CS 120/CSCI E-177: Introduction to Cryptography

Salil Vadhan and Alon Rosen Oct. 3, 2006

Lecture Notes 5:

Private-Key Encryption: Computational Security

Recommended Reading.

- Katz-Lindell 3.2, 3.3

1 Introduction

- Motivation: Recall *statistical security*: for every $m_0, m_1 \in \mathcal{P}$ and set T of ciphertexts,

$$|\Pr[E_K(m_0) \in T] - \Pr[E_K(m_1) \in T]| \leq \varepsilon.$$

  That is, there is no test T that distinguishes the encryptions of any pair of messages with probability better than ε.

  - Still requires $|\mathcal{K}| \geq (1-\varepsilon) \cdot |\mathcal{P}|$.

- *(Computational) indistinguishability*: only consider tests T defined by "feasible" algorithms A, i.e. replace the event "$E_K(m) \in T$" with "$A(E_K(m)) = 1$".

- First Goal: Construct computationally secure encryption schemes that go beyond the Shannon barrier (i.e. have $|\mathcal{K}| \ll |\mathcal{P}|$.

  - Still restricted to "one use" and passive adversary.

- Later: Model and achieve security for multiple messages and active adversaries.

2 Asymptotic formalization

- Need a *security parameter* 1^n: n is chosen by the sender and receiver in advance depending on the level of security they want.

- A "feasible" adversary is any poly(n)-time adversary. We will always allow the adversary to be *nonuniform*, i.e. have a program of size poly(n).

- Require that G, E, D all run in polynomial time (i.e. poly(n)). G now takes n as input (in unary).

- Main point: G, E, D run in some fixed polynomial time (e.g. time n^2) but security must hold against adversaries with even larger running time. Thus, as we set n larger and larger (e.g. as technology improves), the scheme takes much less time to use than it does to break.

- The message space can change with the security parameter: $\mathcal{P} = \bigcup_n \mathcal{P}_n$. For example, $\mathcal{P}_n$ can be $\{0,1\}, \{0,1\}^n, \{0,1\}^*$.

- What should ε be? A function $\varepsilon : \mathbb{N} \to [0,1]$ is *negligible* if for every c, there exists n_0 s.t. $\varepsilon(n) < 1/n^c$ for all $n > n_0$.

Definition 1 (indistinguishable encryptions (asymptotic version)) *Let (G, E, D) be an encryption scheme over $\mathcal{P} = \bigcup_n \mathcal{P}_n$ where all messages in $\mathcal{P}_n$ have the same length. (G, E, D) has* (computationally) indistinguishable encryptions *if for every (nonuniform) PPT A, there is a negligible function ε such that for all $m_0, m_1 \in \mathcal{P}_n$,*

$$|\Pr[A(E_K(m_0)) = 1] - \Pr[A(E_K(m_1)) = 1]| \leq \varepsilon(n),$$

where the probabilities above are taken over $K \stackrel{R}{\leftarrow} G(1^n)$, the coin tosses of E_K, and the coin tosses of A.

In other words, no feasible algorithm/adversary can distinguish the encryptions of any pair of messages with nonnegligible probability (a.k.a. "advantage").

- To handle varying message lengths (e.g. $\mathcal{P}_n = \{0,1\}^*$): only consider pairs (m_0, m_1) with $|m_0| = |m_1| \leq \text{poly}(n)$.

3 Concrete formalization

- feasible adversary = time $\leq t$ on specific computational model (e.g. $t = 2^{100}$ cycles on a Pentium D) using a program of size $\leq t$.
- G, E, D should all run in time $\ll t$.

Definition 2 (indistinguishable encryptions (concrete version)) *Let (G, E, D) be an encryption scheme over $\mathcal{P}$ where all messages in $\mathcal{P}$ have the same length. (G, E, D) is (t, ε)*-secure *if for every probabilistic algorithm A running in time t and for all $m_0, m_1 \in \mathcal{P}$,*

$$|\Pr[A(E_K(m_0)) = 1] - \Pr[A(E_K(m_1)) = 1]| \leq \varepsilon.$$

where the probabilities above are taken over $K \stackrel{R}{\leftarrow} G$, the coin tosses of E_K, and the coin tosses of A.

