COM 402 Final Exam, 13.01.2020

Name: **Anonymous**

Sciper: **123456**

Please wait for instructions before opening this document

- This is a **closed book** exam. Books, notes and electronic devices are not allowed.

Multiple choice questions:

- There are 10 multiple choice questions, counting 1 point each.
- Only one answer is correct, *there is a 0.25pt penalty for wrong answers*
- Make a mark *inside* the box corresponding to your answer
- Use a black or blue pen to mark your answers. Pencils are not allowed.
- Use a white-out fluid or tape if you ticked the wrong answer.
- If you white-out a wrong answer, do not try to re-draw the boxes.

Open text questions:

- There are 19 open text questions, counting 1 point each.
- Please write your answers in the corresponding text boxes.
- Do not write more than three lines. Any text outside of the boxes or after three lines **will be ignored**.
- Do not tick the '0', '0.5', '1' boxes of the top of the text boxes.

Questions

- The supervisors will not answer any questions regarding the content of the exam questions

Question 1 Which of the following threats is an example of a *commodity threat*:

- □ Exploiting an SQL injection on an e-banking web site.
- □ Exploiting a buffer overflow of the `ping` command to become administrator of the local machine.
- □ Hacking web servers to display political messages.
- □ Mass-mailing e-mails falsely pretending to have hacked the recipient and asking for money.

Question 2 In a mandatory access control (MAC) setting aimed at protecting the confidentiality of data, which of the following statements is true?

- □ Subjects can only write documents that have the same or lower level than them.
- □ Subject can only read documents that have the same or higher level than them.
- □ Subjects can not read and write objects of the same level.
- □ Subject can not write to objects that have lower levels than them.

Question 3 To enable recovery of secure backups, a storage provider uses a Hardware Security Module (HSM) to protect encryption keys. This is to ensure that only the data owner can recover her backed up data. In this scenario:

- □ The data owner needs to know the private key of the HSM.
- □ The data owner needs to know the public key of the HSM.
- □ The HSM needs to know the signing key of the data owner.
- □ The HSM encrypts the data on the data owner's device.

Question 4 The goal of a privacy evaluation is to understand whether a mechanism provides a given privacy property. When doing this evaluation:

- □ It is important to model the privacy mechanism as a deterministic transformation between the input and the output.
- □ The prior knowledge of the adversary is irrelevant.
- □ You should assume the adversary does not know the internals of the mechanism.
- □ You should assume the adversary knows the internals of the mechanism.

Question 5 You want to protect the information stored in the database of your application using AES for encryption. You choose to do the encryption in the application (the logic tier). What is an important limitation of this solution?

- □ Data at rest is not encrypted.
- □ Data in motion is not encrypted.
- □ Data can still be accessed by the database administrators.
- □ The database cannot sort or aggregate data.

Question 6 A humanitarian aid organization is providing food aid to beneficiaries. Distributing resources last minute is costly, and the organization would like to use machine learning in order to predict when an area will enter into a crisis based on historical data. This way, the organization can ensure a better allocation of resources and even send teams in advance to ensure a timely response. The company behind NextCrisis claims that the tool offers amazing performance: if a crisis indeed happens, NextCrisis always predicts it correctly, and it only predicts a crisis when it does not happen 5% of the time. You know that crises are a globally rare phenomenon, with only 3 in every 1,000 regions needing resources at a given time. What is the probability that NextCrisis is correct when it predicts that a region will be in crisis soon?
Hint: Remember that Bayes is your friend

□ 0.123

□ 0.057

□ 0.5

□ 0.001

Question 7 We consider a blockchain system using a proof-of-work mechanism to impose a rate of new blocks per minute, as implemented for the homework. The proof-of-work mechanism works as follows:
Let h be the SHA-256 hash of a candidate block as a bit-string of length 256, and
$\texttt{int}(h)$ be the base-10 integer representation of h,
$\texttt{zeros}(h)$ be the number of leading zero bits of h.
A node accepts candidate blocks for which the following *condition* holds: $\texttt{int}(h) < 2^{233}$.

