Notes for Lecture 3 (Draft)

Summary

Last time we introduced the setting of *one-time symmetric key encryption*, defined the notion of *semantic security*, and proved its equivalence to *message indistinguishability*.

Today we complete the proof of equivalence (found in the notes for last class), discuss the notion of *pseudorandom generator*, and see that it is precisely the primitive that is needed in order to have message-indistinguishable (and hence semantically secure) one-time encryption.

1 Pseudorandom Generators And One-Time Encryption

Definition 1 (Pseudorandom Generator) *A function $G : \{0,1\}^k \to \{0,1\}^m$ is a (t,ϵ)-secure pseudorandom generator if for every boolean function T of complexity at most t we have*

$$|\mathbb{P}_{x\sim U_k}[T(G(x)) = 1] - \mathbb{P}_{x\sim U_m} T(x) = 1]| \leq \epsilon \tag{1}$$

(We use the notation U_n for the uniform distribution over $\{0,1\}^n$.)

The definition is interesting when $m > k$ (otherwise the generator can simply output the first m bits of the input, and satisfy the definition with $\epsilon = 0$ and arbitrarily large t). Typical parameters we may be interested in are $k = 128$, $m = 2^{20}$, $t = 2^{60}$ and $\epsilon = 2^{-40}$, that is we want k to be very small, m to be large, t to be huge, and ϵ to be tiny. There are some unavoidable trade-offs between these parameters.

Lemma 2 *If $G : \{0,1\}^k \to \{0,1\}^m$ is $(t, 2^{-k-1})$ pseudorandom with $t = O(m)$, then $k \geq m + 1$.*

Exercise 1 *Prove that if $G : \{0,1\}^k \to \{0,1\}^m$ is (t,ϵ) pseudorandom, and $k < m$, then*

$$t \cdot \frac{1}{\epsilon} \leq O(m \cdot 2^k)$$

Suppose we have a pseudorandom generator as above. Consider the following encryption scheme:

- Given a key $K \in \{0,1\}^k$ and a message $M \in \{0,1\}^m$,

$$Enc(K, M) := M \oplus G(K)$$

- Given a ciphertext $C \in \{0,1\}^m$ and a key $K \in \{0,1\}^k$,

$$Dec(K, C) = C \oplus G(K)$$

(The XOR operation is applied bit-wise.)

It's clear by construction that the encryption scheme is correct. Regarding the security, we have

Lemma 3 *If G is (t, ϵ)-pseudorandom, then (Enc, Dec) as defined above is $(t - m, 2\epsilon)$-message indistinguishable for one-time encryption.*

2 Security for Multiple Encryptions: Vanilla Version

Definition 4 (Message indistinguishability for multiple encryptions) *(Enc, Dec) is (t, ϵ)-message indistinguishable for c encryptions if for every $2c$ messages $M_1, \ldots, M_c$, $M'_1, \ldots, M'_c$ and every T of complexity $\leq t$ we have*

$$\begin{aligned} &\mathbb{P}[T(Enc(K, M_1), \ldots, Enc(K, M_c)) = 1] \\ -&\mathbb{P}[T(Enc(K, M'_1), \ldots, Enc(K, M'_c)) = 1]| \leq \epsilon \end{aligned}$$

Similarly, we define semantic security, and the asymptotic versions.

Exercise 2 *Prove that no encryption scheme (Enc, Dec) in which $Enc()$ is deterministic (such as the scheme for one-time encryption described above) can be secure even for 2 encryptions.*

Encryption in some versions of Microsoft Office is deterministic and thus fails to satisfy this definition. (This is just a symptom of bigger problems; the schemes in those versions of Office are considered completely broken.)

If we allow the encryption algorithm to keep *state* information, then a pseudorandom generator is sufficient to meet this definition. Indeed, usually pseudorandom generators designed for such applications, including RC4, are optimized for this kind of "stateful multiple encryption."