- G doesn't take any input.

4 Examples of Insecure Schemes

- Shift cipher
- Substitution cipher
- Biased one-time pad: $G(1^n)$: for $i = \{1, \ldots, n\}$, set $k_i = \{1 \text{ with pr. } .49; 0 \text{ with pr. } .51\}$. Output $k = k_1 \ldots k_n$. $\mathcal{P} = \{0,1\}^n$, $\mathrm{E}_k(m) = m \oplus k$.

5 Equivalent Definitions

Definition 3 (guessing-indistinguishability (Katz-Lindell)) *Let (G, E, D) be an encryption scheme over $\mathcal{P} = \bigcup_n \mathcal{P}_n$ where all messages in $\mathcal{P}_n$ have the same length. An encryption scheme (G, E, D) has* guessing-indistinguishable encryptions *if for every (nonuniform) PPT A, there is a negligible function ε such that A succeeds in the following game with probability at most $1/2 + \varepsilon(n)$:*

1. *A outputs a pair of messages $m_0, m_1 \in \mathcal{P}_n$.*
2. *A random key $k \stackrel{R}{\leftarrow} G(1^n)$ and a random bit $b \stackrel{R}{\leftarrow} \{0,1\}$ are chosen.*
3. *A is given $c \stackrel{R}{\leftarrow} E_k(m_b)$ and outputs a bit b'.*
4. *A* succeeds *if $b' = b$.*

Proposition 4 *An encryption scheme has indistinguishable encryptions if and only if it has guessing-indistinguishable encryptions.*

Note *reducibility argument*: we show how to convert a poly-time algorithm A violating guessing-indistinguishability into a poly-time algorithm violating indistinguishability. Similar in spirit to the reductions done in **NP**-completeness (but more delicate to analyze, due to probabilities).

Definition 5 *Let (G, E, D) be an encryption scheme over $\mathcal{P} = \bigcup_n \mathcal{P}_n$ where all messages in $\mathcal{P}_n$ have the same length. An encryption scheme (G, E, D) satisfies* semantic security *if for every nonuniform PPT A, there is a nonuniform PPT A' such that for every distribution M on $\mathcal{P}_n$, every function $f : \mathcal{P}_n \to \{0,1\}^*$, and every (nonuniform) PPT A,*

$$\begin{aligned} \Pr[A(E_K(M)) = f(M)] &\leq \Pr\left[A'(1^n) = f(M)\right] + \text{neg}(n) \\ &\leq \max_v \{\Pr[f(M) = v]\} + \text{neg}(n), \end{aligned}$$

where the probabilities are taken over M, $K \stackrel{R}{\leftarrow} G(1^n)$, and the coin tosses of E and A.

- The function f captures the information about the message that the adversary is trying to compute.
- Examples:
  - $f(m) = m$: recovering entire plaintext.
  - $f(m) = m_1$: recovering first bit.
- Semantic security says that the best an adversary can compute f after seeing the ciphertext is essentially the same as before seeing the ciphertext — namely guess the most likely value.

Theorem 6 *An encryption scheme has indistinguishable encryptions if and only if it has semantic security.*

Hence if we assume (or prove) indistinguishability (i.e. distinguishing encryptions is hard), then we can deduce semantic security (i.e. computing information about the message is hard).

Proof: We'll only prove that indistinguishable encryptions implies semantic security.

Let A be any PPT adversary, M a distribution on $\mathcal{P}_n$ and $f : \mathcal{P}_n \to \{0,1\}^*$ any function. Fix any message $m_0 \in \mathcal{P}$, and let $A'(1^n)$ be the algorithm that chooses $k \stackrel{\text{R}}{\leftarrow} G(1^n)$ and runs $A(E_k(m_0))$. Then,

$$\begin{aligned} \Pr\left[A(E_K(M)) = f(M)\right] &\leq \Pr\left[A(E_K(m)) = f(M)\right] + \operatorname{neg}(n) \\ &= \Pr\left[A'(1^n) = f(M)\right] + \operatorname{neg}(n) \\ &\leq \max_v \{\Pr\left[f(M) = v\right]\} + \operatorname{neg}(n) \end{aligned}$$

■