Lecture 3. Symmetric encryption

Introduction to cryptology

Bruno Grenet

M1 INFO, MOSIG & AM

Université Grenoble Alpes – IM²AG

https://membres-ljk.imag.fr/Bruno.Grenet/IntroCrypto.html

Block ciphers are not enough

Block ciphers offer

- One-to-one (deterministic) encryption
- Fixed-size messages

We need

- One-to-many (non-deterministic) encryption
- Variable-size messages

Block ciphers are not enough

Block ciphers offer

- One-to-one (deterministic) encryption
- Fixed-size messages

We need

- One-to-many (non-deterministic) encryption
- Variable-size messages

Symmetric encryption scheme

$$\begin{cases} \mathrm{Enc} : \{0,1\}^{\kappa} \times \{0,1\}^* \to \{0,1\}^* \\ \mathrm{Dec} : \{0,1\}^{\kappa} \times \{0,1\}^* \to \{0,1\}^* \end{cases}$$

- Enc is a *randomized* encryption ~~scheme~~ algorithm
- Dec is a (deterministic) decryption ~~scheme~~ alg.
- *Correctness:* for all $k \in \{0,1\}^{\kappa}$, $m \in \{0,1\}^*$ and $c \leftarrow \mathrm{Enc}_k(m)$, $\mathrm{Dec}_k(c) = m$
- *Efficiency:* for all $k \in \{0,1\}^{\kappa}$, $m \in \{0,1\}^*$ and $c \leftarrow \mathrm{Enc}_k(m)$, $|c| \simeq |m|$

Block ciphers are not enough

Block ciphers offer

- One-to-one (deterministic) encryption
- Fixed-size messages

We need

- One-to-many (non-deterministic) encryption
- Variable-size messages

Symmetric encryption scheme

$$\begin{cases} \mathsf{Enc} : \{0,1\}^\kappa \times \{0,1\}^* \to \{0,1\}^* \\ \mathsf{Dec} : \{0,1\}^\kappa \times \{0,1\}^* \to \{0,1\}^* \end{cases}$$

- Enc is a *randomized* encryption scheme
- Dec is a (deterministic) decryption scheme
- *Correctness:* for all $k \in \{0,1\}^\kappa$, $m \in \{0,1\}^*$ and $c \leftarrow \mathsf{Enc}_k(m)$, $\mathsf{Dec}_k(c) = m$
- *Efficiency:* for all $k \in \{0,1\}^\kappa$, $m \in \{0,1\}^*$ and $c \leftarrow \mathsf{Enc}_k(m)$, $|c| \simeq |m|$

- How to build symmetric encryption schemes?
- What are *good* encryption schemes?

From block ciphers to symmetric encryption schemes

The tool: modes of operations

- Transforms a block cipher into a *symmetric encryption scheme*

$$E : \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^n \rightsquigarrow \begin{cases} \mathrm{Enc} : \{0,1\}^\kappa \times \{0,1\}^* \to \{0,1\}^* \\ \mathrm{Dec} : \{0,1\}^\kappa \times \{0,1\}^* \to \{0,1\}^* \end{cases}$$

- A mode is *good* if it turns *good BCs* into *good encryption schemes*

Another approach: from stream ciphers

- Basic (incomplete) idea:
  - Use one-time pad with *pseudo-random* bits
  - Produce the pseudo-random bits on the fly
- In terms of security:
  - block cipher $\leftrightsquigarrow$ pseudo-random permutation
  - stream cipher $\leftrightsquigarrow$ pseudo-random generator

1. Security notions for symmetric encryption schemes

2. From block ciphers to symmetric encryption schemes: modes of operation

Experiment for chosen-plaintext attack

IND-CPA experiment $\mathrm{Exp}^{\mathrm{IND-CPA}}_{\mathrm{Enc}}(A)$ for $\mathrm{Enc} : \mathcal{K} \times \mathcal{M} \to \mathcal{M}$

Challenger draws $k \twoheadleftarrow \mathcal{K}$

Adversary has *oracle* access to $\mathrm{Enc}_k(\cdot)$: on query x_i, gets $c_i \leftarrow \mathrm{Enc}_k(x_i)$

Adversary creates two equal-length messages m_0 and m_1 and submits them

Challenger draws $b \twoheadleftarrow \{0, 1\}$ and answers with $c \leftarrow \mathrm{Enc}_k(m_b)$

