Corrected

COM 402 Final Exam

Jan 17th, 2022

1
Name: **Anonymous**
Sciper: **123123**

Do NOT open this document until instructed to do so.

Instructions

- This is a **closed book** exam. Books, notes and electronic devices are not allowed, except for up to two A4 pages of notes (this amounts for *one* double-sided A4 sheet or equivalent).
- **No question will be answered during the exam**, except possibly questions related to the understanding of the English language.
- **Once finished leave this document on the table.**
- You can **use the blank pages at the end as scratch paper**. It will not be accounted in the grading.

Multiple choice questions:

- There are multiple choice questions, counting 1 point each.
- Only one answer per question is correct.
- Ticking the correct answer counts for 1 point.
- Ticking an incorrect answer, more than one answer, or no answer counts for 0 point.
- Make a mark **inside** the box corresponding to your answer.
- Use a black or blue pen to mark your final answers. **Pencils are not allowed**.
- If you ticked the wrong answer, use a white-out fluid or tape to erase the box completely. **Do not try to re-draw the box**.

Open text questions:

- There are 13 open text questions, counting 1 point each.
- Please write your answers in the corresponding text boxes. Any text outside of the boxes **will be ignored**.
- **Pencils are not allowed**.
- Do not tick the '0', '0.5', '1' boxes of the top of the text boxes.

Question 1 Which of the following is *false* about a HMAC used for cookies?

- ☐ It uses a shared secret between the client and the server.
- ☐ It does not encrypt the cookies.
- ☐ It provides both integrity and authenticity.
- ■ It provides only integrity.

Question 2 Which of the following is *true* for rainbow tables?

- ☐ The reduction function takes as input a password and returns a hash from the set of possible outputs of the employed hash function.
- ☐ In order to save memory, only the last element of the chain is stored.
- ■ The reduction function takes as input a hash and returns a password from the specific set the rainbow table is tailored to.
- ☐ In order to save memory, only the first element of the chain is stored.

Question 3 Which of the following is *true*?

- ☐ Windows privileges can never act as capabilities.
- ■ Setuid lets simple users execute some well defined privileged actions.
- ☐ Unix has more access rights than Windows.
- ☐ Windows' *admin* is the highest privilege access possible.

Question 4 Which of the following is *true* in the context of a buffer overflow?

- ☐ A canary is a constant value pushed on the stack.
- ☐ Return-oriented programming is used to execute a piece of code on the stack.
- ■ If a canary is placed before the return address, an attacker can still overwrite local variables.
- ☐ Most compilers set the stack to 'non-writable' as a defense mechanism.

Question 5 What is *false* about anonymous credentials?

- ☐ The verifier cannot distinguish two users with the same attributes.
- ☐ A user can disclose only relevant attributes to the verifier.
- ■ A user can create valid credentials for its attributes.
- ☐ The verifier cannot link two showings of the same credential.

Question 6 Consider an attacker performing a gray-box attack against a machine learning-as-a-service (MLaaS) that hosts a non-linear machine-learning model, e.g., a neural network, and outputs the prediction in clear. Which one of the following is *true*?

- ☐ Model stealing will take $d+1$ queries for an input size (number of features) of d.
- ☐ If the model has been trained in a federated learning setting, attribute inference attacks are not possible.
- ■ MLaaS is prone to model stealing, membership inference, and attribute inference attacks.
- ☐ If the model is kept encrypted, model stealing attacks are not possible.

Question 7 Which of the following is *false* about TLS?

- ■ Both client and server use an asymmetric cipher to encrypt the data.
- ☐ The client can optionally be authenticated by a client certificate.
- ☐ Both client and server use a HMAC to guarantee integrity.
- ☐ The server is authenticated with a certificate.

Question 8 Which one of the following is *false* about remote access protection?

- ☐ A VPN can be based on TLS.
- ■ A VPN can be used to protect privacy as all connection locations are always hidden by design.
- ☐ The destination only sees the VPN gateway address.
- ☐ A VPN uses encryption and encapsulation for confidentiality.

