CS 161 Spring 2024

Introduction to Computer Security

Computer Security

Name:

Student ID:

This exam is 110 minutes long.

For questions with circular bubbles, you may select only one choice.

- Unselected option (completely unfilled)
- Only one selected option (completely filled)
- Don't do this (it will be graded as incorrect)

For questions with square checkboxes, you may select one or more choices.

Nou can select

nultiple squares (completely filled)

Anything you write outside the answer boxes or you cross out will not be graded. If you write multiple answers, your answer is ambiguous, or the bubble/checkbox is not entirely filled in, we will grade the worst interpretation.

$Q1$ Honor Code (0 points)

Pre-exam activity (0 points):

It's Leap Day, which means it's time for the quadrennial Leap-Frog Race! Everyone entered their frogs into the contest, but unfortunately the results got mixed up. Evan-Bot needs your help to piece together the standings from eyewitness reports!

- 1. Alice: "Bob's frog was right behind EvanBot's frog."
- 2. Bob: "CodaBot's frog was somewhere ahead of mine."

3. CodaBot: "I couldn't tell whose it was, but there was a frog with a hat somewhere ahead of Mallory's frog."

4. Eve: "My frog was always ahead of Alice's frog."

5. Mallory: "My frog definitely finished before CodaBot's frog. I'm telling the truth, I promise."

Circle the winning frog.

I understand that I may not collaborate with anyone else on this exam, or cheat in any way. I am aware of the Berkeley Campus Code of Student Conduct and acknowledge that academic misconduct will be reported to the Center for Student Conduct and may further result in, at minimum, negative points on the exam.

Read the honor code above and sign your name:

$Q2$ True/False (16 points)

Each true/false is worth 1 point.

Q2.1 The Mallory Security Agency requires 128-bit AES encryption for all data, and "Top Secret" data is encrypted with an additional 256-bit key stored in a separate location.

True or False: This is an example of Defense in Depth.

Q2.2 Devices storing cryptographic keys are often required to be fitted with tamper-evident seals.

True or False: This is an example of Detect If You Can't Prevent.

Q2.3 True or FALSE: All strings in C are terminated with the character $'\n$.

Q2.4 True or False: Enabling non-executable pages prevents the ret2libc attack.

For the next 2 subparts: Suppose we have a little-endian C program with a local variable char buf[8]. Consider the following possible GDB output after running the command $x/4wx$ buf:

0xffffff58: 0x61657270 0x0000006d 0xc0dab0bb 0xbaa15691

Q2.5 True or False: The words 0x61657270 and 0x0000006d correspond to the data in buf.

 \bigcirc (A) True \bigcirc (B) False

Q2.6 True or False: buf[9] is 0xda.

 \bigcirc (A) True \bigcirc (B) False

Q2.7 True or False: Every nondeterministic (randomized) encryption scheme is IND-CPA secure.

 \bigcap (A) True \bigcap (B) False

Q2.8 True or False: If Alice encrypts a message to Bob using ElGamal, an eavesdropper who can solve the discrete log problem would be able to correctly decrypt the message.

 \bigcap (A) True \bigcap (B) False

- Q2.9 True or False: Digital signature schemes often encrypt the message before signing to prevent existential forgery attacks.
	- O (A) True O (B) False
- Q2.10 True or False: AES-CTR does not require the message to be padded.
	- O (A) True O (B) False
- Q2.11 True or FALSE: HMAC is equivalent to NMAC with $K_1 = K_2$.
	- \bigcap (A) True \bigcap (B) False
- Q2.12 True or False: Using the output of a PRNG as the key for a one-time pad provides perfect security, even against an attacker with infinite computational power.
	- \bigcap (A) True \bigcap (B) False
- Q2.13 True or False: Length-extension attacks are often used to break the collision resistance of a hash function.
	- \bigcap (A) True \bigcap (B) False
- Q2.14 True or False: The security of Diffie-Hellman relies on the modulus being difficult to factor.
	- \bigcap (A) True \bigcap (B) False
- Q2.15 True or FALSE: The security of ElGamal relies on the difficulty of finding $g^{ab} \bmod p$ given $g^a,g^b \bmod p.$
	- O (A) True O (B) False
- Q2.16 True or False: Certificates are generally not sent over insecure channels due to the risk of a replay attack.

