Probabilistic Encryption

CS 5158/6058 Data Security and Privacy

Spring 2018

Instructor: Boyang Wang

CPA Security Game

Def. A symmetric-key encryption Π is indistinguishable under chosen-plaintext attacks, or is CPA-secure, if for all PPT adversaries $\mathcal{A}$ there is a negligible function s.t.

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

- Is CPA model really necessary?
  - U.S. knew **`AF`** was the target, suspected Midway
  - U.S. sent "Midway is low on water."
  - Japan sent "**`AF`** is low on water."
  - Practice: Who was Adversary in CPA?

Deterministic v.s. Probabilistic

- Deterministic enc. is not secure under multiple ciphertexts or under CPA
  - $M_0 = (m_{0,1}, m_{0,2})$ and $M_1 = (m_{1,1}, m_{1,2})$
    - $m_{0,1} == m_{0,2}$ and $m_{1,1} != m_{1,2}$
    - Return $C_b = (c_{b,1}, c_{b,2})$
    - If $c_{b,1} == c_{b,2}$, $b' = 0 = b$; else $b' = 1 = b$

- Need <u>probabilistic encryption</u>
  - Output different ciphertexts from a same message

Pseudorandom Function

- Func(m,n): a function family includes all the mappings from $\mathcal{D}=\{0,1\}^m$ —> $\mathcal{R}=\{0, 1\}^n$
  - E.g., m =3 and n =2, one f(d) from Func(3,2)

d	000	001	010	011	100	101	110	111
f(d)	10	11	11	00	10	01	11	01

- |Func(m,n)| = $2^{n \cdot 2^m}$
  - 2^n outputs, each output has 2^m inputs

Keyed Function

- A <u>keyed function</u> *F* mapping from $\mathcal{D}$={0,1}^{m} —> $\mathcal{R}$={0, 1}^{n}:

$$F : \mathcal{K} \times \mathcal{D} \to \mathcal{R}$$
$$\mathcal{K} = \{0,1\}^l, \mathcal{D} = \{0,1\}^m, \mathcal{R} = \{0,1\}^n$$

  - First input is called the key *k*
  - *k* is <u>chosen uniformly</u> from $\mathcal{K}$

$$F_k(x) = F(k,x) = y$$

  - *F* is efficient (i.e., polynomial time)
  - Given a key *k*, F_k is **<u>deterministic</u>**

Pseudorandom Function

- F is a PRF: if F_k is indistinguishable from f
  - k is chosen uniformly from K
  - f is chosen uniformly from Func(m, n)

Def. Let $F : \{0,1\}^l \times \{0,1\}^m \to \{0,1\}^n$ be an efficient keyed function. F is a PRF is for all PPT adversary $\mathcal{A}$, there is a negligible function s.t

$$|\Pr[\mathcal{A}^{F_k(\cdot)}(1^l) = 1] - \Pr[\mathcal{A}^{f(\cdot)}(1^l) = 1]| \leq \mathtt{negl}(l)$$

PRF v.s. Func(m,n)

$$F_k : \{0,1\}^m \to \{0,1\}^n, \quad k \in \{0,1\}^l$$

- PRF *F* is not even close to Func(m, n)
  - $|F| = 2^l$
  - $|\text{Func}(m, n)| = 2^{n \cdot 2^m}$

- <u>Practice:</u> if m = 4, l = 2, and n = 2
  - What is $|F|$? and what is |Func(m, n)|?
  - $|F| = 2^2 = 4$; $|\text{Func}(4, 2)| = 2^{32} = 4{,}294{,}967{,}296$

4 v.s. 4,294,967,296

Pseudorandom Generator

- Pseudorandom Generator (PRG)
  - Efficient (polynomial-time), deterministic function
  - Use a <u>short</u> random string to generate a <u>long</u> pseudorandom string
  - Polynomial-time adversary can only negligibly distinguish PRG's output from random

From PRG to PRF

- Build a PRF from PRG (GGM method, 1984)
  - Goldreich, Goldwasser & Micali (Turing Award'12)

- PRG: $G: \{0,1\}^n \longrightarrow \{0, 1\}^{2n}$
  - $G(x) = G_0(x)||G_1(x)$,
  - $G_0(0)$ is the left half of $G(0)$, $G_1(0)$ is the right half
  - E.g., $G(x) = 0010$ $G_0(x) = 00$, $G_1(x) = 10$
  - E.g., $G(x) = 111000$, $G_0(x) = 111$, $G_1(x) = 000$
  - $G(x) = 01011111$, $G_0(x) = ??$ $G_1(x) = ??$

