COM S 6830 – Cryptography September 29, 2009

Lecture 10: Pseudo-Random Generators and Secure Encryption

Instructor: Rafael Pass *Scribe: Eleanor Birrell*

1 Pseudo-Random Generators

Recall from the previous lecture that we defined a pseudorandom generator to be a (deterministic) function that takes a short string to a longer string that is indistinguishable for random. More specifically:

Definition 1 *A (deterministic) function $G : \{0,1\}^* \to \{0,1\}^*$ is a* pseudorandom generator (PRG) *if the following three properties hold:*

1. *(Efficiency): G is deterministic function that is (probabilistic) polynomial-time computable.*
2. *(Expansion): $|G(x)| = \ell(|x|)$ where $\ell(k) > k$.*
3. *(Pseudorandomness): The ensemble $\{x \leftarrow \{0,1\}^n : G(x)\}_n$ is pseudorandom.*

We also introduced our first example of a PRG (with one-bit expansion).

Example 2 *Let f be a OWP and let B be a hard-core predicate for f, then $G(x) = f(s)||b(s)$ is a PRG.*

We now proceed to demonstrate that the existence of such a PRG guarantees the existence of a general class of PRGs.

Theorem 3 *If there exists a PRG G with 1-bit expension then there exists a PRG G' with polynomial expansion*

Proof. Let G be a PRG $G : \{0,1\}^n \to \{0,1\}^{n+1}$. Define a new (deterministic) function G' by $G'(x_0) = b_1 \dots b_m$ (where $m > n$ and $G(x_i) = x_{i+1}||b_{i+1}$). We claim that G' is a PRG with polynomial expansion.

For simplicity we introduce the following notation: $G^0(x) = \epsilon$, $G^i(s) = b||G^{i-1}(x)$ where $G(s) = x||b$. Observe that, using this notation, $G'(s) = G^m(s)$.

Assume for contradiction that G' is not a PRG, that is that there exists a distinguisher D and a polynomial p such that for infinitely many $n \in \mathbb{N}$, D distinguishes $\{U_m\} and \{G^m(U_n)\}$ with probability $\frac{1}{p(n)}$. For each n, define a sequence of hybrids $H_0, H_1, \dots, H_m$ by $H_i = U_{m-i}||G^i(U_n)$.

It is immediately clear that $H_0 = U_m$ and $H_m = G^m(U_n)$. By the Hybrid Lemma (Lecture 8), there therefore exists a value i such that D distinguishes H_i, H_{i+1} with probability $\frac{1}{p'(n)}$ for some polynomial p'.

$$\begin{aligned} H_i &= \{\ell \leftarrow U_{m-i-1}, b \leftarrow \{0,1\}, r \leftarrow G^i(x) : \ell||b||r\} \\ H_{i+1} &= \{\ell \leftarrow U_{m-i-1}, x||b \leftarrow G(U_n), r \leftarrow G^i(x) : \ell||b||r\} \end{aligned}$$

Define an efficient operation M that on input $y_1, \ldots y_m$ begins by choosing $\ell \leftarrow U_{m-i-1}$, sets $b = y_1$, computes $r = G^i(y_2 \ldots y_m)$, and outputs $\ell||b||r$. Observe that $M(U_{n+1}) = H_i$ and $M(G(U_n)) = H_{i+1}$. Since $U_{n+1} \approx G(U_n)$ and since M is an efficient operation, by the Efficient Operation Lemma (Lecture 8), $H_i \approx H_{i+1}$ which contradicts our assumption. ■

Under this construction any OWF f with a hard-core predicate h gives rise to a PRG $G(x) = h(x)||h(f(x))|| \ldots ||h(f^{m-1}(x))$. If you instead start with a *collection* of one-way permutations $\{f_i\}$, $G(r_1, r_2) = h_i(x)||h_i(f_i(x))|| \ldots ||h_i(f_i^{m-1}(x))$ where r_1 is used to sample $i \in I$ (to determine the function) and r_2 is used to sample x. For concreteness, consider the following two examples.

