Raluca Popa
Spring 2018

CS 161
Computer Security

Homework 2

Due: Wednesday, February 14, at 11:59pm

Instructions. This homework is due **Wednesday, February 14, at 11:59pm**. No late homeworks will be accepted. You *must* submit this homework electronically via Gradescope (not by any other method). When submitting to Gradescope, *for each question* your answer should have each question's answer on its own page. This assignment must be done on your own.

Problem 1 ***True-or-False Questions*** **(40 points)**

Answer each question. You don't need to justify or explain your answer.

(a) TRUE or FALSE: Diffie–Hellman protects against eavesdroppers but is vulnerable to man-in-the-middle attacks.

(b) TRUE or FALSE: Suppose there is a transmission error in a block B of ciphertext using CBC mode. This error propagates to every block in decryption, which means that the block B and every block after B cannot be decrypted correctly.

(c) TRUE or FALSE: The IV for CBC mode must be kept secret.

(d) TRUE or FALSE: The random number r in El Gamal must be kept secret.

(e) TRUE or FALSE: The best way to be confident in the cryptography that you use is to write your own implementation.

(f) TRUE or FALSE: Alice and Bob share a symmetric key k. Alice sends Bob a message encrypted with k stating, "I owe you \$100", using AES-CBC encryption. Assuming AES is secure, we can be confident that an active attacker cannot tamper with this message; its integrity is protected.

(g) TRUE or FALSE: If the daily lottery numbers are truly random, then they can be used as the entropy for a one-time-pad since a one-time-pad needs to be random.

(h) TRUE or FALSE: It is okay if multiple people perform El Gamal encryption with the same modulus p.

(i) TRUE or FALSE *(Optional, 0 points)*: It is okay if an RSA library generates multiple public keys with the same N, as long as it uses different values for e.

(j) TRUE or FALSE *(Optional, 0 points)*: Alice and Bob share a secret symmetric key k which they use for calculating MACs. Alice sends the message $M =$"I (Alice) owe you (Bob) \$100" to Bob along with its message authentication code $\text{MAC}_k(M)$. Bob can present $(M, \text{MAC}_k(M))$ to a judge as proof that Alice owes him \$100 since a MAC provides integrity.

Problem 2 *New Block Cipher Mode* (30 points)

Nick decides to invent a new block cipher mode, called NBC. It is defined as follows:

$$C_i = E_k(C_{i-1}) \oplus P_i$$
$$C_0 = \text{IV}$$

Here $(P_1, \ldots, P_n)$ is the plaintext message, E_k is block cipher encryption with key k.

(a) Given $(C_0, C_1, \ldots, C_n)$ and the key k, explain how to recover the original message $(P_1, \ldots, P_n)$.

(b) As we saw in discussion, CBC mode is vulnerable to a chosen plaintext attack when the IV which will be used to encrypt the message is known in advance. Is NBC vulnerable to the same issue?

(c) Say that Alice means to send the message $(P_1, \ldots, P_n)$ to Bob using NBC mode. By accident, Alice typos and encrypts $(P_1 \oplus 1, \ldots, P_n)$ instead (i.e., she accidentally flips the last bit of the first block).

True or False: after Bob decrypts the resulting ciphertext, every block after the first is incorrect. Explain your answer.

(d) Alice encrypts the message $(P_1, \ldots, P_5)$. Unfortunately, the block C_3 of the ciphertext is lost in transmission, so that Bob recieves $(C_0, C_1, C_2, C_4, C_5)$. Assuming that Bob knows that he is missing C_3, which blocks of the original plaintext can Bob recover?

Problem 3 *Alternate Feedback* (40 points)

Leland has taken inspiration from Nick's cipher mode and decided to create his own scheme, FFM, which in his words is a "new and improved" version of Nick's scheme. Leland chose to not redesign the wheel and keeps the block & key size to be 128b, but believes that by using the inputs in a different way that the scheme could be better. The following is a diagram of the FFM block cipher mode of encryption.

