CS6111 (2025): Assignment 2

Instructor: John Augustine

Due: Sept. 14, 2025 (Noon 12:00 PM IST).

Name: Shri Prathaa M
Roll No: EE22B144

1. (1 mark) Read Chapter 3 of Katz and Lindell, Edition 3, except starred subsections.

Solution: Yes. I have read the starred subsections.

2. (3 marks) The Electronic Code Book (ECB) mode of operation encrypts a message M by partitioning it into n-bit blocks $M_1, M_2, \ldots, M_L$ and computing $C_i = F_k(M_i)$ for each block. Is the ECB mode of operation CPA-secure? If yes, provide proof. If not, describe a concrete chosen-plaintext attack.

Solution:
Not CPA-secure (Not EAV-secure too).

Attack: Choosing two messages of equal length with different block pattern, for example,

$$M^{(0)} = (X, X), \qquad M^{(1)} = (X, Y) \text{ with } X \neq Y.$$

Under ECB, $C^{(0)} = (F_k(X), F_k(X))$ has two identical blocks, while $C^{(1)} = (F_k(X), F_k(Y))$ has distinct blocks with probability $\geq 1 - negl(n)$. In a CPA experiment, first input the $(M^{(0)}, M^{(1)})$; upon receiving the challenge ciphertext $C^\star$, output 0 iff its two blocks are equal, else output 1.

A private-key encryption scheme $\Pi = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ is **CPA-secure** if for all probabilistic polynomial-time adversaries $\mathcal{A}$, there exists a negligible function negl such that:

$$\Pr\left[\mathsf{PrivK}^{\mathsf{cpa}}_{\mathcal{A},\Pi}(n) = 1\right] \leq \frac{1}{2} + \mathsf{negl}(n)$$

However the probability in ECB is $\geq \frac{1}{2}(1) + \frac{1}{2}(1 - negl(n))$ This attack succeeds with probability $1 - \mathsf{negl}(n)$. Thus, ECB mode is not CPA-secure.

3. (3 marks) Can we achieve "perfect" CPA-security? If so, how would you define it? Otherwise, explain why not.

Solution: Perfect CPA-security would require that even an unbounded adversary, given access to an encryption oracle, cannot distinguish between the encryptions of two chosen messages with probability better than $\frac{1}{2}$. That is, for any adversary $\mathcal{A}$ (even computationally unbounded), we would have:

$$\Pr[\mathcal{A} \text{ succeeds}] = \frac{1}{2}.$$

To achieve this (in a setting like One-Time Pad), the encryption scheme would have to satisfy $|K| \geq |M|$ and the key distribution makes the ciphertext distribution uniform for any 2 queried messages (Thus, it has to use a perfectly random key (not generated by a pseudorandom generator). However, in the CPA setting, the same key is reused for many encryption queries. This allows an unbounded adversary to query the encryption oracle on all possible plaintexts and eventually recover the key. Therefore, perfect CPA-security is *impossible* to achieve.

4. (5 marks) Let F be a pseudorandom function. Comment whether the following constructions of functions from F are also pseudorandom, with proof.

   (a) $F'_k(x) = F_k(x)||F_k(F_k(x))$

   Solution:

   Not a PRF.

   Distinguisher: Query x to the new function, F' to get $y_1 \| y_2$. Then query y_1 to F' and get $z_1 \| z_2$. We have $y_1 = F_k(x)$ and $y_2 = F_k(F_k(x)) = F_k(y_1) = z_1$, so $y_2 = z_1$ always. For a truly random function $R : \{0,1\}^n \to \{0,1\}^{2n}$, the probability that $y_2 = z_1$ is 2^{-n}. Hence the comparitive advantage for a distinguisher is $1 - 2^{-n}$.

   (b) $F'_k(x) = F_k(x \oplus c_0)||F_k(x \oplus c_1)$, where c_0 and c_1 are arbitrary but fixed bit strings that the distinguisher knows.

   Solution:

   No, the function F' is not pseudorandom.

   Define a distinguisher $\mathcal{D}$ as follows:

   1. Query the oracle for (x) with input $x = 0$. Parse the $2n$-bit output as (a, b), where a are the first n bits $((c_0))$ and b are the second n bits $((c_1))$.

