CS 120/CSCI E-177: Introduction to Cryptography

Salil Vadhan and Alon Rosen Oct. 31, 2006

Lecture Notes 12:

Security for Multiple Messages and Active Attacks

Recommended Reading.

- Katz–Lindell §3.4.3, 3.5, 6.5.

1 Introduction

Defining the security of a cryptographic primitive involves three aspects:

1. What constitutes a 'break'?
2. What are the adversary's resources?
3. What is the adversary's access to the system?

In this lecture, we will generalize our views of the first and third aspects of private-key encryption.

2 Multiple-Message Security

Definition 1 (multiple-message indistinguishability) *Let (G,E,D) be an encryption scheme over $\mathcal{P} = \bigcup_n \mathcal{P}_n$. (G,E,D) satisfies* multiple-message indistinguishability *if for every (nonuniform) PPT A and every polynomial q, there is a negligible function ε such that for all $\overline{m}_0 = (m_0^1, \ldots, m_0^{q(n)}), \overline{m}_1 = (m_1^1, \ldots, m_1^{q(n)}) \in \mathcal{P}_n^{q(n)}$ such that $\|m_0^i\| = \|m_1^i\|$ for all i, we have*

$$\left| \Pr\left[A(E_K(m_0^1), \ldots, E_K(m_0^{q(n)})) = 1\right] - \Pr\left[A(E_K(m_1^1), \ldots, E_K(m_1^{q(n)})) = 1\right] \right| \leq \varepsilon(n),$$

where the probabilities above are taken over $K \stackrel{R}{\leftarrow} G(1^n)$, the coin tosses of E_K, and the coin tosses of A.

Remark: it suffices to consider $\overline{m}_0$ and $\overline{m}_1$ that differ in at most one component.

A (stateful) construction:

Proposition 2 *There is no secure encryption scheme in which E is deterministic and stateless.*

3 Pseudorandom Functions

- Motivation: *stateless* secure encryption. Two parties share a short key k that allows them to generate *exponentially many* pseudorandom pads. To encrypt, they pick one at random and use it as a one-time pad.

- Define $\mathcal{R}_\ell$ to be the set of *all* functions from $\{0,1\}^\ell$ to $\{0,1\}^\ell$.

  Definition 3 *$\mathcal{F} = \bigcup_n \mathcal{F}_n$, where $\mathcal{F}_n = \{f_k : \{0,1\}^{\ell(n)} \to \{0,1\}^{\ell(n)}\}_{k \in \mathcal{I}_n}$, is a* family of pseudorandom functions (PRFs) *if*

  - *There is a PPT G such that $G(1^n) \in \mathcal{I}_n$.*
  - *Given $k \in \mathcal{I}_n$, and $x \in \{0,1\}^{\ell(n)}$, can evaluate $f_k(x)$ in time* $\mathrm{poly}(n)$*.*
  - *For every PPT D, there is a negligible function ε such that*

    $$\left|\Pr\left[D^{f_K}(1^n) = 1\right] - \Pr\left[D^{f}(1^n) = 1\right]\right| \leq \varepsilon(n),$$

    where the probabilities are taken over $K \stackrel{R}{\leftarrow} G(1^n)$, $f \stackrel{R}{\leftarrow} \mathcal{R}_{\ell(n)}$, and the coin tosses of D.

- Notes:

  - Often (and in KL), $\mathcal{I}_n = \{0,1\}^n$, $G(1^n)$ outputs random n-bit string, and $\ell(n) = n$.
  - The key k is *secret* so the adversary D cannot evaluate f_k on its own (unlike collections of one-way functions).
  - A short key k generates a very large amount shared pseudorandomness: $\ell \cdot 2^\ell$ pseudorandom bits from n-bit key! We can consider a pseudorandom function to be $2^{\ell(n)}$ blocks, each of length $\ell(n)$. The block position corresponds to x and the block contains the value $f_k(x)$. This sequence is too long to be read in polynomial time so the PPT adversary will have random access to the sequence.

- How to understand $f \stackrel{\mathrm{R}}{\leftarrow} \mathcal{R}_\ell$, i.e. a truly random function from $\mathcal{R}_\ell$? The static view is that the values of f are chosen all at once. The dynamic view is the following: when $f(x)$ queried for the first time, it is set to a random value (remembered for future queries). Random values are chosen on the fly for $f(x)$, with the provision that we will get the same answer $f(x)$ if we query twice at the same point x.

