Private-Key Encryption and Pseudorandomness (Part II)

Yu Zhang

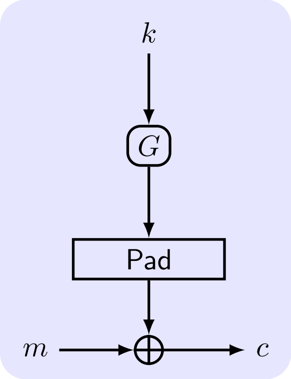
Harbin Institute of Technology

Cryptography, Autumn, 2021

Outline

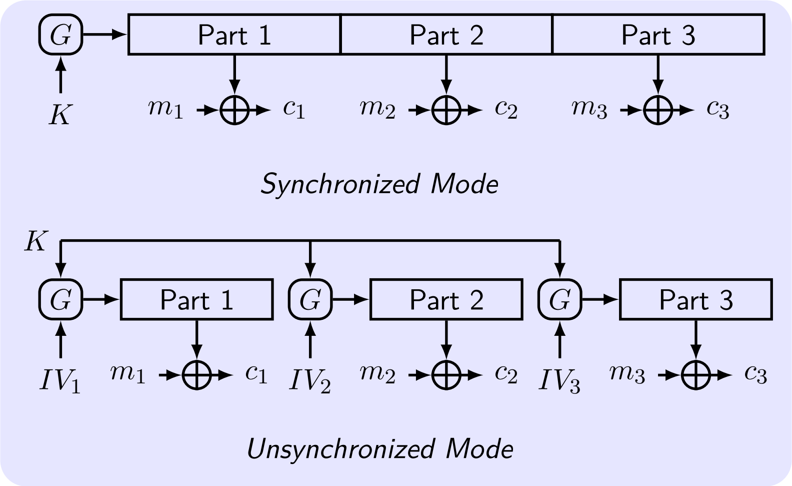
Content

Stream Ciphers



- **Idea**: Generalization of one-time pad
- **Stream cipher**: Enc. by XORing with pseudorandom stream (keystream)
- **Multiple messages**: Be concatenated into a single one and encrypted
- **Keystream**: Generated by a variable-length PRG
- **Strength**: Faster than block cipher
- **Weakness**: Difficult to be secure

Secure Multiple Encryptions Using a Stream Cipher

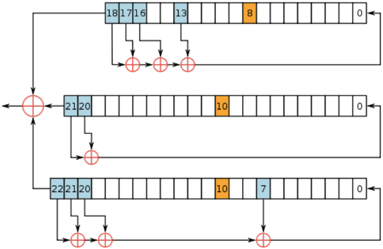


Initial vector IV is chosen *u.a.r* and public

Q: which mode is better in your opinion?

Questionable Security

- **State of the art**: No standardized and popular one. Security is questionable, e.g., RC4 in WEP protocol in 802.11, Linear Feedback Shift Registers (LFSRs) used in A5/1 for GSM.



WARNING

Don't use any stream cipher. If necessary, construct one from a block cipher.

- eStream project worked on secure stream ciphers. Salsa20/12 is a promising candidate.

Related Keys: Real World Cases

Keys (the IV-key pair) for multiple enc. must be independent

Attacks on 802.11b WEP

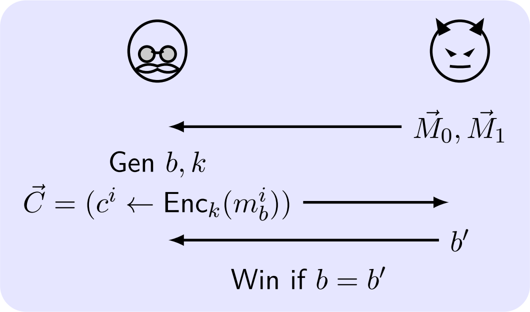
Unsynchronized mode: $\text{Enc}(m_i) := \langle IV_i, G(IV_i \| k) \oplus m_i \rangle$

- Length of IV is 24 bits, repeat IV after $2^{24} \approx$ 16M frames
- On some WiFi cards, IV resets to 0 after power cycle
- $IV_i = IV_{i-1} + 1$. For RC4, recover k after 40,000 frames