On a system of N nodes, the described mechanism results in a rate of R new blocks per minute. If the number of nodes increases to $2N$, which of the following *conditions* would keep the rate at R new blocks per minute ?
Note: Assume that all nodes are equally powerful.

□ $\texttt{zeros}(h) >= 24$.

□ $\texttt{int}(h) < 2^{233}$.

□ $\texttt{int}(h) < 2^{234}$.

□ $\texttt{zeros}(h) < 24$.

Question 8 Regarding Attribute Based Credentials (ABCs), which of the following statements is *wrong*?

□ ABCs are a critical component for homomorphic encryption.

□ Given a transcript of an already-used credential, the verifier cannot identify a user, even if the verifier colludes with the issuer.

□ Attributes can be encoded as numbers.

□ They enable selective disclosure of attributes.

Question 9 Regarding genetic privacy and security, among the following statements, which one is *wrong*?

□ De-identification of personalized-health data sets provides full anonymization.

□ Pharmaceutical companies usually make use of de-identified data.

□ Sequencing the DNA of an individual without consent is illegal in most jurisdictions.

□ Homer's re-identification attack requires notably the knowledge of the genome of the victim (her set of variants).

Question 10 You used to have a pair of keys (private and public) to exchange encrypted and signed messages with your friends. You lost your private key when your hard disk crashed. Which of the following operation is no longer possible:

□ Sending encrypted messages to your friends.

□ Your friends encrypting and sending messages to you.

□ Verifying the signature of messages received from your friends.

□ Signing messages before sending them to your friends.

Software Security

Question 11 The two series of instructions marked with arrows have been added automatically to the code of the function `say_hello` by the compiler. Explain briefly the purpose of the first block (instructions 1 to 3) and the second block (4 to 7):

```
          push   %rbp
          mov    %rsp,%rbp
1. ->     sub    $0x50,%rsp
2. ->     mov    %fs:0x28,%rax
3. ->     mov    %rax,-0x8(%rbp)

         [original function code]

4. ->    mov    -0x8(%rbp),%rcx
5. ->    xor    %fs:0x28,%rcx
6. ->    je     7e7 <say_hello+0x6d>
7. ->    callq  630 <__stack_chk_fail@plt>
         leaveq
         retq
```

□0 □0.5 □1

....

....

....

Secure communications

Question 12 There are two ways to get an ephemeral key in TLS: key agreement using Diffie-Hellman and key transport using RSA. Explain a problem that might arise if key transport is used instead of Diffie-Hellman.

□0 □0.5 □1

....

....

....

Question 13 Some attacks in HTTPS have been fixed by the introduction of the HTTP Strict Transport Security (HSTS) header. Describe one of these attacks: □0 □0.5 □1

....

....

....

Kerberos

Question 14 When a user uses the Kerberos protocol to access a service on a Windows server, the user has to encrypt three different pieces of data with three different keys. List the three keys, and for each key, say where it comes from (which entity generated it), what is encrypted with it and where the encrypted data is sent:

□0 □0.5 □1

....

....

....

Database Security

Question 15 Most databases support fine-grained access control to tables, columns and rows. For example, access to columns of tables can be controlled with a command like

```
grant SELECT,UPDATE,INSERT (name, grade, year) ON com402.students to
Alice;
```

Describe a mechanism that can be used to grant Bob read access to the students of year 2019 only:

Note: *We expect an explanation as the answer, not a SQL statement.* □0 □0.5 □1

....

....

....

Passwords

Question 16 What is the goal of memory-hard password hashing functions? How can they achieve this goal?

□0 □0.5 □1

....

....

....

Firewalls

Question 17 A Web Application Firewall (WAF) can protect a web application against certain types of attacks. Name one type of attack against which WAFs are efficient and describe how the WAF protects the application against this type of attack.

□0 □0.5 □1

....

....

....

Network sniffing

In most operating systems, only administrators have the privilege to sniff network traffic. In recent Linux systems, users that are members of the `Wireshark` group automatically obtain the right to sniff traffic with Wireshark.

Question 18 Name which authorization mechanism is used in this case and explain the configuration needed.

Hint: The mechanism used is neither setuid nor setguid.