From PRG to PRF

- PRG: G: $\{0,1\}^n$ —> $\{0, 1\}^{2n}$
  - $G(x) = G_0(x)||G_1(x)$
- PRF: F: $\{0,1\}^m$ —> $\{0,1\}^n$
  - <u>Recursively call G m rounds</u>
  - Given $x_1x_2\ldots x_m$ and k as input, each x_i is 0 or 1
    - R1: Compute $G_{x1}(k)$
    - R2: Compute $G_{x2}(G_{x1}(k))$
    - R3: Compute $G_{x3}(Gx_2(G_{x1}(k)))$
  - Recursively run G, after m rounds, get PRF's output

From PRG to PRF

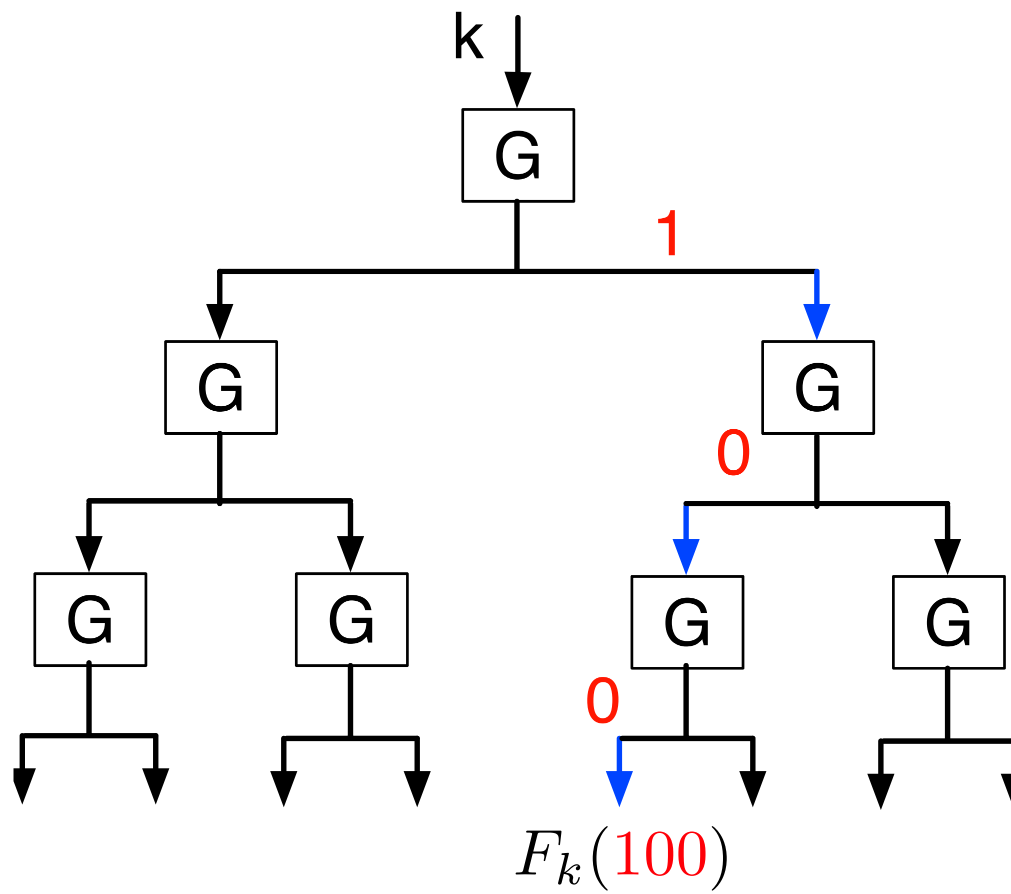
- PRG: G: $\{0,1\}^n \longrightarrow \{0, 1\}^{2n}$
  - $G(x) = G_0(x) || G_1(x)$
- PRF: F: $\{0,1\}^m \longrightarrow \{0,1\}^n$
  - Given $x_1x_2\ldots x_m$ and k as input, each x_i is 0 or 1
  - Recursively call G m rounds
  - Each call use previous output as next input
  - $F_k(x_1x_2\ldots x_m) = G_{x_m}(G_{x_{m-1}}(\ldots\ G_{x_2}(G_{x_1}(k))\ldots))$
  - E.g., $m = 2$, $F_k(01) = G_1(G_0(k))$
  - E.g., $m = 2$, $F_k(10) = G_0(G_1(k))$