Security for Multiple Encryptions

The multiple-message eavesdropping experiment $\mathsf{PrivK}^{\mathsf{mult}}_{\mathcal{A},\Pi}(n)$:

1. $\mathcal{A}$ is given input 1^n, outputs $\vec{M}_0 = (m_0^1, \ldots, m_0^t)$, $\vec{M}_1 = (m_1^1, \ldots, m_1^t)$ with $\forall i, |m_0^i| = |m_1^i|$.
2. $k \leftarrow \mathsf{Gen}(1^n)$, a random bit $b \leftarrow \{0, 1\}$ is chosen. Then $c^i \leftarrow \mathsf{Enc}_k(m_b^i)$ and $\vec{C} = (c^1, \ldots, c^t)$ is given to $\mathcal{A}$.
3. $\mathcal{A}$ outputs b'. If $b' = b$, $\mathsf{PrivK}^{\mathsf{mult}}_{\mathcal{A},\Pi} = 1$, otherwise 0.



Definition of Multi-Encryption Security

Definition 1

Π has **indistinguishable multiple encryptions in the presence of an eavesdropper** if $\forall$ PPT $\mathcal{A}$, $\exists$ negl such that

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

Question:

Does any cipher we have learned so far have indistinguishable multiple encryptions in the presence of an eavesdropper?

Attack On Deterministic Multiple Encryptions

Question:

Generally, if Π's encryption function is **deterministic**, i.e., a plaintext will be always encrypted into the same ciphertext with the same key, is Π multiple-encryption-secure?

Attack:

For the deterministic encryption, the adversary may generate $m_0^1 = m_0^2$ and $m_1^1 \neq m_1^2$, and then outputs $b' = 0$ if $c^1 = c^2$, otherwise $b' = 1$.

Chosen-Plaintext Attacks (CPA)

CPA: the adversary has the ability to obtain the encryption of plaintexts of its choice

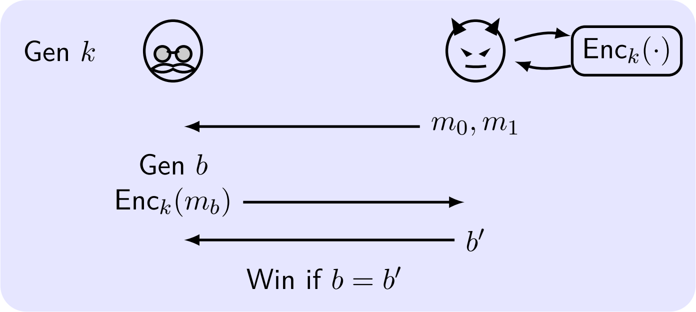
A story in WWII

- Navy cryptanalysts believe the ciphertext “AF” means “Midway island” in Japanese messages
- But the general did not believe that Midway island would be attacked
- Navy cryptanalysts sent a plaintext that the freshwater supplies at Midway island were low
- Japanese intercepted the plaintext and sent a ciphertext that “AF” was low in water
- The US forces dispatched three aircraft carriers and won

CPA Indistinguishability Experiment

The CPA indistinguishability experiment $\mathsf{PrivK}^{\mathsf{cpa}}_{\mathcal{A},\Pi}(n)$:

1. $k \leftarrow \mathsf{Gen}(1^n)$
2. $\mathcal{A}$ is given input 1^n and **oracle access** $\mathcal{A}^{\mathsf{Enc}_k(\cdot)}$ to $\mathsf{Enc}_k(\cdot)$, outputs m_0, m_1 of the same length
3. $b \leftarrow \{0,1\}$. Then $c \leftarrow \mathsf{Enc}_k(m_b)$ is given to $\mathcal{A}$
4. $\mathcal{A}$ **continues to have oracle access** to $\mathsf{Enc}_k(\cdot)$, outputs b'
5. If $b' = b$, $\mathcal{A}$ succeeded $\mathsf{PrivK}^{\mathsf{cpa}}_{\mathcal{A},\Pi} = 1$, otherwise 0



Definition of CPA Security

Definition 2

Π has **indistinguishable encryptions under a CPA (CPA-secure)** if $\forall$ PPT $\mathcal{A}$, $\exists$ negl such that

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

- Q: Is any cipher we have learned so far CPA-secure? Why?

