CS655: Cryptography **Feb 1 and Feb 6, 2007**

Lecture Outline for Symmetric Encryption

4 Symmetric Encryption

Readings: Sections 4.1–4.8 of Bellare&Rogaway

4.1 Symmetric encryption schemes

- A *symmetric encryption scheme* is given by $\mathcal{SE} = (\mathcal{K}, \mathcal{E}, \mathcal{D})$, where
  - $\mathcal{K}$ is the key generation algorithm.
    $\mathcal{K}$ is randomized.
  - $\mathcal{E} : \mathsf{Keys}(\mathcal{SE}) \times \{0,1\}^* \rightarrow \{0,1\}^* \cup \{\perp\}$ is the encryption algorithm.
    $\mathcal{E}$ may be randomized and/or stateful.
  - $\mathcal{D} : \mathsf{Keys}(\mathcal{SE}) \times \{0,1\}^* \rightarrow \{0,1\}^* \cup \{\perp\}$ is the decryption algorithm.
    $\mathcal{D}$ is deterministic.

- The scheme $\mathcal{SE} = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ is said to provide *correct encryption* if for any key $K \in \mathsf{Keys}(\mathcal{SE})$, any sequence of messages $M_1, \cdots, M_q \in \{0,1\}^*$, and any sequence of ciphertext $C_1 \xleftarrow{\$} \mathcal{E}_K(M_1), C_2 \xleftarrow{\$} \mathcal{E}_K(M_2), \cdots, C_1 \xleftarrow{\$} \mathcal{E}_K(M_1)$ that may arise in encrypting $M_1, \cdots, M_q$, we have $D_K(C_i) = M_i$ for each $C_i \neq \perp$.

- The plaintext space of $\mathcal{SE} = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ is given by

$$\mathcal{M} = \left\{ M \,\middle|\, \mathsf{Pr}\left[K \xleftarrow{\$} \mathcal{K}; C \xleftarrow{\$} \mathcal{E}_K(M) : C \neq \perp \right] = 1 \right\}$$

- When encryption is randomized, encrypting the same message twice may get different ciphertexts. This implies that the ciphertext space must be larger than the plaintext space, which often means that length of ciphertext is longer then length of plaintext.

- Encryption may be stateful. This means that encryption is also dependent on a state that is initialized in some pre-specified way. After each invocation of $\mathcal{E}$, the state is updated. Typically, the state is a counter.

- Encryption can be both randomized and stateful. But this is rare.

4.2 Some symmetric encryption schemes

Encryption modes:

- **ECB mode**: Given a block cipher $E : \mathcal{K} \times \{0,1\}^n \rightarrow \{0,1\}^n$, using the ECB mode yields a stateless and deterministic symmetric encryption scheme.

- **CBC$ mode**: CBC with random IV. Stateless, randomized encryption.

- **CBCC mode**: CBC with IV being a counter. Stateful, deterministic encryption. Note that counter is not allowed to reuse. (The modeling of states in Figure 4.3 is problematic. A better way seems to be to require the encryption function to provide an interface with two methods: initialization (with a key) and encryption. There seems to be another bug with counter checking as well.)
- **CTRC mode**: Counter mode. Stateful, deterministic encryption.
- **CTR$ mode**: Counter mode with randomized counter. Stateless, randomized encryption.

4.3 Issues in privacy

- **The main question: what is the security requirement for encryption?**
- What is insecurity? Recover the key means insecure. Recover plaintext means insecure. Recover partial information about plaintext also means insecure.
- What is "ideal encryption"? No adversary can gain non-negligible information about the plaintext from a ciphertext.
  - Compare with perfect secrecy.
  - Normally we do leak some information about plaintext, e.g., length of the plaintext. Very expensive/clumsy to hide length as well.
- Goal: Given a ciphertext, no adversary can gain any information about the plaintext other than the length of the plaintext. We need to formalize this.
- How to formalize this?
  - A symmetric cipher has perfect secrecy if and only if it satisfies the following condition:

    $$\forall_{M_0 \in \mathcal{M}} \forall_{M_1 \in \mathcal{M}} \forall_{C_0 \in \mathcal{C}} \left(\mathsf{Pr}\left[\mathsf{CT} = C_0 \mid \mathsf{PT} = M_0\right] = \mathsf{Pr}\left[\mathsf{CT} = C_0 \mid \mathsf{PT} = M_1\right] \right)$$

  - A symmetric cipher is insecure if and only if

    $$\exists_{M_0 \in \mathcal{M}} \exists_{M_1 \in \mathcal{M}} \exists_{C_0 \in \mathcal{C}} \left(\mathsf{Pr}\left[\mathsf{CT} = C_0 \mid \mathsf{PT} = M_0\right] \neq \mathsf{Pr}\left[\mathsf{CT} = C_0 \mid \mathsf{PT} = M_1\right] \right)$$

  - A symmetric cipher is computationally insecure if and only if there is an adversary that can output two messages M_0, M_1 of equal length, and when given a ciphertext C, can tell whether it is the ciphertext of M_0 or M_1.