From PRG to PRF

- PRG, G: $\{0,1\}^n$ —> $\{0, 1\}^{2n}$, PRF, F: $\{0,1\}^m$ —> $\{0,1\}^n$
  - n=1, G(0) —> 10, G(1) —> 01
  - if m = 2 & k = 0, x_1x_2 = 10, what's output $F_k(x_1x_2)$?
  - $F_k(x_1x_2) = F_k(10) = G_0(G_1(k))$
    1. $G_1(k)$ is the right half of G(k)
    2. Given k = 0, G(0) = 10, $G_1(0) = 0$
    3. $G_0(G_1(k)) = G_0(0)$
    4. $G_0(0)$ is the left half of G(0)
    5. G(0) = 10, $G_0(0) = 1$
    6. $F_k(10) = G_0(G_1(k)) = 1$

From PRG to PRF

- A binary tree, each node is a PRG
- Output's left half is input of left child
- An output of PRF is a leaf's left/right half of the output
- Inputs of PRF decide the path in the tree



$F_k(100) = G_0(G_0(G_1(k)))$

From PRG to PRF

- Practice: GGM Method
- PRG, G: $\{0,1\}^n \longrightarrow \{0, 1\}^{2n}$, PRF, F: $\{0,1\}^m \longrightarrow \{0,1\}^n$
  - $n=1$, $G(0) \longrightarrow 10$, $G(1) \longrightarrow 01$
  - $F_k(x_1x_2x_3) = G_{x3}(\ G_{x2}(\ G_{x1}(k)\ )\ )$
  - $k = 1$, $x_1x_2x_3 = 110$, What is the output of $F_k(110)$?

- $G(k) = G(1) = 0\underline{1}$, $G_{x1}(k) = G_1(1) = 1$;
- $G(G_{x1}(k)) = G(1) = 0\underline{1}$, $G_{x2}(G_{x1}(k)) = G_1(1) = 1$;
- $G(G_{x2}(G_{x1}(k))) = G(1) = \underline{0}1$, $G_{x3}(G_{x2}(G_{x1}(k))) = G_0(1) = 0$
- $F_k(110) = G_{x3}(G_{x2}(G_{x1}(k))) = 0$

Permutation Family

- Perm(n,n): a permutation family includes all the permutations from $\mathcal{D}=\{0,1\}^n \longrightarrow \mathcal{R}=\{0,1\}^n$
  - E.g., n = 3, one f(d) from Perm(3,3)

d	000	001	010	011	100	101	110	111
f(d)	100	110	111	001	101	011	000	010

- $|\text{Perm}(n,n)| = (2^n)!$
  - 2^n outputs, each permutation is a <u>bijection</u>

Permutation Family

- $|\text{Perm}(n,n)| = (2^n)!$
- <u>Example:</u> n = 2, then $|\text{Perm}(2,2)| = 4! = 24$
  - 4 of 24 permutations are listed below

d	00	01	10	11
f1(d)	00	01	10	11
f2(d)	11	00	01	10
f3(d)	10	11	00	01
f4(d)	01	10	11	00

Keyed Permutation

- A <u>keyed permutation</u> *F* from $\mathcal{D}$={0,1}^{n} -> $\mathcal{R}$={0, 1}^{n}:

$$F : \mathcal{K} \times \mathcal{D} \rightarrow \mathcal{R}$$
$$\mathcal{K} = \{0,1\}^l, \mathcal{D} = \{0,1\}^n, \mathcal{R} = \{0,1\}^n$$

  - First input is key, *k* is <u>chosen uniformly</u> from $\mathcal{K}$
  - *F* is efficient (i.e., polynomial time)
  - F_k is deterministic
  - F_k is efficiently <u>invertible</u>

$$F_k(x) = y, \quad F_k^{-1}(y) = x$$

Keyed Permutation

$$F : \mathcal{K} \times \mathcal{D} \rightarrow \mathcal{R}$$
$$\mathcal{K} = \{0,1\}^l, \mathcal{D} = \{0,1\}^n, \mathcal{R} = \{0,1\}^n$$

- Example of keyed permutation *F*: l = 2, n = 2

d	00	01	10	11
k=00,f(d)	11	00	01	10
k=01,f(d)	10	11	00	01
k=10,f(d)	01	10	11	00
k=11,f(d)	00	01	10	11