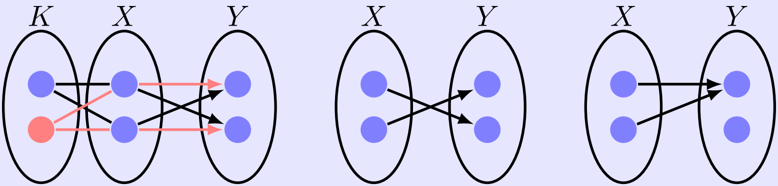
Proposition 3

Any private-key encryption scheme that is CPA-secure also is ***multiple-encryption-secure****.*

- Q: Does **multiple-encryption-security** mean CPA-security? (homework)

Content

Concepts on Pseudorandom Functions



$F : K \times X \to Y$

2D function

$F_k(x) \overset{\text{def}}{=} F(k, x)$

Keyed function

$f : X \to Y$

Look-up table

- **Keyed function** $F : \{0,1\}^* \times \{0,1\}^* \to \{0,1\}^*$
  $F_k : \{0,1\}^* \to \{0,1\}^*$, $F_k(x) \overset{\text{def}}{=} F(k, x)$
- **Look-up table** f: $\{0,1\}^n \to \{0,1\}^n$ with size $= ?$ bits
- **Function family** Func_n: all functions $\{0,1\}^n \to \{0,1\}^n$.
  $|\mathsf{Func}_n| = 2^{n \cdot 2^n}$
- **Length Preserving**: $\ell_{key}(n) = \ell_{in}(n) = \ell_{out}(n)$

Definition of Pseudorandom Function

Intuition: A PRF F generates a function F_k that is indistinguishable from truly random selected function f (look-up table) in Func_n.
However, the function has **exponential length**. Give D the deterministic **oracle access** $D^{\mathcal{O}}$ to the functions $\mathcal{O}$.

Definition 4

An efficient length-preserving, keyed function F is a **pseudorandom function (PRF)** if $\forall$ PPT distinguishers D,

$$\left|\Pr[D^{F_k(\cdot)}(1^n) = 1] - \Pr[D^{f(\cdot)}(1^n) = 1]\right| \leq \mathsf{negl}(n),$$

where f is chosen *u.a.r* from Func_n.

Q: Is the fixed-length OTP a PRF?

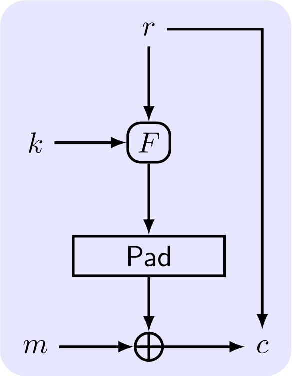
Questions

Q: Without knowing the key and the oracle access, could anyone learn something about the output from the input with a non-negligible probability?

Let $F : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ be a PRF. Is G a PRF?

- $G((k_1, k_2), x) = F(k_1, x) \| F(k_2, x)$
- $G(k, x) = F(k, x \oplus 1^n)$
- $G(k, x) = \begin{cases} F(k, x) & \text{when } x \neq 0^n \\ 0^n & \text{otherwise} \end{cases}$
- $G(k, x) = \begin{cases} F(k, x) & \text{when } x \neq 0^n \\ k & \text{otherwise} \end{cases}$
- $G(k, x) = F(k, x) \bigoplus F(k, x \oplus 1^n)$

CPA-Security from Pseudorandom Function



Construction 5

- *Fresh random string* r*.*
- $F_k(r)$*:* $|k| = |m| = |r| = n$*.*
- Gen*:* $k \in \{0,1\}^n$*.*
- Enc*:* $s := F_k(r) \oplus m$*,*
  $c := \langle r, s \rangle$*.*
- Dec*:* $m := F_k(r) \oplus s$*.*

Theorem 6

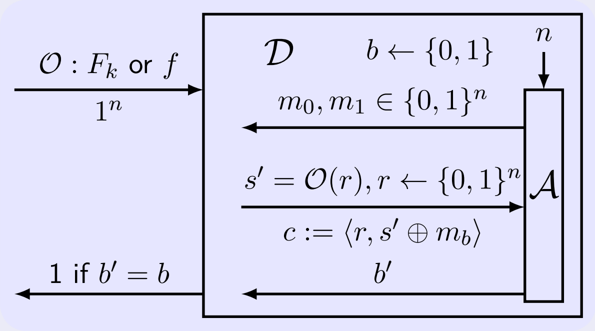
If F *is a PRF, this fixed-length encryption scheme* Π *is CPA-secure.*

Proof of CPA-Security from PRF

Idea: First, analyze the security in an idealized world where f is used in $\tilde{\Pi}$; next, claim that if Π is insecure when F_k was used then this would imply F_k is not PRF by reduction.