Question 9 Consider a blockchain system that uses a proof-of-work mechanism to control the rate of new blocks per minute. The proof-of-work mechanism works as follows:

Let h be the SHA-256 hash of a candidate block, seen as a bit-string of length 256, and
$\texttt{zeros}(h)$ be the number of leading zero bits of h,
$\texttt{int}(h)$ be the base-10 unsigned integer representation of h.
A node accepts candidate blocks for which the following *condition* holds: $\texttt{int}(h) < 2^{233}$.

On a system of N equally powerful nodes (i.e., each node computes the same number of SHA-256 operations per second), the described mechanism results in a rate of R new blocks per minute. Would the number of nodes increase to $2N$, which of the following *conditions* keeps the rate at R new blocks per minute ?

- ☐ $\texttt{int}(h) < 2^{234}$.
- ■ $\texttt{zeros}(h) >= 24$.
- ☐ $\texttt{int}(h) < 2^{233}$.
- ☐ $\texttt{zeros}(h) < 24$.

Question 10 Let p, q be two primes as in the Paillier cryptosystem (i.e., $\gcd(pq, (p-1)(q-1)) = 1$) and $n = pq$. In addition, let Enc be Paillier encryption, $\phi(n) = (p-1)(q-1)$ and m, m' be two messages. Which of the following statements is *true*?

- ☐ The public-key is (p, q).
- ☐ $\mathsf{Enc}(m) + \mathsf{Enc}(m') = \mathsf{Enc}(m + m')$
- ☐ $r^{n\phi(n)} \bmod n^2 = 1 + r \bmod n^2$, for all $r \in \mathbb{Z}_n^*$
- ■ $(1+n)^m r^n \bmod n^2 = (1+mn)r^n \bmod n^2$, for all $r \in \mathbb{Z}_n^*$

Question 11 Which of the following is *true* about the one-time pad cipher?

- ■ The key must be at least as long as the plaintext.
- ☐ The one-time pad is an asymmetric cryptographic algorithm.
- ☐ A key can be used for more than once.
- ☐ The ciphertext produced by a one-time pad is not malleable.

Question 12 Which of the following is *true* about technologies for privacy and security protection?

- ☐ Homomorphic encryption is the most efficient and versatile technology.
- ■ Trusted execution environments are prone to side-channel attacks.
- ☐ Secure multiparty computation minimizes the communication cost.
- ☐ Distributed ledger technologies provide the highest data privacy.

Question 13 Which of the following is NOT protected against in information security?

- ☐ Loss of availability.
- ■ Theft of the computer hardware.
- ☐ Loss of integrity.
- ☐ Loss of confidentiality.

Question 14 Which of the following is *true* about a stack buffer overflow:

- ☐ A stack buffer overflow always crashes the program.
- ■ Randomizing the memory space makes it harder to determine where the executable code that can be used resides.
- ☐ A stack canary sets a limit in the memory space and prevents the execution beyond.
- ☐ Non-executable stack sets the whole memory to non-executable.

Question 15 Assume that an attacker has the following snapshots of two databases where $(\cdot)$ and $[\cdot]$ denotes that a value is encrypted via *probabilistic* and *deterministic* encryption, respectively. Which of the following is *true*?

Database 1

Zipcode	Name	Age
1004	(John)	(41)
1004	(Alice)	(45)
1005	(Bob)	(45)
1006	(Mary)	(45)
1007	(Louise)	(45)

Database 2

Zipcode	Cancer Status	Age
1004	[yes]	[41]
1004	[no]	[45]
1005	[no]	[45]
1006	[no]	[45]
1007	[no]	[45]

- ☐ The attacker can infer that John has cancer.
- ■ The attacker can infer that the people with the same age share the same cancer status.
- ☐ The attacker can infer that the person's age with zipcode 1004 is 41 with 50% probability.
- ☐ The attacker can infer that John's Zipcode is 1004.