 O (A) True O (B) False

Q3 'Tis But a Scratch (16 points)

King Arthur and his knights are searching for the Holy Grail. They find a castle protected by the following vulnerable C code:

Stack at Line 8 RIP of castle (1) (2) (3) cave

The find_grail function writes a secret 16-byte string to holy_grail, and it's your job to make the program output the holy_grail string! You can assume find_grail does not modify the stack in any other way.

Q3.1 (1 point) What values go in blanks (1) through (3) in the stack diagram above?

Q3.2 (1 point) Which vulnerability is present in the code?

 \bigcirc (B) Format string vulnerability (D) Heap overflow In the next two subparts, provide inputs that would cause the program to output the holy_grail string (possibly as part of a larger output).

If a part of the input can be any non-zero value, use 'A'*n to represent the n bytes of garbage.

- Q3.3 (4 points) Input to fread at Line 11:
- Q3.4 (4 points) Input to fread at Line 17:

The next four subparts are independent from each other.

- Q3.5 (1 point) Would it still be possible for your exploit to leak holy_grail if fread(&holy_hand_grenade, 1, 1, stdin) on Line 11 is replaced with gets(&holy_hand_grenade)?
	- \bigcirc (A) Yes, because gets would allow the attacker to overwrite the address of cave.
	- \bigcirc (B) Yes, because holy_hand_grenade is above holy_grail in the stack diagram.
	- \bigcap (C) No, the null terminator added by gets will cause the final printf to terminate before reaching holy_grail.
	- \bigcirc (D) No, because the null terminator added by gets will partially overwrite the SFP of castle.
- Q3.6 (1 point) Would it still be possible for your exploit to leak holy_grail with stack canaries enabled?
	- \bigcap (A) Yes, because the exploit writes around the canary to overwrite values above the canary.
	- \bigcirc (B) Yes, because the exploit never tries overwriting values above the canary.
	- \bigcirc (C) No, because all 4 bytes of the canary are overwritten by garbage.
	- \bigcirc (D) No, because the least-significant byte of the canary is overwritten by a null byte.
- Q3.7 (1 point) Would it still be possible for your exploit to leak holy_grail with non-executable pages enabled?
	- \bigcirc (A) Yes, because the exploit never writes any executable instructions on the stack.
	- \bigcirc (B) Yes, because the malicious instructions being executed are not on the stack.
	- \bigcap (C) No, because the exploit writes executable instructions on the stack.
	- \bigcirc (D) No, because non-executable pages stops the exploit from writing anything on the stack.

Q3.8 (1 point) Would it still be possible for your exploit to leak holy_grail with ASLR enabled?

 \bigcirc (A) Yes, because the exploit does not require knowing any absolute addresses.

 \bigcap (B) Yes, because the printf call will always leak an address on the stack.

 \bigcirc (C) No, because we need to know the address of holy_grail in order to leak its value.

 \bigcirc (D) No, because we need to know the address of castle in order to complete the exploit.

For the following subparts, your goal is now to execute a 72-byte shellcode, rather than leaking holy_grail.

- Q3.9 (1 point) Assuming no memory safety defenses are enabled, would it be possible to exploit this program to execute shellcode?
	- O (A) Yes O (B) No
- Q3.10 (1 point) Assuming canaries are enabled, would it still be possible to exploit this program to execute shellcode?
	- \bigcap (A) Yes, because we can leak the canary and overwrite the canary with itself.
	- \bigcirc (B) Yes, because the printf call can write directly to the RIP, skipping the canary.
	- \bigcap (C) No, because we run the final fread before we can leak the canary.
	- \bigcirc (D) No, because it is not possible to run shellcode at all in this program.