Proof.

Reduce D to $\mathcal{A}$:



□

Proof of CPA-Security from PRF (Cont.)

Proof.

Analyze $\Pr[\mathsf{Break}]$, Break means $\mathsf{PrivK}^{\mathsf{cpa}}_{\mathcal{A},\tilde{\Pi}}(n) = 1$:
$\mathcal{A}$ collects $\{\langle r_i, f(r_i)\rangle\}$, $i = 1, \ldots, q(n)$ with $q(n)$ queries;
The challenge $c = \langle r_c, f(r_c) \oplus m_b\rangle$.

- Repeat: $r_c \in \{r_i\}$ with probability $\frac{q(n)}{2^n}$. $\mathcal{A}$ can know m_b.
- $\overline{\text{Repeat}}$: As OTP, $\Pr[\mathsf{Break}] = \frac{1}{2}$

$$\begin{aligned}\Pr[\mathsf{Break}] &= \Pr[\mathsf{Break} \wedge \mathsf{Repeat}] + \Pr[\mathsf{Break} \wedge \overline{\mathsf{Repeat}}] \\ &\leq \Pr[\mathsf{Repeat}] + \Pr[\mathsf{Break} | \overline{\mathsf{Repeat}}] \\ &\leq \frac{q(n)}{2^n} + \frac{1}{2}.\end{aligned}$$

$\Pr[D^{F_k(\cdot)}(1^n) = 1] = \Pr[\mathsf{PrivK}^{\mathsf{cpa}}_{\mathcal{A},\Pi}(n) = 1] = \frac{1}{2} + \varepsilon(n)$.
$\Pr[D^{f(\cdot)}(1^n) = 1] = \Pr[\mathsf{PrivK}^{\mathsf{cpa}}_{\mathcal{A},\tilde{\Pi}}(n) = 1] = \Pr[\mathsf{Break}] \leq \frac{1}{2} + \frac{q(n)}{2^n}$.
$\Pr[D^{F_k(\cdot)}(1^n) = 1] - \Pr[D^{f(\cdot)}(1^n) = 1] \geq \varepsilon(n) - \frac{q(n)}{2^n}$. $\varepsilon(n)$ is negligible. $\square$

CPA-Security from PRF for Arbitrary-Length

- For arbitrary-length messages, $m = m_1, \ldots, m_\ell$

$$c := \langle r_1, F_k(r_1) \oplus m_1, r_2, F_k(r_2) \oplus m_2, \ldots, r_\ell, F_k(r_\ell) \oplus m_\ell \rangle$$

Corollary 7

If F is a PRF, then Π is CPA-secure for arbitrary-length messages.

- **Efficiency**: $|c| = 2|m|$.

Content

Pseudorandom Permutations

- **Bijection**: F is one-to-one and onto
- **Permutation**: A bijective function from a set to itself
- **Keyed permutation**: $\forall k, F_k(\cdot)$ is permutation
- F is a bijection $\iff$ F^{-1} is a bijection

Definition 8

An efficient, keyed permutation F is a **strong pseudorandom permutation (PRP)** if $\forall$ PPT distinguishers D,

$$\left|\Pr[D^{F_k(\cdot),F_k^{-1}(\cdot)}(1^n) = 1] - \Pr[D^{f(\cdot),f^{-1}(\cdot)}(1^n) = 1]\right| \leq \mathsf{negl}(n),$$

where f is chosen *u.a.r* from the set of permutations on n-bit strings.

If F is a PRP then is it a PRF?

Questions

Let $X = \{0, 1\}$ (1 bit), answer the following questions.

1. What are the functions in the permutation over X?
2. $K = \{0, 1\}$, what is the simplest permutation $F(k, x)$ over X?
3. Is your F a secure PRP?
4. Is your F a secure PRF?
5. What if $X = \{0, 1\}^{128}$ and $K = \{0, 1\}^{128}$?
6. Could you give a (or another) PRP over $X = \{0, 1\}^{128}$?

