Lecture 3. Symmetric encryption

Passive adversaries, small shared secret

Bruno Grenet

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

Introduction to cryptology
Université Grenoble Alpes – IM²AG
M1 INFO, MOSIG & AM

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

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

- Enc is a *randomized* encryption algorithm
- Dec is a (deterministic) decryption algorithm

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 $c \leftarrow \mathsf{Enc}_k(m)$, $|c| \simeq |m|$

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 algorithm
- Dec is a (deterministic) decryption algorithm

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*

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

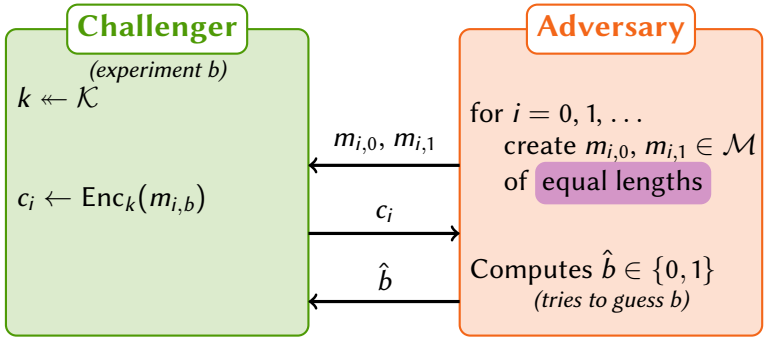
- A mode is *good* if it turns *good BCs* into *good encryption schemes*

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

Contents

Reminder: IND-CPA game



Remarks

- Oracle access to Enc_k during the whole experiment
  - To get $c \leftarrow \mathrm{Enc}_k(m)$, create $m_{i,0} = m_{i,1} = m$
- Equal-length messages $\rightsquigarrow$ message length not hidden!
  - Impossible to hide if messages of any length
  - Use padding beforehand if message length is sensitive

Reminder: IND-CPA advantage

IND-CPA advantage of an adversary $\mathcal{A}$

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

Extremal cases

- Guessing $\hat{b}$ at random $\rightsquigarrow$ advantage 0
- Resource-unbounded $\mathcal{A} \rightsquigarrow$ advantage 1

IND-CPA advantage function

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

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

From lecture 1: Two approaches

IND-CPA secure: **any** *efficient* adversary has *negligible* advantage

Asymptotic security

complexity-theoretic definition

- Expressed in terms of the *security parameter* λ
- Efficient: polynomial in λ $\qquad q = \text{poly}(\lambda),\ t = \text{poly}(\lambda)$
- Negligible: $< 1/p(\lambda)$ for any polynomial p

$$\text{Enc is IND-CPA secure} \iff \text{Adv}^{\text{IND-CPA}}_{\text{Enc}}(\text{poly}(\lambda), \text{poly}(\lambda)) \ll 1/\text{poly}(\lambda)$$

Concrete security

mostly chosen in this course

- No formal definition of *efficient* or *negligible*
- Compare schemes by comparing their advantage functions
- Plug some explicit values:
  - what is an *reasonable* number of operations / number of queries?
  - what is an *acceptable* advantage?

From lecture 1: Orders of magnitude (time)

Computation time: number of elementary operations

- $t \simeq 2^{40}$: $\sim$ 1 day on my laptop
- $t \simeq 2^{60}$: possible on a large CPU/GPU cluster — done in academia
- $t \simeq 2^{80}$: possible with an ASIC cluster — Bitcoin mining
- $t \simeq 2^{128}$: *very* hard

How much time for 2^{128} operations?

- With all 500 fastest super-computers, assuming parallelizability — $10 \cdot 10^9$ gFLOPS
- Time: $2^{128}/(10 \cdot 10^9 \cdot 10^9 \times 365 \cdot 24 \cdot 3600) \approx 10^{12}$ years — $\approx 4\times$ age of Earth

How much power for 2^{128} operations in 100 years?

- With hardware at 10^6 gFLOPS using 1 000W per device — pretty good!
- Number of devices: $2^{128}/(10^6 \cdot 10^9 \times 100 \cdot 365 \cdot 24 \cdot 3600) \approx 10^{14}$
- Power: $10^{14} \times 1000\text{W} \approx 100\,000$ TW — $\approx 65 \cdot 10^6$ EPR

From lecture 1: Orders of magnitude (probabilities)

Probabilities

- $p = \frac{1}{2}$: get a TAIL with a fair coin
- $p = \frac{1}{6}$: get a 6 with a fair die
- $p \approx 2^{-24}$: probability to win at French lottery
- $p \approx 2^{-72}$: probability to win 3 times in a row at French lottery

Examples

- An attack that takes 1 second and has a probability of success of 2^{-60} is expected to have succeeded less than once since the Big Bang
- Errors in CPU due to cosmic rays happen with much larger probabilities!

