CS6111: Foundations of Cryptography

Assignment 1

Instructions

- Deadline is Monday, August 19.
- We encourage submissions by Latex. Paper is also accepted.
- Proofs, where required, must be formal. All answers are concise.
- The Addendum provides information that explains the method behind some questions.

References

- Introduction to Cryptography - Delfs and Knebl
- A Graduate Course in Applied Cryptography - Boneh and Shoup (link)
- Introduction to Modern Cryptography - Katz and Lindell
- Cryptography: Theory and Practice - Douglas Stinson (3rd edition)
- Handout 1 - Entropy Axioms
  - It gives an overview of entropy with many definitions. Page 7 starts talking about the axioms.

1 Perfect Secrecy

1. (8 points) We defined Shannon's entropy in class and mentioned its uniqueness. "Handout1 - Entropy Axioms" shows how the uniqueness of the definition can be derived from axioms.
   (a) (2 points) Explain, with examples of distributions, how Shannon's entropy captures the amount of information in bits we obtain after an event.
   (b) (2 points) Define conditional entropy and relative entropy. Justify why their definitions make sense.
   (c) (4 points) Justify both sets of **axioms** for Shannon's entropy (pg 8,9). Explain how you believe they each capture the essential properties of entropy we require.

   Hint: These "justify" questions do not require rigorous proof and are slightly open-ended in that sense. We expect each axiom to be explained and how it relates to the information-theoretic notions detailed in parts (a), (b). Do not simply rewrite the axioms.

Solution:

2. (5 points) Suppose a cryptosystem (E, D) achieves perfect secrecy for a particular distribution $P = (p_1, \ldots, p_n)$ on a plaintext set $M = \{m_1, \ldots, m_n\}$. Prove that the system is also perfectly secret for all other probability distributions P' on M.

   Hint: Understand the definitions carefully.

3. (5 points) What is the affine cipher? Show that the number of keys (a, b) when $m = 26$ is 312.

   (a) (2 points) Prove prefect secrecy when each key is chosen with equal probability.

   (b) (2 points) Suppose instead that we choose (only valid values of) a according to a given probability distribution P and b is chosen randomly. Prove that perfect secrecy still holds.

   Hint: Note that the message space is just single characters from the alphabet. The rest should follow from the definitions of perfect secrecy.

4. (5 points) Let (Enc, Dec) be a perfectly-secure scheme. Define a new scheme $Enc'((k_1, k_2), m) = (Enc(k_1, k_2), Enc(k_2, m))$. Prove that Enc' is perfectly-secure.

Variations to Perfect Security

Let (Enc, Dec) be an encryption scheme on (M, K, C).

Indistinguishability Attack Game - Experiment b:

1. Attacker Eve chooses $m_0, m_1 \in M$.
2. Challenger chooses a random $k \in K$ and gives Eve $c = Enc_k(m_b)$.
3. Eve outputs her guess b'.

Eve wins if $b' = b$.

Advantage: Let W_b be the probability that Eve outputs 1 in Experiment b. The semantic-security advantage of the attacker is defined as $\mathsf{SSAdv}(Eve, Enc) = |W_0 - W_1|$.

- If the scheme is perfectly secure, then the probability that Eve wins is $1/2$. Equivalently, Eve's advantage is 0.

5. (5 points) Consider a variant of the one-time pad where $M = C = \{0,1\}^L$ but the keyspace $K \subset \{0,1\}^L$ is restricted to strings with an even number of 1s. Prove that this version is not perfectly secure and an adversary can win the indistinguishability game with probability 1.

   Hint: Design a strategy where Eve chooses m_0, m_1 such that, knowing the challenger uses the modified keyspace, Eve will always be able to distinguish correctly.

6. (5 points) Define an encryption scheme to be ϵ-perfectly secure if the probability that Eve wins the game is $1/2 + \epsilon$. Show that ϵ-perfect secrecy can be achieved with $|K| < |M|$ when $\epsilon > 0$. Prove a lower bound on the size of K in terms of M, ϵ.

