Lecture 6 – Symmetric-Key Encryption

COSC-260 Codes and Ciphers

Adam O'Neill

Adapted from http://cseweb.ucsd.edu/~mihir/cse107/

Setting the Stage

- We have studied our first lower-level primitive, blockciphers.
- Today we will study how to use it to build our first higher-level primitive, symmetric-key encryption.

Syntax

A symmetric-key encryption scheme is a triple of algorithms $SE = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ with message-space $MsgSp$ such that:

- $\mathcal{K}$ is the key-generation algorithm that is randomized and outputs a key K
- $\mathcal{E}$ is the encryption algorithm that is randomized and on inputs $K, m \in MsgSp$ outputs a ciphertext c.
- $\mathcal{D}$ is the deterministic decryption algorithm that on inputs K, c outputs a message $m \in MsgSp$ or $\perp$

or "Symbol", "bot", or "bottom"

Correctness

We say that SE is correct if For all k output by $\mathcal{K}$ and all $m \in \mathcal{M}sg_{SP}$

$$\Pr\left[\mathcal{D}_k(\mathcal{E}_k(m)) = m\right] = 1.$$

over coins of $\mathcal{E}$.

Game:

$K \xleftarrow{\$} \mathcal{K}$

$c \xleftarrow{\$} \mathcal{E}(K, m)$

$m' \leftarrow \mathcal{D}(K, c)$

ret $(m' = m)$

Outputs 1 w.p. 1.

Blockcipher Modes of Operation

key-length k e.g AES
block-length n. k=n=128

$E : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^n$ a block cipher

Notation: $x[i]$ is the i-th n-bit block of a string x, so that $x = x[1] \ldots x[m]$ if $|x| = nm$.

Always:

Alg $\mathcal{K}$
$K \xleftarrow{\$} \{0,1\}^k$
return K

Electronic Codebook Mode

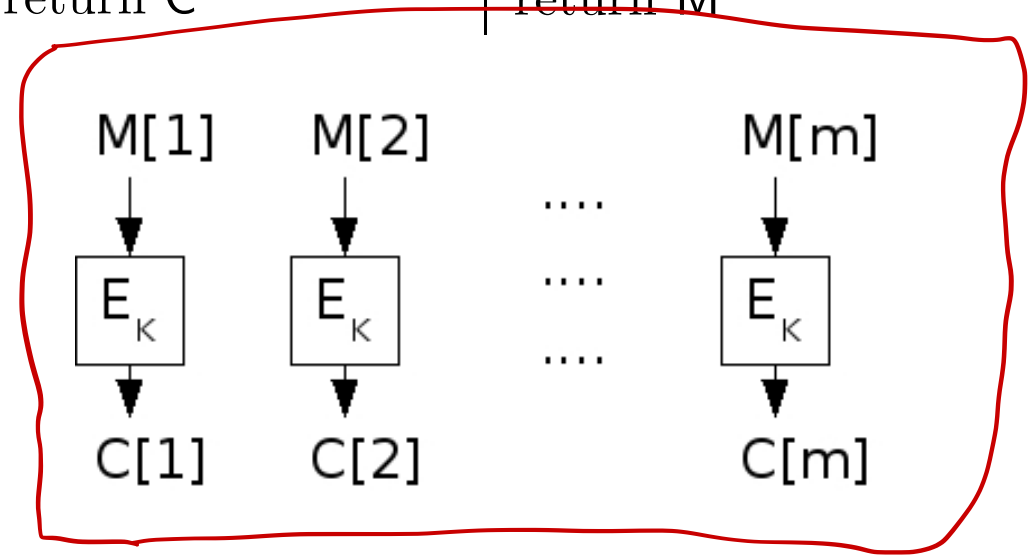
$\mathcal{SE} = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ where:

