CS2107 Midterm Semester 1 AY24/25 October 10, 2024

Instructions:

1. *Use the Answer Sheet. Remember to fill in your id. Use 2B pencil. Submit only the Answer Sheet.*
2. *Total 25 MCQ and 8 pages.*
3. *Note that the numberings of the questions are consecutive starting from 1, even when some questions are logically grouped.*
4. *Only ask administrative questions. If you think that there is ambiguity in the questions, make your own interpretation. We will resolve the ambiguity later.*
5. Choose the most suitable answer. Quoted statements are modifications of phrases extracted from the public domain or other documents. They may not use the same notations in our lecture. Nonetheless, from the context, the meaning is clear.

Assumptions:

1. *Unless otherwise stated, we follow Kerckhoffs's principle, i.e. attackers know the algorithms.*
2. *We assume that it is feasible to carry out 2^{64} primitive cryptographic operations (e.g. hash, one-block AES operation, multiplication of two large integers), whereas it is infeasible to carry out 2^{100} operations.*
3. *We assume that 2^{40} probes are feasible for online dictionary attack, whereas 2^{60} operations are infeasible.*
4. *We assume that 2^{64} operations are feasible for offline dictionary attack, whereas 2^{100} operations are infeasible.*

(Terminologies and general concepts)

1. "*Many systems are open source. In the context of cybersecurity, we can say that these systems are practicing _______.*"

A. accountability
B. agility
C. Kerckhoff's principle
D. security by obscurity
E. absolute transparency

2. An attacker uses an object resembles the coin to buy drink from a vending machine. This illustrates the requirement of ______ in the C-I-A triad.

A. "C"
B. "I"
C. "A"
D. none
E. "C", "I" and "A"

3. A hacker installed a hardware keylogger on the desktop in LT8 and successfully stole a few passwords. This incident best illustrates which security requirement mentioned in class?

A. Confidentiality
B. Integrity
C. Phishing
D. Authenticity
E. Availability

4. *Non-repudiation* can be treated as the requirement of _______ in the C-I-A triad?

   A. "C" B. "I"
   C. "A" D. "C" and "A"
   E. "C", "I" and "A"

5. Which of the following statements on zero-day vulnerability is likely ***wrong***?

   A. Many zero-day vulnerabilities are already documented in the CVE. The difficulty is in patching them and thus difficult to mitigate.
   B. A zero-vulnerability is typically discovered by the attacker before the developer, and thus there is no patch when the attack is launched.
   C. Zero-day vulnerabilities are difficult to find, but some could be simple mistakes that are overlooked for a long period of time.
   D. Zero-day vulnerabilities could be due to implementation bugs, design flaws or even configuration mistakes.

6. "*A good example of________ from everyday life is the withdrawing of money from a cash machine; only the correct combination of a bank card and a PIN allows the transaction to be carried out*."

   A. strong authentication B. multi-factor authentication
   C. unilateral authentication D. challenge and response
   E. physical security

7. "*If you must use a password, a PIN, and your smartphone to log in to your application, then we can say that you are undergoing ____.*"

   A. a multi-step verification, but not a multi-factor authentication.
   B. a multi-factor authentication, but not a multi-step verification.
   C. a 2-step verification or a 3-factor authentication
   D. a 3-step verification or a 2-factor authentication
   E. a 3-step verification or a 3-factor authentication

8. "*Several techniques for __________have been developed. However, some are more effective than others. Perhaps the most common method seen today is eye blinking detection. This seems reasonable; after all, a photo cannot blink.*"

   A. biometric authentication B. identification
   C. passive authentication D. multi-factor authentication
   E. liveness detection

(Encryption)

9. In our lecture note, an image of a zebra is used to illustrate that _____.
   A. a stream cipher does not guarantee integrity.
   B. a stream cipher is vulnerable to padding oracle attack.
   C. a stream cipher does not provide indistinguishability.
   D. the IV of stream cipher must not be predictable.
   E. the IV of stream cipher must not be repeated.

10. “*The disadvantage of* __________ *is a lack of diffusion. Because it encrypts identical plaintext blocks into identical ciphertext blocks, it does not hide data patterns well.*”

    A. ECB mode B. CBC mode
    C. CTR mode D. GCM mode
    E. All mode of operations. It is a property of block ciphers.

11. Lecturer highlighted that a typical attack model would formulate the attacker’s goal and the attacker’s capabilities. In the case of “known plaintext attack”, what is the goal and capability?

