CS2107 —Introduction to Information Security
School of Computing
National University of Singapore

Midterm Test

2025/2026 Semester 1

9 October 2025 **Time allowed:** 1 hour

Instructions (please read carefully):

1. This is a **CLOSED** book assessment, but you are allowed to bring **ONE** double-sided A4 sheet of notes for this assessment.
2. The assessment paper contains **SEVENTEEN (17) questions** and comprises **TEN (10) pages** including this cover page.
3. The time allowed is **1 hour**.
4. Use of calculators is allowed in the test.
5. You are allowed to use pencils, ball-pens or fountain pens. No red color.
6. **Marks may be deducted** for unrecognizable handwriting. All corrections should be cleanly erased or covered with correction fluid or tape.

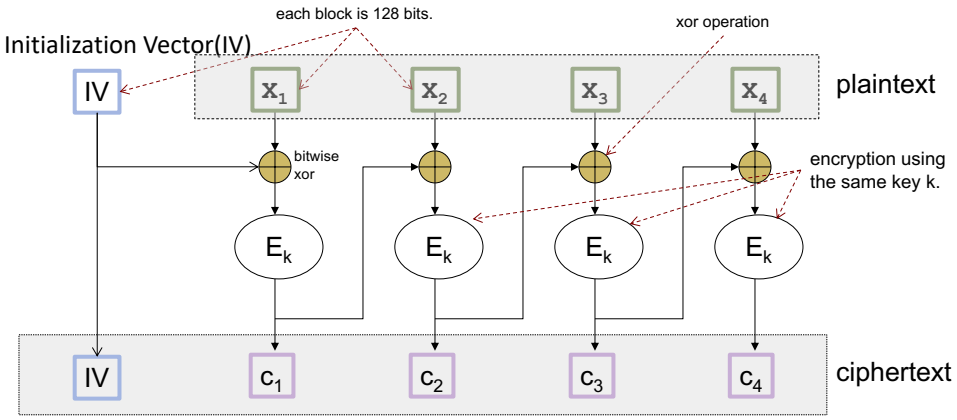
Instructions for ANSWER SHEET:

1. Write down your **student number** in the answer sheet and shade the corresponding circles with dark ink or pencil.
2. DO NOT WRITE YOUR NAME!
3. The answer sheet comprises 2 pages.
4. You must submit only the **ANSWER SHEET** and no other documents. The question set may be used as scratch paper.
5. All answers must be written within the corresponding box provided. **Anything written outside the answer box will not be accepted.**

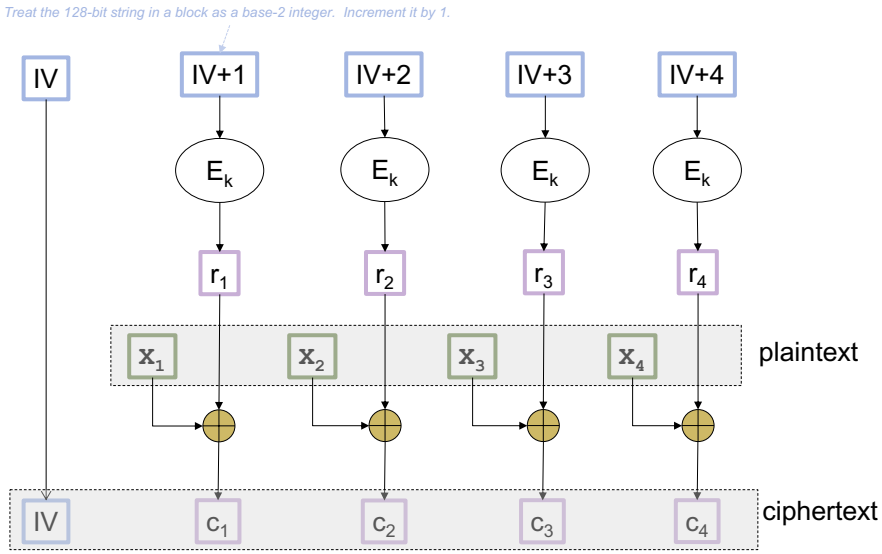
GOOD LUCK!

References

1. Guideline on key strength in this mid-term:
   (a) Exhaustive search over 100-bit keys is feasible. That is, 2^{100} number of cryptographic operations is feasible.
   (b) To secure against online dictionary attack, password entropy of 49 bits is sufficient.
   (c) To secure against offline dictionary attack, password entropy of 128 bits is sufficient.
2. ASCII representation of "`0`", "`1`", and "`2`" in hexadecimal is `30`, `31`, and `32` respectively.
3. The notation || refers to string concatenation. E.g. "`AB`" || "`C`" = "`ABC`".
4. Description of CBC mode encryption from lecture note.