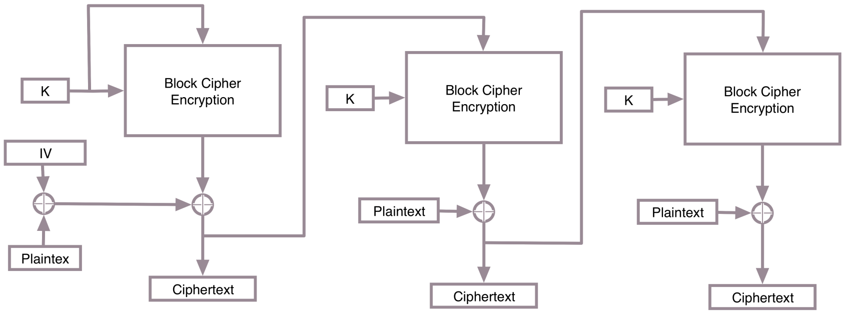

Figure 1: FFM Encryption Mode

(a) Explain/draw what the decryption mode will have to look like.

(b) If you reuse the IV for two secret messages, M_1 and M_2, both using the same key, producing two ciphertexts C_1 and C_2 seen by the eavesdropper, what can the eavesdroper learn? Assume that the first bit of M_1 and M_2 are different but the rest of the bits may or may not be the same.

(c) If the first bit of the ciphertext is corrupted in transmission after the encryption is complete and then decrypted, which bits of the decrypted plaintext will be corrupted?

(d) TRUE or FALSE: The encryption can be parallelized.

(e) TRUE or FALSE: The decryption can be parallelized.

(f) Is this IND-CPA? Why or why not?

Problem 4 ***Finding Common Patients*** **(40 points)**

Caltopia has two hospitals: Bear Hospital and Tree Hospital, each of which has a database of patient medical records. These records contain highly sensitive patient information that should be kept confidential. For both hospitals, each medical record is a tuple (p_i, m_i), where p_i and m_i are strings that correspond to the patient's full name and medical record respectively; assume that every person in Caltopia has a unique full name. Thus, we can think of Bear Hospital's patient database as a list of tuples $(x_1, m_1), (x_2, m_2), ..., (x_n, m_n)$, where m_i is the medical information that Bear Hospital has for patient x_i. Similarly, we can think of Tree Hospital's database as a list $(y_1, m'_1), (y_2, m'_2), ..., (y_m, m'_m)$, where m'_i is a string that encodes the medical information that Tree Hospital has for the patient named y_i. Note that for a given patient, Tree Hospital and Bear Hospital might have different medical information.

The two hospitals want to collaborate on a way to identify which Caltopia citizens are patients at both hospitals. However, due to privacy laws, the two hospitals cannot share any plaintext information about patients (including their names) unless both hospitals know *a priori* that a patient has used both hospitals.

Thus, the two hospitals decide to build a system that will allow them to identify common patients of both hospitals. They enlist the help of Lady Olenna, who provides them with a trusted, third-party server S, which they will use to discover the names of patients who use both hospitals. Specifically, Bear Hospital will take some information from its patient database and transform it into a list $(x_1^*), (x_2^*), ..., (x_n^*)$ (where (x_i^*) is somehow derived from x_i (the patient's full name) and upload it to S. Similarly, Tree Hospital will take information from its patient database, transform it into a list $(y_1^*), (y_2^*), ..., (y_m^*)$, and upload this transformed list to S. Finally, S will compute a set of tuples $P = (i, j) : x_i = y_j$ of all pairs (i, j) such that $x_i^* = y_j^*$ and send P to both Bear Hospital and Tree Hospital. The two hospitals can then take their respective indices from the tuples in P to identify patients who use both hospitals.

We want to ensure three requirements with the above scheme: (1) if $x_i = y_j$, then $(i, j) \in P$, (2) if $x_i \neq y_j$, then it is very unlikely that $(i, j) \in P$, (3) even if Eve (an attacker) compromises S, she cannot learn the name of any patient at either hospital or the medical information for any patient. For this question, assume that Eve is a passive attacker who cannot conduct Chosen Plaintext Attacks; however, she does know the names of everyone in Caltopia, and there are citizens whose full names are a unique length.