2. Query the oracle for (x) with input $x = c_1 \oplus c_2$. Parse the $2n$-bit output as (a', b'), where:

$$a' = (c_0 \oplus x) = (c_0 \oplus (c_1 \oplus c_2)) = (c_1) \quad (\text{since } c_0 \oplus c_0 = 0),$$
$$b' = (c_1 \oplus x) = (c_1 \oplus (c_1 \oplus c_2)) = (c_0).$$

3. The distinguisher $\mathcal{D}$ outputs 1 if and only if $a' = b$ and $b' = a$.

Let us now compute the distinguisher's advantage:

$$\Pr[\mathcal{D}^{(\cdot)}(1^n) = 1] = 1 \quad (\text{by the construction of } F', \text{ the condition always holds}).$$
$$\Pr[\mathcal{D}^{r(\cdot)}(1^n) = 1] = \Pr[a' = b \wedge b' = a] \quad \text{for a truly random function } r.$$
$$= \frac{1}{2^n} \cdot \frac{1}{2^n} = \frac{1}{2^{2n}}.$$

The final equality holds because a' and b' are independent n-bit strings, each uniformly random and independent of the previous outputs a and b.
Therefore

$$\left|\Pr[\mathcal{D}^{(\cdot)}(1^n) = 1] - \Pr[\mathcal{D}^{r(\cdot)}(1^n) = 1]\right| = \left|1 - \frac{1}{2^{2n}}\right| = 1 - \mathsf{negl}(n).$$

Since a distinguisher $\mathcal{D}$ exists that succeeds with probability significantly better than random guessing, the function $F'(x)$ is not pseudorandom.

5. (7 marks) Let F be a strong pseudorandom permutation. For each of the following schemes, the shared key is a uniformly random $k \in \{0,1\}^n$. State how decryption is done in each scheme and whether the scheme has:

   (i) indistinguishable encryptions in the presence of an eavesdropper (EAV-security),
   (ii) CPA-security (chosen-plaintext attack security),

   Support your answer with an attack or a proof of security.

   (a) To encrypt $m \in \{0,1\}^{2n}$, parse m as $m_1 \parallel m_2$ with $|m_1| = |m_2|$. Then choose uniform $r \in \{0,1\}^n$ and send $\langle r, m_1 \oplus F_k(r), m_2 \oplus F_k(r)\rangle$.

Solution:
Decryption: Compute $p = F_k(r)$ and output $m_1 = c_1 \oplus p$, $m_2 = c_2 \oplus p$.
Security: *Not EAV-secure (and not CPA-secure).*
Because $c_1 \oplus c_2 = (m_1 \oplus p) \oplus (m_2 \oplus p) = m_1 \oplus m_2$, an eavesdropper learns the XOR of the two messages. Hence not EAV secure

Distinguishing attack for CPA security: Choose $m^{(0)} = (X, X)$ and $m^{(1)} = (X, Y)$; Output 0 if $c_1 \oplus c_2 = 0^n$, else output 1. For $m^{(0)}$ we surely decrypt it to 0. For $m^{(1)}$, since it is a PRP, it always gives different c_1 and c_2. Thus with probability of 1 we can perform CPA attack. It is not CPA secure.

(b) To encrypt $m \in \{0,1\}^{n/2}$, choose uniformly random bit strings r_1 and r_2 of length $n/2$ each and send the ciphertext $\langle F_k((r_1 \oplus r_2) \parallel m)\rangle$.

Solution:

Decryption: As the function is a PRP, it is a bijective map given the knowledge of k and access to F_k. Thus we can compute the inverse bijective map to get the $((r_1 \oplus r_2) \parallel m)$. The last $\frac{n}{2}$ bits give us m.

In the eavesdropping setting, the adversary sees one ciphertext. Since $s = r_1 \oplus r_2$ is uniformly random, the input $x = s \parallel m$ is uniformly distributed over $\{0,1\}^n$. Thus, $c = F_k(x)$ is computationally indistinguishable from random, f as F is a strong PRP. Thus it is EAV- secure

Under chosen-plaintext attack, the adversary can query the encryption oracle. Consider two queries:

1. Query m: get $c = F_k(s \parallel m)$ with $s = r_1 \oplus r_2$.
2. Query m': get $c' = F_k(s' \parallel m')$ with $s' = r_1' \oplus r_2'$.

