CS255: Cryptography and Computer Security **Winter 2013**

Final Exam

Instructions:
– **Answer all five questions.**
– The exam is open book and open notes. Wireless devices are not allowed.
– Students are bound by the Stanford honor code.
– You have two hours and thirty minutes.

Problem 1. Questions from all over.

a. Let p be a large prime and $g \in \mathbb{Z}_p^*$ of order $p-1$. Is the function $f(x) = g^x$ in $\mathbb{Z}_p$ whose domain is $\{1, \ldots, p-1\}$ a *trapdoor* one-way function? Justify your answer.

b. Briefly explain the main idea for building an authenticated key exchange protocol (secure against man in the middle attacks) from the basic Diffie-Hellman protocol.

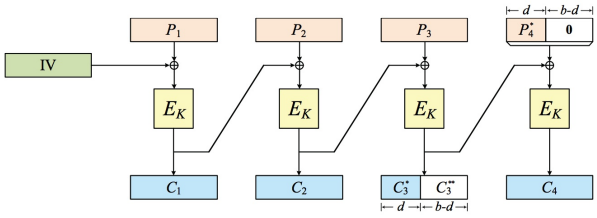
c. Let (N, e) be an RSA public key. Recall that to sign messages using RSA-FDH we use a hash function $H : \{0,1\}^n \to \mathbb{Z}_N$ and compute the signature on a message m as $\sigma \leftarrow H(m)^d$ in $\mathbb{Z}_N$. Suppose the adversary can find two messages m_1, m_2 such that $H(m_1) = H(m_2) \cdot 2^e$ in $\mathbb{Z}_N$. Does this let the adversary break RSA-FDH? That is, can the adversary create an existential forgery using a chosen message attack?

d. Same question as part (c) except that the adversary can find two messages m_1, m_2 such that $H(m_1) = H(m_2) + 1$ in $\mathbb{Z}_N$. Briefly justify your answer.

e. When storing hashed and salted passwords in a password file, what is the purpose of using a slow hash function?

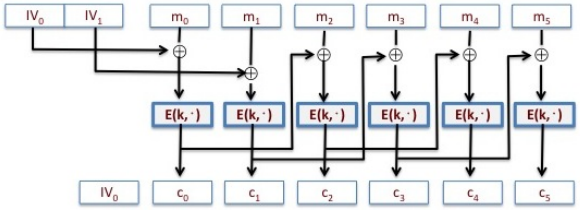
Problem 2. Variants of CBC encryption.

a. One problem with CBC encryption is that messages need to be padded to a multiple of the block length and sometimes a dummy block needs to be added. The following figure describes a variant of CBC that eliminates the need to pad:



The method pads the last block with zeros if needed (a dummy block is never added), but the output ciphertext contains only the shaded parts of C_1, C_2, C_3, C_4. Note that, ignoring the IV, the ciphertext is the same length as the plaintext. This technique is called *ciphertext stealing*. (1) Explain how decryption works. (2) Can this method be used if the plaintext contains only one block?

b. Another problem with CBC encryption is that it cannot be sped up by parallel processing. The following figure shows a variant of CBC that supports 2-way parallelism. It can be sped up by a factor of two using two processors.



Here E is the encryption algorithm of a secure PRP such as AES. Suppose one chooses IV_0 at random and sets $\mathrm{IV}_1 = \mathrm{IV}_0 \oplus B$ for some fixed public constant B (e.g. $B = 1^n$ where n is the block size of E). Is the resulting system CPA-secure? If yes, briefly explain why (it's fine to rely on theorems presented in class). If not, describe an attacker that wins the CPA security game.

c. Suppose one chooses IV_0 at random and sets $\mathrm{IV}_1 = k'$ where k' is part of the secret key. That is, the secret key is (k, k') and this secret key is used to encrypt multiple plaintexts. Is the resulting system CPA-secure? If yes, briefly explain why (it's fine to rely on theorems presented in class). If not, describe an attacker that wins the CPA security game.

d. Suppose one chooses IV_0 and IV_1 independently at random and includes both in the ciphertext. Is the resulting system CPA-secure? If yes, briefly explain why (it's fine to rely on theorems presented in class). If not, describe an attacker that wins the CPA security game.

Problem 3. Let (E, D) be an encryption system that provides authenticated encryption. Here E does not take a nonce as input and therefore must be a randomized encryption algorithm. Which of the following systems provide authenticated encryption? For those that do briefly explain why. For those that do not, present an attack that either breaks CPA security or ciphertext integrity.

a. $E_1(k,m) = \big[c \leftarrow E(k,m),\ \text{output}\ \ (c,c)\ \big]$ and $D_1(k,\ (c_1,c_2)\ ) = D(k,c_1)$

b. $E_2(k,m) = \big[c \leftarrow E(k,m),\ \text{output}\ \ (c,c)\ \big]$ and $D_2(k,\ (c_1,c_2)\ ) = \begin{cases} D(k,c_1) & \text{if } c_1 = c_2 \\ \mathsf{fail} & \text{otherwise} \end{cases}$

c. $E_3(k,m) = \big(\ E(k,m),\ E(k,m)\ \big)$ and $D_3(k,\ (c_1,c_2)\ ) = \begin{cases} D(k,c_1) & \text{if } D(k,c_1) = D(k,c_2) \\ \mathsf{fail} & \text{otherwise} \end{cases}$

To clarify: $E(k,m)$ is randomized so that running it twice on the same input will result in different outputs with high probability.

d. $E_4(k,m) = \big(\ E(k,m),\ H(m)\ \big)$ and $D_4(k,\ (c_1,c_2)\ ) = \begin{cases} D(k,c_1) & \text{if } H(D(k,c_1)) = c_2 \\ \mathsf{fail} & \text{otherwise} \end{cases}$

where H is a collision resistant hash function.