    A. Goal: Not specified;
    Capability: Attacker has pair(s) of plaintext the corresponding ciphertext.
    B. Goal: Not specified;
    Capability: Attacker has access to a decryption oracle.
    C. Goal: Find the secret key;
    Capability: Attacker has pair(s) of plaintext the corresponding ciphertext.
    D. Goal: Find the secret key;
    Capability: Attacker has access to a decryption oracle.
    E. Goal: Find the secret key;
    Capability: Attacker has many ciphertexts.

12. There are many advantages of public key cryptography (PKC) over symmetric key. Which of the following is being highlighted the most by the lecturer? (There could be multiple correct statements. Choose the one that the lecturer highlighted).

    A. PKC is provably secure.
    B. Symmetric key needs secure channel among each pair to establish the key. PKC only needs a secure broadcast channel.
    C. PKC has a homomorphic property which is useful in applications, e.g. blind signature.
    D. PKC doesn’t need mode-of-operation while block cipher is only designed for small block.
    E. PKC can provide authenticity, but symmetric key is only for confidentiality.

13. Certain mode of operations can be parallelized so that two processors can complete the operation in roughly half the time taken by a single processor. Can CBC mode be parallelized?

A. Yes for encryption. No for decryption.
B. No for encryption. Yes for decryption.
C. Yes for both encryption and decryption.
D. No for both encryption and decryption.
E. Depend on the choice of IV.

(Padding Oracle) All the followings are encrypted using CBC mode and each block consists of 16 (sixteen) bytes. Numbers written in the form x00 are in hexadecimal format. E.g. x0F represents the value 15 (fifteen). A single block (x10,x10,…,x10) is considered “Correctly padded”

14. Due to space constraint, in this question, each integer is to be treated as hexadecimal, e.g. “1” represents “x01”.

Suppose a padded two-block message is:
(0,0,0,0, 1,1,1,1, 2,2,2,2, 3,3,3,3), (1,2,3,4, 5,6,7,8, 9,0,1,2, 3,3,3,3)

What is the original message before padding?

A. (0,0,0,0, 1,1,1,1, 2,2,2,2, 3,3,3,3), (1,2,3,4, 5,6,7,8, 9,0,1,2, 3,3,3,3)
B. (0,0,0,0, 1,1,1,1, 2,2,2,2, 3,3,3,3), (1,2,3,4, 5,6,7,8, 9,0,1,2, 3)
C. (0,0,0,0, 1,1,1,1, 2,2,2,2, 3,3,3,3), (1,2,3,4, 5,6,7,8, 9,0,1,2)
D. (0,0,0,0, 1,1,1,1, 2,2,2,2, 3), (1,2,3,4, 5,6,7,8, 9,0,1,2, 3)
E. (0,0,0,0, 1,1,1,1, 2,2,2,2), (1,2,3,4, 5,6,7,8, 9,0,1,2)

15. Suppose the attacker knew that a 16-byte plaintext was of the following form:

$(\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3, \boldsymbol{b}_4, \boldsymbol{b}_5, \boldsymbol{b}_6, \boldsymbol{b}_7, \boldsymbol{b}_8, \boldsymbol{b}_9, \boldsymbol{b}_{10}, \boldsymbol{b}_{11}, \boldsymbol{b}_{12}, \boldsymbol{b}_{13}, \boldsymbol{b}_{14}, \text{x03}, \text{x01})$

The attacker did not know the value of $\boldsymbol{b}_i$'s. Let $\boldsymbol{v}$ be the one-block IV and $\boldsymbol{c}$ be the one-block ciphertext of the above plaintext.

$\mathbf{v} = (v_1, v_2, v_3, \ldots, v_{16})$
$\mathbf{c} = (c_1, c_2, c_3, \ldots, c_{16})$

The attacker is given $\boldsymbol{v}$, $\boldsymbol{c}$ and wants to find out the value of $\boldsymbol{b}_{14}$. Following oracle attack in the lecture, what could be the next query to be sent?

A. $(\mathbf{v},\mathbf{c'})$, where $\mathbf{c}'=(c_1,c_2,c_3,c_4,c_5,c_6,c_7,c_8,c_9,c_{10},c_{11},c_{12},c_{13},c_{14},c_{15}\oplus\text{x03},c_{16}\oplus\text{x03})$
B. $(\mathbf{v},\mathbf{c'})$, where $\mathbf{c}'=(c_1,c_2,c_3,c_4,c_5,c_6,c_7,c_8,c_9,c_{10},c_{11},c_{12},c_{13},c_{14},c_{15}\oplus\text{x01},c_{16}\oplus\text{x03})$
C. $(\mathbf{v'},\mathbf{c})$, where $\mathbf{v}'=(v_1,v_2,v_3,v_4,v_5,v_6,v_7,v_8,v_9,v_{10},v_{11},v_{12},v_{13},v_{14},v_{15}\oplus\text{x03},v_{16}\oplus\text{x03})$
D. $(\mathbf{v'},\mathbf{c})$, where $\mathbf{v}'=(v_1,v_2,v_3,v_4,v_5,v_6,v_7,v_8,v_9,v_{10},v_{11},v_{12},v_{13},v_{14},v_{15}\oplus\text{x01},v_{16}\oplus\text{x03})$
E. $(\mathbf{v'},\mathbf{c})$, where $\mathbf{v}'=(v_1,v_2,v_3,v_4,v_5,v_6,v_7,v_8,v_9,v_{10},v_{11},v_{12},v_{13},v_{14},v_{15},v_{16}\oplus \text{x02})$