Fill in your solutions below. Your solution can use the cryptographic hash SHA-256 and/or AES with one of the three block cipher encryption modes discussed in class; keep in mind that Eve can also compute SHA-256 hashes and use AES with any block cipher mode. You can assume that Bear Hospital and Tree Hospital share a key k that is not known to anyone else. You *cannot* use public-key cryptography or modular arithmetic.

(a) In the collaboration scheme described above, how should Bear Hospital compute x_i^* (as a function of x_i)? How should Tree Hospital compute y_i^* (as a function of y_i)? Specifically, your solution should define a function F that Bear Hospital will use to transform x_i into x_i^*, and if relevant, a function G that Tree Hospital will use to transform y_i into y_i^*.

(b) Explain why requirement (1) is met by your solution, i.e., explain why it is guaranteed that if $x_i = y_j$, then $x_i^* = y_j^*$ will hold. Explain your answer in one or two sentences.

(c) Explain why requirement (2) is met by your solution, i.e., if $x_i \neq y_j$, explain why it is unlikely that $x_i^* = y_j^*$. Explain your answer in one or two sentences.

(d) Explain why requirement (3) is met by your solution, i.e., if S is compromised by Eve, then the information known to S does not let Eve learn any patient information (neither the names of patients at a particular hospital nor the medical history for any patient). Explain your answer in one or two sentences.

Problem 5 ***Hashing Functions*** **(20 points)**

Recall the definition of "one-way functions" and "collision-resistance" from lecture. We say a function f is one-way if given $f(x)$ it is hard to find x' such that $f(x') = f(x)$. Likewise, we say a function f is "collision-resistant" if it is hard to find two inputs x, y such that $f(x) = f(y)$ but $x \neq y$. For each of the given functions H below, determine if it is one-way or not, and if it is collision-resistant or not. (State any assumptions that you make.)

(a) $H(x) = x$. One-way? Collision-resistant?

(b) $H(x) = x \bmod 2$. One-way? Collision-resistant?

(c) $H(x) = E_k(x)$, where E_k is a secure block cipher with a known and published key k. One-way? Collision-resistant?

(d) $H(x) = g^x \bmod p$, where p is a prime, $2 \leq g < p - 1$, and $0 \leq x < p - 1$. Both g and p are public. One-way? Collision-resistant?

(e) $H(x) = \text{SH}(x)$, where SH is a cryptographically-secure hash function which takes $\Theta(x^2)$ time to compute.

Problem 6 ***El Gamal Encryption*** **(30 points)**

Recall the definition of El Gamal encryption from lecture. Bob publishes a large prime p, and an integer g with $1 < g < p-1$. To generate a key, Bob chooses a random value $0 \leq b \leq p-2$, and computes $B = g^b \bmod p$. Bob's public key is B, and his private key is b. If Alice wants to send a message m to Bob, she begins by generating a random r such that $0 \leq r \leq p-2$, and creates the ciphertext $(c_1, c_2) = (g^r \bmod p, m \cdot B^r \bmod p)$. To decrypt the ciphertext, Bob calculates $c_1^{-b} c_2 \equiv m \pmod{p}$.

(a) Suppose you intercept a ciphertext (c_1, c_2) that Alice has encrypted for Bob, which is the encryption for some message m. Explain how to construct a ciphertext (c_1', c_2') which is the encryption of $2m$.

(b) Suppose you intercept two ciphertexts (c_1, c_2) and (c_1', c_2') that Alice has encrypted for Bob. Assume they are encryptions of some unknown messages m_1 and m_2. Show how you can construct a ciphertext which is a valid El Gamal encryption of the message $m_1 \cdot m_2 \bmod p$.

(c) Consider a new scheme where the value r is not generated randomly every time. Instead, Alice begins by randomly generating an initial value r_0, and then simply incrementing r_0 by 1 every time she needs to encrypt another message. Is the resulting encryption scheme IND-CPA?

Problem 7 ***Feedback*** **(0 points)**

Optionally, feel free to include feedback. What's the single thing we could do to make the class better? Or, what did you find most difficult or confusing from lectures or the rest of class, and what would you like to see explained better?