Question 16 Although caches are ubiquitous in modern CPU architectures, they introduce a serious side-channel attack vector such as...

- ☐ Power Analysis Attacks.
- ☐ Electromagnetic Attacks.
- ☐ Sound-Based Attacks.
- ■ Timing-Based Attacks.

Question 17 In order to thwart power analysis attacks, implementations of cryptographic algorithms seek to make the measurements as noisy as possible. What is a sensible way to achieve this?

- ☐ Making sure the algorithm is always executed concurrently with other programs.
- ■ Adding randomness into the computation.
- ☐ Executing the algorithm multiple times.
- ☐ Encrypting any output with a separate key.

Blockchain

Question 18 Explain the differences between proof-of-work (POW) and proof-of-stake (POS) with their advantages and limitations.

☐0 ☐0.5 ■1

POW: need to solve a complex cryptographic puzzle: + anyone can join - waste of energy. POS: assign consensus shares + address energy waste problem - need public randomness

Question 19 Explain the countermeasures used to prevent double spending attacks.

☐0 ☐0.5 ■1

Use of a proof-of-work (hash power) makes it hard to perform as one needs more than 50% of the network to cheat and additionally merchants can simply wait for confirmations (at least 6 as a rule of thumb)

Differential Privacy

Question 20 Consider a vaccination database curated by the Department of Health of a country. This database consists of a single table named 'vaccination_status' with two columns: the unique user ID (user_id) comprising alphanumeric values and the number of vaccine doses received by this user (num_doses) comprising positive integer values. The department relies on the Laplacian mechanism to enable a differentially private query-interface. An analyst has been allotted with a total privacy budget of B to query the vaccination status database.
Every day, the analyst performs two queries:

1. Q1: "how many people have been vaccinated with at least one dose" (i.e., `SELECT count(user_id) FROM vaccination_status WHERE num_doses > 0;`). This query uses a budget ε_1.
2. Q2: "how many doses have been delivered" (i.e., `SELECT sum(num_doses) FROM vaccination_status;`). This query uses a budget ε_2.

(a) Give and explain what is the sensitivity of Q1 and the parameter of the distribution that the mechanism needs to sample for Q2.
(b) What is the amount of privacy budget remaining after k days?
Answer and explain both (a) and (b) using the distribution $Lap(\cdot)$, B, k, user_id, num_doses, ε_1, and ε_2.

☐0 ☐0.5 ■1

(a.1) the sensitivity of Q1 is $\Delta_1 = 1$.
(a.2) the noise required for Q2 is sampled from $Lap(max(\text{num\_doses})/\varepsilon_2)$.
(b)$B - k \cdot \varepsilon_1 - k \cdot \varepsilon_2$

Side-Channels & Rowhammer

Question 21 The first operation of the AES encryption algorithm involves the XORing of the secret key bytes with the plaintext bytes, i.e., $x_i \oplus k_i$, for $0 \leq i \leq 15$ and $x_i, k_i \in \{0, 1, \ldots, 255\}$. A side-channel adversary has access to an unprotected AES software implementation whose power consumption is freely measurable. Let $P(x_i \oplus k_i)$ be the power consumed by the XORing of the i-th key and plaintext byte. For simplicity's sake, let us assume that $P(x_i \oplus k_i) > P(x'_i \oplus k_i)$ if and only if $\mathrm{HW}(x_i \oplus k_i) > \mathrm{HW}(x'_i \oplus k_i)$ where HW(z) is the Hamming weight of the byte z, i.e., number of bits set to 1. Analogously, $P(x_i \oplus k_i) < P(x'_i \oplus k_i)$ if and only if $\mathrm{HW}(x_i \oplus k_i) < \mathrm{HW}(x'_i \oplus k_i)$. **Using exactly 256 encryptions/measurements of $P(x_i \oplus k_i)$ with varying x_i give an algorithm that uniquely recovers the i-th key byte k_i.**