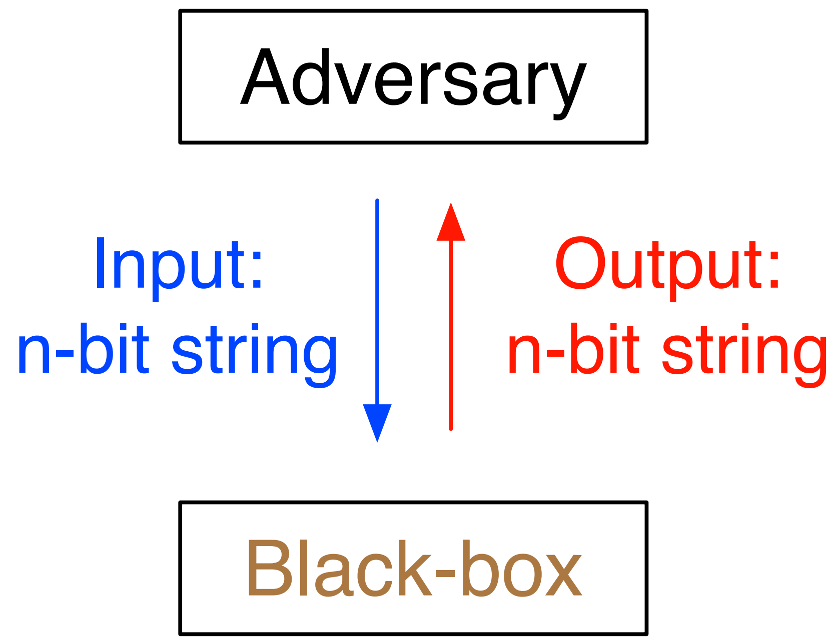
Keyed Permutation

$$F : \mathcal{K} \times \mathcal{D} \to \mathcal{R}$$
$$\mathcal{K} = \{0,1\}^l, \mathcal{D} = \{0,1\}^n, \mathcal{R} = \{0,1\}^n$$

- $|F|$: the no. of permutations in keyed permutation F:
  - Given each key, F_k is <u>deterministic</u>
  - $|F|$ is equal to the number of keys 2^l
  - E.g., n = 2, l = 2, then $|F| = 2^l = 2^2 = 4$
- Permutation family: Perm(n,n) $\mathcal{D}=\{0,1\}^n$ -> $\mathcal{R}=\{0, 1\}^n$
  - $|\text{Perm}(n,n)| = (2^n)!$
  - E.g., n = 2, then $|\text{Perm}(2, 2)| = 4! = 24$

Pseudorandom Permutation

- Adversary A interacts with a black-box (either function F_k or a permutation f)
  - A cannot tell which one it is, if F is a PRP



Pseudorandom Permutation

- PRP F is not even close to Perm(n, n)
  - $|F| = 2^n$ and $|\text{Perm}(n, n)| = (2^n)!$
  - <u>Practice:</u> n=3, $|F|$ = ?? and $|\text{Perm}(n, n)|$ = ??

8 v.s. 40320

- <u>A PRP is a PRF, if n is large</u>: a random permutation is indistinguishable from a random function

Secure Enc. from PRF

- A straightforward solution from PRF (PRP)
    - $\mathsf{KeyGen}(1^n) : k \xleftarrow{u} \{0,1\}^n$
    - $\mathsf{Enc}_k(m) : c \leftarrow F_k(m)$
    - $\mathsf{Dec}_k(c)\ m \leftarrow F_k^{-1}(c)$
- Efficient to compute
- Ciphertext c does not directly reveal plaintext m
- But still <u>deterministic</u>, not CPA-secure

CPA-Secure Enc. from PRF

- Add a random, and send it in ciphertext
  - $\mathsf{KeyGen}(1^n) : k \xleftarrow{u} \{0,1\}^n$
  - $\mathsf{Enc}_k(m) : c \leftarrow \langle r, F_k(r) \oplus m \rangle,\ r \xleftarrow{u} \{0,1\}^n$
  - $\mathsf{Dec}_k(c)$ given $c = \langle r, s \rangle,\ m \leftarrow F_k(r) \oplus s$
- E.g., n = 2, r = 00, F_k(r) = 10, m = 01, c = (00, 11)
- Practice: r = 10, F_k(r) = 11, m = 01 what is c = ??
- c = (r, s) = (10, 10)

- $\mathsf{KeyGen}(1^n) : k \xleftarrow{u} \{0,1\}^n$
- $\mathsf{Enc}_k(m) : c \leftarrow \langle r, F_k(r) \oplus m \rangle,\ r \xleftarrow{u} \{0,1\}^n$
- $\mathsf{Dec}_k(c)$ given $c = \langle r, s \rangle,\ m \leftarrow F_k(r) \oplus s$

x	00	01	10	11
k=00,F(x)	11	00	01	10
k=01,F(x)	10	11	00	01
k=10,F(x)	01	10	11	00
k=11,F(x)	00	01	10	11

- <u>Practice:</u> given m=01, r=10, k=10, what is c=??