16. This is a continuation of Q14. Suppose after the query was sent, the oracle replied with “Correctly padded”, what was $\boldsymbol{b}_{14}$?

A.	x00	B.	x01
C.	x02	D.	x03
E.	x04		

17. Consider a single block plaintext **x**. Let the IV and it’s ciphertext be

$\mathbf{v} = (v_1, v_2, v_3, \ldots, v_{16})$

$\mathbf{c} = (c_1, c_2, c_3, \ldots, c_{16})$

(S1): On query (**v**, **c**), the padding oracle returns “*Correctly padded*”;
(S2): On query (**v’**,**c**), the padding oracle returns “*Wrongly padded*”, where

$\mathbf{v'} = (v_1,v_2,v_3,v_4,v_5,v_6,v_7,v_8,v_9,v_{10},v_{11},v_{12},v_{13},v_{14},\ \text{x07} \oplus v_{15},\ \ \text{x07} \oplus v_{16}\ )$

We do not have other information on **x**. From (S1) and (S2), we can eliminate some possibilities and derive the possible padding sizes of **x**. Let ***S*** be the set of possible padding sizes. What is ***S***?

A. {1,2}
B. {5,6}
C. { 3,4,5,6,7,8,9,10,11,12,13,14,15,16 }
D. {1,2,3,4, 7,8,9,10,11,12,13,14,15,16}
E. {1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16}

18. This question is same as the previous question (Q17), except that in step **(S2)**, the padding oracle returns “*Correctly padded*”. What is ***S***?

A. {1,2}
B. {5,6}
C. {1,2,5,6}
D. {1,2,3,4, 7,8,9,10,11,12,13,14,15,16}
E. {1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16}

19. An attacker has access to a padding oracle and is given **v** (the IV) and **c** (the ciphertext) of a plaintext.

    v = (v_1, v_2, v_3, …, v_{16})
    c = (c_1, c_2, c_3, …, c_{16})

    No other information on the plaintext is given to the attacker. The attacker wants to determine the plaintext. Suppose on query (**v**, **c**), the padding oracle returns "*Wrongly padded*". Note that the oracle attack in our lecture note & tutorial only start from a given (**v**, **c**) that is correctly padded. Which of the statement is correct and most appropriate?

    A. The attacker is unable to carry on because the original plaintext is not correctly padded. So, it is not possible to employ the padding oracle to find the plaintext.
    B. The attacker can find the plaintext. The attacker iterates on ***r***, starting from r=x00, and queries the oracle with (**v'**,**c**) where

       v'= (v_1, …, $v_{16} \oplus \boldsymbol{r}$).

       If the output is "Correctly padded", the attacker now has a starting point and can proceed with the padding oracle attack. Otherwise, increments ***r*** and repeats the process. Eventually, the iteration will halt.
    C. Same as the method in (B) except that the query sent is (**v**, **c'**) where

       c' = (c_1, …, $c_{16} \oplus \boldsymbol{r}$).
    D. Same as the method in (B) except that the query sent is (**v'**, **c'**) where **v'**, **c'** is respectively defined in (B) and (C).
    E. The methods in (B), (C) and (D) are all correct.

20. Bob made the following remark about padding oracle and decryption oracle.

    "*(1) With access to a padding oracle, the attacker can essentially decrypt any given ciphertext by adaptively querying the oracle.*
    (2) Hence, having multiple adaptive accesses to a padding oracle is equivalent to having access to a decryption oracle."

    What do you think?

    A. Do not agree. An attacker with access to a padding oracle can get more information on the padding than from a decryption oracle.
    B. Do not agree. An attacker with access to a decryption oracle will always obtain the plaintext. However, with padding oracle, there is a small probability that the attacker fails to decrypt it.
    C. Do not agree. The first statement does not hold. This is because if the plaintext is not correctly padded, then it is not possible to find it using padding oracle.
    D. Do not agree. While the first statement is true, this does not imply the second statement.
    E. Agree.

