FORENSICS CYBER-SECURITY

MEIC, METI

2021/2022 1st Semester

2nd Exam February 24, 2022 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. (2 + 0.5 + 0.5 + 0.5 + 0.5 = 4 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. ____: Cybercrime has been defined by the EC-Council as “any illegal act involving a computer, its systems, or its applications”.

   b. ____: In digital forensics, preserving the determinism of digital evidence is essential to avoid the risk of the evidence being considered inadmissible in court.

   c. ____: Cookies, browsing histories, and web server logs are examples of digital evidence that an attacker may take from a digital crime scene.

   d. ____: During a case resolution process, the job of a forensic analyst ends with the production of admissible evidence.

   e. ____: Unlike internal investigations, civil and criminal investigations require the involvement of courts.

   f. ____: Regarding integrity checking methods, algorithms CRC-32 and MD5 are quite comparable with respect to the security properties they provide.

   g. ____: Dynamism of system state is an important obstacle to the collection of consistent memory dumps of running processes.

   h. ____: In regards to methods for extracting data from a storage device, bit-stream copy is a good option because it normally preserves the order of volatility while gathering evidence.

2. 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.

   “Mr. Levy Fran Velucci is now a suspect for the crimes of embezzlement against the Red Hawks Football Club. The police authorities obtained a search warrant to collect relevant evidence from the club’s facilities while Mr. Velucci was at work. Agent Sharp, the forensic analyst accompanying the police, took over the operations. First, he went for the email server. Since he did not want to disrupt the club’s email service, he created a full copy of a backup file from the backup server. This backup file had been created on the day before the search and contains a complete snapshot of all the corporate email exchanged by every employee of the club over the past five years. Second, agent Sharp went for the club’s file server infrastructure comprising four storage servers. Given the stored data was sizeable, Mr. Sharp copied only the documents owned or shared by Mr. Velucci (in total 186GB). The email backups and file server documents were double copied onto several forensically sound external SSD drives. The hashes of the copies where computed and verified. Lastly, the agent confiscated Mr. Velucci’s laptop and smartphone, who acquiesced by revealing the password and pin number, respectively, of these devices. After verifying that these credentials managed to unlock the laptop and smartphone, the agent switched off both devices and bagged them inside evidence bags. He then filled in a chain of custody form containing hashes, IDs of all the collected hardware, date and time, and his handwritten signature, and took all this material to the forensic lab for future analysis.”

3. In certain cases, it may be necessary to perform live forensics to collect digital evidence.

   a. Which of the following commands would be more useful to list the open files in a live Linux system? Justify your answer.

      (1) `netstat -nap` (2) `lsof -nDr` (3) `dd if=/dev/mem bs=2k`

   b. “Live forensics will always cause alterations in the persistent state of the analyzed device.” Do you agree with this statement? Explain your reasons.

II. (0.5 + 0.5 + 1.5 + 1 + 0.5 + 0.5 + 1 + 0.5 + 0.5 + 0.5 + 0.5 + 0.5 = 8 points)

1. Steganalysis has been broadly defined as “the art and science of detecting hidden data”.

   a. What is the main difference between *targeted steganalysis* and *blind steganalysis*?

   b. Indicate *one* technical approach to performing structural detection of hidden content inside image files.

2. Consider the following unallocated disk space containing blocks of six deleted files. Empty blocks filled with zeros are indicated with symbol “—”.

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

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

   The body of the HTML file F contains the string “GIF87a”, which corresponds to the byte sequence “0x47 0x49 0x46 0x38 0x37 0x61”. Your job is to recover the deleted files from unallocated disk space using a single-pass structure-based file carver that supports three file formats:

   - GIF: “0x47 0x49 0x46 0x38 0x37 0x61” header, “0x00 0x3B” footer
   - JPEG: “0xFF 0xD8” header and “0xFF 0xD9” footer
   - BMP: “BM” header and no footer

   Which files can the tool recover: *fully*, i.e the original file, *partially*, i.e., part of the file, or *unsuccessfully*, i.e. none of the file content? Justify your answer.

3. You were called to perform a memory dump of the Node.js server hosting the website of the popular online bookstore “Books4All”. The Node.js server consists of a process running on a Linux-powered machine maintained by “Books4All”. The suspicion is that the website might have been compromised by a code injection vulnerability and your task is to retrieve volatile state that can help diagnose this problem. You cannot disrupt the service or install new hardware, but you have remote root access to the server. Suggest a technique to collect the memory dump. Justify.

4. While analysing a forensic image of an Ext3 file system (ext3.dd), the forensic analyst executed the command indicated below. Answer the following questions:

```
# icat -f linux-ext3 ext3.dd 69457 | xxd
0000000: 510f 0100 0c00 0102 2e00 0000 00d0 0000 Q...............
0000016: 0c00 0202 2e2e 0000 520f 0100 2800 0b01 ........R...(...
0000032: 6162 6364 6566 672e 7478 7400 530f 0100 abcdefg.txt.S...
0000048: 1400 0c01 6669 6c65 2074 776f 2e64 6174 ....file two.dat
0000064: 540f 0100 1000 0702 7375 6264 6972 3100 T.......subdir1.
0000080: 550f 0100 b003 0801 5253 5455 5657 5859 U.......RSTUVWXY
0000096: 0000 0000 0000 0000 0000 0000 0000 0000 ................
[REMOVED]
```

a. What is the purpose of the *icat* tool and what is the meaning of parameter 69457?

b. Based on the returned output, make an informed guess as to whether the analyzed object corresponds to a regular file or to a directory. Justify.