- Shared Random Function Paradigm
  - Design scheme where all honest parties share a truly random function.
  - Prove it secure in this case.
  - Replace truly random function with pseudorandom function.
  - Use definition of PRF to deduce that it remains secure.
  - Important: adversary does not share the function!

4 Encryption from PRFs

- The encryption scheme is as follows:
  - $E_k(m)$ for $m \in \{0,1\}^\ell$: Choose $r \stackrel{\text{R}}{\leftarrow} \{0,1\}^\ell$. Output $c = (r, f_k(r) \oplus m)$.
  - $D_k((r,s)) = f_k(r) \oplus s$.
- **Theorem 4** *If a pseudorandom function family with $\ell(n) = n$ is used, the above encryption scheme is secure.*
- **Proof Sketch:**

  The proof is similar to the one with PRGs. Given two sequences $\overline{m}_0, \overline{m}_1$ of messages, we have the distributions $\mathbf{Real}_0$ and $\mathbf{Real}_1$ of encryptions these two sequences. We define $\mathbf{Ideal}_0$ and $\mathbf{Ideal}_1$ but where a truly random function f is used.

  The only way to distinguish $\mathbf{Ideal}_0$ from $\mathbf{Ideal}_1$ is if the same r is chosen twice (otherwise it is just like independent one-time pads), which happens with probability at most $\leq \binom{q(n)}{2}/2^{\ell(n)} = \text{neg}(n)$.

  $\mathbf{Real}_i$ is indistinguishable from $\mathbf{Ideal}_i$ by pseudorandomness of PRF. □

5 Constructing PRFs

- Let G be a length-doubling pseudorandom generator. Write $G(x) = G_0(x)G_1(x)$, where $\|G_0(x)\| = \|G_1(x)\| = \|x\|$.
- Define $\mathcal{F}$ by $f_k(x_1 \cdots x_n) = G_{x_n}(G_{x_{n-1}}(\cdots G_{x_1}(k)))$.
  - Think of this as binary tree of depth n with root labelled k. If a node has label x, left child is labelled $G_0(x)$, right child is labelled $G_1(x)$. Labels at leaves are values of PRF.
- **Theorem 5** *If G is a PRG, the $\mathcal{F}$ is a family of pseudorandom functions.*

  Proof: Hybrid argument over levels of tree. For details, see Katz–Lindell. ■

6 Security Against Active Attacks

- **Chosen-plaintext attacks:** adversary can (adaptively) request encryptions of messages of its choice.

- **Definition 6 (indistinguishability under chosen-plaintext attack)** *Let* (G, E, D) *be an encryption scheme over* $\mathcal{P} = \bigcup_n \mathcal{P}_n$. (G, E, D) *satisfies* indistinguishability under chosen-plaintext attack *if for every (nonuniform) PPT A, there is a negligible function* ε *such that the probability that A outputs 1 in Experiments 0 and 1 differ by at most* $\varepsilon(n)$*, where Experiment i is defined as follows:*

  1. $k \stackrel{R}{\leftarrow} G(1^n)$.
  2. *Let* $(m_0, m_1) \stackrel{R}{\leftarrow} A^{E_k(\cdot)}(1^n)$.
  3. *Let* $c \stackrel{R}{\leftarrow} E_k(m_i)$.
  4. *Run* $A^{E_k(\cdot)}(c)$.

  Both E_k *(if stateful) and A maintain state between various calls to them in the experiment.*

- **Proposition 7** *Indistinguishability under chosen-plaintext attack implies multiple-message indistinguishability (indeed, even multiple-message indistinguishability under chosen-plaintext attack).*

- **Proposition 8** *The two encryption schemes from earlier this lecture satisfy indistinguishability under chosen-plaintext attack. In particular, if one-way functions exist, then there are encryption schemes secure under chosen-plaintext attack.*

- **Chosen-ciphertext attacks:** also give adversary access to a decryption oracle $D_k(\cdot)$ which it can query at any point except the challenge ciphertext c.

  1. "Gold standard" for secure encryption.
  2. Can also be achieved based on PRFs (and hence OWFs).