Proposition 9

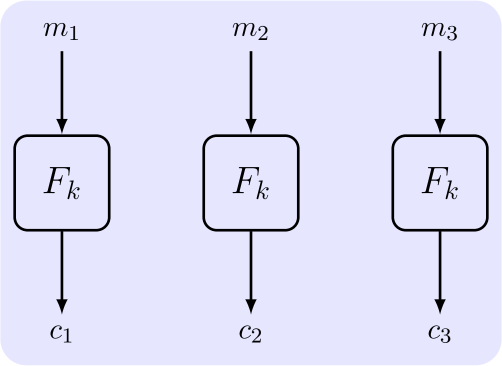
Switching Lemma IF F is a PRP and additionally $\ell_{in}(n) \geq n$, then F is also a PRF.

A random lookup table and a random permutation are indistinguishable. So PRP is also PRF.

Modes of Operation:

- A way of encrypting arbitrary-length messages using a PRP or PRF
- A way of constructing a PRG from a PRP or PRF

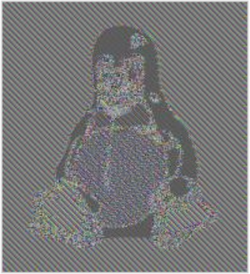
Electronic Code Book (ECB) Mode



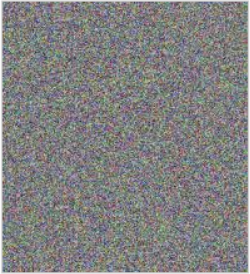
- Q: is it indistinguishable in the presence of an eavesdropper?
- Q: can F be any PRF?

Attack on ECB mode

Original image

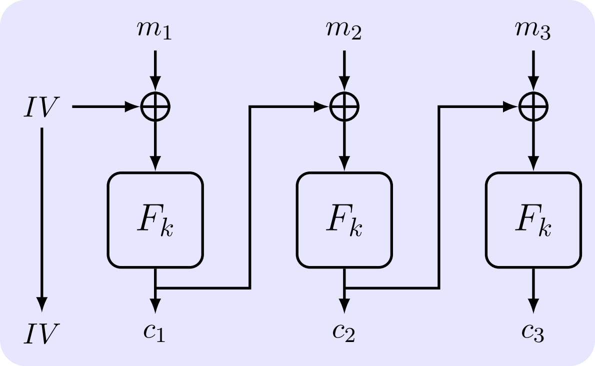


Encrypted using ECB mode



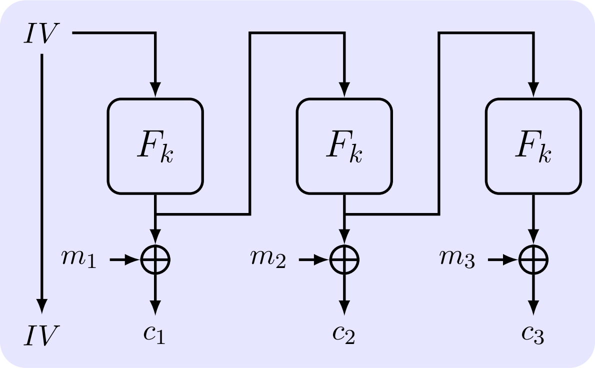
Modes other than ECB result in pseudo-randomness

Cipher Block Chaining (CBC) Mode



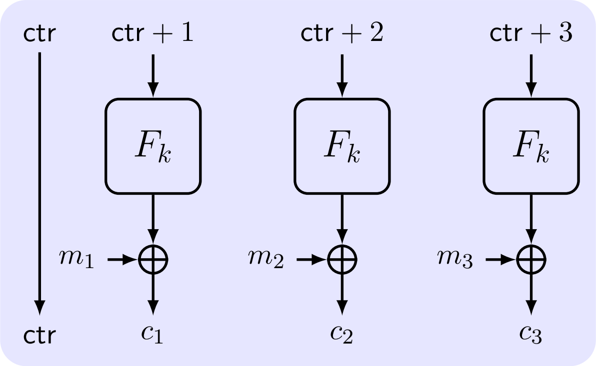
- IV: initial vector, a fresh random string.
- Q: is it CPA-secure? what if IV is always 0?
- Q: is the encryption parallelizable, i.e., outputting c_2 before getting c_1?
- Q: can F be any PRF?