Example 4 (Modular Exp. PRG (Blum-Micali)) *Use the random seed to generate (p, g, x) where p is a prime greater than 2, g is a generator for Z_p^*, and $x \in Z_p^*$. Output $half_{p-1}(x)||half_{p-1}(g^x \mod p)||half_{p-1}(g^{g^x} \mod p) \ldots$ where $half$ is defined to be the predicate that determines whether an argument is in the upper half or lower half of the range.*

Example 5 (RSA PRG) *Use the random seed to generate n-bit primes p, q, define $N = pq$, choose $e \in_r Z_{\phi(N)}, x \in_R Z_N^*$. Let $lsb(x)$ denote the least significant bit, and output $lsb(x)||lsb(x^e \mod N)||lsb((x^e)^e \mod N)|| \ldots$.*

2 Secure Encryption

Recall that we consider an encryption scheme to be secure if the encryptions of any two messages (over a random key) are indistinguishable.

Definition 6 *(Gen, Enc, Dec) is a* single-message secure encryption *over message space M if for all nonuniform PPT algorithms A there exists a negligible function ε such that for all $n \in \mathbb{N}$ and for all $m_0, m_1 \in M$,*

$$|\Pr[k \leftarrow Gen(1^n : A(Enc_k(m_0)) = 1] - \Pr[k \leftarrow Gen(1^n) : A(Enk_k(m_1)) = 1]| \leq \varepsilon(n).$$

This differs from Shannon's definition of secure encryption in that A is restricted to polynomial time and we allow negligible error.

Another way of defining secure encryption would be to require that an adversary, given a sample from either $\{Enc(m_0)\}$ or $\{Enc(m_1)\}$ cannot predict which distribution the sample came from with probability better than $\frac{1}{2} + \varepsilon(n)$ for some negligible function ε. By the Prediction Lemma (Lecture 8), these notions are equivalent.

Theorem 7 *Let $G(s)$ be a length-doubling PRG, and define an encryption scheme (Enc, Gen, Dec) by*

$$\begin{aligned} &Gen(1^n) : s \leftarrow \{0,1\}^{n/2}, \text{ output } k = s \\ &Enc_k(m) = m \oplus G(k) \\ &Dec_k(c) = c \oplus G(k). \end{aligned}$$

(Enc, Gen, Dec) is a single-message secure encryption scheme.

Proof. Assume that (Enc, Gen, Dec) is not a single-message secure encryption scheme, i.e. that there exists a nonuniform PPT A and a polynomial p such that for infinitely many n there exists $m_n^0, m_n^1 \in M_n$ such that A distinguishes $\{k \leftarrow Gen(1^n) : m_n^0 \oplus G(k)\}$ frow $\{k \leftarrow Gen(1^n) : m_n^1 \oplus G(k)\}$ with probability $\frac{1}{p(n)}$.

Define a sequence of Hybrids

$$\begin{aligned} H_n^0 &= \{k \leftarrow U_{n/2} : m_n^0 \oplus G(k)\} \\ H_n^1 &= \{k \leftarrow U_n : m_n^0 \oplus k\} \\ H_n^2 &= \{k \leftarrow U_n : m_n^1 \oplus k\} \\ H_n^3 &= \{k \leftarrow U_{n/2} : m_n^1 \oplus G(k)\} \end{aligned}$$

By construction, D distinguishes H_0, H_3 with probability $1/p(n)$ for infinitely many $n \in \mathbb{N}$, therefore by the Hybrid Lemma (Lecture 8) for infinitely many $n \in \mathbb{N}$ D also distinguishes between two consecutive hybrids with probability $1/4p(n)$. However, we will proceed to show that this gives rise to a contradiction.

Define a pair of nonuniform PPT machines M^0, M^1 by $M^i(X_n) = \{r \leftarrow X_n : m_n^i \oplus r\}$. It is immediately clear that M^i is an efficient operation. Observe now that $H_n^0 = M^0(\{k \leftarrow U_{n/2} : G(k)\})$ and $H_n^1 = M^0(U_n)$. Since the input ensembles are indistinguishable, by the Efficient Operations Lemma (Lecture 8), $H_n^0 \approx H_n^1$. Similarly, $H_n^3 = M^1(\{k \leftarrow U_{n/2} : G(k)\})$ and $H_n^2 = M^1(U_n)$. Since the input ensembles are still indistinguishable, by the Efficient Operations Lemma (Lecture 8) $H_n^2 \approx H_n^3$.

It follows that in order for our assumption to be consistent, D must be able to distinguish between H_n^1 and H_n^2. However, by the perfect secrecy of the one-time pad, we know that H_n^1 and H_n^2 are *identically* distributed, which yields the desired contradiction.■