COM S 6830 – Cryptography Oct 6, 2009

Lecture 12: Definitions of Message Security

Instructor: Rafael Pass *Scribe: Gabriel Bender*

1 Multimessage-Secure Encryption

Last time, we proved that no stateless encryption scheme is multimessage secure. We can get around this problem by making use of a pseudorandom function.

Proposition 1 *Let $\{f_s : \{0,1\}^{|s|} \to \{0,1\}^{|s|}\}$ be a family of pseudorandom functions. Then the following encryption scheme is multimessage-secure:*

$$\begin{aligned} Gen(1^n) &= (s \leftarrow \{0,1\}^n : s) \\ Enc_k(m) &= (r \leftarrow \{0,1\}^n : r \mid\mid m \oplus f_k(r)) \\ Dec_k(r||c) &= (c \oplus f_k(r)) \end{aligned}$$

Proof. Suppose not. Then there exist destinguisher $\mathcal{D}$ and messages $m_0, \ldots, m_{q(n)}$ and $m'_0, \ldots, m'_{q(n)}$ s.t. $\mathcal{D}$ distinguishes the following two sets with non-negligible probability:

$$\{k \leftarrow Gen(1^n) : Enc_k(m_0), Enc_k(m_1), \ldots, Enc_k(m_{q(n)})\}$$
$$\{k \leftarrow Gen(1^n) : Enc_k(m'_0), Enc_k(m'_1), \ldots, Enc_k(m'_{q(n)})\}$$

In particular, there exists a polynomial $q(n)$ s.t. for infinitely many $n \in \mathbb{N}$, $\mathcal{D}$ distinguishes the two sets given above. For fixed n, we apply the Hybrid lemma with the following hybrids:

$$\begin{aligned} H_1 &= \{s \leftarrow \{0,1\}^n; r_0, \ldots, r_{q(n)} \leftarrow \{0,1\}^n : \\ &\quad r_0 \mid\mid m_0 \oplus f_s(r_0), \ldots, r_{q(n)} \mid\mid m_{q(n)} \oplus f_s(r_{q(n)})\} \\ H_2 &= \{RF \leftarrow (\{0,1\}^n \to \{0,1\}^n); r_0, \ldots, r_q \leftarrow \{0,1\}^n : \\ &\quad r_0 \mid\mid m_0 \oplus RF(r_0), \ldots, r_{q(n)} \mid\mid m_{q(n)} \oplus RF(r_{q(n)})\} \\ H_3 &= \{r_0, \ldots, r_{q(n)} \leftarrow \{0,1\}^n; P_0, \ldots, P_{q(n)} \leftarrow \{0,1\}^n : \\ &\quad r_0 \mid\mid m_0 \oplus P_0, \ldots, r_{q(n)} \mid\mid m_{q(n)} \oplus P_{q(n)}\} \\ H_4 &= \{r_0, \ldots, r_q \leftarrow \{0,1\}^n; P_0, \ldots, P_{q(n)} \leftarrow \{0,1\}^n : \\ &\quad r_0 \mid\mid m'_0 \oplus P_0, \ldots, r_{q(n)} \mid\mid m'_{q(n)} \oplus P_{q(n)}\} \\ H_5 &= \{RF \leftarrow (\{0,1\}^n \to \{0,1\}^n); r_0, \ldots, r_q \leftarrow \{0,1\}^n : \\ &\quad r_0 \mid\mid m'_0 \oplus RF(r_0), \ldots, r_{q(n)} \mid\mid m'_{q(n)} \oplus RF(r_{q(n)})\} \\ H_6 &= \{s \leftarrow \{0,1\}^n; r_0, \ldots, r_{q(n)} \leftarrow \{0,1\}^n : \\ &\quad r_0 \mid\mid m'_0 \oplus f_s(r_0), \ldots, r_{q(n)} \mid\mid m'_{q(n)} \oplus f_s(r_{q(n)})\} \end{aligned}$$

H_1 and H_2 are indistinguishable because they can be viewed as the output of the same oracle Turing Machine, with oracle f_s for H_1 and RH for H_2. By the definition of a pseudorandom function, H_1 and H_2 are therefore indistinguishable. By the same argument, it can distinguish H_6 from H_5 with no better than negligible probability.

When all the r_i are distinct, all the $RF(r_i)$ in H_2 are selected independently and at random, so that this distribution is identical to that of H_3. The probability that there exists $i = j$ s.t. is bounded above by $\binom{q(n)}{2}/2^n$, a union bound over pairs of messages that both messages in a pair are equal. This is a negligible function. So we are unable to distinguish between H_2 and H_3 except with negligible probability. The same argument shows that H_5 and H_4 are indistinguishable.

The indistinguishability of H_3 and H_4 follows from the security of the one-time pad cipher: roughly speaking, given an encryption, all plaintext decryptions are equally likely unless we have access to a key. This concludes our proof. ■

2 Stronger Definitions of Security

We might also wish to consider definitions of security that are stronger than multi-message security.

Let $\Pi = (Gen, Enc, Dec)$ be an encryption scheme. Let A be a non-uniform PPT and $n \in \mathbb{N}, b \in \{0, 1\}$. We define a random variable

$$\begin{aligned} \text{IND}_b^{O_1,O_2}(\Pi, A, n) \quad &= \quad k \leftarrow Gen(1^n); m_0, m_1, \sigma \leftarrow A^{O_1(k)}(1^n); \\ & \quad\quad c \leftarrow Enc_k(m_b) : A^{O_2}(C, \sigma) \end{aligned}$$

Each definition below requires that

$$\{IND_0^{O_1,O_2}(\Pi, A, n)\}_{n \in \mathbb{N}} \approx \{IND_1^{O_1,O_2}(\Pi, A, n)\}_{n \in \mathbb{N}}$$

However, the oracles O_1 and O_2 that are available to an adversary depend on the definition:

- Chosen-Message (Chosen-Plaintext) Attack/CPA Security: O_1 provides access to Enc_k and O_2 always returns 0 and therefore provides no useful information.

- CCA1/Lunch Time Attack: O_1 provides access to both Enc_k and Dec_k; O_2 always returns 0.

- CCA2: O_1 provides access to both Enc_k and Dec_k; O_2 also provides access to both Enc_k and Dec_k. In this case, we only quantify over Turing Machines A that never invoke the decryption oracle of O_2 on the encrypted input message c.

The encryption scheme we proposed at the beginning of the lecture is CPA- secure because knowing the encryption of a message $(r \mid\mid m \oplus f_k(r))$ does us no good unless the selected value of r is the same as for the input ciphertext $(c = r_c \mid\mid m \oplus f_k(r))$. This happens with probability $\frac{1}{2^n}$ for each message that is encrypted by O_1, and O_1 is allowed to query at most a polynomial number of messages, so the likelihood that it our distinguisher queries $f_s(r_c)$ is negligible. By exactly the same argument, our encryption scheme is CCA1-secure. However, it is not CCA-2 secure because, given an encrypted message, we could query the decryption oracle of O_2 on input $(r_c \mid\mid 0)$ to obtain the value of $f_s(r_c)$.