4.4 Indistinguishability under chosen-plaintext attack

- **The IND-CPA game**: The game is between the Challenger and an adversary.
  1. The Challenger chooses $K \xleftarrow{\$} \mathcal{K}$.
  2. The adversary chooses two equal-length messages M_0 and M_1, and sends them to the Challenger.
  3. The Challenger chooses $b \xleftarrow{\$} \{0, 1\}$, and sends $C \xleftarrow{\$} \mathcal{E}_K(M_b)$ to the adversary.

4. The adversary A outputs $b' \in \{0,1\}$ and wins if $b' = b$.

The advantage is defined to be $|\Pr[A \text{ wins }] - 0.5|$.

If the advantage is small, it means that the adversary cannot distinguish whether a ciphertext is the encryption of M_0 or the encryption of M_1, even if M_0 and M_1 are chosen by the adversary

- **Definition in the notes**: Consider two experiments:
  - $\mathbf{Exp}^{\textsf{ind-cpa-1}}_{\mathcal{SE}}(A) : \{\ K \xleftarrow{\$} \mathcal{K};\ d \xleftarrow{\$} A^{\mathcal{E}_K(\mathrm{LR}(\cdot,\cdot,1))};$ Return $d\ \}$.
    - ∗ where $\mathcal{E}_K(\mathrm{LR}(\cdot,\cdot,\mathrm{b}))$ is the Left-or-right encryption oracle, defined as follows if $|M_0| \neq |M_1|$ then return $\perp$, else returns $\mathcal{E}_K(M_b)$
  - $\mathbf{Exp}^{\textsf{ind-cpa-0}}_{\mathcal{SE}}(A) : \{\ K \xleftarrow{\$} \mathcal{K};\ d \xleftarrow{\$} A^{\mathcal{E}_K(\mathrm{LR}(\cdot,\cdot,0))};$ Return $d\ \}$.

  The *IND-CPA advantage* of A is defined as

$$\mathbf{Adv}^{\text{ind-cpa}}_{\mathcal{SE}}(A) = \mathsf{Pr}\left[\mathbf{Exp}^{\textsf{ind-cpa-1}}_{\mathcal{SE}}(A) = 1\right] - \mathsf{Pr}\left[\mathbf{Exp}^{\textsf{ind-cpa-0}}_{\mathcal{SE}}(A) = 1\right]$$

  - The choice of world is made only once.

- The IND-CPA game is formalized as the following experiment:
  - $\mathbf{Exp}^{\textsf{ind-cpa-cg}}_{\mathcal{SE}}(A) : \{\ b \xleftarrow{\$} \{0,1\}; K \xleftarrow{\$} \mathcal{K};\ b' \xleftarrow{\$} A^{\mathcal{E}_K(\mathrm{LR}(\cdot,\cdot,\mathrm{b}))};$ Return $b == b'\ \}$.
    - ∗ One difference from the informal description of the IND-CPA game is that here the adversary is given an oracle encrypting either left or right, rather than just the ability to invoke it only once.
    - ∗ The advantage can be defined as $\mathsf{Pr}\left[\mathbf{Exp}^{\textsf{ind-cpa-cg}}_{\mathcal{SE}}(A) = 1\right] - /1/2$.
  - We have the following:

$$\mathbf{Adv}^{\textsf{ind-cpa}}_{\mathcal{SE}}(A) = 2 \cdot \mathsf{Pr}\left[\mathbf{Exp}^{\textsf{ind-cpa-cg}}_{\mathcal{SE}}(A) = 1\right] - 1.$$

4.5 Example chosen-plaintext attacks

- *Attack on ECB*
- *Any deterministic, stateless scheme is insecure*
- *CBC with counter IV is insecure*

4.6 Semantic security

- SEM-CPA security: captures the idea that a secure encryption scheme should hide all computable information about an unknown plaintext.
- Formalization of SEM-CPA using two experiments $\mathbf{Exp}^{\textsf{ss-cpa-1}}_{\mathcal{SE}}(A)$ and $\mathbf{Exp}^{\textsf{ss-cpa-0}}_{\mathcal{SE}}(A)$.

