0368.4162: Introduction to Cryptography, Fall 2010 Instructors: Ran Canetti and Iftach Haitner

Problem Set 3

January 3, 2011 Due: Monday January 17 2011 in class

1. **[30 points]** Let $H = \{H^n\}_{n\in\mathbf{N}}$, $H^n = \{H_h : \{0,1\}^* \to \{0,1\}^n\}_{h\in\{0,1\}^n}$ be a function ensemble. Say that H is target collision resistant (TCR) if any polytime adversary A wins in the following game only with negligible probability (in n):

   - h is chosen at random from $\{0,1\}^n$
   - A gets h, generates m
   - A gets r chosen at random from $\{0,1\}^n$
   - A generates m', r' and wins if $m \neq m'$ and $H_h(m,r) = H_h(m',r')$.

   (Note that this notion is different than the relaxation discussed in class of collision resistance.)

   (a) **[10 points]** Show that if there exist TCRs then there exist TCRs which are not CRFs.

   (b) **[20 points]** Let $S = (GEN, SIG, VER)$ be a signature scheme that's EU-CMA secure when applied to messages of fixed length n. (EU-CMA security is the unforgability defined in class. That is, any PPT attacker that has access to a signing oracle can generate a (message,signature) pair that pass the verification test, where the message was not signed before by the signing oracle, only with negligible probability.) Consider the following scheme $S' = (GEN', SIG', VER')$:

   - GEN' runs GEN, and in addition chooses $h \in \{0,1\}^n$ and adds h to both the signing key and the verification key.
   - $SIG'((sk,h),m) = SIG(sk, h(m,r)), r$. Here $r \leftarrow \{0,1\}^n$ and sk is the signing key of SIG.
   - $VER'((vk,h),m,(s,r)) = VER(vk, H_h(m,r), s)$. Here vk is the verification key of VER.

   Show that S' is EU-CMA secure for messages with arbitrary length.

   Direction: Note that the security of S' depends on two primitives: the EU-CMA unforgeable scheme S and the TCR ensemble H. One way to handle this situation is to first reduce the security of S to that of S', under the assumption that the TCR never breaks, and then bound the probability that the TCR breaks. To do that, you need to precisely specify the meaning of the "CTR never breaks" event. (This is not the only way to go about this proof, but it might be the simplest.)

2. **[30 points]** Let $F = \{F^n\}_{n\in\mathbf{N}}$, $F^n = \{F_k : \{0,1\}^{2n} \to \{0,1\}^{2n}\}_{k\in\{0,1\}^n}$ be a PRP ensemble. Consider the following shared key encryption scheme for message space $M_n = \{0,1\}^n$:

   - For $m \in \{0,1\}^n$, $ENC(k,m) = f_k(m \circ r)$, where $r \leftarrow U_n$. (Here $\circ$ denotes concatenation.)
   - $DEC(k,c)$ outputs the first n bits of $f_k^{-1}(c)$.

   Is (ENC, DEC) IND-CPA secure, with respect to message space M and leakage function $l(m) = |n|$? Either prove based on the assumption that the underlying ensemble F is a PRP, or provide a counter example.

3. A popular approach for encrypting long messages using public-key encryption is to first encrypt a (relatively short) key k using asymmetric encryption, and then to encrypt the message using a symmetric encryption scheme with key k. (The main gain here is efficiency, since symmetric schemes are significantly faster than asymmetric schemes.) The goal of this exercise is to prove the security of this approach, called *hybrid encryption.*

   Let $E^a = (GEN^a, ENC^a, DEC^a)$ be an asymmetric encryption scheme, and let $E_s = (GEN^s, ENC^s, DEC^s)$ be a symmetric encryption scheme. Let $E_h = (GEN^h, ENC^h, DEC^h)$ be the following asymmetric encryption scheme:

- $GEN^h(1^n)$ computes $(ek, dk) \leftarrow GEN^a(1^n)$.
- $ENC^h(1^n, ek, m)$ computes $k \leftarrow GEN^s(1^n)$, and then outputs c_1, c_2 where $c_1 = ENC^a(ek, k)$, $c_2 = ENC^s(k, m)$. (Recall that both Enc^a and Enc^s are randomized algorithms.)
- $DEC^h(dk, c = (c_1, c_2)) = DEC^s(DEC^a(dk, c_1), c_2)$

(a) **[20 points]** Show that if E^a and E^s are a CPA secure than E^h is CPA secure.