Q4 I Sawed This Shellcode In Half! (16 points)

Consider the following vulnerable C code:

```
1 void boat (void * shellcode_first_half, void * shellcode_second_half) {
2 // fp contains the address of the fgets function
3 uintptr t fp = (uintptr t) fgets;
4
5 char [32] buf;
6 char* buf_ptr = &buf;
7
8 fgets (buf, 32, stdin);
9 printf (buf);
10
11 fgets (buf, 32, stdin);
12 printf (buf);
13 }
```
This is the result of running disas fgets in GDB:

```
10x08076030: push %ebp
2 \vert 0 \times 08076034: mov %esp, %ebp
3 \vert 0 \times 08076038: sub %ebp, 20
4 \vert \ldots50 \times 08076050: mov %ebp, %esp
60x08076054: pop %ebp
7 \vert 0 \times 08076058 : ret
```
shellcode_first_half is a pointer to the first half of shellcode and shellcode_second_half is a pointer to the second half of shellcode. Both halves of shellcode have a ret instruction at the end.

Assumptions:

- ASLR is enabled, but all other memory safety defenses are disabled.
- The program can print an arbitrarily large number of bytes using printf.
- All addresses are at least 0x01000000.
- esp is not modified by the shellcode.
- Q4.1 (2 points) Which of these inputs to the fgets on Line 8 will always leak the values of all the local variables in the boat stack frame? Select all that apply.

Now that we have the values of all the local variables in the boat stack frame, we need to compute two values to use later in our exploit.

You can assume that the values outputted from the printf on Line 9 have all been converted to integers that you can perform arithmetic on.

Q4.2 (2 points) Which of these values do we need to use later in our exploit?

&buf is the address of buf, which you leaked in Q4.1.

Hint: What is the address in memory we want to write to?

Q4.3 (2 points) Which of these values do we need to use later in our exploit?

fp is the address of the fgets function, which you leaked in Q4.1.

Hint: What is the value we want to write into memory?

 $O(A)$ fp O (B) fp + 0x4 (C) fp + 0x8 (E) fp + 0x24 O (D) fp + 0x20 O (F) fp + 0x28

Let x be the value you computed in Q4.2, and y be the value you computed in Q4.3.

You can perform arithmetic on these values (e.g. $x + 5$), and you can assume that the resulting value is converted to the proper format (e.g. raw bytes, or decimal representation), so you don't have to manually do any conversions.

If a part of the input can be any non-zero value, use 'A'*n to represent the n bytes of garbage.

Q4.4 (8 points) Provide an input to the fgets call on Line 11 that will execute shellcode.

- Q4.5 (1 point) If canaries are enabled, would it still be possible to exploit this program to execute shellcode?
	- \bigcirc (A) Yes, because the exploit writes directly to the RIP instead of smashing through the canary.
	- O (B) Yes, because overwriting the RIP directly disables the canary check altogether.
	- \bigcirc (C) No, because all 4 bytes of the canary are overwritten by garbage.
	- \bigcirc (D) No, because the ret instruction at the end of the first shellcode will fail the canary check.
- Q4.6 (1 point) What happens to the exploit if boat is called with the arguments reversed?
	- In other words, instead of boat(shellcode_first_half, shellcode_second_half), what if we called boat(shellcode_second_half, shellcode_first_half)?
		- \bigcirc (A) The instructions of shellcode would all execute in reverse order.
		- \bigcirc (B) The shellcode would execute unchanged.
		- \bigcirc (C) The second half of shellcode executes, followed by the first half of shellcode.
		- (D) The program would crash after returning from boat, but before executing shellcode.

Q5 Challenging Constructions (16 points)

Consider the following encryption scheme:

The encryption formula for this scheme is

$$
C_i = E_K(M_i + IV + i)
$$

Q5.1 (1 point) Select the correct decryption formula for the given scheme.