The inputs to F_k are $s \parallel m$ and $s' \parallel m'$. Since the outputs are distinct for a PRP (being a bijective map), for $m \neq m'$, we get distinct outputs. For purely random the probability of getting distinct output for both queries is $1 - 2^{-n}$. The distinguishing advantage thus is $\leq 2^{-n}$. With computational bound on query to F_k it becomes impossible to identify even if the message was sent again (due to randomness of $r_1 \oplus r_2$). Thus, we can only distinguish the 2 messages with negligible probability making it CPA secure.

6. (5 marks) Let $F : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ be a pseudorandom function. Consider a function $H : \{0,1\}^{2n} \times \{0,1\}^{2n} \to \{0,1\}^{2n}$ that takes $2n$-bit inputs $L_0 \parallel R_0$ and $2n$-bit keys $k_1 \parallel k_2$ (where L_0, R_0, k_1, k_2 are all n-bit strings). For $i = 1, 2$, define L_i and R_i as

$$L_i = R_{i-1} \ ; \ R_i = L_{i-1} \oplus F_{k_i}(R_{i-1})$$

and output the $2n$-bit string $L_2 \parallel R_2$. So, $H_{k_1 \parallel k_2}(L_0 \parallel R_0) = L_2 \parallel R_2$. Is H a pseudorandom function? Specify a distinguisher or a proof that it is indeed pseudorandom.

Solution:

$$L_1 = R_0, \qquad R_1 = L_0 \oplus F_{k_1}(R_0),$$
$$L_2 = R_1 = L_0 \oplus F_{k_1}(R_0), \qquad R_2 = L_1 \oplus F_{k_2}(R_1) = R_0 \oplus F_{k_2}(L_0 \oplus F_{k_1}(R_0)).$$

Query two inputs with the same right half R_0 and different left halves L_0, L_0'. Let outputs be (L_2, R_2) and (L_2', R_2'). Then

$$L_2 \oplus L_2' = (L_0 \oplus F_{k_1}(R_0)) \oplus (L_0' \oplus F_{k_1}(R_0)) = L_0 \oplus L_0'.$$

So the XOR of the left halves of the outputs equals the XOR of the left halves of the inputs *always*. A truly random function on $2n$-bit inputs exhibits this with probability about 2^{-n}. Thus on querying 2 messages L_0, R_0 and L_0', R_0, we can distinguish it from uniformly random function by a probability $\geq 1 - negl(n)$.Therefore H is not pseudorandom.

[The function H is a permutation because it is invertible: given output (L_2, R_2) and keys k_1, k_2, the input (L_0, R_0) can be uniquely recovered via:

$$L_1 = R_2 \oplus F_{k_2}(L_2), \quad R_0 = L_1, \quad L_0 = L_2 \oplus F_{k_1}(L_1).$$

Since H is a permutation, it is injective. Thus, for any distinct inputs $x_1, \ldots, x_q$, the outputs $H(x_1), \ldots, H(x_q)$ are distinct. In contrast, a random function $\mathcal{R}$ maps distinct inputs to independent random outputs, so the probability that q outputs are distinct is approximately $\exp(-q^2/2^{2n+1})$.

The probability that all q outputs of a random function $\mathcal{R}$ are distinct is:

$$\prod_{i=0}^{q-1} \left(1 - \frac{i}{2^{2n}}\right).$$

Using the approximation $1 - x \leq e^{-x}$ for $x \geq 0$, we get:

$$\prod_{i=0}^{q-1} \left(1 - \frac{i}{2^{2n}}\right) \leq \exp\left(-\sum_{i=0}^{q-1} \frac{i}{2^{2n}}\right).$$

Since $\sum_{i=0}^{q-1} i = \frac{(q-1)q}{2}$, we have:

$$\exp\left(-\frac{(q-1)q}{2^{2n+1}}\right) \approx \exp\left(-\frac{q^2}{2^{2n+1}}\right).$$

For $q = 2^n + 1$, this gives $\exp(-1/2) \approx 0.6065$.

A distinguisher that outputs "PRF" if all outputs are distinct thus has advantage $\approx 1 - 0.6065 = 0.3935$, which is non-negligible. Hence, H is not pseudorandom.]

Acknowledgment

I would like to thank Swathi Shree N (EE22B149) for their useful discussion on solving the problems.