(Hash, Mac, Signature)

21. IoT devices typically have low computing resources. Bob is designing an IoT application, and the application needs a cryptographic secure hash. To save storage space, Bob uses a hash H(m) as follow:

$$H(\mathbf{m}) = L_{110}(\text{SHA3}(\mathbf{m})) \oplus T_{110}(\text{SHA3}(\mathbf{m}))$$

where $L_n()$ and T_n () is function that gives the leading n-bit substring and trailing n-bit substring respectively. For instance L_3 ('110011') = '110', and T_3('110011') = '011'. Here, digest size of SHA3() is 256 bits.

Which of the following is an appropriate statement on the security of H().

A. H() is secure. Since output of SHA3() consists of 256 bits, thus L_{110}(SHA3(m)) and T_{110}(SHA3(m)) do not overlap. Hence there is no cancellation and thus the output remains pseudo-random.

B. H() is secure. If we can find a collision of H(), then it is also a collision of SHA3(). This contradicts the assumption that SHA3() is collision resistant.

C. Not secure. 110-bit digests are too short to counter birthday attack.

D. Not secure. There are many ***m*** where L_{110}(SHA3(***m***)) = T_{110}(SHA3(***m***)). For those ***m***'s, their digest will be string of zeros. Thus not collision-resistant.

E. Not secure. Choose m_1 be a 256-bit string of all zeros, and m_2 be a 256-bit string of all ones. H() will give the same digest and thus we have a collision.

22. A lecturer needed to assign a 16-byte sequence to each student. Those sequences were to be made public and broadcasted. The lecturer chose the sequence to be the 16 leading bytes of SHA3(***x***) where ***x*** was the student's NUS student ID. The lecturer next posted the assigned sequences in Canvas. The lecturer didn't use the IDs directly due to privacy concerns. It was against some policy to reveal personal information in Canvas. What do you think?

A. The method would not work. This is because there is a good chance that two students being assigned a same 16-byte sequence.

B. The method would not work. This is because a student ID is less than 16-byte and thus too short for SHA3().

C. Lecturer violated the policy. To conform with the policy, the lecture should use salted hash. Specifically, the assigned number should be in the form of (***r*** || L_{64} (SHA3(***r***||***x***))), where ***r*** is the 8-byte random salt, and L_{64}() keeps the leading 8 bytes of the digest.

D. Lecturer violated the privacy policy. This is because the published sequences can be inverted by exhaustively searching the ID.

E. Lecturer didn't violate the privacy policy. Although not truly random, SHA3 is cryptographically pseudo-random. Hence, we are unable to distinguish the published numbers from truly random numbers.

23. There are 2^{10} Pokémon in the world. Alice selects 2^4 unique Pokémons. Bob uniformly and randomly chooses (with replacement) ***M*** number of Pokémons. We say that Bob wins if he has chosen a Pokémon that is also selected by Alice. What is the smallest ***M*** so that the probability that Bob wins is larger than 0.5?

    A. 2^1 B. 2^4
    C. 2^7 D. 2^9
    E. Unable to decide. This is because the probability depends on how Alice chooses the Pokémons, which the question does not specify.

24. Bob designed a hash function H() which outputs 256-bit digests. Alice wrote a program that, when given a digest ***d***, the program can output a word document ***D*** within 5 hours, where the leading 50 bits of ***d*** is the same as the leading 50 bits of H(***D***). What is the implication?

    A. H() is not collision resistant.
    B. H() is not one-way.
    C. H() is not 2nd pre-image resistant.
    D. H() is not cryptographically pseudo-random.
    E. No implication on the security of H().

25. Alice claims that it is feasible to find two JPEG files $\boldsymbol{m}_1$, $\boldsymbol{m}_2$ s.t.

    $$L_{110}\,(\mathrm{SHA3}(\boldsymbol{m}_1)) = T_{110}\,(\mathrm{SHA3}(\boldsymbol{m}_2))$$

    Do you agree with her?
    (Note: The notation $L_{110}()$ and $T_{110}()$ as defined in Q21)

    A. Yes. It is feasible to employ birthday attack to find such $\boldsymbol{m}_1$ and $\boldsymbol{m}_2$.
    B. Yes. Since a JPEG file can be more than 110 bits, certainly there many JPEG files meeting the above requirement.
    C. No. It is infeasible to invert a 110-bit digest to find a match.
    D. No. Since SHA3 is collision resistant, it is infeasible to find a match.
    E. It depends on the choice of SHA3. If the digest size is 512 bits, then it is not feasible. If the digest size is 224 bits, then it is feasible.

------------ END of Paper ------------