Lecture 06: Private-key Encryption (Definition & Security of One-time Pad)

Objective

- First, we shall define the correctness and the security of private-key encryption schemes
- We shall argue that the one-time pad is correct and secure

Private-key Encryption

- Three algorithms
  - Key Generation: Generate the secret key sk
  - Encryption: Given the secret key sk and a message m, it outputs the cipher-text c (Note that the encryption algorithm can be a randomized algorithm)
  - Decryption: Given the secret key sk and the cipher-text c, it outputs a message m' (Note that the decryption algorithm can be a randomized algorithm)

Story of the Private-key Encryption Process

- Yesterday Alice and Bob met and generated a secret key $\mathsf{sk} \sim \mathsf{Gen}()$
  - Read as: the secret key sk is sampled according to the distribution $\mathsf{Gen}()$
- Today Alice wants to encrypt a message m using the secret key sk. Alice encrypts $c \sim \mathsf{Enc}_{\mathsf{sk}}(m)$
  - Read as: the cipher-text c is sampled according to the distribution $\mathsf{Enc}_{\mathsf{sk}}(m)$
- Then Alice sends the cipher-text c to Bob. An eavesdropper gets to see the cipher-text c
- After receiving the cipher-text c Bob decrypts it using the secret key sk. Bob decrypts $m' \sim \mathsf{Dec}_{\mathsf{sk}}(c)$
  - Read as: the decoded message m' is sampled according to the distribution $\mathsf{Dec}_{\mathsf{sk}}(c)$

Correctness

- We want the decoded message obtained by Bob to be identical to the original message of Alice with high probability
- We insist

$$\mathbb{P}\left[\mathbb{M} = \mathbb{M}'\right] = 1$$

  - Recall we use capital alphabets to represent the random variable corresponding to the variable (so, $\mathbb{M}$ is the random variable for the message encoded by Alice and $\mathbb{M}'$ is the random variable for the message recovered by Bob)

Security

- We want to say that the cipher-text c provides the adversary no additional information about the message
- We insist that, for all message m, we have

$$\mathbb{P}\left[\mathbb{M} = m | \mathbb{C} = c\right] = \mathbb{P}\left[\mathbb{M} = m\right]$$

Cropping any Constraint makes the Problem Trivial

- Suppose we insist only on correctness and not on security
  - The trivial scheme where $\mathsf{Enc}_{\mathsf{sk}}(m) = m$, i.e. the encryption of any message m using any secret key sk is the message itself, satisfies correctness. But is completely insecure!
- Suppose we insist only on security and not on correctness
  - The trivial scheme where $\mathsf{Enc}_{\mathsf{sk}}(m) = 0$, i.e. the encryption of any message m using any secret key sk is 0, satisfies this security. But Bob cannot correctly recover the original message m with certainty!
- So, the non-triviality is to simultaneously achieve correctness and security

One-time Pad

- Let $(G, \circ)$ be a group
- Secret-key Generation:

$\mathsf{Gen}()$:   • Return $\mathsf{sk} \xleftarrow{\$} G$

- Encryption:

$\mathsf{Enc}_{\mathsf{sk}}(m)$:   • Return $c := m \circ \mathsf{sk}$

- Decryption:

$\mathsf{Dec}_{\mathsf{sk}}(c)$:   • Return $m' := c \circ \mathsf{inv}(\mathsf{sk})$

- Note that Encryption and Decryption is deterministic
- The only randomized step is the choice of sk during the secret-key generation algorithm

Correctness of One-time Pad

- It is trivial to see that

$$\mathbb{P}\left[\mathbb{M} = \mathbb{M}'\right] = 1$$

- So, one-time pad is correct!

