Notes for Lecture 18

Summary

Today we discuss the three ways in which definitions of security given in class differ from the way they are given in the Katz-Lindell textbook,.

Then we study the security of *hybrid* encryption schemes, in which a public-key scheme is used to encode the key for a private-key scheme, and the private-key scheme is used to encode the plaintext.

We also define RSA and note that in order to turn RSA into an encryption scheme we need a mechanism to introduce randomness.

1 Definitions of Security

There are three ways in which the definitions of security given in class differ from the way they are given in the textbook. The first one applies to all definitions, the second to definitions of encryption, and the third to CPA and CCA notions of security for encryption:

1. Our definitions usually refer to schemes of fixed key length and involve parameters t, ϵ, while the textbook definitions are asymptotic and parameter-free.

   Generally, one obtains the textbook definition by considering a family of constructions with arbitrary key length (or, more abstractly "security parameter") k, and allowing t to grow like any polynomial in k and requiring ϵ to be negligible. (Recall that a non-negative function $\nu(k)$ is negligible if for every polynomial p we have $\lim_{k\rightarrow\infty} p(k) \cdot \nu(k) = 0$.)

   The advantage of the asymptotic definitions is that they are more compact and make it easier to state the result of a security analysis. (Compare "if one-way permutations exist, then length-increasing pseudorandom generators exist" with "if $f : \{0,1\}^n \rightarrow \{0,1\}^n$ is a (t, ϵ) one-way permutation computable in time $\leq r$, then there is a generator $G : \{0,1\}^{2n} \rightarrow \{0,1\}^{2n+1}$ computable in time $r + O(n)$ that is $(t\epsilon^4/(n^2 \log n) - r\epsilon^{-4}n^3 \log n, \epsilon/3)$ pseudorandom")

   The advantage of the parametric definitions is that they make sense for fixed-key constructions, and that they make security proofs a bit shorter.

(Every asymptotic security proof starts from "There is a polynomial time adversary A, an infinite set N of input lengths, and a polynomial $q()$, such that for every $n \in N \ldots$.)

2. Definitions of security for an encryption algorithm $E()$, after the proper quantifications, involve an adversary A (who possibly has oracles, etc.) and messages m_0, m_1; we require

$$|\,\mathbb{P}[A(E(m_0)) = 1] - \mathbb{P}[A(E(m_1)) = 1]| \leq \epsilon$$

   while the textbook usually has a condition of the form

$$\mathop{\mathbb{P}}_{b \in \{0,1\}}[A(E(m_b)) = b] \leq \frac{1}{2} + \frac{\epsilon}{2}$$

   The two conditions are equivalent.

3. Definitions of CPA and CCA security, as well as all definitions in the public-key setting, have a different structure in the book. The difference is best explained by an example. Suppose we have a public-key scheme (G, E, D) such that, for every valid public key pk, $E(pk, pk)$ has some distinctive pattern that makes it easy to distinguish it from other ciphertexts. This could be considered a security weakness because an eavesdropper is able to see if a party is sending a message that concerns the public key.

   You can show as an exercise that if a secure public-key encryption scheme exists, then there is a public-key encryption scheme that is secure according to our definition from last lecture but that has a fault of the above kind.

   The textbook adopts a two-phase definition of security, in which the adversary is allowed to choose the two messages m_0, m_1 that it is going to try and distinguish, and the choice is done after having seen the public key.

2 Hybrid Encryption

Let (G_1, E_1, D_1) be a public-key encryption scheme and (E_2, D_2) a private-key encryption scheme.

Consider the following *hybrid* scheme (G, E, D):

- $G()$: same as $G_1()$
- $E(pk, m)$: pick a random key K for E_2, output $(E_1(pk, K), E_2(K, m))$

- $D(sk, (C_1, C_2))$: output $D_2(D_1(sk, C_1), C_2)$

Theorem 1 *Suppose* (G_1, E_1, D_1) *is* (t, ϵ_1)*-secure for one encryption and* (E_2, D_2) *is* (t, ϵ_2)*-secure for one encryption. Suppse also that* E_1, E_2 *have running time* $\leq r$*. Then* (G, E, D) *is* $(t - 2r, 2\epsilon_1 + \epsilon_2)$*-secure for one encryption.*