Hint: This means that Eve can choose m_0, m_1 such that there is an ϵ additional probability that, seeing c (and knowing nothing about k), she can figure out which m_i was used to encrypt c. What does this say about the mapping? (See Addendum)

7. (5 points) Assume that we only require an encryption scheme Enc, Dec to satisfy the following: For all $m \in M$ we have $\mathsf{Pr}[Dec_k(Enc_k(m)) = m] \geq 2^{-t}$ (for some parameter t). The probability is taken over the choice of key and randomness used in the encryption algorithm. Show that perfect secrecy can now be achieved with $|K| < |M|$ when $t \geq 1$. Prove a lower bound on the size of K in terms of M and t.

   Hint: After seeing c, even with the key k used in the encryption $c = E_k(m^*)$, there is only a $1/2^t$ probability that the now-probabilistic decryption algorithm gives the intended m^*. Think about this as a table mapping (described in the Addendum).

> Let (Enc, Dec) be an encryption scheme on (M, K, C).
>
> **Parity Prediction Game**:
>
> 1. Challenger chooses a random $m \in M$ and $k \in K$ and gives Eve $c = Enc_k(m)$.
> 2. Eve guesses $parity(m)$ and wins if she is correct.
>
> **Advantage:** Let W be the probability that Eve outputs the correct parity. The advantage of the attacker is defined as $\mathsf{ParityAdv}(Eve, Enc) = |W - 1/2|$.

8. (5 points) Consider the above Parity Prediction Game. Prove, by reduction, that a semantically secure scheme is also secure against parity prediction: that is, show that if there exists an attacker with advantage ϵ in predicting parity, then we can get advantage 2ϵ in the Indistinguishability game.

   Hint: This is known as a blackbox reduction, which you will see all over cryptography. We have one attacker E who can (only) win the parity prediction game. We want to create another attacker E' who can win the Indistinguishability game. E' will use E as a **black box***: that is, E' now acts as a Parity-Challenger to E and can choose any m it desires in step 1. It receives the guess of E from step 2. Now, like in Question 5, design a strategy for E' to adversarially choose m_0, m_1 to win the indistinguishability game against a challenger.*

2 Randomized Algorithms Warmup

1. (4 points) Start with rolling an n sided die. If it lands on $r > 1$, roll an r-sided die and repeat. If it lands on $r = 1$, then halt. What is the expected number of rolls before halting?

   Hint: Obtain a recurrence relation and solve it. Solving it involves algebraic manipulation.

2. (3 points) Prove the following "success amplification" theorem for Monte Carlo randomized algorithms (one whose output might be incorrect but will always halt in polynomial time).

   Let P and Q be polynomials and A be a probabilistic algorithm computing a function f such that: $\Pr[A(x) = f(x)] > \frac{1}{2} + \frac{1}{P(|x|)}$. *Then, by repeating the computation, prove that you can amplify the success probability to* $\Pr[A(x) = f(x)] > 1 - \frac{1}{Q(|x|)}$. *How many repetitions does it take?*

   Hint: Bernouilli trials.

3 Addendum

What is an encryption scheme?

Questions 6 and 7 ask you to provide encryption schemes that satisfy modified security descriptions. An encryption scheme is simply a table listing all possible messages $m \in M$, all possible ciphertexts $c \in C$, and a valid key $k \in K$ connecting a given m and c such that $E_k(m) = c$. That is, (m, k, c) is an entry in this table. This table (encryption scheme) is public. But the key k used by two parties is private to them. Note that this table also abstracts the randomness used in the encryption (and decryption) algorithms.

- For example, the One-Time Pad is described as an algorithm using the $\oplus$ operator. But it can equivalently be represented as a table over $M = K = C = \{0,1\}^l$. It includes all entries of the form $(m, k, m \oplus k)$ for all m, k.

So in your answer you can simply describe a mapping between M and C using keys K. How many messages are connected to how many ciphertexts via how many keys? Ensure that every m and k has a corresponding c. Using this, prove the required security property.

- Hence, you need not necessarily use the One-Time Pad or detail a specific encryption algorithm.