Combining orders of magnitude

If $\mathsf{Adv}^{\text{IND-CPA}}_{\text{Enc}}(2^{128}, 2^{128}) < 2^{-60}$, Enc is pretty (IND-CPA) secure!

From lecture 1: Orders of magnitude (Probabilities)

Probabilities

- $p = \frac{1}{2}$: get a TAIL with a fa
- $p = \frac{1}{6}$: get a 6 with a fair
- $p \approx 2^{-24}$: probability to w
- $p \approx 2^{-72}$: probability to w

h lottery

Examples

- An attack that takes 1 sec
- have succeeded less than

f success of 2^{-60} is expected to

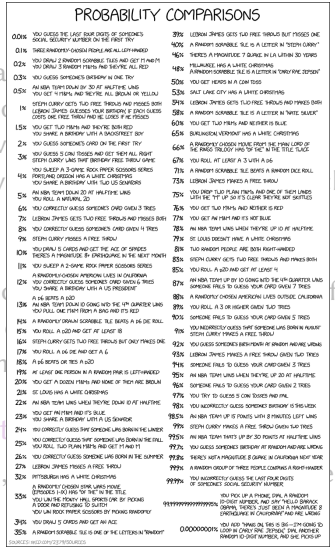
- Errors in CPU due to cosm

larger probabilities!

Combining orders of magnit

If $\mathrm{Adv}^{\mathrm{IND\text{-}CPA}}_{\mathrm{Enc}}(2^{128}, 2^{128}) < 2^{-60}$

cure!



https://xkcd.com/2379/

17%	YOU ROLL A D6 DIE AND GET A 6
18%	A D6 BEATS OR TIES A D20
19%	AT LEAST ONE PERSON IN A RANDOM PAIR IS LEFT-HANDED
20%	YOU GET A DOZEN M&MS AND NONE OF THEM ARE BROWN
21%	ST. LOUIS HAS A WHITE CHRISTMAS
22%	AN NBA TEAM WINS WHEN THEY'RE DOWN 10 AT HALFTIME
23%	YOU GET AN M&M AND IT'S BLUE YOU SHARE A BIRTHDAY WITH A US SENATOR
24%	YOU CORRECTLY GUESS THAT SOMEONE WAS BORN IN THE WINTER
25%	YOU CORRECTLY GUESS THAT SOMEONE WAS BORN IN THE FALL YOU ROLL TWO PLAIN M&MS AND GET M AND M
26%	YOU CORRECTLY GUESS SOMEONE WAS BORN IN THE SUMMER
27%	LEBRON JAMES MISSES A FREE THROW
32%	PITTSBURGH HAS A WHITE CHRISTMAS
33%	A RANDOMLY CHOSEN STAR WARS MOVIE (EPISODES I-IX) HAS "OF THE" IN THE TITLE YOU WIN THE MONTY HALL SPORTS CAR BY PICKING A DOOR AND REFUSING TO SWITCH YOU WIN ROCK PAPER SCISSORS BY PICKING RANDOMLY
34%	YOU DRAW 5 CARDS AND GET AN ACE
35%	A RANDOM SCRABBLE TILE IS ONE OF THE LETTERS IN "RANDOM"

93%	LEBRON JAMES MAKES A FREE THROW GIVEN TWO TRIES
94%	SOMEONE FAILS TO GUESS YOUR CARD GIVEN 3 TRIES
95%	AN NBA TEAM WINS WHEN THEY'RE UP 20 AT HALFTIME
96%	SOMEONE FAILS TO GUESS YOUR CARD GIVEN 2 TRIES
97%	YOU TRY TO GUESS 5 COIN TOSSES AND FAIL
98%	YOU INCORRECTLY GUESS SOMEONE'S BIRTHDAY IS THIS WEEK
98.5%	AN NBA TEAM UP 15 POINTS WITH 8 MINUTES LEFT WINS
99%	STEPH CURRY MAKES A FREE THROW GIVEN TWO TRIES
99.5%	AN NBA TEAM THAT'S UP BY 30 POINTS AT HALFTIME WINS
99.7%	YOU GUESS SOMEONE'S BIRTHDAY AT RANDOM AND ARE WRONG
99.8%	THERE'S NOT A MAGNITUDE 8 QUAKE IN CALIFORNIA NEXT YEAR
99.9%	A RANDOM GROUP OF THREE PEOPLE CONTAINS A RIGHT-HANDER
99.99%	YOU INCORRECTLY GUESS THE LAST FOUR DIGITS OF SOMEONE'S SOCIAL SECURITY NUMBER
99.999999999999995%	YOU PICK UP A PHONE, DIAL A RANDOM 10-DIGIT NUMBER, AND SAY "HELLO BARACK OBAMA, THERE'S JUST BEEN A MAGNITUDE 8 EARTHQUAKE IN CALIFORNIA!" AND ARE WRONG
0.000000001%	YOU ADD "HANG ON, THIS IS BIG—I'M GOING TO LOOP IN CARLY RAE JEPSEN," DIAL ANOTHER RANDOM 10-DIGIT NUMBER, AND SHE PICKS UP