  1. The Challenger picks $K \xleftarrow{\$} \mathcal{K}$, initialize the cipher.

2. The adversary picks a message space $\mathcal{M}_i$ (an algorithm that samples a message from the space) and sends to the Challenger.
3. The Challenger draws two random messages M_i and M_i' from the space, makes sure that they have the same length, encrypts M_i in $\mathbf{Exp}_{\mathcal{SE}}^{\text{ss-cpa-1}}(A)$ and M_i' in $\mathbf{Exp}_{\mathcal{SE}}^{\text{ss-cpa-0}}(A)$, and returns the ciphertext to the adversary.
4. Repeat the previous two steps for a total of q times.
5. The adversary outputs a function f and a value Y.
6. The adversary wins if and only if $f(M_1, \cdots, M_q) = Y$.

The SEM-CPA advantage of A is defined as

$$\mathbf{Adv}_{\mathcal{SE}}^{\text{sem-cpa}}(A) = \mathsf{Pr}\left[\mathbf{Exp}_{\mathcal{SE}}^{\text{ss-cpa-1}}(A) = 1\right] - \mathsf{Pr}\left[\mathbf{Exp}_{\mathcal{SE}}^{\text{ss-cpa-0}}(A) = 1\right]$$

- **Theorem: IND-CPA security implies SEM-CPA security**

4.7 Security of CTR modes

- **Security of CTRC mode:** Let $F : \mathcal{K} \times \{0,1\}^n \to \{0,1\}^n$ be a family of functions and let $\mathcal{SE} = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be the corresponding CTRC symmetric encryption scheme using F. Then given any A that attacks the IND-CPA security of $\mathcal{SE}$, there exists B that attacks the PRF security of F such that

$$\mathbf{Adv}_{\mathcal{SE}}^{\text{ind-cpa}}(A) \leq 2 \cdot \mathbf{Adv}_{\mathcal{SE}}^{\text{prf}}(B)$$

  - First show that for any adversary A attacking $\mathcal{SE}[\mathsf{Func}(n,n)]$,

$$\mathbf{Adv}_{\mathcal{SE}[\mathsf{Func}(n,n)]}^{\text{ind-cpa}}(A) = 0.$$

    This is true because $\mathcal{SE}[\mathsf{Func}(n,n)]$ is essentially one-time pad.

  - Then show that A can be used to construct B that attacks PRF security of F. The basic idea is that when B gets a random function g, then no algorithm can have any advantage attacking encryption done using g in CTRC mode, and when B gets a function drawn from F, A has an advantage attacking the encryption. Then by betting on A's success, one can distinguish a random function from a function drawn from F.

- **The PRF/PRP switching lemma (in Section 3.9 of the Bellare-Rogawar notes):** Let $E : \mathcal{K} \times \{0,1\}^n \to \{0,1\}^n$ be a function family. Let A be an adversary that asks at most q oracle queries. Then we have

$$\left|\mathbf{Adv}_E^{\text{prf}}(A) - \mathbf{Adv}_E^{\text{prp}}(A)\right| \leq \frac{q(q-1)}{2^{n+1}}$$

- **Security of CTR\$ mode:** Let $F : \mathcal{K} \times \{0,1\}^n \to \{0,1\}^n$ be a family of functions and let $\mathcal{SE} = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be the corresponding CTR\$ symmetric encryption scheme using F. Then given any A that attacks the IND-CPA security of $\mathcal{SE}$ that makes at most σ queries, there exists B that attacks the PRF security of F such that

$$\mathbf{Adv}_{\mathcal{SE}}^{\text{ind-cpa}}(A) \leq 2 \cdot \mathbf{Adv}_{\mathcal{SE}}^{\text{prf}}(B) + \frac{0.5\sigma^2}{2^n}$$

4.8 Security of CBC with a random IV

- **Security of CBC\$ mode:** Let $E : \mathcal{K} \times \{0,1\}^n \rightarrow \{0,1\}^n$ be a block cipher and let $\mathcal{SE} = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be the corresponding CBC\$ symmetric encryption scheme using E. Then given any A that attacks the IND-CPA security of $\mathcal{SE}$ that makes at most σ queries, there exists B that attacks the PRF security of F such that

$$\mathbf{Adv}^{\mathsf{ind\text{-}cpa}}_{\mathcal{SE}}(A) \leq 2 \cdot \mathbf{Adv}^{\mathsf{prf}}_{\mathcal{SE}}(B) + \frac{0.5\sigma^2}{2^n}$$

  - First show that for any adversary A attacking $\mathsf{CBC\$[Func}(n,n)]$,

$$\mathbf{Adv}^{\mathsf{ind\text{-}cpa}}_{\mathsf{CBC\$[Func}(n,n)]}(A) \leq \frac{\sigma^2}{2^n}.$$