CS 600.442 – Modern Cryptography September 23^{rd} 2015

Lecture 7: Secret-Key Encryption

Instructor: Abhishek Jain *Scribe: Samantha Siow*

1 Indistinguishability Security

We concluded the last lecture by looking at the syntax for secret-key encryption and defining correctness property. We now turn to define security for secret-key encryption. We start by considering an indistinguishability-based notion of security. This definition only guarantees security for one-message. Later on, we will consider security for multiple messages.

Definition 1 (Indistinguishability Security (for one message)) *A secret key encryption scheme (Gen, Enc, Dec) is secure if for all n.u. PPT adversaries $\mathcal{A}$, there is a negligible function $\mu(\cdot)$ s.t.:*

$$Pr\left[\begin{array}{l} s \leftarrow Gen(1^n) \\ (m_0, m_1) \longleftarrow \mathcal{A}(1^n) : \mathcal{A}(Enc(s, m_b)) = b \\ b \overset{\$}{\leftarrow} \{0,1\} \end{array}\right] \leq \frac{1}{2} + \mu(n)$$

One can view the above definition as the following game between a challenger and an adversary:

1. Challenger samples a key $s \leftarrow Gen(1^n)$.
2. $\mathcal{A}$ chooses two messages m_0 and m_1 from the message space.
3. Challenger chooses a bit b uniformly at random and sends $Enc(s, m_b)$ to $\mathcal{A}$.
4. If $\mathcal{A}$ can determine the bit b, then she wins. We say that the encryption secure is indistinguishably secure if $\mathcal{A}$ wins with probability at most $\frac{1}{2} + \mu(n)$ for some negligible function $\mu(\cdot)$.

Definition 2 (Indistinguishability Security (Alternative Definition)) *A secret key encryption scheme (Gen, Enc, Dec) is secure if $\forall$ n.u. PPT distinguisher D, there exists a negligible function μ s.t. for all message pairs (m_0, m_1) from the message space, we have:*

$$\left|\Pr\left[D\left(1^n, Enc\left(s, m_0\right)\right) = 1 : s \leftarrow Gen\left(1^n\right)\right] - \Pr\left[D\left(1^n, Enc\left(s, m_1\right)\right) = 1 : s \leftarrow Gen\left(1^n\right)\right]\right| \leq \mu(n)$$

2 Secret-key Encryption using PRGs

We can encrypt messages of polynomial length using a poly-stretch PRG.

- $Gen(1^n) := s \overset{\$}{\leftarrow} \{0,1\}^n$
- $Enc(s, m) := m \oplus PRG(s)$
- $Dec(s, c) :=$ Output $c \oplus PRG(s)$

Security Guarantee:

$$Enc(s \xleftarrow{\$} \{0,1\}^n, m_0) \approx Enc(s \xleftarrow{\$} \{0,1\}^n, m_1)$$

Proof. We start by defining a sequence of hybrid experiments:

- $H_1 : Enc(s \xleftarrow{\$} \{0,1\}^n, m_0)$
- $H_2 : m_0 \oplus r$, where $r \xleftarrow{\$} \{0,1\}^{|m_0|=|m_1|}$
- $H_3 : m_1 \oplus r$, where $r \xleftarrow{\$} \{0,1\}^{|m_0|=|m_1|}$
- $H_4 : Enc(s \xleftarrow{\$} \{0,1\}^n, m_1)$

Note that in H_2, we modify the encryption algorithm to a one-time pad (OTP). In particular, instead of computing the encryption of m_0 using the PRG with seed s, we instead compute an OTP of m_0 using randomness r. In H_3, we switch from computing an OTP of m_0 to computing an OTP of m_1.

Now, we make the following simple claims:

- From the security of PRG, we have that $H_1 \approx H_2$.
- Next, by the security of OTP, we have that $H_2 \approx H_3$.
- From the security of PRG, we have that $H_3 \approx H_4$.

By using the transitivity property of computational indistinguishability, we have that $H_1 \approx H_4$, as required. ■

Now we have constructed an encryption scheme for polynomially long messages. However, the main drawback is that it only guarantees security of one message. We now tackle the question of security for multiple messages. We start by giving a security definition and then constructing a multi-message secure encryption scheme.