SOURCES: XKCD.COM/2379/SOURCES

https://xkcd.com/2379/

Comments on IND-CPA security

- No formal definition of IND-CPA secure, only a measure *(but in asymptotic security)*
- IND-CPA $\Rightarrow$ non-determinism
- IND-CPA $\Rightarrow$ the adversary cannot compute any single bit of the message
- IND-CPA $\Rightarrow$ the adversary can compute *very few* information on the message

Stronger security notion: IND-CCA

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)

Contents

Goal

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

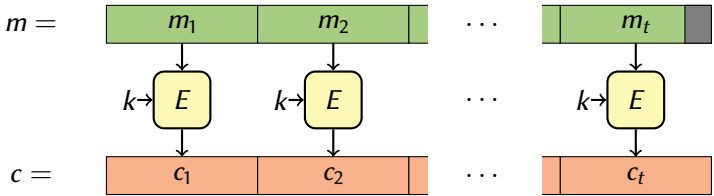
- E is made to encrypt one block of data
- Enc should encrypt any number of blocks

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

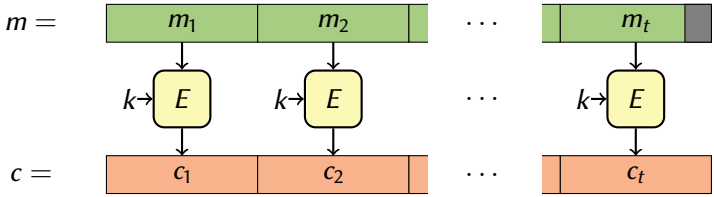
Desired properties

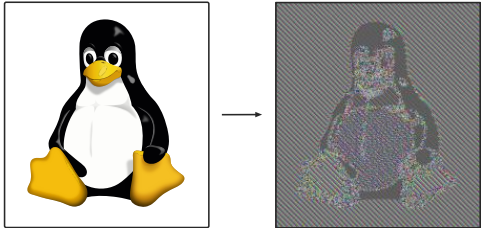
- Security:
  - E *secure* $\implies$ Enc *secure*
  - Low (S)PRP advantage $\implies$ low IND-CPA advantage
- Efficiency:
  - Efficient encryption and decryption — if E is efficient
  - Ciphertext not too large compared to message

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



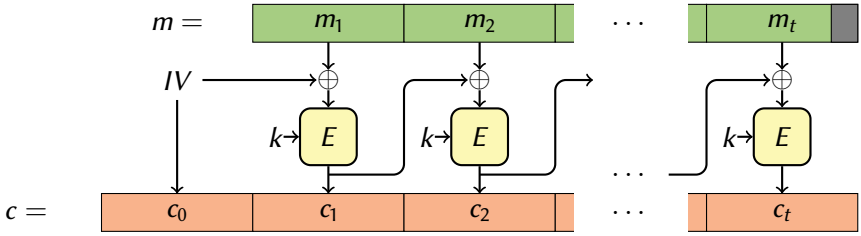
Obvious (bad) idea: Electronic Code Book (ECB)



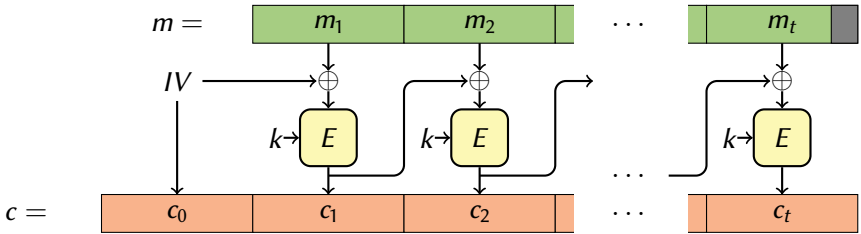


Source : Wikipédia (modif.)

