CSA E0 235: Cryptography March 26, 2015

Chalk & Talk 1 : Semantic Security and Indistinguishability

Instructor: Arpita Patra *Submitted by: Arshed Nabeel*

1 Formal Definitions of Security

Recall the notions of *semantic security* and *indistinguishability* of encryption schemes.

Semantic security tries to capture the notion that, no matter what prior information the adversary has, the knowledge of the ciphertext will not afford him any additional knowledge.

Definition 1.1. *An encryption scheme* $\Pi = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *is* semantically secure *(under ciphertext-only attack) if, for every PPT algorithm cA, and polynomial-time computable functions* $f(\cdot)$ *and* $g(\cdot)$*, there exists another PPT algorithm* $\mathcal{A}'$ *such that*

$$\left| \mathsf{Pr}\left[\mathcal{A}(1^n, \mathsf{Enc}_k(m), h(m)) = f(m)\right]_{\substack{k \leftarrow \mathsf{Gen}(1^n) \\ m \in_R \mathcal{M}}} - \mathsf{Pr}\left[\mathcal{A}'(1^n, |m|, h(m)) = f(m)\right]_{m \in_R \mathcal{M}} \right|$$

is negligible.

However, this definition is technical and is clumsy to work with while proving security of cryptosystems. Thus, we use an alternate definition, based on indistinguishabiity of messages, as our working definition.

Definition 1.2. *An encryption scheme* $\Pi = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *is* indistinguishable *under ciphertext-only attack if, for every probabilistic polynomial time algorithm* $\mathcal{A}$ *and two arbitrary messages* m_0, m_1*,*

$$\mathsf{Pr}[\mathcal{A}(1^n, \mathsf{Enc}_k(m_b)) = b] \leq \frac{1}{2} + \epsilon(n)$$

where ϵ *is a negligible function.*

2 Equivalence of the Security Definitions

Theorem 2.1. *Definitions 1.1 and 1.2 are equivalent: i.e., an encryption scheme is semantically secure under a ciphertext-only attack if and only if it is indistinguishable for any cipher-text only adversary.*

Claim 2.1. *If an encryption scheme* $\Pi = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *is indistinguishable for all ciphertext-only adversaries, then* Π *is semantically secure under ciphertext-only attack.*

Proof. Consider an arbitrary message m, $k \leftarrow \mathsf{Gen}(1^n)$, $c = \mathsf{Enc}_k(m)$ and some polynomial-time computable functions $h(\cdot)$ and $f(\cdot)$. Consider any p.p.t. algorithm $\mathcal{A}$ that takes as input $(1^n, c, h(m))$ and outputs $f(m)$ with probability $p(n)$. We construct another algorithm $\mathcal{A}'$, that takes as input $(1^n, |m|, h(m))$ and outputs $f(m)$ with probability $p'(n)$ such that $p(n) - p'(n)$ is negligible if Π satisfies indistinguishability property.

The algorithm $\mathcal{A}'$, on input $(1^n, |m|, h(m))$, does the following:

- Compute $k' \leftarrow \mathsf{Gen}(1^n)$ and $c' = \mathsf{Enc}_{k'}(1^{|m|})$. c' is a dummy ciphertext, the fact that k and k' may be different doesn't matter, we only need some string that 'looks' like a valid ciphertext.
- Run $\mathcal{A}(1^n, c', h(m))$ and output whatever $\mathcal{A}$ outputs.

If $p(n) - p'(n)$ is not negligible, we could use $\mathcal{A}$ to construct a distinguisher $\mathcal{D}$ as follows. What we now know is that there exists $m, h(\cdot)$ and $f(\cdot)$ such that

$$\mathsf{Pr}\left[\mathcal{A}(1^n, \mathsf{Enc}_k(m), h(m)) = f(m)\right] - \mathsf{Pr}\left[\mathcal{A}'(1^n, |m|, h(m)) = f(m)\right]$$

is non-negligible. Our distinguisher $\mathcal{D}$ works as follows:

- Set $m_0 = m$ and $m_1 = 1^{|m|}$, and send m_0, m_1 to the challenger. The challenger returns $c \leftarrow \mathsf{Enc}_k(m_b)$ for a randomly chosen b.
- Run $\mathcal{A}(1^n, c, h(m))$. If $\mathcal{A}(1^n, c, h(m)) = f(m)$, return 0, else return 1.

We have,

$$\mathsf{Pr}\left[\mathcal{A}(1^n, c, h(m)) = f(m) \mid c = \mathsf{Enc}_k(m_0)\right] = p(n)$$

and

$$\mathsf{Pr}\left[\mathcal{A}(1^n, c, h(m)) = f(m) \mid c = \mathsf{Enc}_k(m_1)\right] \leq p'(n)$$

since if $c = \mathsf{Enc}_k(m_1)$, c doesn't give any extra information than a dummy ciphertext gives.

Therefore, $\mathsf{Pr}\left[\mathcal{D} \text{ succeeds}\right] \geq \frac{1}{2}(p(n) - p'(n))$. If the difference is negligible, $\mathcal{D}$ wins with non-negligible probability. □

Claim 2.2. *If an encryption scheme* $\Pi = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *is semantically secure under ciphertext-only attack, then* Π *is indistinguishable for all ciphertext-only adversaries.*

Proof. To prove a contradiction, assume that Π is efficiently distinguishable, and there exists a two messages m_0, m_1 and a distinguisher $\mathcal{D}$ that can distinguish between the encryptions of m_0 and m_1 with non-negligible probabiltiy. We use $\mathcal{D}$ to show that Π is not semantically secure.

Define the function f as follows:

$$f(m) = \begin{cases} 0 & \text{if } m = m_0 \\ 1 & \text{if } m = m_1 \\ \perp & \text{otherwise.} \end{cases}$$

The function h could be chosen to be any computable function. Now, construct an algorithm $\mathcal{A}$ that works as follows:

- $\mathcal{A}$ takes as input $(1^n, \mathsf{Enc}_k(m), h(m))$.
- $\mathcal{A}$ runs the distinguisher $\mathcal{D}(1^n, \mathsf{Enc}_k(m)$, and outputs whatever $\mathcal{D}$ outputs.

It is easy to see that, for $m = m_0$

$$\begin{aligned} \mathsf{Pr}\left[\mathcal{A}(1^n, \mathsf{Enc}_k(m), h(m)) = f(m)\right] &= \mathsf{Pr}\left[\mathcal{D}(1^n, \mathsf{Enc}_k(m_0)) = 0\right] \\ &\geq \frac{1}{2} + \epsilon(n) \end{aligned}$$

which is non-negligibly more than $\frac{1}{2}$. On the other hand, for any algorithm that doesn't have access to $\mathsf{Enc}_k(m)$, the best bet is a random guess, and the probability of success is $\frac{1}{2}$. Hence, Π violates the definition of semantic security. □