5. Description of CTR mode encryption from lecture note.



Multiple Choice Questions [32 marks]

(Each question is worth 2 marks and has only one correct answer).

1. ___ is a public repository containing discovered vulnerabilities. It is a list of entries, each containing an identification number, a description, and at least one public reference.

   A) NSA

   B) RFC

   C) NIST

   D) CVE

2. A threat model formulates the attacker's capability and goal. For encryption, which of the followings is considered to be the most difficult goal from the attacker's perspective.

   A) Total break, i.e. finding the key.

   B) Recovering the plaintext.

   C) Deciding whether the plaintext is a JPEG image or Microsoft Power Point file.

   D) Deciding whether the plaintext is the 3-character string "`YES`" or "`NIL`".

3. In security analysis, we often assume that in the worst case, the adversary knows all the procedures and algorithms, except the short secret. This is known as __.

   A) Kerckhoffs's principle

   B) Security by obscurity

   C) Multi-factor authenticity

   D) Strong security model

4. Consider an attacker, Mallory, who has access to an encryption oracle that employs AES in CBC-mode. Unlike usual CBC-mode, this oracle always set the IV to be the string of all zeros. Mallory receives a ciphertext c and knows that its plaintext is either m_0 or m_1. Could Mallory determine the plaintext?

   A) No. AES CBC mode is believed to be secure against encryption oracle, thus Mallory cannot distinguish whether the ciphertext correspondes to m_0 or m_1.

   B) Yes. Mallory can request for the ciphertext of m_0 from the oracle. If it is the same as c, then the plaintext must be m_0, otherwise it is m_1.

   C) Yes. By asking the oracle to decrypt the ciphertext c.

   D) Yes. Since the IV is always the string of all zeros, the Zebra's example in the lecture note can be applied. Mallory can now determine $m_0 \oplus m_1$ and then infer the plaintext.

5. Consider the following 16-byte ASCII message m:

   `Password:1234567`

   The message m is then padded (following the method in the lecture note) and then encrypted using

   (*) AES CTR mode

   with a random IV. The resulting ciphertext is a 3-block (v, c_1, c_2). Mallory obtains (v, c_1, c_2) and also knows that it corresponds to the plaintext m. Mallory wants to modify the ciphertext (v, c_1, c_2) so that, upon decryption, the plaintext is changed to

   `Password:_CS2107`

   and the padding format is still correct.

   Which one of the following approaches allows Mallory to achieve this?

   A) Modify the bytes in v corresponding to the password (positions 10 to 15) by XORing the ASCII value of the original and the desired characters.

   B) Modify the bytes in the c_1 corresponding to the password (positions 10 to 15) by XORing the ASCII value of the original and the desired characters.

   C) Modify the bytes in c_2 corresponding to the password (positions 10 to 15) by XORing the ASCII value of the original and the desired characters.

   D) It is not possible, since any modification likely leads to an incorrect padding.

6. This question is same as Q5 except that the method in (*) is changed to AES CBC mode, and the choices of answer are the followings.

   A) Same as Q5 option (A).

   B) Same as Q5 option (B).

   C) Same as Q5 option (C).

   D) It is not possible. Since CBC mode is non-malleable, any modification will spread to the whole plaintext and thus the padding will likely become incorrect.

7. Which of the following three options should not be considered as a 2-factor authentication to carry out transaction. If all are 2-factor authentication, choose the last option.

   A) The user carries out the followings: (1) Login to the bank website (using userid and password) via browser; (2) Next, enters an OTP received from SMS; (3) Selects the transaction.

   B) The user carries out the followings: (1) Login to the bank website (using userid and password) via browser; (2) Next, enters an OTP displayed on a hardware token; (3) Selects the transaction.

   C) During registration, the user installs an online banking app onto a mobile phone p. To complete registration, the user shows his/her face to the camera, and then enters userid and password via the app. Later, to carry out transaction with 2-factor authentication, the user carries out the following: (1) Opens the registered app in the phone p by showing his/her face to the camera; (2) Selects the transaction.

   D) All three choices are 2-factor authentication.