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



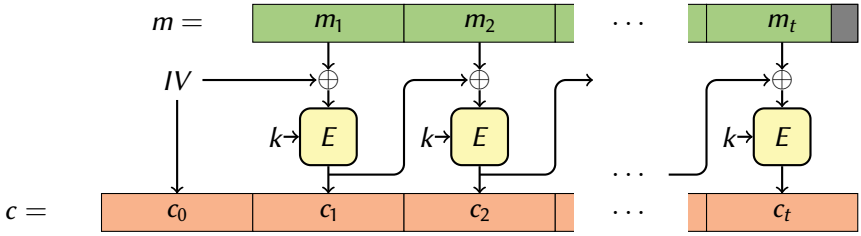
First (real) example of mode of operation: Cipher Block Chaining (CBC)



- IV: *random* initialization vector in $\{0,1\}^n$
- Input: $m = m_1 \| \cdots \| m_t$ — padding if needed
- Output: $c = c_0 \| c_1 \| \cdots \| c_t$ — size $n(t+1)$

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

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



Adversary when IV is not uniform

1. One-block query: $m \to r \| c = r \| E_k(m \oplus r)$ — *r is IV*
2. *Guess* the next IV: r'
3. Challenges: $\to r'' \| c_b = r \| E_k(m_b \oplus r'')$
   - $m_0 = m \oplus r \oplus r'$
   - m_1 uniform

   $E_k(m_0 \oplus r'') = E_k(m \oplus r \oplus r' \oplus r'')$
4. If $r' = r''$ (*correct guess*), return $\begin{cases} b = 0 & \text{if } c = c_b \\ b = 1 & \text{otherwise} \end{cases}$, otherwise answer at random

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

- Two identical ciphertext blocks $c_i = c'_j$ reveal information on m_i and m'_j:

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

- 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

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

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

- $\Pr[\exists\, i \neq j,\ y_i = y_j] = \Pr\left[\bigvee_{i \neq j} y_i = y_j\right] \underset{\text{union bound}}{\leq} \sum_{i \neq j} \Pr[y_i = y_j]$
- $\Pr[y_i = y_j] = \frac{1}{N}$
- $\Pr[\exists\, i \neq j,\ y_i = y_j] \leq \sum_{i \neq j} \frac{1}{N} = \frac{1}{N}\binom{q}{2}$

Proof of the birthday lower bound

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

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

- $p = \Pr[\forall i \neq j, y_i \neq y_j] = \Pr[y_2 \notin \{y_1\} \wedge y_3 \notin \{y_1, y_2\} \wedge \cdots \wedge y_q \notin \{y_1, \ldots, y_{q-1}\}]$
- $E_i =$ "$y_i \notin \{y_1, \ldots, y_{i-1}\}$" : $p = \Pr[E_2 \wedge E_3 \wedge \cdots \wedge E_q]$

$= \Pr[E_2] \cdot \Pr[E_3 | E_2] \cdot \Pr[E_4 | E_2 \wedge E_3] \cdot \cdots \cdot \Pr[E_q | E_2 \wedge \cdots \wedge E_{q-1}]$

- $\Pr[E_i | E_2 \wedge \cdots \wedge E_{i-1}] = \frac{N-(i-1)}{N} = 1 - \frac{i-1}{N}$.

$\Rightarrow p = \prod_{i=2}^{q} \left(1 - \frac{i-1}{N}\right) \leq \prod_{i=2}^{q} e^{-\frac{i-1}{N}} = e^{-\sum_{i=2}^{q} \frac{i-1}{N}} = e^{-\frac{q(q-1)}{2N}}$.

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

The birthday attack against CBC

Adversary $\mathcal{A}_{\text{BIRTHDAY}}$

- 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

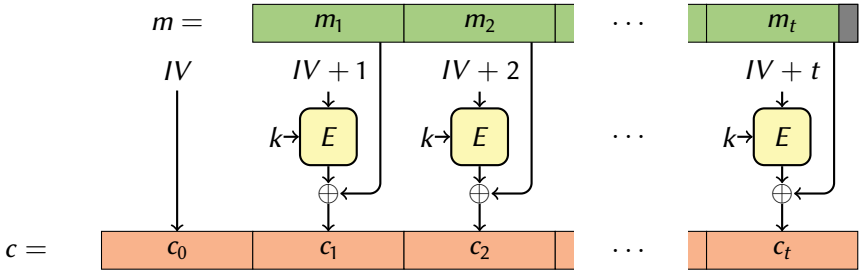
- Correct answer if there exists $i \neq j$ s.t. $c_i = c_j$, since $c_{i-1} \oplus c_{j-1} = m_i \oplus m_j$
- $\Pr[\exists i \neq j, c_i = c_j] \gtrsim \frac{1}{4} \rightarrow \text{advantage} \gtrsim \frac{1}{4}$
- Time to find collisions: $O(2^{n/2})$

