, and R_3, not necessarily by this order. Packet P was intercepted arriving in one of the relay nodes.

$$P = \{\mathrm{R}_1, k'\}_{\mathrm{K}^+_{R_3}} \{\{\mathrm{B}, k''\}_{\mathrm{K}^+_{R_1}} \{\{M\}_{\mathrm{K}^+_B}\}_{k''}\}_{k'}$$

   a. In which relay node was P captured: R_1, R_2, or R_3? Justify.

   b. Tell if this relay is: the entry node, the middle node, or the exit node. Justify.

2. Why are botnets very effective at launching distributed denial of service attacks?

3. A malware sample was obtained from a compromised server. To analyze its behavior, a forensic analyst ran *strace* against the malware binary. Below is a fragment of the output:

```
...
socket(PF_INET, SOCK_STREAM, IPPROTO_IP) = 3
connect(3, {sa_family=AF_INET, sin_port=htons(80),
    sin_addr=inet_addr("146.193.41.139")}, 16) = 0
write(3, "GET /~pirate/serverlist.html "..., 143) = 143
read(3, "HTTP/1.1 200 OK\r\nDate: Thu, 24 J"..., 377) = 377
read(3, "<!DOCTYPE html PUBLIC \"-//W3C//D"..., 8192) = 8192
...
```

   a. What type of malware analysis method is being used? Justify.

   b. Suggest a hypothesis about what the malware might be trying to achieve? Justify.

4. "A forensic investigator that manages to intercept the traffic between the Bitcoin wallet software running on a user's device and the nodes of the Bitcoin network is able to deanonymize the transactions issued by that user." Do you agree with this statement? Justify your answer.

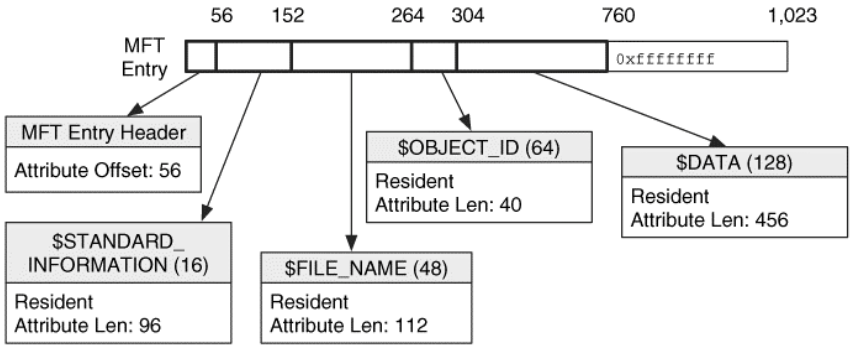
5. Ken sent an email to Bruce with the goal of blackmailing him. To avoid being tracked, Ken sent the email from a fake gmail account using his laptop. He used his smartphone as a hotspot to connect to the Internet. Ken has contract for a data plan with a mobile carrier. Bruce went to the police who retrieved the header (see below) of the offending email. Eventually, the police reached Ken and arrested him. Suggest how the police might have done this, indicating *in detail* the forensic evidence that could have been retrieved and linked with Ken.

```
Return-path: <boogeyman@gmail.com>
Delivery-date: Wed, 23 Jan 2019 18:33:27 +0000
Received: from smtp.inesc-id.pt ([146.193.32.2]) by mail.inesc-id.pt with esmtps
    id 1gmNL5-0000hJ-2d; Wed, 23 Jan 2019 18:33:27 +0000
Received: from mail-wr1-f53.google.com by mail.inesc-id.pt with ESMTP
    id x0NIg1D3012939; Wed, 23 Jan 2019 18:42:07 GMT
Received: from smtp.gmail.com by mail-wr1-f53.google.com with SMTP
    id p7so3758655wru.0; Wed, 23 Jan 2019 10:42:06 -0800 (PST)
Received: from [188.140.58.213] by smtp.gmail.com with ESMTPSA
    id x10sm122579127wrn.29.2019.01.23.10.41.59; Wed, 23 Jan 2019 10:42:00 -0800 (PST)
Date: Wed, 23 Jan 2019 18:41:57 +0000
Subject: Pay or you will suffer!
Message-ID: <yns84mpg476uips877vgcp4b.1548268917203>
From: Boogeyman <boogeyman@gmail.com>
To: bruce@inesc-id.pt
```

6. Describe the steps to repackage a legitimate Android app so as to deploy a Trojan horse in it.

7. When recovering deleted files from an Ext2 file system, what is the role – if any – of these two file system data structures: inode bitmap and block bitmap?

8. During an investigation of a Windows system, we find that the disk volume was formatted a day before it was acquired. By searching for the ASCII signature "FILE" inside the unallocated space of the current file system, we managed to retrieve one first MFT entry from the previous file system. The layout of the recovered MFT entry is depicted below:



Do you agree with the following statements? Justify your answer.

a. "The contents of the file referred to by this MFT entry can be entirely recovered."

b. "This file contains no Alternate Data Streams."

c. "This file had been deleted by the time the file system was reformatted."

d. "To get more MFT entries, carve out subsequent fixed-sized sequences of 760 bytes."