Adversary tries to guess b and outputs $\hat{b}$

Remarks

- Oracle access during the whole experiment
- Equal-length messages $\rightsquigarrow$ message length not hidden!
    - Impossible to hide if messages of any length
    - Use padding beforehand if message length is sensitive

Chosen-plaintext attack advantage

IND-CPA advantage of an adversary A

$$\mathsf{Adv}^{\mathsf{IND-CPA}}_{\mathsf{Enc}}(A) = \left|\Pr\left[\mathsf{Exp}^{\mathsf{IND-CPA}}_{\mathsf{Enc}}(A) = 1 \middle| b = 1\right] - \Pr\left[\mathsf{Exp}^{\mathsf{IND-CPA}}_{\mathsf{Enc}}(A) = 1 \middle| b = 0\right]\right|$$

- Equal to $|\Pr[\hat{b} = 1 | b = 1] - \Pr[\hat{b} = 1 | b = 0]|$ and to $|2\Pr[\hat{b} = b] - 1|$
- Extremal cases:
  - Guessing $\hat{b}$ at random $\rightsquigarrow$ advantage 0 $\longrightarrow$ $\mathrm{Adv}^{\mathrm{IND-CPA}}_{\mathrm{Enc}}(0, 1) \gtrsim 0$
  - Resource-unbounded A $\rightsquigarrow$ advantage 1 $\longrightarrow$ $\mathrm{Adv}^{\mathrm{IND-CPA}}_{\mathrm{Enc}}(\text{"}{+\infty}\text{"}, \text{"}{+\infty}\text{"}) = 1$

$\mathrm{Adv}^{\mathrm{IND-CPA}}_{\mathrm{Enc}}(0, \text{"}{+4}\text{"}) = 1$

IND-CPA advantage

$$\mathsf{Adv}^{\mathsf{IND-CPA}}_{\mathsf{Enc}}(q, t) = \max_{A_{q,t}} \mathsf{Adv}^{\mathsf{IND-CPA}}_{\mathsf{Enc}}(A_{q,t})$$

where $A_{q,t}$ is an alg. that runs in time $\leq t$ and makes $\leq q$ queries to the challenger

Comments on IND-CPA security

- No formal definition of IND-CPA secure, only a measure *(but in asymptotic security)*
- IND-CPA $\implies$ non-determinism (A can query $\mathrm{Enc}_k(m_0)$ and $\mathrm{Enc}_k(m_1)$)
- IND-CPA $\implies$ the adversary cannot compute any single bit of the message
- IND-CPA $\implies$ the adversary can compute *very few* information on the message

Stronger security notions

- Indistinguishability under chosen ciphertext attack
  - Access to both an encryption oracle and a decryption oracle
  - 2 variants: non-adaptative (IND-CCA) or adaptative (IND-CCA2)
- Indistinguishability for multiple encryptions either CPA or CCA
  - Challenger draws $b \twoheadleftarrow \{0,1\}$
  - Adversary submits *pairs* of challenges (m_i^0, m_i^1) and gets $c_i \leftarrow \mathrm{Enc}_k(m_i^b)$
  - Adversary must find b

1. Security notions for symmetric encryption schemes

2. From block ciphers to symmetric encryption schemes: modes of operation

Goal

$$E : \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^n \rightsquigarrow \begin{cases} \mathrm{Enc} : \{0,1\}^\kappa \times \{0,1\}^* \to \{0,1\}^* \\ \mathrm{Dec} : \{0,1\}^\kappa \times \{0,1\}^* \to \{0,1\}^* \end{cases}$$

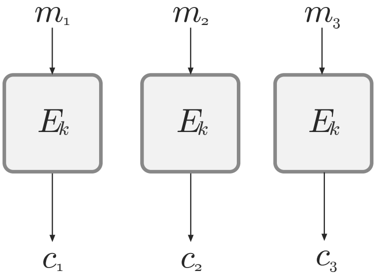
- E is made to encrypt one block of data
- Enc should encrypt any number of blocks

