0368.3049 **Fall 2013**

Introduction to Modern Cryptography

Benny Chor and Nir Bitansky

Assignment 3

Published December 19, 2013. Due January 9, in mailbox 372 (Schreiber building).
Submission in pairs is encouraged (submission in threes or more is not allowed). A 5-point bonus will be given to typed (as opposed to handwritten) submissions.

Remark: As usual, in all questions, assume that algorithms may be probabilistic, and in all probability statements, the probability is also taken over their random coin tosses. For $n \in \mathbb{N}$, we denote by $[n]$ the set of integers $\{1, \ldots, n\}$.

Problem 1: Public-key encryption is not perfectly secure. Let (K, E, D) be a public-key encryption scheme with $1 - \varepsilon$ correctness; that is, for any $m \in \{0,1\}^*$,

$$\Pr_{\substack{sk,pk \leftarrow K \\ c \leftarrow E_{pk}(m)}} [D_{sk}(c) = m] \geq 1 - \varepsilon \ .$$

(a) **7 Points:** Show that if the scheme is perfectly correct, i.e. $\varepsilon = 0$, there exists an unbounded attacker A such that, for any $m \in \{0,1\}^*$,

$$\Pr_{\substack{sk,pk \leftarrow K \\ c \leftarrow E_{pk}(m)}} [A(c, pk) = m] = 1 \ .$$

(b) **5 point Bonus:** Show that more generally, in the case that $\varepsilon \in [0,1]$, there exists an unbounded attacker A such that, for any $m \in \{0,1\}^*$,

$$\Pr_{\substack{sk,pk \leftarrow K \\ c \leftarrow E_{pk}(m)}} [A(c, pk) = m] \geq 1 - \varepsilon \ .$$

Since some of the people were distracted by length issues: If you want assume for simplicity that $|m| = |pk| = |sk| = |r|$, where r is the randomness used by E_{pk}, are all of fixed size n. This is not essential to the question, and should'nt vary your answer.

Problem 2: Public-key encryption from Trapdoor functions (14 Points). Let $F : \{0,1\}^n \to \{0,1\}^n$ be a trapdoor function with a key sampling algorithm K. Recall that K samples a pair of keys sk, pk, such that given sk, it is possible to compute $x = F_{pk}^{-1}(F_{pk}(x))$. On the other hand, given only pk, F_{pk} is one-way. Assume also that F has an ε-hardcore-bit $B : \{0,1\}^n \to \{0,1\}$. Show how to construct from F a public-key bit-encryption scheme. Prove it is 2ε-semantically-secure.
Hint: $E_{pk}(b; r) = F_{pk}(r), B(r) \oplus b$.
Note: ε-HCB means: $pk, F_{pk}(x), B(x) \approx_{c,\epsilon} pk, F_{pk}(x), u$, where $pk \leftarrow K$, $x \leftarrow U_n$, $u \leftarrow U_1$.

Problem 3: Hybrid arguments and CPA encryption. In this question, we will introduce the "hybrid distributions" technique, and apply it to deduce multiple-message security from single-message security, for CPA encryption.

(a) Let A be an algorithm and let $D_0, \ldots, D_t$ be distributions, such that

$$\Pr_{d \leftarrow D_0}[A(d) = 1] - \Pr_{d \leftarrow D_t}[A(d) = 1] \geq \varepsilon \ .$$

Show that

i. (7 points) There exists $j^* \in [t]$ such that

$$\left| \Pr_{d \leftarrow D_{j^*-1}}[A(d) = 1] - \Pr_{d \leftarrow D_{j^*}}[A(d) = 1] \right| \geq \varepsilon/t \ .$$

ii. (2 point bonus)

$$\mathbb{E}_{j \leftarrow [t]} \left[\Pr_{d \leftarrow D_{j-1}}[A(d) = 1] - \Pr_{d \leftarrow D_j}[A(d) = 1] \right] \geq \varepsilon/t \ .$$

