Alin Tomescu, CSE408
Tuesday, February 10th, Lecture #4

PRNGs continued

Data Processing Inequality

Last class' theorem: If $D_0 \overset{t}{\underset{\varepsilon}{\sim}} D_1$ and f is a function running in time t' then $f(D_0) \overset{t-t'}{\underset{\varepsilon}{\sim}} f(D_1)$.

Proof by contra-positive

Instead of proving $A \Rightarrow B$, we will prove $\sim B \Rightarrow \sim A$.

Suppose $f(D_0)$ and $f(D_1)$ are not $(t-t', \varepsilon)$-computationally indistinguishable. Then there exists an algorithm A running in time $\leq t - t'$ with $Adv\ A = \varepsilon = \Pr[A(x) = 0 | x \leftarrow f(D_o)] - \Pr[A(x) = 0 | x \leftarrow f(D_1)]$.

We will construct another algorithm A' that will distinguish between D_0 and D_1 in time $\leq t$ proving our theorem.

$$A'(x) = A\big(f(x)\big)$$

$$\Pr[A'(x) = 0 | x \leftarrow D_0] = \Pr[A(x) = 0 | x \leftarrow f(D_0)]$$

$$\Pr[A'(x) = 0 | x \leftarrow D_1] = \Pr[A(x) = 0 | x \leftarrow f(D_1)]$$

$$Adv\ A' = \Pr[A(x) = 0 | x \leftarrow f(D_0)] - \Pr[A(x) = 0 | x \leftarrow f(D_1)] = Adv\ A = \varepsilon$$

Also, A will run in time $\leq t - t' + t'$, so it will run in time $\leq t$. QED.

Theorem about PRNGs

If $G: \{0,1\}^l \rightarrow \{0,1\}^{l+1}$ is a (t, ε)-PRNG running in time t', then $G': \{0,1\}^l \rightarrow \{0,1\}^{l+2}$ is a $(t - t', 2\varepsilon)$-PRNG. G' uses two consecutive calls to G to generate a pseudo-random string of length $l + 2$, but in the second call the last bit is dropped and appended to the result, since G can only take inputs of size l.

We can prove this using the DPI theorem because G' is pretty much G applied to G itself: $G' \approx G \circ G$ with a few minor alterations to the input and output of the second call.

Proof

By definition, G is a (t, ε)-PRNG, which means that $G(U_l) \overset{t}{\underset{\varepsilon}{\sim}} U_{l+1}$

Let us *slowly* define G' formally. $G'(s_0) = s_2$, where:

- $s_2 = G(s_1 - last\ bit\ of\ s_1) + last\ bit\ of\ s_1$
- $s_1 = G(s_0)$

So $G'(s) = G\big(G(s) - last\ bit\ of G(s)\big) + last\ bit\ of\ G(s)$. Therefore, $G'(s) = f\big(G(s)\big)$, where $f(x) =$ removes and remembers the last bit of x, computes G on the trimmed version of x and appends the last bit of x to the result. Note that f will run in time t' since it makes one call to G which runs in time t'

Since, $G(U_l) \underset{\varepsilon}{\overset{t}{\sim}} U_{l+1}$ then, using DPI, we get $f\big(G(U_l)\big) \underset{\varepsilon}{\overset{t-t'}{\sim}} f(U_{l+1}) \Leftrightarrow G'(U_l) \underset{\varepsilon}{\overset{t-t'}{\sim}} f(U_{l+1})$

We will now prove that $f(U_l + 1) \underset{\varepsilon}{\overset{t-t'}{\sim}} U_{l+2}$. By transitivity, it will follow that $G'(U_l) \underset{2\varepsilon}{\overset{t-t'}{\sim}} U_{l+2}$

Note that $f(U_{l+1}) = G(U_l) \parallel U_1$.

Now, let $h(x) = x \parallel U_1$.

It follows that $f(U_{l+1}) = G(U_l) \parallel U_1 = h\big(G(U_l)\big)$.

Note that $h(U_{l+1}) = U_{l+1} \parallel U_1 = U_{l+2}$.