Output Feedback (OFB) Mode



- Q: is it CPA-secure?
- Q: is the encryption parallelizable?
- Q: can F be any PRF?

Counter (CTR) Mode



- ctr is an IV
- Q: is it CPA-secure?
- Q: is the encryption parallelizable?
- Q: can F be any PRF?

CTR Mode Is CPA-secure

Theorem 10

If F is a PRF, then randomized CTR mode is CPA-secure.

Proof.

The message length and the number of query are $q(n)$.
Overlap: the sequence for the challenge overlaps the sequences for the queries from the adversary.
ctr*: ctr in the challenge. ctr$_i$: ctr in the queries, $i = 1, \ldots, q(n)$.
Overlap: $\mathsf{ctr}_i - q(n) < \mathsf{ctr}^* < \mathsf{ctr}_i + q(n)$.

$$\Pr[\text{Overlap}] \leq \frac{2q(n) - 1}{2^n} \cdot q(n)$$

□

Proof of CPA-secure CTR Mode (Cont.)

Proof.

See proof of theorem 6. (1) Analyze Break : $\mathsf{PrivK}^{\mathsf{cpa}}_{\mathcal{A},\tilde{\Pi}}(n) = 1$.

$$\begin{aligned}
\Pr[\mathsf{Break}] &= \Pr[\mathsf{Break} \wedge \mathsf{Overlap}] + \Pr[\mathsf{Break} \wedge \overline{\mathsf{Overlap}}] \\
&\le \Pr[\mathsf{Overlap}] + \Pr[\mathsf{Break} | \overline{\mathsf{Overlap}}] \\
&\le \frac{2q(n)^2}{2^n} + \frac{1}{2}.
\end{aligned}$$

(2) Reduce D to $\mathcal{A}$

$$\Pr[D^{f(\cdot)}(1^n) = 1] = \Pr[\mathsf{PrivK}^{\mathsf{cpa}}_{\mathcal{A},\tilde{\Pi}}(n) = 1] \le \frac{2q(n)^2}{2^n} + \frac{1}{2}$$

$$\Pr[D^{F_k(\cdot)}(1^n) = 1] = \Pr[\mathsf{PrivK}^{\mathsf{cpa}}_{\mathcal{A},\Pi}(n) = 1] \le \frac{1}{2} + \varepsilon(n)$$

If F is PRP, $\varepsilon(n)$ is negligible.

□

IV Should Not Be Predictable

If IV is predictable, then CBC/OFB/CTR mode is not CPA-secure.
Q: Why? (homework)

Bug in SSL/TLS 1.0

IV for record $\#i$ is last CT block of record $\#(i-1)$.

API in OpenSSL

```
void AES_cbc_encrypt (
    const unsigned char *in,
    unsigned char       *out,
    size_t              length,
    const AES_KEY       *key,
    unsigned char       *ivec,    User supplies IV
    AES_ENCRYPT or AES_DECRYPT);
```

Non-deterministic Encryption

Three general methods of non-deterministic encryption for CPA security.
Enc: $s := F_k(r) \oplus m$, $c := \langle r, s \rangle$.

- **Randomized**: r is chosen $u.a.r$, as Construction 5
  - more entropy needed, and long ciphertext
- **Stateful**: r is a counter, like CTR mode
  - synchronization on the counter between two parties
- **Nonce-based**: r is a nonce (number used only once)
  - make sure that nonces are distinct , and long ciphertext

Content

Security Against CCA

The CCA indistinguishability experiment $\mathsf{PrivK}^{\mathsf{cca}}_{\mathcal{A},\Pi}(n)$:

1. $k \leftarrow \mathsf{Gen}(1^n)$.
2. $\mathcal{A}$ is given input 1^n and oracle access $\mathcal{A}^{\mathsf{Enc}_k(\cdot)}$ and $\mathcal{A}^{\mathsf{Dec}_k(\cdot)}$, outputs m_0, m_1 of the same length.
3. $b \leftarrow \{0,1\}$. $c \leftarrow \mathsf{Enc}_k(m_b)$ is given to $\mathcal{A}$.
4. $\mathcal{A}$ continues to have oracle access **except for c**, outputs b'.
5. If $b' = b$, $\mathcal{A}$ succeeded $\mathsf{PrivK}^{\mathsf{cca}}_{\mathcal{A},\Pi} = 1$, otherwise 0.