(b) (7 points) Let (K, E, D) be a symmetric-key probabilistic bit-encryption scheme. Let A be a polytime algorithm that can ε-distinguish between encryptions of some two t-bit messages $m, m' \in \{0,1\}^t$:

$$\Pr_{\substack{sk \leftarrow K \\ c_i \leftarrow E_{sk}(m_i)}}[A(c_1, \ldots, c_t) = 1] - \Pr_{\substack{sk \leftarrow K \\ c'_i \leftarrow E_{sk}(m'_i)}}[A(c'_1, \ldots, c'_t) = 1] \geq \varepsilon \ .$$

Consider the distributions $D_0, \ldots, D_t$, where D_i is a distribution on t bit-encryptions $c_1, \ldots, c_i, c'_{i+1}, \ldots, c'_t$ such that the first i plaintext bits are $m_1, \ldots, m_i$ and the last $t - i$ are $m'_{i+1}, \ldots, m'_t$. Let $j^* \in [t]$ as exists by the previous item.
Use A to construct A' (with roughly the same running time as A) that, given m, m', j^* can $\frac{\varepsilon}{t}$-distinguish encryptions of 0 from 1, in a CPA attack; that is,

$$\left| \Pr_{\substack{sk \leftarrow K \\ c^* \leftarrow E_{sk}(0)}}[A'^{E_{sk}(\cdot)}(c^*, j^*, m, m') = 1] - \Pr_{\substack{sk \leftarrow K \\ c^* \leftarrow E_{sk}(1)}}[A'^{E_{sk}(\cdot)}(c^*, j^*, m, m') = 1] \right| \geq \varepsilon/t \ ,$$

where $E_{sk}(\cdot)$ is an oracle that given any bit b returns a random encryption of b.

(c) (5 point bonus) Use A, from the previous item, to construct A' (with roughly the same running time as A) that, given m, m' **but not** j^* can $\frac{\varepsilon}{t}$-distinguish encryptions of 0 from 1, in a CPA attack; that is,

$$\left| \Pr_{\substack{sk \leftarrow K \\ c^* \leftarrow E_{sk}(0)}}[A'^{E_{sk}(\cdot)}(c^*, m, m') = 1] - \Pr_{\substack{sk \leftarrow K \\ c^* \leftarrow E_{sk}(1)}}[A'^{E_{sk}(\cdot)}(c^*, m, m') = 1] \right| \geq \varepsilon/t \ ,$$

where $E_{sk}(\cdot)$ is an oracle that given any bit b returns a random encryption of b.
Guidance: Consider A' that given a single bit encryption c^*:

- samples a random $j \leftarrow [t]$,
- uses the oracle $E_{sk}(\cdot)$ to sample encryptions $c_1, \ldots, c_j, c'_{j+1}, \ldots, c'_t$ according to D_j,
- if $m_j \neq m'_j$, replaces c_j with c^*.
- A' runs A on the resulting sample, obtains the resulting bit b, and returns $b \oplus m_j$.

(d) (7 points) Explain, in no more than four sentences, how to extend the previous items to the case that A only distinguishes encryptions of m and m' in a CPA attack; that is, when it is given an oracle $E_{sk}(\cdot)$.

Problem 4: Partially-homomorphic encryption and private information retrieval.

(a) **14 Points:** In the recitation, we have seen how to construct, from any bit-encryption that is homomorphic to addition modulo 2, a PIR with communication complexity $O(N^{\frac{1}{2}}k)$, where N is the date-base size and k is the size of a single bit encryption. Show how to construct a PIR with communication complexity $O(N^{\frac{1}{3}}k^2)$, under the same assumption.