  $\to$ Use E several times to encrypt a message $m \in \{0,1\}^*$

Desired properties

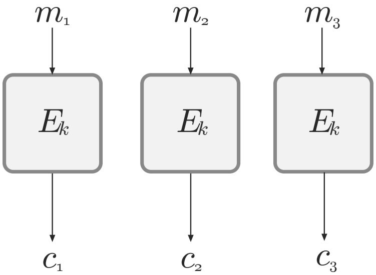
- Security:
  - E *good* $\implies$ Enc *good*
  - Low (S)PRP-advantage $\implies$ low IND-CPA advantage
- Efficiency:
  - Efficient encryption and decryption if E is efficient
  - Ciphertext not too large compared to message

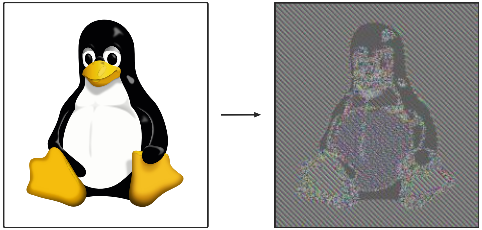
Obvious (bad) idea: Electronic Code Book (ECB)



Source : J. Katz, Y. Lindell. Introduction to modern cryptography. 3rd ed, CRC Press, 2021. (modif.)

Obvious (bad) idea: Electronic Code Book (ECB)

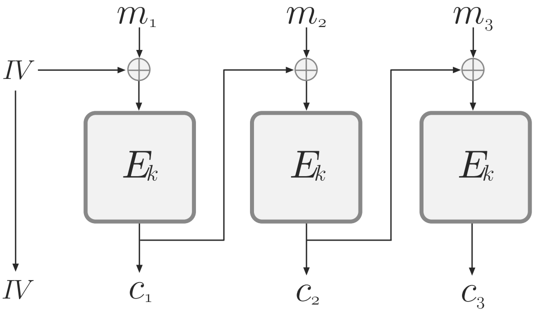




Source : J. Katz, Y. Lindell. Introduction to modern cryptography. 3rd ed, CRC Press, 2021. (modif.)

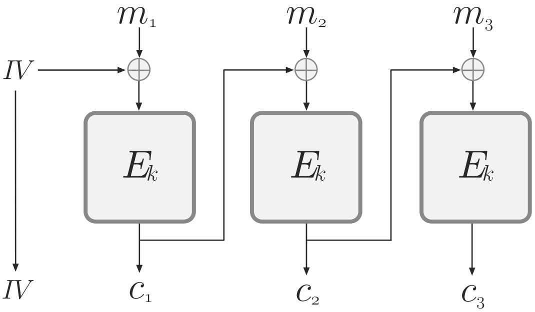
Source : Wikipédia (modif.)

First (real) example of mode of operation: Cipher Block Chaining (CBC)



Source : J. Katz, Y. Lindell. Introduction to modern cryptography. 3rd ed, CRC Press, 2021. (modif.)

First (real) example of mode of operation: Cipher Block Chaining (CBC)

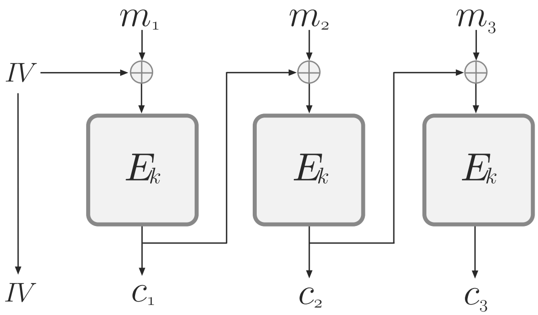


- IV: *random* initialization vector in $\{0,1\}^n$
- Input: $m = m_1 \| \cdots \| m_\ell$ — padding if needed
- Output: $c = IV \| c_1 \| \cdots \| c_\ell$ — size $n(\ell + 1)$

- IND-CPA security if E is a good PRP and IV truly random

Source : J. Katz, Y. Lindell. Introduction to modern cryptography. 3rd ed, CRC Press, 2021. (modif.)

First (real) example of mode of operation: Cipher Block Chaining (CBC)



Adversary when IV is not uniform

1. One-block query m: $r \| c \leftarrow r \| E_k(m \oplus r)$ — *r is IV*
2. *Guesses* the next IV: r'
3. Challenges $m_0 = m \oplus r \oplus r'$ and m_1 uniform: $r'' \| c_b \leftarrow r'' \| E_k(m_b \oplus r'')$
4. If $r' = r''$, return $b = 0$ if $c = c_b$, $b = 1$ otherwise
5. If $r' \neq r''$, *failure*

Source : J. Katz, Y. Lindell. Introduction to modern cryptography. 3rd ed, CRC Press, 2021. (modif.)