3 Multi-Message Secure Encryption

Definition 3 (Multi-message Secure Encryption) *A secret-key encryption scheme (Gen, Enc, Dec) is multi-message secure if ∀ n.u. PPT adversaries $\mathcal{A}$, for all polynomials $q(\cdot)$, there exists a negligible function $\mu(\cdot)$ s.t.*

$$Pr\left[\begin{array}{l} s \xleftarrow{\$} Gen(1^n) \\ \{(m_0^i, m_1^i)\}^{q(n)} \longleftarrow \mathcal{A}(1^n) : \mathcal{A}(\{Enc(s, m_b)\}^{q(n)}) = b \\ b \xleftarrow{\$} \{0,1\} \end{array}\right] \leq \frac{1}{2} + \mu(n)$$

Remark 1 (OTP is not multi-message secure) *It is easy to see that a one-time pad does not support this definition, because an adversary can choose challenges (m_0^1, m_1^1) and (m_0^2, m_1^2) s.t. $m_0^1 = m_0^2$ but $m_1^1 \neq m_1^2$. In this case, upon receiving the challenge ciphertexts, the adversary can simply compare the ciphertexts to see whether they are the same or not. If they are the same, then he guesses $b = 0$, otherwise $b = 1$. It is easy to see that this is a valid strategy and this adversary predicts b correctly with probability 1.*

Remark 2 (Non-adaptive vs adaptive adversary) *In Definition 3, adversary, $\mathcal{A}$ chooses all the challenge messages together, and then receives all the challenge ciphertexts together. Such an adversary is called* non-adaptive.

We can consider a stronger, adaptive *adversary who insists on seeing the challenge cipherhtext corresponding to a query before making the next query. In other words, such an adversary is given oracle access to the encryption algorithm whom it can query adaptively, and then must make its prediction of the challenge bit b. We write this definition below:*

Definition 4 (Multi-message Secure Encryption against Adaptive Adversaries) *A secret-key encryption scheme (Gen, Enc, Dec) is multi-message secure against adaptive adversaries if $\forall$ n.u. PPT adversaries $\mathcal{A}$, there exists a negligible function $\mu(\cdot)$ s.t.*

$$Pr\left[\begin{array}{c} s \xleftarrow{\$} Gen(1^n) \\ b \xleftarrow{\$} \{0,1\} \end{array} : \mathcal{A}^{Enc^b(s,\cdot)}(1^n) = b\right] \leq \frac{1}{2} + \mu(n)$$

where the oracle $Enc^b(s,\cdot)$ takes as input a pair of messages (m_0^i, m_1^i) and outputs $Enc(s, m_b^i)$.

How to Construct Multi-message Secure Encryption? One approach to construct a multi-message secure encryption scheme is to use a poly-stretch PRG. Suppose we want to encrypt a sequence of messages $m_1, m_2, \ldots$, where each message is n-bits long. Then we can evaluate a poly-stretch PRG over a seed s to compute $PRG(s) = S_1, \ldots, S_2, \ldots$, where each chunk S_i is of length n. We can then use chunk S_i to encrypt the message m_i, ensuring that each chunk S_i is used to encrypt only one message.

The problem with this approach is that it yields a *stateful* encryption scheme. Ideally, we would like to construct a *stateless* multi-message secure encryption scheme.

Towards that end, we make the following observation:

Theorem 1 (Necessity of Randomized Encryption) *A multi-message secure encryption scheme cannot be both deterministic and stateless.*

In other words, a stateless multi-message secure encryption scheme must be *randomized.*

The proof the above theorem follows easily. In particular, we can use the same attack that we used to show insecurity of one-time pads against two message queries.

4 Multi-message Secure Encryption using PRFs

We will now construct a multi-message secure encryption scheme using PRFs. In keeping with the above theorem, our encryption scheme will be randomized. In particular, the encryption algorithm will make explicit use of randomness. Intuitively, by making the encryption algorithm randomized, we will ensure that even if the same message is encrypted twice, it will yield a different ciphertext. This easily fails the attacks on deterministic encryption discussed earlier.

Construction. Let $\{f_s : \{0,1\}^n \rightarrow \{0,1\}^n\}$ be a family of PRFs (which are keyed, deterministic algoritihms).

- $Gen(1^n)$: $s \xleftarrow{\$} \{0,1\}^n$ (sampling from the PRF family).