should be the same

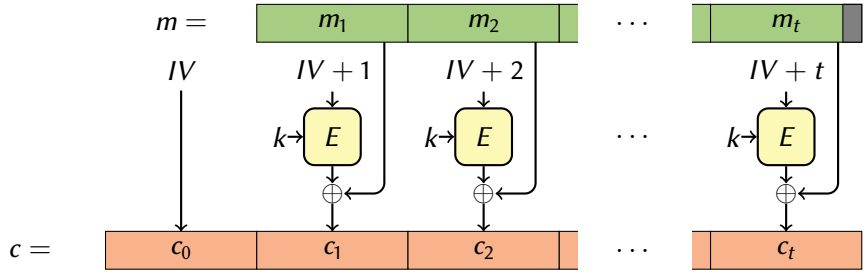
Conclusion

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

Second example of mode of operation: Counter (CTR)



Second example of mode of operation: Counter (CTR)



- IV: *random* initialization vector in $\{0,1\}^n$
- Input: $m = m_1 \| \cdots \| m_t$
- Output: $c = c_0 \| c_1 \| \cdots \| c_t$ size $n(t+1)$

- 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

1. PRF Security

on the board.

IND-CPA security for CTR: sketch of the proof

2. Proof Given an IND-CPA adv. A_0 for CTR[E], build a PRF adv. for E

- A_0 : adversary in the IND-CPA game with q queries and time t

Queries: $m_{i,0} = m_{i,0}^1 \| m_{i,0}^2 \| \dots \| m_{i,0}^{t_i}$

$m_{i,1} = m_{i,1}^1 \| m_{i,1}^2 \| \dots \| m_{i,1}^{t_i}$

Answers: $c_i = c_i^0 \| c_i^1 \| \dots \| c_i^{t_i}$ where $\begin{cases} c_i^0 = IV_i \\ c_i^s = E_k(IV_i + s) \oplus m_{i,b}^s \end{cases}$

let $\boxed{Q = \sum_{i=0}^{q+1} t_i}$

IND-CPA security for CTR: sketch of the proof

- Consider $\mathrm{CTR}[f]$ ~> counter mode with a uniform function f instead of F_k.

$$\left| \Pr\left[A^{\mathrm{CTR}[E_n]} \to 1 \mid b\right] - \Pr\left[A^{\mathrm{CTR}[f]} \to 1 \mid b\right] \right| \leq \mathrm{Adv}^{\mathrm{PRF}}_{E}(Q, t)$$

- Assume (in $\mathrm{CTR}[f]$) that there is no collision $IV_i + s = IV_{i'} + s'$ then all the $f(IV_i + s)$ are uniform and independent
  → $\mathrm{CTR}[f]$ is the one-time pad under this assumption

(under the assumption) $\left| \Pr\left[A^{\mathrm{CTR}[f]} \to 1 \mid b=1\right] - \Pr\left[A^{\mathrm{CTR}[f]} \to 1 \mid b=0\right] \right| = 0$

IND-CPA security for CTR: sketch of the proof

- $\Pr\left[\exists\, i, i', s, s' ,\ |V_i + s = |V_{i'} + s'\right] \leq \frac{Q(Q-1)}{2 \cdot 2^n}$ (birthday bound)

3. Conclusion

$$\mathrm{Adv}^{\mathrm{IND\text{-}CPA}}_{\mathrm{CTR}[E]}(Q, t) \left| \Pr\left[\mathcal{A}^{\mathrm{CTR}[E]} \longrightarrow 1 \mid b = 1\right] - \Pr\left[\mathcal{A}^{\mathrm{CTR}[E]} \longrightarrow 1 \mid b = 0\right] \right)$$

$$\leq \left| \Pr\left[\mathcal{A}^{\mathrm{CTR}[f]} \longrightarrow 1 \mid b = 1\right] - \Pr\left[\mathcal{A}^{\mathrm{CTR}[f]} \longrightarrow 1 \mid b = 0\right] \right| + 2 \cdot \mathrm{Adv}^{\mathrm{PRF}}_{E}(Q, t)$$

$$\leq \frac{Q(Q-1)}{2^{n+1}} + 2\, \mathrm{Adv}^{\mathrm{PRF}}_{E}(Q, t)$$

Finally

Modes of operations

- *good* mode of operation + *secure* block cipher $\Rightarrow$ *secure* symmetric encryption scheme
- Different mode of operations require different security for the block cipher
  - PRP — *pseudo-random permutation*
  - PRF — *pseudo-random function*
  - 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

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!**