Generic CBC collision attack

Observation

- For fixed k, E_k is a permutation $\rightarrow E_k(x) = E_k(y) \iff x = y$
- In CBC, inputs to E_k are of the form $m_i \oplus c_{i-1}$ $\qquad c_0 = IV$

$$E_k(m_i \oplus c_{i-1}) = E_k(m'_j \oplus c'_{j-1}) \iff m_i \oplus c_{i-1} = m'_j \oplus c'_{j-1}$$

Consequence

- Assume we get two identical ciphertext blocks:

$$\begin{aligned} c_i = c'_j &\iff E_k(m_i \oplus c_{i-1}) = E_k(m'_j \oplus c'_{j-1}) \\ &\iff m_i \oplus c_{i-1} = m'_j \oplus c'_{j-1} \\ &\iff c_{i-1} \oplus c'_{j-1} = m_i \oplus m'_j \end{aligned}$$

- $c_{i-1} \oplus c'_{j-1}$ reveals information about m_i and m'_j
  $\Rightarrow$ breaks IND-CPA security *no matter how good E!*

Probability to get collisions?

Assumption

The distribution of the $(m_i \oplus c_{i-1})$ is approx. uniform

- If c_0 is the IV, it has to be approx. uniform
- If c_{i-1} is a ciphertext, non (approx.) uniformity would imply an attack

Birthday bound

Draw $y_1, \ldots, y_q$ uniformly from a size-N set, with $q \leq \sqrt{2N}$. Then

$$\frac{q(q-1)}{4N} \leq 1 - e^{-q(q-1)/2N} \leq \boxed{\Pr[\exists i \neq j, y_i = y_j]} \leq \frac{q(q-1)}{2N}$$

Consequence

- Collision found w.h.p. if $q \simeq \sqrt{N}$
- For CBC: Collision w.h.p. after observing $\simeq 2^{n/2}$ ciphertext blocks
- Note: does not depend on key size κ

Proof of the birthday upper bound

If $y_1, \ldots, y_q \twoheadleftarrow S$ with $|S| = N$, then $\Pr[\exists i \neq j, y_i = y_j] \leq \frac{q(q-1)}{2N} = \frac{\binom{q}{2}}{N}$

$$\Pr[\exists i \neq j \;\; y_i = y_j] = \Pr\Big[\bigvee_{i \neq j} y_i = y_j\Big] \underset{\substack{\uparrow \\ \text{union} \\ \text{bound}}}{\leq} \sum_{i \neq j} \Pr[y_i = y_j]$$

$\Pr[y_i = y_j] = \frac{1}{N}$ since once y_i is chosen, the prob. that y_j is chosen with the same value is $\frac{1}{\#\text{poss. values}} = \frac{1}{N}$.

$$\Rightarrow \Pr[\exists i \neq j, y_i = y_j] \leq \binom{q}{2} \times \frac{1}{N}$$

Proof of the birthday lower bound

If $y_1, \ldots, y_q \twoheadleftarrow S$ with $|S| = N$, then $\Pr[\exists i \neq j, y_i = y_j] \geq 1 - e^{-\frac{q(q-1)}{2N}}$

$(\boxed{\geq} \frac{q(q-1)}{4N}$ if $q \leq \sqrt{2N})$

$(*) = \Pr[\forall i \neq j, y_i \neq y_j] = \Pr[y_2 \neq y_1 \wedge y_3 \notin \{y_1, y_2\} \wedge \cdots \wedge y_q \notin \{y_1, \ldots y_{q-1}\}]$

E_i : "$y_i \notin \{y_1, \ldots, y_{i-1}\}$ given that no collision occured between $y_1, \ldots, y_{i-1}$"

$(*) = \Pr[E_2 \wedge E_3 \wedge \cdots \wedge E_q]$

$(1 + x \leq e^x)$

$\Pr[E_i] = \frac{N-i+1}{N}$

$(*) = \prod_i \Pr[E_i] = \prod_i \frac{N-i+1}{N} = \prod_i \left(1 - \frac{i-1}{N}\right) \leq \prod_i e^{-\frac{i-1}{N}} = e^{-\frac{q(q-1)}{2N}}$ $\square$