Definition 11

Π has **indistinguishable encryptions under a CCA (CCA-secure)** if $\forall$ PPT $\mathcal{A}$, $\exists$ negl such that

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

Understanding CCA-security

- In real world, the adversary might conduct CCA by influencing what gets decrypted
  - If the communication is not authenticated, then an adversary may send certain ciphertexts on behalf of the honest party
- CCA-security implies "**non-malleability**"
- None of the above scheme is CCA-secure

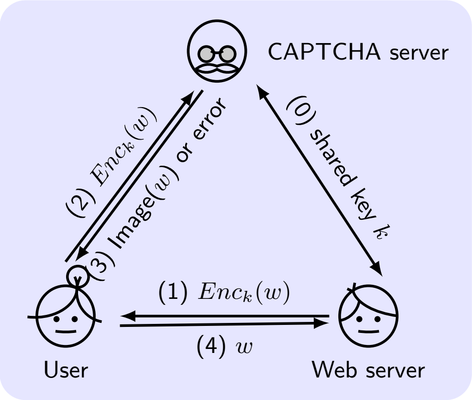
CCA against Construction 5

$\mathcal{A}$ gives m_0, m_1 and gets $c = \langle r, F_k(r) \oplus m_b \rangle$, and then queries c' which is the same with c except that a single bit is flipped. The $m' = c' \oplus F_k(r)$ should be the same with m_b except ____?

Q: Show that the above modes (CBC, OFB and CTR) are also not CCA-secure. (homework)

Padding-Oracle Attacks: Real-world Case

Padding-oracle attacks are originally published in 2002. It can be used to automatically obtain the CAPTCHA text, as CAPTCHA server will return an error (as decryption oracle) when deciphering the CT of a CAPTCHA text received from a user.

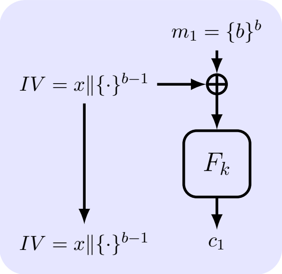


Padding-Oracle Attacks

PKCS #5 Padding: append b bytes of b to the message in order to make the total length a multiple of the block length (append a dummy block if needed). The decryption server will return a **Bad Padding Error** for incorrect padding.

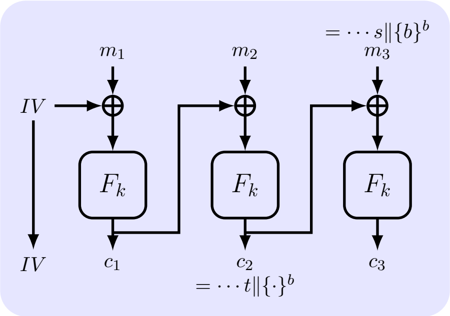
Padding-Oracle Attacks:

- In a one-block CBC, by modifying the 1st byte of IV, attacker can learn whether m is NULL. If yes, error will occur.



- append $\{b\}^b$ as a dummy block if m is NULL
- change the 1st byte of IV from x to y, get decrypted block $(x \oplus y \oplus b)\|\{b\}^{b-1}$, and trigger an error
- If no error, learn whether m is 1 byte by modifying the 2nd byte of IV and so on

Padding-Oracle Attacks (Cont.)



- Once learn the length of m, learn the last byte of m (s) by modifying the one before the last block in the ciphertext
- $m_{last} = \cdots s\|\{b\}^b$, $c_{last-1} = \cdots t\|\{\cdot\}^b$
- modify c_{last-1} to $c'_{last-1} = \cdots u\|(\{\cdot\}^b \oplus \{b\}^b \oplus \{b+1\}^b)$
- Q: If no padding error, then $s =$?

Summary

- Definitions: CPA, CCA (padding-oracle attack)
- Primitives: PRG, PRF, PRP
- Constructions: stream cipher, block cipher, EBC, CBC, OFB, CTR