We know that $G(U_l) \underset{\varepsilon}{\overset{t}{\sim}} U_{l+1}$, so by DPI it follows that:

$$h\big(G(U_l)\big) \underset{\varepsilon}{\overset{t}{\sim}} h(U_{l+1}) \Leftrightarrow f(U_{l+1}) \underset{\varepsilon}{\overset{t}{\sim}} U_{l+2}$$

We proved that $f(U_{l+1}) \underset{\varepsilon}{\overset{t}{\sim}} U_{l+2}$, we also know that $G'(U_l) \underset{\varepsilon}{\overset{t-t'}{\sim}} f(U_{l+1})$ therefore it follows that $G'(U_l) \underset{2\varepsilon}{\overset{t-t'}{\sim}} U_{l+2}$

Transitivity property

If $D_0 \underset{\varepsilon}{\overset{t}{\sim}} D_1 \underset{\varepsilon'}{\overset{t}{\sim}} D_2$ then $D_0 \underset{\varepsilon+\varepsilon'}{\overset{t}{\sim}} D_2$.

Proof

$\forall A$ running in time t

$$Adv\ A = |\Pr[A(x) = 0 | x \leftarrow D_0] - \Pr[A(x) = 0 | x \leftarrow D_2]| =$$

$$= |\Pr[A(x) = 0 | x \leftarrow D_0] - \Pr[A(x) = 0 | x \leftarrow D_1] + \Pr[A(x) = 0 | x \leftarrow D_1] - \Pr[A(x) = 0 | x \leftarrow D_2]| =$$

Using the property of absolute value $|a + b| \leq |a| + |b|$, we get:

$$|\Pr[A(x) = 0 | x \leftarrow D_0] - \Pr[A(x) = 0 | x \leftarrow D_1]| + |\Pr[A(x) = 0 | x \leftarrow D_1] - \Pr[A(x) = 0 | x \leftarrow D_2]| \leq \varepsilon + \varepsilon'$$

Therefore $D_0 \underset{\varepsilon+\varepsilon'}{\overset{t}{\sim}} D_2$.

Concatenation theorem

Theorem: If $G_1 : \{0,1\}^{l_1} \to \{0,1\}^{L_1}$ is (t_1, ε_1)-secure PRNG running in time t_1' and $G_2 : \{0,1\}^{l_2} \to \{0,1\}^{L_2}$ is (t_2, ε_2)-secure PRNG running in time t_2' then $G_1 \parallel G_2$ is $(t_3, \varepsilon_1 + \varepsilon_2)$-secure PRNG, with $t_3 = \min(t_1 - t_2', t_2)$.

Proof: $G_1(U_{l_1}) \underset{\varepsilon_1}{\overset{t_1}{\sim}} U_{L_1}$.

Let $f(x) = x \parallel G_2(y), where\ y \leftarrow U_{l_2}$. Then f will run in time t_2'.

Using DPI, we get $f\left(G_1(U_{l_1})\right) \underset{\varepsilon_1}{\overset{t_1 - t_2'}{\sim}} f(U_{L_1}) \Leftrightarrow G_1(U_{l_1}) \parallel G_2(U_{l_2}) \underset{\varepsilon_1}{\overset{t_1 - t_2'}{\sim}} U_{L_1} \parallel G_2(U_{l_2})$

But $G_2(U_{l_2}) \underset{\varepsilon_2}{\overset{t_2}{\sim}} U_{L_2}$. Therefore, by transitivity $G_1(U_{l_1}) \parallel G_2(U_{l_2}) \underset{\varepsilon_1 + \varepsilon_2}{\overset{\min(t_1 - t_2', t_2)}{\sim}} U_{L_1} \parallel U_{L_2}$

Examples of secure PRNGs

If $AES: \{0,1\}^{128}(\text{key}) \times \{0,1\}^{128}(\text{msg}) \rightarrow \{0,1\}^{128}(ctxt)$ is secure then $G: \{0,1\}^{128} \rightarrow \{0,1\}^{\text{L}}$, $G(x) = (AES(x,0), AES(x,1), AES(x,2) \dots)$ is a secure PRNG.

If RSA is secure then $G(x) =$ use x as a random source for generating 2048-bit RSA modulus $N = pq$ and exponent e and output $\left(f\left(b^{\frac{1}{e}}\ mod\ N\right)\right)_{b=2}^{100000}$ and $f(x) =$ 11 least significant bits of x.