8. Consider RSA where the modulo $n = 77$. What is $\phi(n)$?

   A) (11, 7)

   B) 60

   C) 77

   D) 49

9. Suppose the RSA modulo is $n = 77$ and the encryption key is $e = 11$, what is the decryption key?

   A) 8

   B) 1

   C) 11

   D) 5

10. We know that the public key of RSA includes the modulo n and the exponent e. Lecture note does not explicitly mention whether $\phi(n)$ is to be made public. Which statement is the most appropriate?

    A) $\phi(n)$ must be made public, otherwise it is not possible to encrypt.

    B) Although not necessary, but by Kerckhoffs's principle, it is recommended to make $\phi(n)$ public.

    C) $\phi(n)$ must not be made public, otherwise security will be compromised

    D) Currently, there is no known method that uses knowledge of $\phi(n)$ to extract any information of plaintext from the ciphertext. However, there is no proof that it is infeasible to do so. Hence, it is recommended not to make $\phi(n)$ public.

11. A ____ requires two inputs: a message and a secret key known only to the originator of the message and its intended recipients(s). This allows the recipient to verify the authenticity of the message.

    A) ciphertext

    B) digest

    C) mac

    D) signature

12. Consider a 160-bit hash digest. To find a collision using birthday attack, the expected number of hash operations required is approximately

    A) 2^{40}

    B) 2^{80}

    C) 2^{160}

    D) 2^{320}

13. A company conducted a lucky draw and published NRIC numbers of the 100 winners on their website so that the participants can check whether they had won. This led to significant public backlash on social media, as publishing NRIC numbers could reveal the identities of the winners and compromise their privacy. Later, someone suggested a privacy-preserving alternative: instead of publishing raw NRIC, the company should publish the digests of the NRIC, e.g., $\texttt{SHA3}(x)$, where x is the NRIC encoded in ASCII. This way, a participant could still check whether they had won, while their identity would remain protected. However, this proposal sparked further debate, many call it "flaw". An acceptable system should be secure (do not leak information) and usable (enable participants to verify).

    Which of the following options is the most appropriate assessment of the above proposal?

    (Note: NRIC is Singapore identification system. It consists of 7 digits and two uppercase alphabets, e.g. `S0000000A`.)

    A) Indeed it is flaw. `SHA3` produces a pseudo random digest, which is effectively equivalent to displaying random values on the website. A participant would be unable to determine whether they had won and thus the proposal is not usable.

    B) Indeed it is flaw. Publishing raw hashes of NRICs can still lead to privacy breaches. A secure and usable approach would use a salted hash. That is, publishing $\texttt{SHA3}(r||x)$ where x is the NRIC and r is a randomly chosen 20-character alphanumeric salt.

    C) Indeed it is flaw. The proposal in (B) is not usable. The company should publish both the salt and digest, that is, publish $(r, \texttt{SHA3}(r||x))$ where x is the NRIC and r is a randomly chosen 20-character alphanumeric salt.

    D) Indeed it is flaw. The proposal in (B) is not usable and (C) is still not secure.

14. For simplicity, let us consider padding oracle (using CBC mode and padding method given in lecture) where the block size is 4 bytes. Let us define the following 8 blocks

$$\begin{aligned} v_0 &= (\texttt{00},\texttt{00},\texttt{00},\texttt{00}), \quad c_0 = (\texttt{00},\texttt{00},\texttt{00},\texttt{00}) \\ v_1 &= (\texttt{00},\texttt{FF},\texttt{FF},\texttt{00}), \quad c_1 = (\texttt{00},\texttt{00},\texttt{00},\texttt{00}) \\ v_2 &= (\texttt{00},\texttt{FF},\texttt{02},\texttt{03}), \quad c_2 = (\texttt{00},\texttt{00},\texttt{00},\texttt{00}) \\ v_3 &= (\texttt{FF},\texttt{00},\texttt{01},\texttt{02}), \quad c_3 = (\texttt{FF},\texttt{FF},\texttt{FF},\texttt{FF}) \end{aligned}$$

where each byte is in hexadecimal representation. Let (v_0, c_0) be a 2-block ciphertext under CBC-mode encryption, where v_0 is the IV. In addition, let (x_4, x_3, x_2, x_1) be the corresponding 4-byte plaintext of (v_0, c_0).

Suppose on query (v_0, c_0), the padding oracle replies "correctly padded". On query (v_1, c_1), the padding oracle also replies "correctly padded". What can we infer?

A) x_1 is `01`.

B) x_1 is either `01` or `02`.

C) x_1 is either `01`, `02`, or `03`.