☐0 ☐0.5 ■1

Note that the Hamming weight of 0 is unique in the set of bytes $\{0, \ldots, 255\}$ hence we iterate over all values and compare their Hamming weights until we find the plaintext $x_i = k_i$.

```
1 :  x_i ← a
2 :  p_ref ← P(x_i ⊕ k_i)
3 :  for b ∈ {0, 1, ..., 255} \ {a}
4 :      x'_i ← b
5 :      p_curr ← P(x'_i ⊕ k_i)
6 :      if p_curr < p_ref
7 :          x_i ← b
8 :          p_ref ← p_curr
9 :  return x_i
```

Question 22 Consider the following simple block cipher for an initial plaintext x and key k:

```
1 :  for i ← 1 to N
2 :      x ← KeyAddition(x, k)
3 :      x ← Substitution(x)
4 :      x ← Permutation(x)
```

The operations KeyAddition and Permutation are linear operations in x and k whereas Substitution is non-linear. Using the Rowhammer exploit, an adversary managed to divert the program flow in such a way that Substitution is skipped in each round and thus not executed. This is akin to Substitution being the identity function. Explain, why it is now possible to recover k by knowing a plaintext x and its corresponding ciphertext.

Hint. A superficial argument is sufficient for this question.

☐0 ☐0.5 ■1

The entire algorithm has been rendered completely linear. Hence, an adversary can construct a system of linear equations with the key bits as unknowns and solve it using his favorite equation solver, e.g., Gaussian elimination.

Machine Learning

Question 23 Consider a federated learning setting with several hospitals where the data of the hospitals remain in their own trusted servers, but a global neural network model is trained collaboratively using *multiparty homomorphic encryption*. As such, the model and any intermediate value exchanged among the servers remain encrypted throughout the process and no party alone can decrypt the model.
State two limitations of using multiparty homomorphic encryption in this setting. Assuming that the servers are malicious, explain one possible attack that can be achieved in the training phase.

☐0 ☐0.5 ■1

Two limitations: computational cost, scalability in terms of complexity of models, limits the computation of non-linear functions, the servers cannot trace the model training (whole process is under encryption). Possible attacks: the servers can poison the model, install backdoors, adversarial examples, etc.

Question 24 Consider a server that provides machine learning-as-a-service (MLaaS) with a pre-trained model. A client is able to query the server to obtain a prediction result on some evaluation data. The server and the client rely on secure multiparty computation to protect confidentiality: the server does not see the client's data nor the end result, and the client only receives the prediction result. What are the possible threats posed by (i) the client and (ii) the server?

☐0 ☐0.5 ■1

(i) The client can perform membership inference attacks, or steal the machine learning model. (ii) The server can give wrong end-results.

SQL Injections

Question 25 Your website lets users authenticate themselves through a password, before handing them private data. In order to do this, the server makes the following SQL query to the database.

```
SELECT user_data FROM database WHERE password LIKE '$pwd';
```

where `$pwd` is the variable containing the password input by the user, without any type of quotes (e.g., ', `, ") to prevent simple SQL injections.

(a) Assume the database contains the data of a single user, who has a password of at most 9 alphabetical ([a-z,A-Z]) characters. Design an attack that fully recovers the password with less than 500 attempts. Explain.

(b) Assume the SQL query is correctly sanitized (i.e., no injection is possible). Give two reasons why authentication should never be handled like that.

☐0 ☐0.5 ■1

(a) % is not removed, one can bruteforce each character at a time (try pwd 'a%', 'b%', etc.). (b) 1. two users sharing the same pwd would get data from the other 2. can try a pwd against all user passwords 3. no hashing

Memory Vulnerabilities

Question 26 Common defenses against memory vulnerabilities are non-executable memory and ASLR. Please describe (1) how an attacker could bypass the non-exec memory. (2) how an attacker could bypass the ASLR.