\n
$$
\text{O (A)} \ M_i = D_K(C_i - IV - i)
$$
\n

\n\n $\text{O (C)} \ M_i = D_K(C_i) + IV + i$ \n

\n\n $\text{O (D)} \ M_i = D_K(C_i + IV - i)$ \n

\n\n $\text{O (D)} \ M_i = D_K(C_i + IV - i)$ \n

To show this scheme is insecure, you want to provide a strategy that always wins the IND-CPA game.

Adversary (you)	Challenger
M and M'	First, the adversary (that's you!) sends two different challenge messages, $M \neq M'$, to the challenge. For your strategy, you can assume M and M' are each two blocks long.
$C = (C_0, C_1)$	Then, the challenge randomly encrypts either M or M' . The resulting two-block ciphertext $C = (C_0, C_1)$ is returned to you.
G M or M'	Finally, you guess whether M or M' was encrypted.

In this strategy, the query phase is not needed (i.e. you never have to ask the challenger to encrypt messages of your choosing beforehand).

Assume that the second challenge message $M' = (?, ?)$ is chosen completely at random.

Clarification during exam: Assume that each "?" value is independently randomly generated, so (?, ?) would not necessarily be two identical values.

Q5.2 (2 points) What must be true of M for this strategy to work?

(Note: ? denotes a randomly-chosen value)

Q5.3 (3 points) The challenger encrypts one of M, M' from the previous subparts and returns $C =$ (C_0, C_1) .

Explain how you would determine whether M or M' was encrypted. Your answer can use these values:

- $M=(M_0,M_1)$ and $M^\prime=(M^\prime_0,M^\prime_1)$
- $C = (C_0, C_1)$. Recall: C is either the encryption of M or M'.

An example of how you could describe your strategy, that has nothing to do with this question: If $C_0 + 161 = M_0$, guess M. Else, guess M'.

- Q5.4 (1 point) What is the probability (excluding negligible factors) that an optimal attacker (i.e. using the correct answer to Q5.3) wins the IND-CPA game?
	- O (A) 50% O (B) 66.7% O (C) 75% O (D) 100%

Now consider a modified version of the previous encryption scheme:

The encryption formula for this scheme is

$$
C_i = E_K(M_i \oplus (IV + i))
$$

You want to devise a strategy to win the IND-CPA game.

Assume that the second challenge message $M' = (?, ?)$ is chosen completely at random.

Q5.5 (1 point) What value should you pick for M ?

Q5.6 (5 points) The challenger encrypts one of M, M' from the previous subparts and returns $C =$ $(C_0, C_1).$

Explain how you would determine whether M or M' was encrypted. Your answer can use these values:

- $M=(M_0,M_1)$ and $M^\prime=(M^\prime_0,M^\prime_1)$
- $C = (C_0, C_1)$. Recall: C is either the encryption of M or M'.

Q5.7 (3 points) What is the probability (excluding negligible factors) that an optimal attacker (i.e. using the correct answer to Q5.6) wins the IND-CPA game?

Justify your answer.

Q6 Authentic Auctions (19 points)

EvanBot wants to hold a charity auction and needs your help! They have designed a secure bidding protocol as follows:

- 1. Each party chooses a bid x .
- 2. Each party applies a function COMMIT to x and broadcasts COMMIT (x) to every other party.
- 3. Once everyone has posted their commitment, each party reveals their bid and any other information relevant to the scheme.

There are two key properties we want to guarantee:

- 1. **Binding**: If a party chooses x, sends COMMIT(x), and then reveals $x' \neq x$, other parties can detect that x' is invalid with overwhelming probability.
- 2. **Hiding:** COMMIT(x) should not leak more than a negligible amount of information about x.

For each of the following schemes, select whether the scheme provides binding, hiding, both, or neither.

You should assume that the range of all possible bids is small enough to brute-force, each subpart is independent, and the auction is run exactly once. If a reveal step is not specified, assume that the reveal step is to send x .

Hint: EvanBot says you should reread the above assumptions and make sure you understand them before proceeding!

