Tutorial 3: Security against Chosen Plaintext Attacks

Submission Guidelines All problems must be solved in class today. Searching on the internet for solutions is strictly discouraged.

1. (a) Formally define (based on a security game) a notion called perfect secrecy under chosen plaintext attacks or perfect IND-CPA-security for symmetric encryption schemes (similar to perfect secrecy in the presence of an eavesdropper).

A: We first define a game called p-ind-cpa between a challenger and an adversary $\mathscr{A}$.

Setup: The challenger generates a key $k \xleftarrow{R} \mathsf{Gen}()$ and initiates the game with $\mathscr{A}$.

Encryption Queries: $\mathscr{A}$ makes encryption queries *adaptively* – for a query on message m, the challenger computes $c = \mathsf{Enc}(k, m)$ and returns the resulting ciphertext to $\mathscr{A}$.

Challenge: $\mathscr{A}$ provides two (equal-length) messages m_0, m_1. The challenger generates $\beta \xleftarrow{U} \{0,1\}$, computes $c^* = \mathsf{Enc}(k, m_\beta)$ and sends the challenge ciphertext c^* to $\mathscr{A}$.

Guess: $\mathscr{A}$ returns its guess β' of β.

$\mathscr{A}$ wins if $\beta = \beta'$. Let $\mathsf{Adv}^{\mathsf{p\text{-}ind\text{-}cpa}}_{\mathscr{A},\mathcal{E}} = \Pr[\beta = \beta']$. Define $\mathcal{E}$ to be t-P-IND-CPA secure if for every adversary $\mathscr{A}$ running in time at most t,

$$\mathsf{Adv}^{\mathsf{p\text{-}ind\text{-}cpa}}_{\mathscr{A},\mathcal{E}} = \frac{1}{2}.$$

(b) Show that there cannot exist an encryption scheme that achieves perfect IND-CPA security.

A: Let $\mathcal{E} = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ be an encryption scheme. We may assume that Enc is randomised for otherwise there is a trivial attack in the above model – just query the encryption of m_0 and test if the resulting ciphertext is c^*. Furthermore, suppose that Enc uses k random coins i.e., let the randomness used by Enc be a k-bit string. Consider the following strategy: after the challenge phase, $\mathscr{A}$ makes an encryption query on m_0. Suppose that the challenge ciphertext is encrypted with random coins r^* and the query is answered with a ciphertext generated using random coins r. If r happens to be equal to r^*, then $\mathscr{A}$ can just compare the two ciphertexts and determine the value of β. Otherwise it guesses β at random. Let E denote the event that $r = r^*$. We have,

$$\begin{aligned}
\Pr[\beta = \beta'] &= \Pr[\beta = \beta' | \beta = 0] \Pr[\beta = 0] + \Pr[\beta = \beta' | \beta = 1] \Pr[\beta = 1] \\
&= \frac{1}{2}(\Pr[\beta = \beta' | \beta = 0] + \frac{1}{2}) \\
&= \frac{1}{2}(\Pr[\beta = \beta' | \beta = 0, \mathsf{E}] \Pr[\mathsf{E}] + \Pr[\beta = \beta' | \beta = 0, \neg\mathsf{E}] \Pr[\neg\mathsf{E}]) + \frac{1}{4} \\
&= \frac{1}{2}\left(1 \cdot \frac{1}{2^k} + \frac{1}{2} \cdot \left(1 - \frac{1}{2^k}\right)\right) + \frac{1}{4} \\
&= \frac{1}{2} + \frac{1}{2^{k+2}} \\
&> \frac{1}{2}
\end{aligned}$$

Therefore, $\mathcal{E}$ is not P-IND-CPA-secure. Since this holds for an arbitrary encryption scheme, there exists no encryption scheme that is P-IND-CPA-secure.

2. Let $\mathcal{E}_1 = (\mathsf{Gen}_1, \mathsf{Enc}_1, \mathsf{Dec}_1)$ and $\mathcal{E}_2 = (\mathsf{Gen}_2, \mathsf{Enc}_2, \mathsf{Dec}_2)$ be two symmetric encryption schemes over $\mathcal{M}_1, \mathcal{K}_1, \mathcal{C}_1$ and $\mathcal{M}_2, \mathcal{K}_2, \mathcal{C}_2$ respectively. You know that one of them is IND-CPA-secure and the other is not but do not know which one is actually IND-CPA-secure. Show how to build an encryption scheme

$\mathcal{E} = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ from $\mathcal{E}_1$ and $\mathcal{E}_2$ that is guaranteed to be IND-CPA-secure. You must prove that your construction is indeed IND-CPA-secure. Assume that $\mathcal{M}_1 = \mathcal{M}_2 = \{0,1\}^n$.

A: Let $\mathcal{E}$ be defined as follows. Set $\mathcal{M} = \{0,1\}^n$; $\mathcal{K} = \mathcal{K}_1 \times \mathcal{K}_2$; $\mathcal{C} = \mathcal{C}_1 \times \mathcal{C}_2$.

$\mathsf{Gen}()$**:** The key is given by $k = (k_1, k_2)$ where $k_1 \xleftarrow{R} \mathsf{Gen}_1()$ and $k_2 \xleftarrow{R} \mathsf{Gen}_2()$.

$\mathsf{Enc}(k, m)$ Choose $r \xleftarrow{U} \{0,1\}^n$. Compute ciphertext as $c = (c_1, c_2)$ where

$$c_1 = \mathsf{Enc}_1(k_1, m \oplus r), \qquad c_2 = \mathsf{Enc}_2(k_2, r).$$

$\mathsf{Dec}(k, c)$**:** Let $c = (c_1, c_2)$. Recover the message as

$$m = \mathsf{Dec}_2(k_2, c_2) \oplus \mathsf{Dec}_1(k_1, c_1).$$

It is rather straightforward to show the correctness of the above scheme.