☐0 ☐0.5 ■1

(1) Non-exec memory: Instead of writing your code on the stack, search for pieces of code (gadgets) in the program that end with a return instruction. Put the addresses of these gadgets on the stack. The gadgets will be executed in sequence. (2) ASLR: ASLR typically works by shifting addresses by a constant value. If we are lucky, the program will leak the address of a function or a variable. From this information, we can calculate other addresses

Access Control

Question 27 In mandatory access control: (1) What is the goal of no *write-down*? (2) What is the goal of no *write-up*?

☐0 ☐0.5 ■1

(1) No write-down: When protecting confidentiality, we don't want users to write to a lower level. (2) No write-up: We don't want users from lower level to write into higher levels to protect integrity.

Networking

Question 28 (a) Explain what is a DMZ and give an example of a threat that is mitigated by a DMZ.
(b) Explain the difference between DMZ and Zero Trust network. Give an advantage of each one over the other.

☐0 ☐0.5 ■1

(a) DMZ separates network into zones. Compromising internet facing server (email, VPN) does not compromise sensitive internal service (e.g., finance) (b) DMZ creates zones (i.e., firewalls between zones). Zero trust network treats each machine as a threat (i.e., firewall on every machine). DMZ simpler (no need to configure each machine). ZTN potentially more secure as security policy enforced on each machine.

Question 29 The version 1.2 of the Transport Layer Security (TLS) protocol is standardized by RFC 5246. Below is an extract of the RFC:

```
7.4.7.  Client Key Exchange Message
    When this message will be sent:
        This message is always sent by the client.  It MUST immediately
        follow the client certificate message, if it is sent.  Otherwise, it
        MUST be the first message sent by the client after it receives the
        ServerHelloDone message.
    Meaning of this message:
        With this message, the premaster secret is set, either by direct
        transmission of the RSA-encrypted secret or by the transmission of
        Diffie-Hellman parameters that will allow each side to agree upon the
        same premaster secret. [...]
    [...]
```

(a) explain the difference between the two possible ways used to set the `premaster secret`, (b) indicate why one of the two methods (which one?) should be preferred and (c) describe a possible attack against the method indicated in (b).

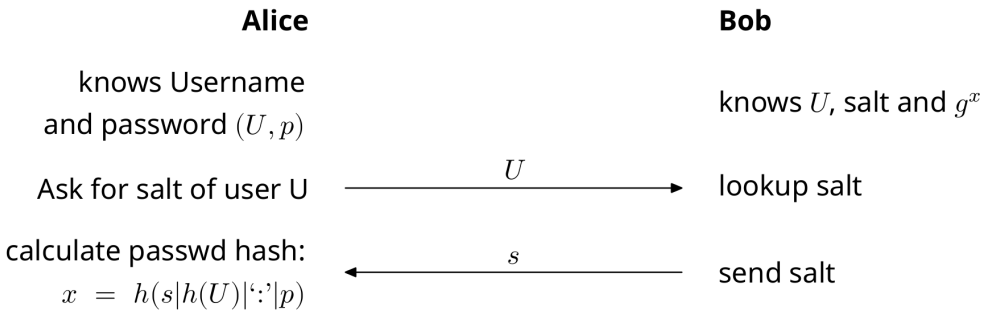
☐0 ☐0.5 ■1

(a) DH key exchange vs. client samples random symmetric key and send to server encrypted with RSA public key of the server. (b) DH key exchange because of (P)FS. . . . . (c) MitM, quantum, can accept also attacks from "Imperfect Forward Secrecy: How Diffie-Hellman Fails in Practice" if some student read the paper.
. . . . . .

. . . . . .

Authentication

Question 30 In Homework 2, you implemented the Secure Remote Password protocol (SRP), a Password Authenticated Key Exchange (PAKE). The first steps of SRP are shown below.



Briefly answer the following questions.

1. List the **two** goals of a PAKE.
2. Discuss a possible attack carried out by an eavesdropper Eve against the SRP protocol.

☐0 ☐0.5 ■1

(a) verify the password of a remote party and exchange a key (b) get salt and username to bruteforce x. .....

.....

.....

Corrected

Corrected

Corrected

Corrected

Corrected