0368.4162 Foundations of Cryptography Dec 1, 2017

Problem Set 1, Reference Solution

1. (25 pts) In class, we saw that for any encryption scheme (E, D) for messages of length ℓ, with keys of length $n \leq \ell - 10$, if E is deterministic, there exist two messages m_0, m_1 and an inefficient A such that

$$\Pr\left[A(ct) = m_b \;\middle|\; \begin{array}{l} sk \leftarrow \{0,1\}^n \\ b \leftarrow \{0,1\} \\ ct \leftarrow E_{sk}(m_b) \end{array} \right] > 0.99 \ .$$

Show that the same holds even if E also tosses, say n, coins (on top of the key). That is, an encryption of any message $m \in \{0,1\}^\ell$ is drawn from a distribution $\{ct \mid sk \leftarrow \{0,1\}^n, r \leftarrow \{0,1\}^n, ct = E_{sk}(m;r)\}$.

Solution: Fix any message m_0 and let C_0 be the distribution over encryptions of m_0, where sk and the encryption randomness r are chosen at random (this is different from class, where C_0 was the set of all such encryptions). Note that for a random message m_1 any fixed ciphertext ct decrypts to m_1 under some key sk with probability at most $2^{|sk|-|m_1|} \leq 2^{-10}$. By averaging the same is true when sampling ct from any distribution; in particular,

$$\Pr_{ct \leftarrow C_0, m_1 \leftarrow \{0,1\}^\ell} [\exists sk : \mathsf{Dec}_{sk}(ct) = m_1] \leq 2^{-10} \ .$$

In particular, we can fix some m_1 that satisfies the above.

This gives rise to the following simple adversary A. Given a ciphertext ct, A tests whether ct decrypts to m_1 under some key, and if so declares that m_1 was encrypted and otherwise it declares that m_0 was encrypted. By the above, if m_0 is indeed encrypted, i.e. ct is sampled from C_0, the test passes with probability at most 2^{-10} whereas if m_1 is encrypted, the test always passes. This implies, as we've seen in class, that A guesses correctly with probability at least .99.

3. Let S_0 and S_1 be two non-uniform PPT algorithms, and let $X = \{X_n\}_{n\in\mathbb{N}}$ be a distribution ensemble. Assume that

$$X, S_0(X) \approx_c X, S_1(X) \ .$$

Here $(X, S_b(X))$ denotes the distribution ensemble $\{X_n, S_b(X_n)\}_{n\in\mathbb{N}}$, where a sample (x, y) is given by first sampling $x \leftarrow X_n$ and then sampling $y \leftarrow S_b(x)$ (note that S_b is randomized and may toss additional coins of its own).

Let p be any polynomial. For $b \in \{0,1\}$, consider a new ensemble $Y_b = \{Y_{b,n}\}_{n\in\mathbb{N}}$, given by

$$Y_{b,n} = (X_n, \overbrace{S_b(X_n), \ldots, S_b(X_n)}^{p(n) \text{ times}}) \ ,$$

where a sample $(x, y_1, \ldots, y_{p(n)})$ is given by sampling $x \leftarrow X_n$ and then independently sampling each $y_i \leftarrow S_b(x)$.

(a) (30 pts) Show that $Y_0 \approx_c Y_1$.

(b) (**Bonus**: 10 pts) show that if S_0, S_1 may be **inefficient**, and there exists computationally-secure (secret-key) encryption, the previous claim is not true.

Solution: Given a n.u. PPT distinguisher A that distinguishes Y_0 from Y_1 with advantage $\varepsilon = \varepsilon(n)$, we'll construct a n.u. PPT distinguisher A' that distinguishes $X, S_0(X)$ from $X, S_1(X)$ with advantage ε/p.

Consider $p+1$ hybrids, where for $i \in \{0, \ldots, p\}$,

$$H_i = (X_n, \overbrace{S_1(X_n), \ldots, S_1(X_n)}^{i \text{ times}}, \overbrace{S_0(X_n), \ldots, S_0(X_n)}^{p-i \text{ times}}) \ ,$$

Note that H_0 corresponds to Y_0 and H_p to Y_1. Then, by the triangle inequality there exists an i such that A distinguishes H_{i-1} and H_i with advantage ε/p.

Our distinguisher A' will work as follows. Given a sample x, y, it will sample the first $i-1$ instances from $S_0(x)$ and the last $p-i$ from $S_1(x)$, and will plant y as the ith instance to produce a sample:

$$(x, \overbrace{S_1(x), \ldots, S_1(x)}^{i-1 \text{ times}}, y, \overbrace{S_0(x), \ldots, S_0(x)}^{p-i \text{ times}}) \ ,$$

It will then run A on the produced sample and output whatever A outputs. Note that A' is indeed efficient as the samplers S_0 and S_1 are. It is left to note that if y was sampled from $S_0(x)$, A's view is identical to H_{i-1}, whereas if y was sampled from $S_1(x)$, it is identical to H_i. Thus A' advantage is at least ε/p.

A Uniform Solution: so far, in class, most of the time we've first established the existence of i, and then our new algorithm A' had this i non-uniformly hardwired in its code. I would like to note here that there's also a uniform solution that is rather natural, which is to sample i at random. Showing that it works, however, is a bit more subtle. The potential issue is that $A's$ distinguishing gap for $j \neq i$ now also affects the advantage, and may not always be positive.[1] Still we can show that in our context, it works out due to cancelations:

$$\mathbb{E}_i\left[A(H_{i-1})\right] - \mathbb{E}_i\left[A(H_i)\right] = \frac{1}{p}\sum_{i\in[p]} \mathbb{E}\left[A(H_{i-1})\right] - \mathbb{E}\left[A(H_i)\right] = \frac{1}{p}\left(\mathbb{E}\left[A(H_0)\right] - \mathbb{E}\left[A(H_p)\right]\right) = \pm\varepsilon/p \ .$$

The Bonus: Let (E, D) be a computationally secure encryption scheme with keys of length n and messages of length $2n$. Then, similarly to the first question, the exists two messages m_0, m_1 and an unbounded attacker A that for a random b, can predict b from an encryption of m_b with probability $1 - 2^{-\Omega(n)}$. Let X be the distribution given by encrypting m_b for a random b. Let S_0 be that attacker A (this is where we use the fact that S_0 could be unbounded). Let S_1 be an algorithm that outputs a random bit.

First note that

$$X, S_0(X) \approx_s E_{sk}(m_b), b \approx_c E_{sk}(m_{b'}), b \equiv X, S_1(X) \ ,$$

where $sk \leftarrow \{0,1\}^n$ and $b, b' \leftarrow \{0,1\}$, and are all independent. The first statistical closeness, follows by the fact that S_0 recovers b with overwhelming probability, and the second computational indistinguishability follows from the security of the encryption scheme.

Second, note that two samples from $S_0(X)$ are easy to distinguish from two samples from $S_1(X)$. In the first case the samples will be identical with probability at least $1 - 2^{-\Omega(n)}$, whereas in the second they will be identical with probability $1/2$.

[1] Try to construct for instance two pairs of distributions A_0, B_0 and A_1, B_1 such that there a distinguisher that distinguishes A_0 from B_0, as well as A_1 from B_1 with advantage 1, but cannot distinguish $\{a \leftarrow A_i \mid i \leftarrow \{0,1\}\}$ from $\{b \leftarrow B_i \mid i \leftarrow \{0,1\}\}$.