Assignment 2

PRFs, CPA Security and Modes of Operation

Instructor: Sruthi Sekar *TAs:* Lakshya Gadhwal, Ravi B Prakash

Problem 1: PRF or not?

Let $\mathcal{F}^{(1)} = \{f_k^{(1)} : \{0,1\}^n \to \{0,1\}^n\}_{k\in\{0,1\}^n}$ and $\mathcal{F}^{(2)} = \{f_k^{(2)} : \{0,1\}^n \to \{0,1\}^n\}_{k\in\{0,1\}^n}$ be length-preserving pseudorandom functions (cf. Def 1 or 2, Lec7). For the following constructions, state whether $\mathcal{F}'$ is a pseudorandom function. If yes, prove it; if not, show an attack.

a) $\mathcal{F}' = \{f'_k : \{0,1\}^{n-1} \to \{0,1\}^{2n}\}_{k\in\{0,1\}^n}$, where $f'_k(x) := f_k^{(1)}(0||x)||f_k^{(1)}(1||x)$.

b) $\mathcal{F}' = \{f'_k : \{0,1\}^{n-1} \to \{0,1\}^{2n}\}_{k\in\{0,1\}^n}$, where $f'_k(x) := f_k^{(1)}(0||x)||f_k^{(1)}(x||1)$.

c) $\mathcal{F}' = \{f'_k : \{0,1\}^{n} \to \{0,1\}^{n}\}_{k\in\{0,1\}^n}$, where $f'_k(x) = f_k^{(1)}(f_k^{(2)}(x))$.

d) $\mathcal{F}' = \{f'_k : \{0,1\}^{n} \to \{0,1\}^{n}\}_{k\in\{0,1\}^n}$, where $f'_k(x) := \overline{f_k^{(1)}(x)}$ (overline denoted bit-string complement).

e) $\mathcal{F}' = \{f'_k : \{0,1\}^{n} \to \{0,1\}^{n}\}_{k\in\{0,1\}^n}$, where $f'_k(x) := f_k^{(1)}(\overline{x})$.

f) Recall the tree-based construction of PRF from length-doubling PRF (cf. GGM, Lec 7). Recall that the value $f_k(x)$, $x \in \{0,1\}^n$, was computed by taking the key k as the seed of the root PRG, and computing the leaf output y_x (taking the path given by x). Consider the "dual" construction $\mathcal{F}' = \{f'_k : \{0,1\}^n \to \{0,1\}^n\}_{k\in\{0,1\}^n}$: to compute $f'_k(x)$, you use the input x as the seed of the root PRG, and then output the leaf value y_k.

Problem 2: EAV, CPA, or Both?

Let F be a pseudorandom function (cf. Def 1 or 2, Lec7) and G be a pseudorandom generator with expansion factor $\ell(n) = n+1$ (cf. Def 1 or 2, Lec 4). For each of the following encryption schemes, state whether the scheme has indistinguishable encryptions in the presence of an eavesdropper (cf. Def 7, Lec 3) and whether it is CPA-secure (cf. Def 3, Lec 7). (In each case, the shared key is a uniform $k \in \{0,1\}^n$.) Explain your answer.

a) To encrypted $m \in \{0,1\}^{n+1}$, choose uniform $r \in \{0,1\}^n$ and output the ciphertext $\langle r, G(r) \oplus m\rangle$.

b) To encrypt $m \in \{0,1\}^n$, output the ciphertext $m \oplus F_k(0^n)$.

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

Problem 3: Understanding CPA Security

For a symmetric-key encryption (SKE) scheme, refer to Def 7, Lec 3 for indistinguishability against an eavesdropper, [1, Definition 3.19] for security of multiple encryptions against an eavesdropper, and Def 3, Lec 7 for CPA security.

A) **Gap b/w Multiple-EAV- and CPA-Securities:** Assuming the existence of pseudorandom functions, prove that there is an encryption scheme that has indistinguishable multiple encryptions in the presence of an eavesdropper, but is not CPA-secure.
Hint: The scheme need not be "natural". You will need to use the fact that in a chosen-plaintext attack, the adversary can choose its queries to the encryption oracle adaptively.

B) **Insecurity of Deterministic encryption:** Let $\Pi = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ be a SKE scheme with *deterministic* encryption. Show that Π cannot be CPA-secure.

C) **Combiners:** Let $\Pi_1 = (\mathsf{Gen}_1, \mathsf{Enc}_1, \mathsf{Dec}_1)$ and $\Pi_2 = (\mathsf{Gen}_2, \mathsf{Enc}_2, \mathsf{Dec}_2)$ be two SKE schemes for which it is known that at least one is CPA-secure (and we don't know which one). Construct a SKE scheme Π that is CPA-secure as long as Π_1 or Π_2 is secure. (Such a construction is called a "*combiner*".)
Hint: it is instructive to first think about constructing such schemes against eavesdroppers.

Problem 4: Custom Counter Mode

Let F be a pseudorandom permutation (cf. Lec 7). Consider the mode of operation (variant of CTR mode) in which a uniform value $\mathtt{ctr} \in_R \{0,1\}^n$ is chosen, and the i^{th} ciphertext block c_i is computed as $c_i := F_k(\mathtt{ctr} + i + m_i)$. Show that this scheme does not have indistinguishable encryptions in the presence of an eavesdropper.