The birthday attack against CBC

Adversary A_{BIRTHDAY}

- Sends two messages with $\simeq 2^{n/2}$ blocks each
  - m_0 with only zeroes
  - m_1 with pairwise distinct blocks
- Gets back $c = \text{Enc}_k(m_b)$
  - If there are two blocks $c_i = c_j$, return 0 if $c_{i-1} \oplus c_{j-1} = 0 \cdots 0$, 1 otherwise
  - If not, return 0 or 1 at random

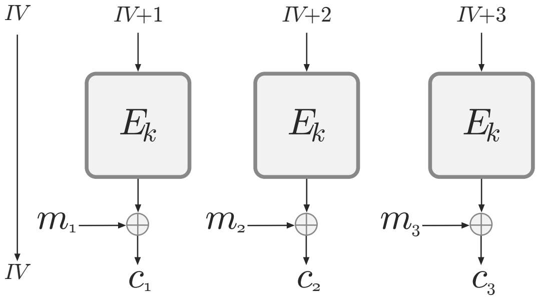
Analysis

- Correct answer if there exists $i \neq j$ s.t. $c_i = c_j$, since $c_{i-1} \oplus c_{j-1} = m_i \oplus m_j$
- $\Pr[\exists i \neq j, c_i = c_j] \gtrsim \frac{1}{4} \rightarrow$ advantage $\gtrsim \frac{1}{2}$
- Time to find collisions: $O(2^{n/2})$

Conclusion

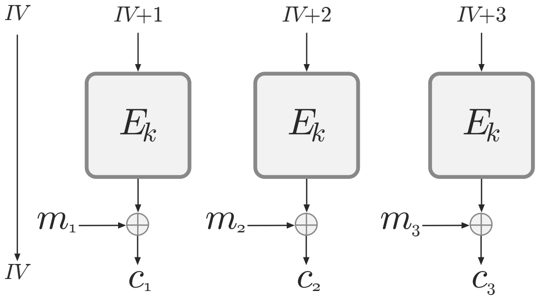
- $\text{Adv}^{\text{IND-CPA}}_{\text{Enc}-CBC}(2^{n/2}, 2^{n/2}) \geq \text{Adv}^{\text{IND-CPA}}_{\text{Enc}-CBC}(A_{\text{BIRTHDAY}}) \gtrsim \frac{1}{2}$
- CBC mode should not be used for too long with the same key!

Second example of mode of operation: Counter (CTR)



Source : J. Katz, Y. Lindell. Introduction to modern cryptography. 3rd ed, CRC Press, 2021. (modif.)

Second example of mode of operation: Counter (CTR)



- IV: *random* initialization vector in $\{0,1\}^n$
- Input: $m = m_1 \| \cdots \| m_\ell$
- Output: $c = IV \| c_1 \| \cdots \| c_\ell$ — size $n(\ell+1)$
- Parallel encryption (fast!) — similar to a *stream cipher*
- Also sensitive to birthday bound
- IND-CPA security from PRF security — variant of PRP security

IND-CPA security for CTR: sketch of the proof

Skipped during the class

Finally

Modes of operations

- A *good* mode of operation turns a *good* block cipher into a *good* symmetric encryption scheme
- Different mode of operations require different quality for the block cipher
  - *Good* PRP
  - *Good* PRF
  - Ideal Block Cipher
- Proofs of security $\rightarrow$ reductions between problems
- Usually: need more $\rightarrow$ *ad hoc* analysis of the resulting system

Other symmetric encryption schemes

- Other modes of operations — OFB, CFB
- Stream ciphers — *Wifi, 5G, ...*

Conclusion

Symmetric encryption, as we saw it

- Two ingredients:
  - a block cipher — fixed-size, deterministic
  - a mode of operation — variable-size, non-deterministic
- Security notions:
  - PRP advantage — block cipher
  - IND-CPA advantage — symmetric encryption
- More advanced security definitions:
  - strong PRP adv., (strong) PRF adv., ideal block cipher
  - IND-CCA, IND-CCA2, multiple encryptions

In practice

- Block cipher: mainly AES, with key size 128 bits
- Modes of operations: *e.g.* extension of CTR in TLS

Final words: **Definitions and proofs are important!**