The underlying idea for the construction is to split the message into two parts, neither of which reveal any information about the message but combined together completely reveal the message. We now need to show that if one of the two schemes $\mathcal{E}_1, \mathcal{E}_2$ is IND-CPA-secure, then $\mathcal{E}$ is IND-CPA-secure.

Lemma 0.1. *Suppose that $\mathcal{E}_1$ is (q, t, ε)-IND-CPA-secure. Then $\mathcal{E}$ is (q, t', ε)-IND-CPA-secure where $t' = t - O(q)$.*

Proof To prove this, we show that an adversary $\mathscr{A}$ for $\mathcal{E}$ can be used to build an adversary $\mathscr{B}_1$ for $\mathcal{E}_1$ with the same advantage.

Description of $\mathscr{B}_1$

- $\mathscr{B}_1$ first generates $k_2 \xleftarrow{R} \mathsf{Gen}_2()$ while its challenger has generated $k_1 \xleftarrow{R} \mathsf{Gen}_1()$ and initiates the ind-cpa game with $\mathscr{A}$.
- On an encryption query on a message m from $\mathscr{A}$, $\mathscr{B}_1$ does the following: pick $r \xleftarrow{U} \{0,1\}^n$, query its challenger for an encryption of $m \oplus r$; on receiving $c_1 = \mathsf{Enc}_1(k_1, m \oplus r)$, compute $c_2 \leftarrow \mathsf{Enc}_2(k_2, r)$ and return (c_1, c_2) to $\mathscr{A}$.
- $\mathscr{A}$ provides two messages m_0, m_1 in the challenge phase. $\mathscr{B}_1$ chooses $r^* \xleftarrow{U} \{0,1\}^n$ and provides $m_0 \oplus r^*, m_1 \oplus r^*$ to its challenger. $\mathscr{B}_1$'s challenger picks $\beta \xleftarrow{R} \{0,1\}$, computes $c_1^* \leftarrow \mathsf{Enc}_1(k_1, m_\beta \oplus r^*)$ and sends it to $\mathscr{B}_1$. Upon receiving the challenge ciphertext c_1^*, $\mathscr{B}_1$ computes $c_2^* = \mathsf{Enc}_2(k_2, r^*)$ and returns $c^* = (c_1^*, c_2^*)$ to $\mathscr{A}$.
- $\mathscr{A}$ sends its guess β' of β and $\mathscr{B}_1$ just relays β' to its challenger.

Clearly $\Pr[\mathscr{B}_1 \text{ wins}] = \Pr[\mathscr{A} \text{ wins}]$ from which it follows that they have the same advantage in winning the respective games. Since $\mathscr{B}_1$ has to compute q encryptions using the Enc_2 algorithm, we have $t = t' + O(q)$. □

Lemma 0.2. *Suppose that $\mathcal{E}_2$ is (q, t, ε)-IND-CPA-secure. Then $\mathcal{E}$ is (q, t', ε)-IND-CPA-secure where $t' = t - O(q)$.*

Proof Again, we show this building an adversary $\mathscr{B}_2$ against $\mathcal{E}_2$ using an adversary $\mathscr{A}$ against $\mathcal{E}$ with the same advantage.

Description of $\mathscr{B}_2$

- $\mathscr{B}_2$ first generates $k_1 \xleftarrow{R} \mathsf{Gen}_1()$ while its challenger has generated $k_2 \xleftarrow{R} \mathsf{Gen}_2()$; $\mathscr{B}_2$ initiates the ind-cpa game with $\mathscr{A}$.
- On an encryption query on a message m from $\mathscr{A}$, $\mathscr{B}_2$ does the following: pick $r \xleftarrow{U} \{0,1\}^n$, query its challenger for an encryption of r; on receiving $c_2 = \mathsf{Enc}_2(k_2, r)$, compute $c_1 \leftarrow \mathsf{Enc}_1(k_1, m \oplus r)$ and return (c_1, c_2) to $\mathscr{A}$.

- $\mathscr{A}$ provides two messages m_0, m_1 to $\mathscr{B}_2$ in the challenge phase. $\mathscr{B}_2$ chooses $\hat{r} \xleftarrow{\text{U}} \{0,1\}^n$ and provides $m_0 \oplus \hat{r}, m_1 \oplus \hat{r}$ to its challenger. $\mathscr{B}_2$'s challenger picks $\beta \xleftarrow{\text{R}} \{0,1\}$, sends $c_2^* \leftarrow \mathsf{Enc}_2(k_2, m_\beta \oplus \hat{r})$ to $\mathscr{B}_2$. Upon receiving the challenge ciphertext c_2^*, $\mathscr{B}_2$ computes $c_1^* = \mathsf{Enc}_1(k_1, \hat{r})$ and returns $c^* = (c_1^*, c_2^*)$ to $\mathscr{A}$. Here, $\mathscr{B}_2$ implicitly sets $r^* = \hat{r} + m_\beta$ and does not know the value of r^* itself. Since $\hat{r}$ is uniformly distributed in $\{0,1\}^n$, so is r^* thus implying that the ciphertext would be distributed as in the real security game.
- $\mathscr{A}$ sends its guess β' of β and $\mathscr{B}_2$ just relays β' to its challenger.

Clearly $\Pr[\mathscr{B}_2 \text{ wins}] = \Pr[\mathscr{A} \text{ wins}]$ from which it follows that they have the same advantage in winning the respective games. Since $\mathscr{B}_2$ has to compute q encryptions using the Enc_1 algorithm, we have $t = t' + O(q)$. □