Problem 5: Pseudorandom Permutations (Hard)

Recall the definitions of pseudorandom functions (PRFs), pseudorandom permutations (PRPs) and strong pseudorandom permutations (SPRPs) from Lec 7.

A) Assume PRPs exist. Show that there exists an F' that is a PRP but not an SPRP.
Hint: Construct an F' such that $F'_k(k) = 0^{|k|}$.

B) Let $F = \{f_k : \{0,1\}^n \to \{0,1\}^n\}_{k \in \{0,1\}^n}$ be a PRP, defined a fixed-length encryption scheme $\Pi = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ as:

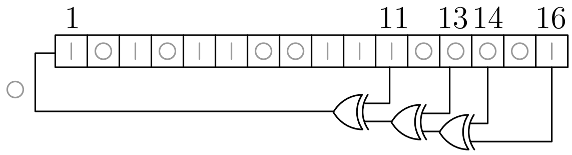
- $\mathsf{Gen}(1^n)$: Pick and output $k \in_R \{0,1\}^n$.
- $\mathsf{Enc}(k, m)$: Pick $r \in_R \{0,1\}^{n/2}$. For $m \in \{0,1\}^{n/2}$, compute and output $c := f_k(r||m)$.

Show how to decrypt, and prove that this scheme is CPA-secure for messages of length $n/2$.

Advanced Problem: Attack on LFSRs

Linear Feedback Shift Registers (LFSRs) are a type of shift register where the input bit is a linear function of its previous state, typically involving XOR operations on specific bits of the register. Due to their simplicity and efficiency, LFSRs are widely used in cryptography, particularly in stream ciphers, pseudo-random number generation, and scrambling of data for error detection and correction. Their deterministic nature allows for the generation of sequences with desirable properties like long periods and good statistical characteristics, making them suitable for generating key streams in stream ciphers such as A5/1, used in GSM mobile communications. However, the linearity of LFSRs also makes them vulnerable to certain cryptanalytic attacks, leading to the development of more complex, non-linear variants in modern cryptographic applications.

In this problem, you will investigate you can mount an attack on a simple LFSR model. Let us take the example of a simple (Fibonacci) LFSR:



The symbols in the boxes represent the current state of the LFSR; in particular, the current state of the LFSR above is 1010110011100001. Since the state consists of 16 bits, this is called a 16-bit LFSR. We index the positions on the state from left to right with indices ranging from 1 to 16. Certain positions are said to be "tapped"; in the above example, the positions at indices $11, 13, 14$, and 16 are tapped. This LFSR can be used as a "PRNG" as follows:

- `Init`: It takes in a 16-bit seed s and initialises the state $\texttt{st}_0$ be the bits of the seed written left-to-right on the LFSR body. For instance, if the above is the starting state of the LFSR, then $\texttt{st}_0 = 1010110011100001$.
- `GetBits`: The bit y output is the bit at the rightmost position of the LFSR, i.e., the bit at index 16. For instance, in the above example, the output bit would be 1. If the current state is $\texttt{st}_t$, then the next state, $\texttt{st}_{t+1}$ would be obtained by shifting each bit one position to the right (the rightmost bit would "fall off" the LFSR as output), and replacing the first (leftmost) bit with the XOR of all the bits which were at the tapped position in $\texttt{st}_0$. In the above example, you can see that the bit that would get replaced in the first position in the next state would be given by $1 \oplus 0 \oplus 0 \oplus 1 = 0$. Thus, the next state of the LFSR would be 0101011001110000.

Now assume that someone uses a n-bit LFSR (where $n \geq 2$) with an unknown seed (initial state).

a) You know the taps to be at positions marked by the distinct indices $[I_1, I_2, \ldots, I_T]$ where $2 \leq T \leq n$ and $1 \leq I_j \leq n \; \forall j \in \{1, 2, \ldots, T\}$. Suppose you are given the first n output bits of the LFSR in order. Design an attack that would enable you to predict all the subsequent output bits of the LFSR with 100% accuracy.

b) Now assume that you do not know where the taps are (however, you do know there are ≥ 2 tapped positions). Suppose you are given the first $2n$ output bits of the LFSR in order. Design an attack that would enable you to predict all the subsequent output bits of the LFSR with 100% accuracy.

Historical Note: The Content Scrambling System (CSS) is a digital rights management (DRM) and encryption system used to protect content on DVD-Video discs. Introduced in 1996, CSS was designed to prevent unauthorized copying and distribution of DVD content by scrambling the data on the disc, making it accessible only to authorized devices with the appropriate decryption keys. CSS employs a pair of LFSRs to generate the pseudo-random keystream that is XORed with the data on the DVD to encrypt it. The LFSRs used in CSS are relatively simple: each LFSR is a 17-bit register with a specific set of taps. The two LFSRs work together, with one controlling the operation of the other, to produce the final keystream. However, the simplicity of these LFSRs and the overall design of CSS made it vulnerable to cryptanalysis, and CSS was eventually broken, leading to the development of more secure content protection systems.

References

[1] Jonathan Katz and Yehuda Lindell. *Introduction to Modern Cryptography.* Chapman & Hall/CRC Cryptography and Network Security Series. Chapman and Hall/CRC, 2nd edition, 2014.