Clarification during exam: Assume that all parties have access to the trusted public keys of every other party.

- Q6.5 (1 point) COMMIT $(x) = H(x) \oplus K$, where K is a randomly-generated one-time-pad key. To reveal, each party sends x, K .
	- \bigcirc (A) Binding \bigcirc (B) Hiding \bigcirc (C) Both \bigcirc (D) Neither

The properties, repeated for your convenience:

- 1. **Binding**: If a party chooses x, sends COMMIT(x), and then reveals $x' \neq x$, other parties can detect that x' is invalid with overwhelming probability.
- 2. **Hiding:** COMMIT(x) should not leak more than a negligible amount of information about x.
- Q6.6 (1 point) COMMIT $(x) = H(x \oplus K)$, where K is a randomly-generated one-time-pad key. To reveal, each party sends x, K .
	- \bigcirc (A) Binding \bigcirc (B) Hiding \bigcirc (C) Both \bigcirc (D) Neither
- Q6.7 (1 point) COMMIT(x) = HMAC(K, x), where K is a publicly-known symmetric key shared by everyone.
	- \bigcirc (A) Binding \bigcirc (B) Hiding \bigcirc (C) Both \bigcirc (D) Neither
- Q6.8 (1 point) COMMIT $(x) = \mathsf{Sign}(SK, x)$, where $\mathsf{Sign} = x^d \bmod N$ (naive RSA signature) and SK is the secret key for the user sending the bid.
	- \bigcap (A) Binding \bigcap (B) Hiding \bigcap (C) Both \bigcap (D) Neither
- Q6.9 (1 point) COMMIT $(x) =$ Sign (SK, x) , where Sign $= H(x)^d \mod N$ (hash-based RSA signature).
	- \bigcirc (A) Binding \bigcirc (B) Hiding \bigcirc (C) Both \bigcirc (D) Neither
- Q6.10 (1 point) COMMIT $(x) = g^x \mod p$, where g is a generator and p is large prime (like in Diffie-Hellman).
	- \bigcap (A) Binding \bigcap (B) Hiding \bigcap (C) Both \bigcap (D) Neither

Q6.11 (1 point) Note: (x, y) denotes a tuple of two values, x and y.

COMMIT(*x*) =
$$
\begin{cases} (x, x + 1) & \text{with probability 0.5} \\ (x - 1, x) & \text{with probability 0.5} \end{cases}
$$

\n① (A) Binding

\n① (B) Hiding

\n① (C) Both

\n① (D) Neither

The properties, repeated for your convenience:

- 1. **Binding**: If a party chooses x, sends COMMIT(x), and then reveals $x' \neq x$, other parties can detect that x' is invalid with overwhelming probability.
- 2. Hiding: COMMIT (x) should not leak more than a negligible amount of information about x.
- Q6.12 (8 points) Design a commitment scheme that was not mentioned in a previous subpart. Your scheme should be both binding and hiding.

Assume you have access to an unlimited source of random bits.

Given x, how do you compute $\text{COMMIT}(x)$?

List of values sent in the reveal step (including x):

List of steps to verify $\text{COMMIT}(x)$ given the values from the reveal step:

Q7 Ephemeral Exchanges (17 points)

Consider the following authentication protocol. pwd is a standard-strength password (i.e. vulnerable to brute-force). PK_{server} is a long-term, trusted public key for the server. Assume there's only a single user/password stored on the server.

Here is an equivalent description of the protocol:

- 1. Both the client and server derive $K_{pwd} = H(pwd)$.
- 2. The client generates a random symmetric key K_{sym} .
- 3. The client sends $Enc(K_{pwd}, \text{RSAEnc}(PK_{server}, K_{sym}))$ to the server.
- 4. The server decrypts the message from the previous step to get K_{sym} .
- 5. The client sends a randomly generated number R_C to the server (challenge message).
- 6. The server replies with $\mathsf{Enc}(K_{sym}, R_C)$.
- 7. The client verifies that the server's response is valid.
- 8. The server sends a randomly generated number R_S to the client (challenge message).
- 9. The client replies with $Enc(K_{sym}, R_S)$.
- 10. The server verifies that the client's response is valid.

