COS 433 — Cryptography — Homework 3.

Boaz Barak

Total of 130 points. Due February 24, 2010.

Exercise 0 (0 points)**.** 1. Read the analysis of the length extension theorem for PRG's in Katz-Lindell (pages 215–220 excerpt on web and handed out in class) and compare it to the proof given in class, you can also compare it to the proof in the Boneh-Shoup book in Section 3.4.2

2. Read the analysis of the constructions of PRF's from PRG's in Boneh-Shoup book (Section 4.6, pages 113–117) and compare it to the proof given in class.

Exercise 1 (25 points, impossibility of statistically testing randomness)**.** Let $T_1, \ldots, T_M : \{0,1\}^n \to \{0,1\}$ be a collection of function that are supposed to be statistical tests for randomness. Prove that if $M < 2^{2^{n/30}}$ there exists a distribution X that passes all these tests but is very far from the uniform distribution. Concretely, show that there exists a random variable X over $\{0,1\}^n$ such that:

- For every $i \in [M]$, $\left|\Pr[T_i(U_n) = 1] - \Pr[T_i(X) = 1]\right| < 2^{-n/50}$
- $\Delta(X, U_n) \geq 1 - 2^{-n/50}$ (where Δ denotes the statistical distance).

(The constants above are fairly arbitrary and are set to allow a lot of slackness.) See footnote for hint.[1]

Based on this exercise, what do you believe can we say about a distribution X if it passes the FIPS 140-2 testing suite for randomness?

Exercise 2 (25 points)**.** 1. Prove the *polynomial hybrid principle*. That is prove that for every m distributions $X_1, \ldots, X_m$ over $\{0,1\}^n$ such that $\Delta_T(X_i, X_{i+1}) \leq \epsilon$ for every $i \in \{1..m-1\}$, $\Delta_T(X_1, X_m) \leq m\epsilon$.

The above claim is used to show that if $X_i \approx X_{i+1}$ for all i, then $X_1 \approx X_m$.

2. Show that there is no *exponential* hybrid principle in the following sense: show that there exist 2^n distributions $X_1, \ldots, X_{2^n}$ over $\{0,1\}^n$ such that (a) for every i, X_i and X_{i+1} are computationally (or even statistically!) indistinguishable, i.e. $\Delta(X_i, X_{i+1}) < 2^{-n/10}$ but (b) X_1 and X_{2^n} are easy to distinguish - there is a linear time algorithm T such that $\mathbb{E}[T(X_1)] > 0.9$ but $\mathbb{E}[T(X_{2^n})] < 0.6$.

Exercise 3 (25 points)**.** In these two questions you'll show that if we have a pseudorandom function family with particular input and output sizes, we can easily obtain a family that handles larger inputs and outputs. (It's easy to handle smaller outputs and inputs by truncation and padding.)

[1] **Hint:** Define X as the uniform distribution over some set $S = \{s_1, \ldots, s_m\}$ where $m = 2^{n/15}$. Show that for every such set S, just because of S's small size it will have very large statistical distance from the uniform distribution. Now, $\Pr[T_i(X) = 1] = \frac{1}{m}\sum_{j=1}^{m} T_i(s_j)$ (*). Now use the Chernoff bound from probability to argue that if you chose the elements $s_1, \ldots, s_m$ at random, then for every i, the probability that the RHS deviates from $\mathbb{E}[T_i(U_n)]$ is very small, and in fact small enough that you can do a union bound over all $i \in \{1..M\}$.

Figure 1: RSA SecurID Device.

1. *(Changing PRFs output size)* Prove that if there exists a collection $\{f_s\}$ of pseudorandom functions with $f_s : \{0,1\}^{|s|} \to \{0,1\}$ (i.e., one-bit output) then there exists a collection $\{f'_s\}$ with $f'_s : \{0,1\}^{|s|} \to \{0,1\}^{|s|}$. See footnote for hint.[2]

2. *(Changing PRFs input size)* Prove that if there exists a collection $\{f_s\}$ of pseudorandom functions with $f_s : \{0,1\}^{|s|} \to \{0,1\}^{|s|}$ then there exists a collection $\{f'_s\}$ with $f'_s : \{0,1\}^* \to \{0,1\}^{|s|}$ (i.e., f'_s for a random $s \in \{0,1\}^n$ is indistinguishable from a random function from $\{0,1\}^*$ to $\{0,1\}^n$. See footnote for hint[3]

Exercise 4 (25+30 points)**.** The RSA SecurID card (see Figure 1) is a credit-card sized device that displays 6 digits that change every minute. The idea is that when you log into your account remotely (say when you want to log into your UNIX account in Princeton from an Internet Cafe) then you have to type the numbers that appear in the card in addition to your PIN or password.

1. What is the security advantage of such a card over traditional password? That is, what sort of attack can this card resist which cannot be resisted using a standard password mechanism? (When making this comparison, assume that it's possible for users to remember a 6-digits PIN or a password with similar security.)

2. Describe how you would implement such a scheme using pseudorandom functions.

   Assume that the PRF family takes a seed of size n to map n bits to n bits, and that the number of possible devices is M (for $M < 2^n$). How many bits of storage does your implementation use at the server and each of the devices? (there is an implementation that uses at most $O(n)$ bits in each place).

3. (15 extra points) *Define* what it means that such a scheme is *secure*. That is, write down a list of desired security properties that any such identification scheme should satisfy. Then, make a formal definition of security based on a game in which the scheme is secure if the probability that the adversary "wins" is small. Any scheme that satisfies the formal definition should have the desired security properties.

4. (15 extra points) *Prove* that your construction above satisfies the definition you made. Say how the security depends on n - the number of bits that the device stores in memory (where its running time is polynomial in n) and on k - the number of digits that it displays to the user. You'll get partial credit for a proof sketch or proof idea, as long as it's clearly written, even if you don't have a full formal proof.

[2] **Hint:** First come up with a pseudorandom family with output longer than 1 but shorter than $|s|$. For example, if $s \in \{0,1\}^{n^2}$ then the output can be n. Then show that existence of PRF implies existence of pseudorandom generators and use that to expand your output.

[3] **Hint:** (This is definitely not the only approach to do this.) First note that such a PRF family implies immediately a family where $f_s : \{0,1\}^{|s|} \to \{0,1\}^{|s|/2}$. Then try to use this to get a function f'_s that works only for inputs whose size is a multiple of $|s|/2$. Then try to get a function that works for every finite length string.