Problem 4. Two-time secure encryption. Recall that the one-time-pad is a one-time encryption system that is secure against infinitely powerful adversaries. Our goal in this question is to design a *2-time* secure encryption against infinitely powerful adversaries. If the encryptor can be stateful then the problem is trivial — simply use two one-time pads. Here, we design a stateless 2-time secure system: every encryption is done independently of the other encryptions.

a. Give a precise definition for what it means for a symmetric encryption system to be semantically secure when a secret key is used to encrypt at most two messages. Make sure to define two experiments EXP(0) and EXP(1) as in the definitions of semantic security and CPA security. Keep in mind that the adversary can be adaptive: its choice for a second message to encrypt may depend on the first ciphertext it receives.

b. Show that the one time pad is insecure under your definition from part (a). Can any deterministic encryption system (without a nonce) be secure under your definition?

c. Let p be a 128-bit prime and consider the following encryption system: the secret key is a random pair $(x, y) \in (\mathbb{Z}_p)^2$ and to encrypt a message $m \in \mathbb{Z}_p$ do:

$$\text{choose a random } r \xleftarrow{R} \mathbb{Z}_p \text{ and output the ciphertext } (r,\ xr + y + m) \in \mathbb{Z}_p^2.$$

Explain how to decrypt a given ciphertext (c_1, c_2) using the secret key (x, y).

d. Show that this system is 2-time secure (using your definition from part a.) against infinitely powerful adversaries.

d.1. Let S be the set of all 4-tuples (x, y, r_0, r_1) in $\mathbb{Z}_p^4$ such that $r_0 \neq r_1$. First argue that if the tuple (x, y, r_0, r_1) is uniform in S then the tuple $(xr_0 + y, xr_1 + y, r_0, r_1)$ is also uniform in S. To do so, show that the following mapping from S to S is one-to-one:

$$(x, y, r_0, r_1) \to (xr_0 + y,\ xr_1 + y,\ r_0,\ r_1)$$

All you need to do is show that this mapping is invertible.

d.2. Use (d.1) to argue 2-time security. In particular, show that the adversary's advantage is at most $1/p$ in distinguishing EXP(0) from EXP(1).
Hint: There are two cases:

- First argue that if the the nonces r_0, r_1 in the two ciphertexts given to the adversary are distinct then the adversary has advantage 0 in distinguishing EXP(0) from EXP(1). To show this observe that the property from (d.1) implies that when encrypting two messages m_0 and m_1 in $\mathbb{Z}_p$ with *distinct* nonces $r_0 \neq r_1$ the resulting ciphertexts $(r_0,\ xr_0 + y + m_0)$ and $(r_1,\ xr_1 + y + m_1)$ are distributed as (r_0, s_0) and (r_1, s_1) where s_0, s_0 are uniform random variables in $\mathbb{Z}_p$ *independent* of m_0 and m_1 (i.e. (s_0, s_1) are distributed the same for all (m_0, m_1)).
- if the nonces r_0, r_1 in the two ciphertexts given to the adversary are the same then the adversary can distinguish EXP(0) from EXP(1). However, observe that $r_0 = r_1$ happens only with probability $1/p$.

e. Show that the system from part (c) is not 3-time secure. That is, show that the adversary distinguish EXP(0) from EXP(1) after making three chosen plaintext queries.

Problem 5. Encryption-based challenge-response identification. In class we discussed MAC-based and signature-based challenge-response identification. Recall that the purpose of challenge-response identification is to defeat attackers capable of mounting an active attack on the identification system. In this question we consider a variant of challenge-response identification based on an encryption scheme rather than a MAC.

Let (E, D) be a symmetric encryption system defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$. The identification system works as follows:

- **setup**: generate a random key $k \in \mathcal{K}$ and set sk $\leftarrow k$ and vk $\leftarrow k$. The same key k will be used for all runs of the identification protocol.
- **identification**: the verifier generates a random message $m \in \mathcal{M}$ and sends $c \leftarrow E(k, m)$ to the prover. The prover responds with $m' \leftarrow D(k, c)$. The verifier accepts if $m = m'$ and rejects otherwise.

a. For each of the following encryption schemes determine if this identification method is secure. If it is secure explain why. If not, present an attack.

a.1. (E, D) is the one-time pad with $\mathcal{K} = \mathcal{M} = \mathcal{C} = \{0,1\}^{128}$.

a.2. (E, D) is AES-based CBC encryption with a random IV where the message space $\mathcal{M}$ is $\{0,1\}^{128}$.

a.3. (E, D) is AES-based GCM encryption where the message space $\mathcal{M}$ is $\{0,1\}^{128}$.

b. Suppose the key k is derived from the user's password (i.e. k is computed via a public deterministic function applied to the password). Can an eavesdropper who obtains the transcript of a successful identification carry out a dictionary attack to expose k? If so explain why. If not, explain why not.

c. As is, the protocol assumes that vk is kept secret on the server. Can you propose a modification to the protocol so that making vk public would not affect security?