- $\mathsf{KeyGen}(1^n) : k \xleftarrow{u} \{0,1\}^n$
- $\mathsf{Enc}_k(m) : c \leftarrow \langle r, F_k(r) \oplus m \rangle,\ r \xleftarrow{u} \{0,1\}^n$
- $\mathsf{Dec}_k(c)$ given $c = \langle r, s \rangle,\ m \leftarrow F_k(r) \oplus s$

x	00	01	10	11
k=00,F(x)	11	00	01	10
k=01,F(x)	10	11	00	01
k=10,F(x)	01	10	11	00
k=11,F(x)	00	01	10	11

- Practice: given m=01, r=10, k=10, what is c=??
- $F_k(r) = F_{10}(10) = 11$, $F_k(r)$ xor m = 10, c = (10, 10)

- KeyGen(1^n) : $k \xleftarrow{u} \{0,1\}^n$
- Enc$_k(m)$: $c \leftarrow \langle r, F_k(r) \oplus m \rangle$, $r \xleftarrow{u} \{0,1\}^n$
- Dec$_k(c)$ given $c = \langle r, s \rangle$, $m \leftarrow F_k(r) \oplus s$

x	00	01	10	11
k=00,F(x)	11	00	01	10
k=01,F(x)	10	11	00	01
k=10,F(x)	01	10	11	00
k=11,F(x)	00	01	10	11

- <u>Practice:</u> given m=01, k=10, what is c=??

x	00	01	10	11
k=00,F(x)	11	00	01	10
k=01,F(x)	10	11	00	01
k=10,F(x)	01	10	11	00
k=11,F(x)	00	01	10	11

- Practice: given m=01, k=10, what is c=??
  - if r = 00, $F_k(r)$ = 01, $F_k(r)$ xor m = 00, c = (00, 00)
  - if r = 01, $F_k(r)$ = 10, $F_k(r)$ xor m = 11, c = (01, 11)
  - if r = 10, $F_k(r)$ = 11, $F_k(r)$ xor m = 10, c = (10, 10)
  - if r = 11, $F_k(r)$ = 00, $F_k(r)$ xor m = 01, c = (11, 01)
  - Probabilistic: same message, different ciphertexts

- KeyGen(1^n) : $k \xleftarrow{u} \{0,1\}^n$
- Enc$_k(m)$: $c \leftarrow \langle r, F_k(r) \oplus m \rangle$, $r \xleftarrow{u} \{0,1\}^n$
- Dec$_k(c)$ given $c = \langle r, s \rangle$, $m \leftarrow F_k(r) \oplus s$

x	00	01	10	11
k=00,F(x)	11	00	01	10
k=01,F(x)	10	11	00	01
k=10,F(x)	01	10	11	00
k=11,F(x)	00	01	10	11

- <u>Practice:</u> given c = (11, 00), k = 01, what is m=??

- $\mathsf{KeyGen}(1^n): k \xleftarrow{u} \{0,1\}^n$
- $\mathsf{Enc}_k(m): c \leftarrow \langle r, F_k(r) \oplus m \rangle,\ r \xleftarrow{u} \{0,1\}^n$
- $\mathsf{Dec}_k(c)$ given $c = \langle r, s \rangle,\ m \leftarrow F_k(r) \oplus s$

x	00	01	10	11
k=00,F(x)	11	00	01	10
k=01,F(x)	10	11	00	01
k=10,F(x)	01	10	11	00
k=11,F(x)	00	01	10	11

- Practice: given c = (11, 00), k = 01, what is m=??
- $F_k(r) = F_{01}(11) = 01$, m = $F_k(r)$ xor s = 01

CPA-Secure Enc. from PRF

- $\mathsf{KeyGen}(1^n) : k \xleftarrow{u} \{0,1\}^n$
- $\mathsf{Enc}_k(m) : c \leftarrow \langle r, F_k(r) \oplus m \rangle,\ r \xleftarrow{u} \{0,1\}^n$
- $\mathsf{Dec}_k(c)$ given $c = \langle r, s \rangle,\ m \leftarrow F_k(r) \oplus s$

- Probabilistic encryption, CPA-secure
- <u>Message length is fixed</u>
  - n=2, can encrypt 2 bits, what if m = 1 or m =101
- We need a scheme for arbitrary message size

Additional Reading

Chapter 3, *Introduction to Modern Cryptography, Drs. J. Katz and Y. Lindell, 2nd edition*