- $Enc(s, m)$: Pick $r \xleftarrow{\$} \{0,1\}^n$. Output $(r, m \oplus f_s(r))$. (Note that since we pick r at random, even if we encrypt the same message twice, the randomness r used for encryption will be different except with exponentially small probability.)
- $Dec(s, (r, c))$: Output $c \oplus f_s(r)$.

The correctness of the construction is easy to see. Note that if $c = m \oplus f_s(r)$, then $c \oplus f_s(r) = m \oplus f_s(r) \oplus f_s(r)$, which is m.

Theorem 2 (Encryption from PRF) *(Gen, Enc, Dec) is a multi-message secure encryption scheme.*

Proof. For simplicity, we will focus on proving security against non-adaptive adversaries. We note that if the underlying PRF scheme is secure against adversaries, then we can also extend the proof of security of the encryption scheme to hold against adversaries.

We need to show that:

$$\{Enc(m_0^1)...Enc(m_0^{q(n)})\} \approx \{Enc(m_1^1)...Enc(m_1^{q(n)})\}$$

Towards this, we will consider the following hybrid experiments:

- H_1 : $Enc(s \xleftarrow{\$} \{0,1\}^n, m_0^i) = (r_i, m_0^i \oplus f_s(r_i))$ for every i, where r_i is a randomly chosen string.
- H_2 : We now change the encryption procedure. Instead of using a PRF f_s, we use a random function $f \xleftarrow{\$} F_n$. That is, for every i, encryption of m_0^i is computed as $(r_i, m_0^i \oplus F(r_i))$, where r_i is a random string.
- H_3 : Once again, we change the encryption procedure. Now, for every i, encryption of m_0^i is computed as $(r_i, m_0^i \oplus p_i)$ where both r_i and p_i are random strings.
- H_4 : Same as above, except that instead of encryption m_o^i, we now encrypt m_1^i, for every i.
- H_5 : Same as experiment H_2, except that here we encrypt m_1^i for every i.
- H_6 : Same as H_1, except that we encrypt m_1^i for every i.

We now make the following observations:

- From the security of PRF, it follows that $H_1 \approx H_2$.
- H_2 and H_3 are statistically indistinguishable.
- From the security of one time pad, $H_3 \approx H_4$
- H_4 and H_5 are statistically indistinguishable.
- By the security of PRF, $H_5 \approx H_6$

From the transitivity property of computational indistinguishability, we have that $H_1 \approx H_6$, as required. ■

5 Semantic Security

We now consider a different security notion for encryption schemes that is referred to as semantic security. We start by defining it below:

Definition 5 ***Semantic Security*** *A secret key encryption scheme (Gen, Enc, Dec) is semantically secure if there exists a PPT simulator algorithm S such that the following two distributions are computationally indistinguishable:*

$$\left\{ \begin{array}{l} s \leftarrow Gen(1^n) \\ (m, z) \longleftarrow \mathcal{M}(1^n) \\ Output \overset{(}{\leftarrow} Enc(s, m), z) \end{array} \right\} \approx \left\{ \begin{array}{l} (m, z) \longleftarrow \mathcal{M}(1^n) \\ Output \leftarrow S(1^n, z) \end{array} \right\}$$

Here $\mathcal{M}$ is a machine that randomly samples a message from the message space and z is auxiliary information for the adversary.

We now establish that semantic security and indistinguishability security are equivalent notions of security. Since semantic security captures the intuition that an adversary does not learn anything about a message from its encryption, we can now say that the same is true for indistinguishability security as well.

Theorem 3 *Indistinguishability Security ⇔ Semantic Security*

Proof. We consider two cases:
Case 1: We first show that semantic security ⇒ indistinguishability security. This follows because of the following:

$$Enc(s, m_0) \approx S(1^n) \approx Enc(s, m_1)$$

By transitivity, we have that $Enc(s, m_0) \approx Enc(s, m_1)$. ■
Case 2: We now show that indistinguishability security ⇒ semantic security. We build S such that the output of $S(1^n)$ is indistinguishable to the left experiment in the definition of semantic security (for simplicity, we ignore auxiliary input here, but the proof easily extends to this case).
Simulator S:

- $s \leftarrow Gen(1^n)$
- Choose a random message $m' \leftarrow \{0, 1\}$ of appropriate length.
- Output $Enc(s, m')$

From indistinguishability security, it follows that $Enc(s, m) \approx Enc(s, m')$. Thus, semantic security follows as well. ■