D) x_1 is either `01`, `02`, `03`, or `04`.

15. This is a continuation of Q14. Suppose the padding oracle replies "correctly padded" on all 3 queries (v_0, c_0),(v_1, c_1), and (v_2, c_2). What can we infer?

A) x_2 is `03`.

B) x_2 is `02`.

C) x_2 is `01`.

D) x_2 is `00`.

16. This is a continuation of Q14 and Q15. Suppose the padding oracle replies "correctly padded" on all 4 queries (v_0, c_0),(v_1, c_1), (v_2, c_2), and (v_3, c_3). What can we infer?

A) x_3 is `03`.

B) x_3 is `0D`.

C) Nothing can be inferred about x_3.

D) This is impossible. The four oracle's replies contradict each other.

Short Answer Section [18 marks]

17. A lecturer wants to collect votes from a class of 100 students after the lecture. Each student is assigned a unique 7-digit ID, which they must keep secret. A vote m is a 16-byte ASCII string in the format:

    (7-digit ID)_(date)_(preference)

    where the preferences is either "`0`", "`1`", or "`2`". For example, a vote m can be

    `1234567_091025_2`

    The lecturer has published a 2048-bit RSA modulo n and the exponent e is fixed to be 65537. After the lecture, every student must cast the vote by performing the following steps:

    S1. Encrypts the vote m to obtain c.

    S2. Encodes c as an ASCII string $\widetilde{c}$ using a binary-to-string encoding scheme, then posts $\widetilde{c}$ to the course forum page as an anonymous user.

    The next day, the lecturer decrypts and verifies each post. The lecturer deletes any post that:

    - Is not in the correct format.
    - Contains an invalid ID, date or preference (e.g the preference is "`3`", or an ID not belong to any student).
    - Has conflicting preference. If multiple posts contain the same ID and date but with different preferences, all such post are deleted.
    - Is a duplicate. If multiple identical votes are posted, only the first one is kept.

    The attacker can view and post anonymously to the forum, but may post at most 5 messages. The attacker cannot delete or modify existing posted vote. The attacker can observe which post being deleted by the lecturer. Since this is the first lecture, the attacker has not observed any post with earlier date.

    An authenticity attack is considered successful if the attacker can, with high probability (greater than 0.5), causes the lecturer to accept a forged vote, or delete an authentic vote. A confidentiality attack is considered successful if the attacker can exclude certain preference of a vote with high confidence, for instance, the attacker can infer that, with high probability (greater than 0.5), that the preference is not "`2`".

    The lecturer has proposed 3 different encryption methods for step S1. For each method, determine whether confidentially and/or authenticity can be compromised.

    - Answer "YES" if the property can be compromised, and briefly (max 20 words) explain why.
    - Answer "NO" if the property cannot be compromised. No explanation is needed.

    Each "YES"/"NO" question is worth 1 mark, making a total of 6 marks. The total for the explanation is 12 marks.

- Method **A**. The encryption is carried out as follow:
  (a) Treats m as an integer. Computes $c = m^e \bmod n$. This is textbook RSA encryption.
  (b) The final ciphertext is c.

- Method **B**. The encryption is carried out as follow:
  (a) Randomly selects a 256-bit AES key k.
  (b) Computes $x = PKCS1(k, e, n)$. This is the optimally padded RSA encryption in the standard PKCS#1.
  (c) Encrypts the 16-byte m using AES CTR mode with the key k to obtain the AES ciphertet (v, y) where v is the IV.
  (d) The final ciphertext is $c = (v \mid\mid y \mid\mid x)$.

- Method **C**. The encryption is carried out as follow:
  (a) Randomly selects a 256-bit key k.
  (b) Computes $x = PKCS1(k, e, n)$.
  (c) Encrypts the 16-byte m using AES CTR mode with the key k to obtain the AES ciphertet (v, y) where v is the IV.
  (d) Compute HMAC of (v, y) with the key k. Let $t = \text{HMAC}(k, v \mid\mid y)$.
  (e) The final ciphertext is $c = (v \mid\mid y \mid\mid x \mid\mid t)$.

  After decryption, the lecturer verifies whether $\text{HMAC}(k, v \mid\mid y)$ is indeed same as t. If not, the lecturer will delete the post.

(Hints: Hexadecimal representations of the ASCII characters `0`, `1`, and `2` are shown in page 2. Take special note of their last 2 bits).

— END OF QUESTIONS —

This page is intentionally left blank for you to use as scratch paper.

— END OF PAPER —