(b) **[10 points]** Assume that E^a is CCA secure and E^s is CPA secure. Is E^h CCA secure? (Either prove or show a counter example.)

(c) **[Bonus: 20 points]** Let $M = (GEN^c, AUTH, VER)$ be a secure MAC scheme, and consider the following variant of E^h, denoted $E'^h = (GEN'^h, ENC'^h, DEC'^h)$:

- $GEN'^h = GEN^h$
- $ENC'^h_h(1^n, ek, m)$ computes $k_s \leftarrow GEN^s(1^n)$ and $k_c \leftarrow GEN^c(1^n)$, and then outputs c_1, c_2, t where $c_1 = ENC^a(ek, k_s \circ k_c)$, $c_2 = ENC^s(k_s, m)$, $t = AUTH(k_c, c_2)$.
- $DEC'^h(dk, c = (c_1, c_2, t))$: let $(k_s, k_c) = DEC^a(dk, c_1)$, $m = DEC^s(k_s, c_2)$; if $VER(k_c, c_2) = 1$ then output m; else output $\perp$.

Assume that E^a is CCA secure, E^s is CPA secure, and M is a secure MAC scheme. Is E'^h CCA secure? If so, prove it. Else, show a counter example and then show how to modify E'^h in a minimal way so that it becomes CCA secure.

4. Recall the definitions of security of encryption schemes, as discussed in class. A symmetric encryption scheme $E = (Gen, Enc, Dec)$ is correct with respect to message domain $M = \{M_n\}_{n \in \mathbf{N}}$ if $\text{Prob}(k \leftarrow Gen(1^n), Dec(k, Enc(k, m)) \neq m)$ is negligible in n.

IND security. E is IND-CPA secure for message domain M if it is correct and any PPT adversary B wins the IND-CPA game with probability at most $1/2 + \nu(n)$, where ν is a negligible function. The IND-CPA proceeds as follows, given security parameter n:

- A key $k \leftarrow Gen(1^n)$ and a bit $b \in \{0, 1\}$ are chosen.
- Whenever B outputs a pair of messages m_0, m_1 in M_n with $|m_0| = |m_1|$, a value $c = Enc(k, m_b)$ is computed and given back to B.
- When B outputs a bit b' the game ends. B wins if $b' = b$.

If B is restricted to generate only a single pair of messages then the resulting notion of security is called IND. If B is restricted to choose all pairs before receiving any ciphertext then the notion of security is called IND-KPA.

SEM security. E is SEM-CPA secure if it is correct and for any adversary A there exists an adversary A' such that for any environment algorithm Env we have $\text{IDEAL}(Env, A') \approx_c \text{REAL}(Env, Enc, A)$, where:

- $\text{REAL}(Env, Enc, A') = \{\text{REAL}_n(Env, Enc, A')\}_{n \in \mathbf{N}}$, and $\text{IDEAL}_n(Env, Enc, A')$ describes the output of Env from the following interaction:
  - Env gets input 1^n
  - The following process is repeated until Env outputs a decision value in $\{0, 1\}$: Env generates a message $m \in M_n$, and obtains $v = A(Enc(k, m))$ for $k \rightarrow Gen(1^n)$.
- $\text{IDEAL}(Env, A') = \{\text{IDEAL}_n(Env, A')\}_{n \in \mathbf{N}}$, and $\text{IDEAL}_n(Env, A')$ describes the output of Env from the following interaction:
  - Env gets input 1^n
  - The following process is repeated until Env outputs a decision value in $\{0, 1\}$: Env generates a message $m \in M_n$, and obtains $v = A'(|m|)$.

If Env is restricted to generate only a single message m then the resulting notion of security is called SEM. If Env is restricted to choose all messages before receiving any output from the adversary then the notion of security is called SEM-KPA.

(a) **[10 points]** We have shown in class that a symmetric encryption scheme is IND for message domain M if and only if it is SEM for message domain M. show that a scheme is IND-CPA for message domain M if and only if it is SEM-CPA for this message domain.

(b) **[10 points]** Show that if a symmetric encryption scheme is IND then it is IND-KPA.

(c) **[10 points]** Show that if there exist IND-KPA schemes then there exist IND-KPA schemes which are not IND-CPA.