5. “The best way to obtain a complete snapshot of the Windows Registry is to collect a copy of the Registry’s hive files from the file system.” Do you agree with this statement? Justify.

6. Give one example of a numerical property that is useful for statistically analyzing network traffic.

7. An attacker operating from a Russian location (IP1: 80.92.32.132) escalated privileges into a company's mail server and performed a port scan in their internal network (192.168.30.0/24). The company exposes a single public IP address (IP2: 176.31.84.249) and forwards all incoming SMTP traffic to the mail server (IP3: 192.168.30.52). The collected *tcpdump* trace is listed below:

```
13:21:45.012014 X.1090 > 192.168.30.27.80: S 92946:92946(0) win 8192
13:21:45.013095 X.1092 > 192.168.30.27.25: S 92932:92932(0) win 8192
13:21:45.014107 X.1093 > 192.168.30.27.22: S 93094:93094(0) win 8192
13:21:45.015865 X.1095 > 192.168.30.27.443: S 93016:93016(0) win 8192
13:21:45.016763 X.1096 > 192.168.30.27.110: S 93106:93106(0) win 8192
```

   a. How can we determine the geographic location of the company that was attacked?

   b. Replace the value "`X`" indicated in the trace with the correct IP address. Explain your choice.

   c. If the attacker now performs a ping sweep (e.g., using "`nmap -sn 192.168.30.0/24`"), what would you expect to observe differently in the network trace?

8. What is the typical strategy to locate a mobile device connected to a cellular network?

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

1. G&S is a medical laboratory that suffered a major cybersecurity attack resulting in massive loss of clinical data. The laboratory runs an intranet operated by their employees for management purposes (e.g., recording patients' analysis). The intranet's URL is `https://intranet.gs.pt/`. Below you find a fragment of the intranet's Apache server log around the time the attack was carried out.

```
192.168.1.34 - - [14/Feb/2022:11:58:25 +0100]
    "GET /home.php?page=index&code=pt HTTP/1.1" 200 36312
192.168.1.36 - - [14/Feb/2022:11:58:26 +0100]
    "GET /patient.php?req=40350032663&code=pt HTTP/1.1" 200 27140
192.168.1.34 - - [14/Feb/2022:11:58:31 +0100]
    "GET /home.php?page=index&code=pt HTTP/1.1" 200 30745
192.168.1.36 - - [14/Feb/2022:11:58:29 +0100]
    "GET /home.php?page=news&section=13 HTTP/1.1" 200 36312
192.168.1.36 - - [14/Feb/2022:11:58:42 +0100]
    "GET /patient.php?req=0;debugResetDB() HTTP/1.1" 200 2219
192.168.1.34 - - [14/Feb/2022:11:59:29 +0100]
    "GET /home.php?page=news&section=9 HTTP/1.1" 200 73141
```

   a. Can you formulate a hypothesis for what may have caused the loss of data?

   b. The IP addresses 192.168.1.34 and 192.168.1.36 are associated with the workstations of two employees – Ana and Bruno – respectively. Assuming the attack was originally mounted through a phishing attack, what sort of evidence would you expect to find in their emails?

2. In conventional search engines, some content cannot be found through URL traversal. Can you give two reasons for why such content cannot be found?

3. Indicate which anonymity properties can an HTTPS proxy give to a client (Alice) who's using it to anonymously access `https://buyweapons.ru`. Answer this question in these four scenarios:

   A - the forensic analyst can intercept the traffic between Alice's computer and the proxy;

   B - the forensic analyst can intercept the traffic between the proxy and the website;

   C - the forensic analyst can entirely control the HTTPS proxy's internal state;

   D - the forensic analyst planted a kernel-level rootkit on Alice's computer controlling the entire system.

4. Botnets are powerful tools for cybercrime.

   a. Give two examples of large-scale attacks that are commonly mounted through botnets.

   b. In a centralized botnet architecture, what is the role of the command and control server(s)?

   c. A recent trend that aims for securing the command and control servers it to deploy the servers behind Tor Onion Services. What might be the reason for this?

5. To analyze a piece of malware, a forensic analyst performed the following sequence of operations:

   (a) ran the malware through a virus scan,
   (b) listed strings inside the binary,
   (c) dumped the binary's symbol information,
   (d) inspected the malware using a debugger.

   Indicate which of these steps correspond to static or dynamic analysis. Justify your answer.

6. Bitcoin accounts are linked with a public-key pair. For forensic analysis purposes, what is the value of knowing: 1) the association between public key and real user identity, and 2) the private key?

7. Regarding mobile forensics, consider the following statements and indicate whether you agree or disagree with them, justifying your answers:

   a. "Most of the data stored on a mobile device is located on internal data storage, which consists of volatile random-access memory technology."

   b. "Safeguarding network isolation while seizing a mobile device is very important, but faraday bags have several limitations for this job."