Q7.1 (1 point) Which equation does the client use in Step 7 to verify the server's response from Step 6 (denoted STEP6)?

Q7.2 (1 point) Which option best explains why the two parties don't use $H(pwd)$ for the final shared symmetric key K_{sym} ?

Clarification during exam (for 7.2, 7.3, 7.4): Assume that the attacker also records messages encrypted and MAC-d with K_{sym} after the protocol finishes.

- \bigcirc (A) The attacker can brute force values of $H(pwd)$ and check candidate keys using messages from the resulting secure channel.
- \bigcirc (B) H outputs too many bits to be used as a symmetric key.
- \bigcap (C) $H(\mathsf{pwd})$ would be the same across different authentications.
- \bigcirc (D) $H(\mathsf{pwd})$ would require a salt to prevent dictionary attacks.
- Q7.3 (2 points) Is it possible for an eavesdropper to dictionary attack pwd given $Enc(K_{pwd}, \text{RSAEnc}(PK_{server}, K_{sym}))$ from Step 3?
	- \bigcirc (A) Yes, they can try decrypting $Enc(K_{pwd}, \text{RSAEnc}(PK_{server}, K_{sym}))$ with each guess of pwd.
	- \bigcirc (B) Yes, because anyone can evaluate RSAEnc(PK_{server}, K_{sym}) and check that $Enc(K_{pwd}, \text{RSAEnc}(PK_{server}, K_{sym}))$ evaluates to the existing value.
	- \bigcap (C) No, because there is no way to tell when a test decryption is successful.
	- \bigcirc (D) No, because RSAEnc is slow to evaluate, preventing brute force.

Q7.4 (2 points) An attacker records all messages sent between a client and server during a successful login.

Later, the attacker learns the value of SK_{server} , the server's secret key (corresponding to PK_{server}). Can the attacker learn the value of K_{sym} from the earlier recorded login?

- \bigcirc (A) Yes, because the attacker can try all possible values of pwd and find K_{sym} .
- \bigcirc (B) Yes, because the attacker can try all possible values of K_{sym} once SK_{server} is leaked.
- \bigcirc (C) No, because K_{sym} is randomly generated for each authentication and deleted immediately afterwards.
- \bigcirc (D) No, because there is no way to tell when a test decryption is successful, even if we leak SK_{server} .

Q7.5 (8 points) Design an attack to successfully pass Step 10 (server verification) without knowing pwd. Assumptions:

- 1. You start two simultaneous connections to the server: Connection A and Connection B. The same server is shown twice in the diagram below.
- 2. You do not need to verify the server's response in Step 7.
- 3. In each of the six boxes below, you can write:
	- $A_{Step X}$ or $B_{Step X}$ to denote the value from step X of Connection A/B.
	- If any value would work for a box, write the word anything.

Q7.6 (1 point) In which of the two connections does the server successfully verify the challenge response sent by the client in Step 9?

 \bigcirc (A) Connection A \bigcirc (B) Connection B \bigcirc (C) Both connections

Q7.7 (2 points) Which of the following modifications would prevent an optimal attack from Q7.5? Select all that apply.

 \Box (A) Adding an HMAC keyed with K_{sym} on the challenge messages (Steps 5, 8).

 \Box (B) Having the server send its challenge message first (i.e. swap steps 5-7 with steps 8-10).

□ (C) Replacing Step 6 with "The server replies with $Enc(K_{sym}, R_C||R_S)$ "

 \Box (D) Generating the challenge numbers using the output of a PRNG seeded with pwd.

 \Box (E) None of the above

Nothing on this page will affect your grade.

Post-Exam Activity

What is Mallory cooking?

Comment Box

Congratulations for making it to the end of the exam! Feel free to leave any thoughts, comments, feedback, or doodles here: