Notes for Lecture 5

Scribed by Manohar Jonnalagedda, posted February 13, 2009

Summary

Having introduced the notion of CPA security in the past lecture, we shall now see constructions that achieve it. Such constructions shall require either *pseudorandom functions* or *pseudorandom permutations*. We shall see later how to construct such objects.

1 Pseudorandom Functions

To understand the definition of a pseudorandom function, it's good to think of it as a pseudorandom generator whose output is *exponentially long*, and such that each bit of the output is efficiently computable given the seed. The security is against efficient adversaries that are allowed to look at at any subset of the exponentially many output bits.

Definition 1 (Pseudorandom Function) *A function* $F : \{0,1\}^k \times \{0,1\}^m \to \{0,1\}^m$ *is a* (t,ϵ)*-secure pseudorandom function if for every oracle algorithm* T *that has complexity at most* t *we have*

$$|\mathop{\mathbb{P}}_{K\in\{0,1\}^k}[T^{F_K}() = 1] - \mathop{\mathbb{P}}_{R:\{0,1\}^m\to\{0,1\}^m}[T^{R}() = 1]| \leq \epsilon$$

Intuitively, this means that an adversary wouldn't be able to distinguish outputs from a purely random function and a pseudorandom function (upto a certain ϵ additive error). Typical parameters are $k = m = 128$, in which case security as high as $(2^{60}, 2^{-40})$ is conjectured to be possible.

As usual, it is possible to give an asymptotic definition, in which $\epsilon(k)$ is required to be negligible, $t(k)$ is allowed to be any polynomial, and F is required to be computable in polynomial time.

2 Encryption Using Pseudorandom Functions

Suppose $F : \{0,1\}^k \times \{0,1\}^m \to \{0,1\}^m$ is a pseudorandom function. We define the following encryption scheme.

- $Enc(K, M)$: pick a random $r \in \{0,1\}^m$, output $(r, F_K(r) \oplus M)$
- $Dec(K, (C_0, C_1)) := F_K(C_0) \oplus C_1$

This construction achieves CPA security.

Theorem 2 *Suppose F is a (t, ϵ) secure pseudorandom function. Then the above scheme is $(\frac{t}{O(m)}, 2\epsilon + t \cdot 2^{-m})$-secure against CPA.*

The proof of Theorem 2 will introduce another key idea that will often reappear in this course: to first pretend that our pseudorandom object is truly random, and perform our analysis accordingly. Then extend the analysis from the pseudorandom case to the truly random case.

Let us therefore consider a modified scheme $(\overline{Enc}, \overline{Dec})$, where instead of performing $F_K(r) \oplus M$, we do $R(r) \oplus M$, where $R : \{0,1\}^m \to \{0,1\}^m$ is a truly random function. We need to look at how secure this scheme is. In fact, we will actually prove that

Lemma 3 *$(\overline{Enc}, \overline{Dec})$ is $(t, \frac{t}{2^m})-CPA$ secure.*

Proof:

In the computation $T^{\overline{Enc}}(\overline{Enc}(r, C))$ of algorithm T given oracle $\overline{Enc}$ and input the ciphertext (r, C), let us define REPEAT to be the event where T gets the messages $(r_1, C_1), \ldots, (r_t, C_t)$ from the oracle, such that r equals one of the r_i.

Then we have

$$\begin{aligned} \mathbb{P}[T^{\overline{Enc}}(\overline{Enc}(M)) = 1] &= \mathbb{P}[T^{\overline{Enc}}(\overline{Enc}(M)) = 1 \wedge REPEAT] \\ &\quad + \mathbb{P}[T^{\overline{Enc}}(\overline{Enc}(M)) = 1 \wedge \neg REPEAT] \end{aligned}$$

similarly,

$$\begin{aligned} \mathbb{P}[T^{\overline{Enc}}(\overline{Enc}(M')) = 1] &= \mathbb{P}[T^{\overline{Enc}}(\overline{Enc}(K, M')) = 1 \wedge REPEAT] \\ &\quad + \mathbb{P}[T^{\overline{Enc}}(\overline{Enc}(M')) = 1 \wedge \neg REPEAT] \end{aligned}$$

so

$$
\begin{aligned}
&|\,\mathbb{P}[T^{\overline{Enc}}(\overline{Enc}(K,M))=1]-\mathbb{P}[T^{\overline{Enc}}(\overline{Enc}(K,M'))=1]|\leq\\
&\quad|\,\mathbb{P}[T^{\overline{Enc}}(\overline{Enc}(M))=1\wedge REPEAT]-\mathbb{P}[T^{\overline{Enc}}(\overline{Enc}(M'))=1\wedge REPEAT]|+\\
&\quad|\,\mathbb{P}[T^{\overline{Enc}}(\overline{Enc}(M))=1\wedge\neg REPEAT]-\mathbb{P}[T^{\overline{Enc}}(\overline{Enc}(M'))=1\wedge\neg REPEAT]|
\end{aligned}
$$

Now the first difference is the difference between two numbers which are both between 0 and $P[REPEAT]$, so it is at most $P[REPEAT]$, which is at most $\frac{t}{2^m}$.

The second difference is zero, because with a purely random function there is a 1-1 mapping between every random choice (of $R, r, r_1, \ldots, r_t$) which makes the first event happen and every random choice that makes the second event happen. □

We have shown that with a purely random function, the above encryption scheme is CPA-secure. We can now turn our eyes to the pseudorandom scheme (Enc, Dec), and prove Theorem 2.

PROOF: Consider the following four probabilities, for messages M, M', and algorithm T :

(a) $\mathbb{P}_K[T^{Enc(K,\cdot)}(Enc(K,M))=1]$

(b) $\mathbb{P}_K[T^{Enc(K,\cdot)}(Enc(K,M'))=1]$

(c) $\mathbb{P}_R[T^{\overline{Enc}(\cdot)}(\overline{Enc}(M))=1]$

(d) $\mathbb{P}_R[T^{\overline{Enc}(\cdot)}(\overline{Enc}(M'))=1]$

From the previous proof, we have $|c-d|\leq\frac{t}{2^m}$. If we are able to show that $|a-c|\leq\epsilon$, $|b-d|\leq\epsilon$, then we have $|a-b|\leq 2\epsilon+\frac{t}{2^m}$.

So, it remains to show that

$$
|\,\underset{K}{\mathbb{P}}[T^{Enc(K,\cdot)}(Enc(K,M))=1]-\underset{R}{\mathbb{P}}[T^{\overline{Enc}(\cdot)}(\overline{Enc}(M))=1]|\leq\epsilon \tag{1}
$$

Suppose, by contradiction, this is not the case. We will show that such a contradiction implies that F is not secure, by constructing an oracle algorithm T' that distinguishes F from a truly random function.

For an oracle G, we define T'^G to be the following algorithm:

- pick a random $r\in\{0,1\}^m$ and compute $C:=(r,G(r)\oplus M)$
- simulate $T(C)$; every time C makes an oracle query M_i, pick a random r_i and respond to the query with $(r_i, G(r_i)\oplus M)$

Note that if T' is given the oracle F_K, then the computation T'^{F_K} is exactly the same as the computation $T^{Enc}(Enc(M))$, and if T' is given the oracle R, where R is a random function, then the computation $T^{\overline{Enc}}(\overline{Enc}(M))$.

Thus, we have

$$\mathop{\mathbb{P}}_{K\in\{0,1\}^k}[T'^{F_K}() = 1] = \mathop{\mathbb{P}}_{K}[T^{Enc(K,\cdot)}(Enc(K,M)) = 1] \tag{2}$$

$$\mathop{\mathbb{P}}_{R:\{0,1\}^m\to\{0,1\}^m}[T'^{R}() = 1] = \mathop{\mathbb{P}}_{R}[T^{\overline{Enc}(\cdot)}(\overline{Enc}(M)) = 1] \tag{3}$$

which means that

$$\left|\mathop{\mathbb{P}}_{K\in\{0,1\}^k}[T'^{F_K}() = 1] - \mathop{\mathbb{P}}_{R:\{0,1\}^m\to\{0,1\}^m}[T'^{R}() = 1]\right| > \epsilon \tag{4}$$

The complexity of T' is at most the complexity of T times $O(m)$ (the time needed to translate between oracle queries of T and oracle queries of T'), and so if T has complexity $t/O(m)$ then T' has complexity $\leq t$. This means that (4) contradicts the assumption that F is $(t,\epsilon)-$secure. $\square$