(b) **10 Point bonus:** In this part of the question, you shall construct (keyed) collision-resistant hash functions from private information retrieval scheme.

i. A function $ECC : \mathbb{Z}_p^n \to \mathbb{Z}_p^m$ is said to be an error-correcting code with distance δ (over alphabet $\mathbb{Z}_p$), if for any two distinct $x, y \in \mathbb{Z}_p^n$, $ECC(x)$ and $ECC(y)$ disagree on at least δm of their coordinates. For $n \in \mathbb{Z}$, let p be a prime such that $p/2 \geq n$), and set $m = p$. Consider the function $ECC_I : \mathbb{Z}_p^n \to \mathbb{Z}_p^p$ that given $y = (y_1, \ldots, y_n) \in \mathbb{Z}_p^n$, computes by interpolation a degree $\leq n-1$ polynomial P_y that passes through the points $(1, y_1), \ldots, (n, y_n)$, and outputs the values $\{P_y(i) : i \in \mathbb{Z}_p\}$ of this polynomial on every point in the field $\mathbb{Z}_p$. Show that ECC_I has distance at least $1/2$.

ii. Assume you are given a PIR scheme (E, A, D) for databases $DB \in \mathbb{Z}_p^p$; namely, DB has p entries, and each entry is an element of $\mathbb{Z}_p$. Further assume that any answer $a = A(DB, c)$ for any encrypted query $c = E(i)$, where $i \in [p]$ is in $\mathbb{Z}_p^2$.[1] We next define a keyed hashing family $H = \left\{H_k : \mathbb{Z}_p^n \to \mathbb{Z}_p^{n/2}\right\}$. A key k for the function is sampled by sampling $n/4$ encrypted queries $k \leftarrow (E(i_1), \ldots, E(i_{n/4}))$, where $i_1, \ldots, i_{n/4}$ are chosen uniformly at random from $[p]$. Given input $y \in \mathbb{Z}_p^n$, H_k first computes $DB_y := ECC_I(y)$, and then computes the PIR answers $a_1, \ldots, a_{n/4}$ for the encrypted queries given in k, with respect to the database DB_y. The output of the function is $(a_1, \ldots, a_{n/4})$.

Show that any A that, given k sampled as above, finds a collision in H_k with probability ε, can be converted into A' that can distinguish, with gap $\geq \varepsilon - 2^{-n/4}$,

$$(E(i_1), \ldots, E(i_{n/4})), i_1, \ldots, i_{n/4}, i'_1, \ldots, i'_{n/4}$$

from

$$(E(i'_1), \ldots, E(i'_{n/4})), i_1, \ldots, i_{n/4}, i'_1, \ldots, i'_{n/4} \ ,$$

where $i_1, \ldots, i_{n/4}, i'_1, \ldots, i'_{n/4}$ are chosen independently and uniformly at random from $[p]$, and either the first or the second tuple is PIR-encrypted.

Problem 5: Digital Signatures (28 Points). Recall that a signature scheme $(K, Sign, Ver)$ is (ε, t)-existentially-unforgeable if, for any polytime A,

$$\Pr_{sk,vk \leftarrow K}\left[A^{Sign_{sk}(\cdot)}(vk) = (m^*, \sigma^*) : \begin{array}{l} A \text{ obtains from } Sign_{sk}(\cdot) \text{ signatures} \\ \text{on } \{m_i\}_{i\in[t]} \text{ of his (adaptive) choice, and} \\ m^* \notin \{m_i\}_{i\in[t]}, Ver_{pk}(m^*, \sigma^*) = 1 \end{array}\right] \leq \varepsilon \ .$$

(a) Let $H : \{0,1\}^* \to \{0,1\}^k$ be an ε-collision-resistant hash function. Let $(K, Sign, Ver)$ be a signature scheme for messages of length k with signatures of size k that is (ε, t)-existentially unforgeable.

[1] The answer a is encoded in $\mathbb{Z}_p^2$ just to simplify notation. The important thing is that it is much shorter than the DB size.

Describe a new signature scheme $(K', Sign', Ver')$ for messages of arbitrary length, but still with signature size k. Show that it is $(2\varepsilon, t)$-existentially-unforgeable.