□0 □0.5 □1

....

....

....

Question 19 Describe an attack that can be prevented if this mechanism is used instead of setuid or setgid.

□0 □0.5 □1

....

....

....

e-Voting

Question 20 In some e-voting protocols, encrypted votes are mixed before being decrypted, in order to obtain anonymity of the votes. Explain how homomorphism with regard to multiplication can be used to provide anonymity in that step.

□0 □0.5 □1

....

....

....

Privacy properties

Agree or disagree with the following statements and justify your answer.

Question 21 Consider a scenario where a user makes use of an anonymous communication network (*e.g.,* Tor + the Tor Browser) to connect to some website. Over several days, this client visits multiple times the same website.

For the administrator of the website, the multiple visits are unlinkable between each others.

□0 □0.5 □1

....

....

....

Question 22 Consider a location privacy mechanism that, instead of the real location, outputs a fake location placed on a circumference of radius 100m centered on the real location. This mechanism provides plausible deniability regarding the real location of the user towards anyone seeing the obfuscated location.

□0 □0.5 □1

....

....

....

Trusted computing

Question 23 You are setting up your startup company and you decide to secure your accounting information by using a Trusted Platform Module (TPM)'s sealing capabilities to ensure that the information is confidential when your office computer shuts down overnight. Your employee Jamie has used some company money to buy himself a new iPhone. To hide his misbehavior, Jamie hacks into your computer and changes the booting program to disable the password that protects the accounting file. Will the TPM unseal the data and allow Jamie to make modifications? Justify your answer.

□0 □0.5 □1

....

....

....

Machine learning

Question 24 Due to the large number of incoming students, the Hogwarts school of magic decided to automate the sorting hat procedure (for a given student, the sorting hat decides into which one out of four houses the student will be assigned). They are using an online deep-learning tool now: a student uploads their picture on the new Hogwarts website, and the model decides to which school the student will be assigned. There was a leak and you learned the architecture of the model, and you also get access to a webpage listing the pictures of all current and past Hogwarts students for each of the four houses. How would you maximize your chances of getting into your favorite house, Slytherin? Justify your answer.
[*Note: the automated sorting will check for duplicates, i.e., it will not sort a person twice.*]
[*Hint: Think about how you can use the pictures of the other students*]

□0 □0.5 □1

....

....

....

Question 25 A bank makes public an API for users to consult the likelihood that they get a loan at 10 CHF per query. Initially, the bank trains a Neural Network with 5000 parameters and opens the API for the users. Three months later, the bank discovers that a linear model with 1000 parameters is better in terms of interpretability and is much cheaper, so they open a new API for this model. The engineer releasing the new model forgets to shut down the Neural Network API. You decide that it is a great idea to steal *one of the* model and open your own API, charging 5 CHF per user query. Explain which model you would steal, and how the attack would work. Justify your choice.

□0 □0.5 □1

....

....

....

Privacy-preserving cryptography

Question 26 Consider the following cryptosystem:
Let p and q be two primes defining a modulus $n = p \cdot q$. Consider the plaintext space $\mathcal{P}$ and cipher space $\mathcal{C}$ such that $\mathcal{P} = \mathcal{C} = \mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$ the group of integers modulo n with order $\phi(n) = (p-1)(q-1)$.
Let (a, b) be a private/public key pair such that $a \cdot b = 1 \mod \phi(n)$.
For $\mathcal{K} = (n, p, q, a, b)$:

- Encrypt(x): $e_{\mathcal{K}}(x) = x^b \mod n$
- Decrypt(y, a): $d_{\mathcal{K}}(y) = y^a \mod n$

Is this scheme homomorphic with respect to additions and/or multiplications? If yes, provide a mathematical argument.

□0 □0.5 □1

Blockchains

Question 27 Explain the difference between Bitcoin and generic smart contracts in terms of the state transition.

□0 □0.5 □1

Privacy and health

Question 28 What are the privacy risks when an individual willingly shares his/her sequenced genome with direct-to-consumer companies such as 23andme.org, beyond the risks for themselves?

□0 □0.5 □1