Security of One-time Pad I

- We want to simplify the probability

$$\mathbb{P}\left[\mathbb{M} = m | \mathbb{C} = c\right]$$

- Using Bayes' Rule, we have

$$= \frac{\mathbb{P}\left[\mathbb{M} = m, \mathbb{C} = c\right]}{\mathbb{P}\left[\mathbb{C} = c\right]}$$

- Using the fact that $\mathbb{P}\left[\mathbb{C} = c\right] = \sum_{x \in G} \mathbb{P}\left[\mathbb{M} = x, \mathbb{C} = c\right]$, we get

$$= \frac{\mathbb{P}\left[\mathbb{M} = m, \mathbb{C} = c\right]}{\sum_{x \in G} \mathbb{P}\left[\mathbb{M} = x, \mathbb{C} = c\right]}$$

Security of One-time Pad II

- We will prove the following claim later

Claim

For any $x, y \in G$, *we have*

$$\mathbb{P}\left[\mathbb{M} = x, \mathbb{C} = y\right] = \mathbb{P}\left[\mathbb{M} = x\right] \cdot \frac{1}{|G|}$$

- Using this claim, we can simplify the expression as

$$\begin{aligned} &= \frac{\mathbb{P}\left[\mathbb{M} = m\right] \cdot \frac{1}{|G|}}{\sum_{x \in G} \mathbb{P}\left[\mathbb{M} = x\right] \cdot \frac{1}{|G|}} \\ &= \frac{\mathbb{P}\left[\mathbb{M} = m\right]}{\sum_{x \in G} \mathbb{P}\left[\mathbb{M} = x\right]} \end{aligned}$$

Security of One-time Pad III

- Using the fact that $\sum_{x \in G} \mathbb{P}[\mathbb{M} = x] = 1$, we get that the previous expression is

$$= \mathbb{P}[\mathbb{M} = m]$$

- This proves that $\mathbb{P}\left[\mathbb{M} = m | \mathbb{C} = c\right] = \mathbb{P}[\mathbb{M} = m]$, for all m and c. This proves that the one-time pad encryption scheme is secure!

Proof of Claim 1

- You will prove the following statement in the homework: If there exists sk such that $x \circ \mathsf{sk} = y$ then sk is unique (i.e., there does not exist $\mathsf{sk}' \neq \mathsf{sk}$ such that $x \circ \mathsf{sk}' = y$)
- Using this result, we get the following. Suppose $z \in G$ be the unique element such that $x \circ z = y$. Then we have:

$$\mathbb{P}\left[\mathbb{M} = x, \mathbb{C} = y\right] = \mathbb{P}\left[\mathbb{M} = x, \mathbb{SK} = z\right]$$

- Note that the secret-key is sample independent of the message x. So, we have

$$\mathbb{P}\left[\mathbb{M} = x, \mathbb{SK} = z\right] = \mathbb{P}\left[\mathbb{M} = x\right] \cdot \mathbb{P}\left[\mathbb{SK} = z\right]$$

- Note that sk is sampled uniformly at random from the set G. So, we have

$$\mathbb{P}\left[\mathbb{M} = x, \mathbb{SK} = z\right] = \mathbb{P}\left[\mathbb{M} = x\right] \cdot \frac{1}{|G|}$$

Example I

- Encrypting bit messages
  - Consider $(G, \circ) = (\mathbb{Z}_2, + \mod 2)$

Example II

- Encrypting n-bit strings
  - Consider $G = \{0,1\}^n$
  - Define $(x_1, \ldots, x_n) \circ (y_1, \ldots, y_n) = (x_1 + y_1 \mod 2, \ldots, x_n + y_n \mod 2)$

Example III

- Encrypting an alphabet
  - Consider $G = \mathbb{Z}_{26}$
  - Define $\circ$ as $+ \mod 26$
- You will construct one more scheme in the homework by interpreting the set of alphabets as $\mathbb{Z}_{27}^*$

Example IV

- Encrypting *n*-alphabet words
  - Consider $G = \mathbb{Z}_{26}^{n}$
  - Define $\circ$ as the coordinate-wise $+ \mod 26$