(b) Let $f : \{0,1\}^k \to \{0,1\}^m$ be a one-way function. Recall the one-time signature discussed in the recitation, where the secret key sk consists of $2n$ random inputs $\left\{(x_i^0, x_i^1) \leftarrow \{0,1\}^{k\times 2}\right\}_{i\in[n]}$ for f, the public verification key vk consists of $\left\{(f(x_i^0), f(x_i^1))\right\}_{i\in[n]}$. A signature $\sigma(m)$, on $m = m_1, \ldots, m_n \in \{0,1\}^n$, consists of $x_1^{m_1}, \ldots, x_n^{m_n}$, and verification is done by applying f to each of the n pieces and comparing to the relevant n pieces of vk.

Show that the scheme is 1-existentially-unforgeable. Specifically, show that an algorithm A that breaks $(\varepsilon, 1)$-existential-unforgeablility, can be used to construct A' that runs roughly in the same time as A, and inverts $f(x)$ for a random x, with probability $\varepsilon/2n$.

(c) Recall the sequential multi-message stateful signature scheme described in the recitation, based on a one-time signature scheme $(K, Sign, Ver)$.

- Initially one-time keys are sampled $(sk_0, vk_0) \leftarrow K$. The initial state of the signer is $st_0 = sk_0$, and the public verification key is $vk = vk_0$.
- Before signing a message the i-th message m_i, the signer's state st_{i-1} includes
  i. All previous messages $\mathbf{m}_{i-1} = m_1, \ldots, m_{i-1}$.
  ii. Previous one-time signing and verification keys $\mathbf{sk}_{i-1} = sk_0, sk_1, \ldots, sk_{i-1}$ and verification keys $\mathbf{vk}_{i-1} = vk_0, vk_1, \ldots, vk_{i-1}$.
  iii. Previous one-time signatures $\boldsymbol{\sigma}_{i-1} = \sigma_1, \ldots, \sigma_{i-1}$.

  To sign m_i, the signer first samples a new pair of one-time keys $(sk_i, vk_i) \leftarrow K$. It then, computes a signature $\sigma_i = Sign_{sk_{i-1}}(m_i, vk_i)$. It then publishes as the signature $\{vk_j, m_j, \sigma_j\}_{j\le i}$, and adds $(sk_i, vk_i, m_i, \sigma_i)$ to the current state st_{i-1}, resulting in a new state st_i.
- The signature is verified by verifying all signatures along the chain: $\left\{Ver_{pk_{j-1}}(m_j, vk_j, \sigma_j)\right\}_{j\le i}$.

Show that any attacker A that breaks (ε, t)-existential-unforgeability of the scheme, can be converted to A' that runs roughly in the same time as A, breaks $(\varepsilon/(t+1), 1)$-existential-unforgeability of the underlying one-time scheme.

(d) **Reusing Randomness in El Gamal Signature Scheme.** Recall that in El Gamal signature scheme, discussed in lecture 8, the signature generation is *randomized*. The signer is supposed to choose a new, independent random k for signing each message. Show that if the signer uses the same k repeatedly, then it is possible to extract the secret key, x, from two (signature, message) pairs and the public key.

Problem 6: Implementing a toy example of Shamir's secret sharing (16 Points). Using Sage (or a different package of your choice), set up a system for 3-out-of-5 secret sharing scheme over the finite field $\mathbb{Z}_7$. Generate two different polynomials$f[x], g[x]$ that have different free terms $(f(0) \neq g(0)$, yet $(f(i) = g(i)$ for $i = 1, 2)$.

In class, we argued that the secret can be expressed as a *linear combination* of the shares. Demonstrate this for two sets of participants: $\{1,2,3\}$ and $\{1,2,5\}$. For each set, compute explicitly the coefficients for extracting the secret. For example, in case of the first set, you should find the coefficients b_1, b_2, b_3 such that $h(0) = b_1 h(1) + b_2 h(2) + b_3 h(3)$ for every degree 2 polynomial. Find such coefficients c_3, c_4, c_5 for the second set of participants as well. Demonstrate that for $f[x], g[x]$ chosen above, your linear combinations indeed work.