Alg $\mathcal{E}_K(M)$

for $i = 1, \ldots, m$ do
    $C[i] \leftarrow E_K(M[i])$
return C

Alg $\mathcal{D}_K(C)$

for $i = 1, \ldots, m$ do
    $M[i] \leftarrow E_K^{-1}(C[i])$
return M



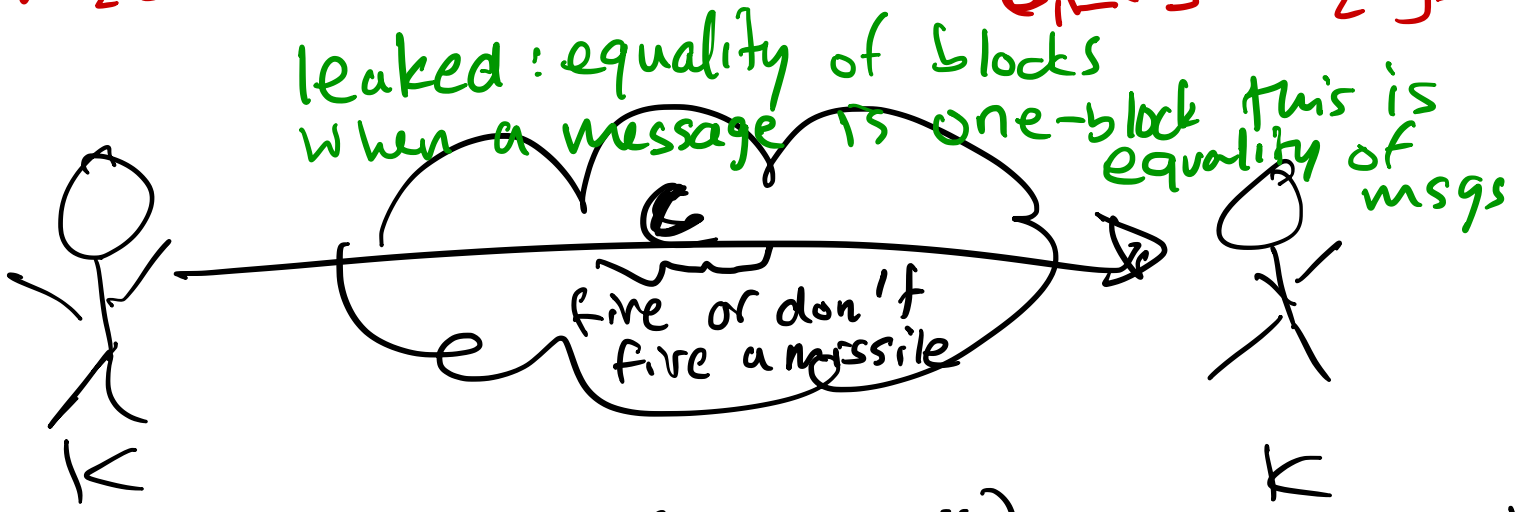
Weakness of ECB

the same block of the message will encrypt to the same block in the ciphertext

$m_1[1] \ldots m_1[n]$

$m_2[1] \ldots m_2[n]$

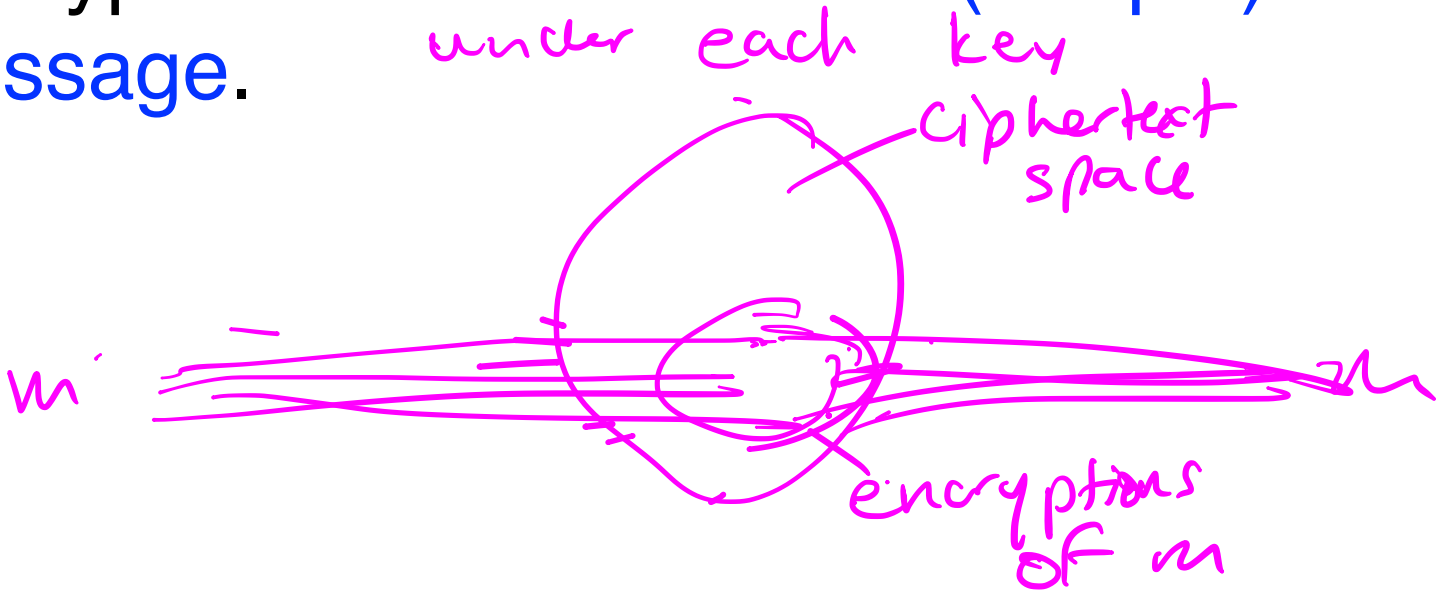
if $m_1[i] = m_2[j]$ then $c_1[i] = c_2[j]$



one-block (e.g. one bit) message is sent

Introducing Randomized Encryption

- Encryption algorithm flips coins.
- Many possible ciphertexts for each message (using the same key).
- Decryption still recovers the (unique) message.



CBC-$:
Cipher-block Chaining Mode with Random IV

$\mathcal{SE} = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ where:

called IV or initialization vector

Alg $\mathcal{E}_K(M)$

$C[0] \xleftarrow{\$} \{0,1\}^n$

for $i = 1, \ldots, m$ do

$\quad C[i] \leftarrow E_K(M[i] \oplus C[i-1])$

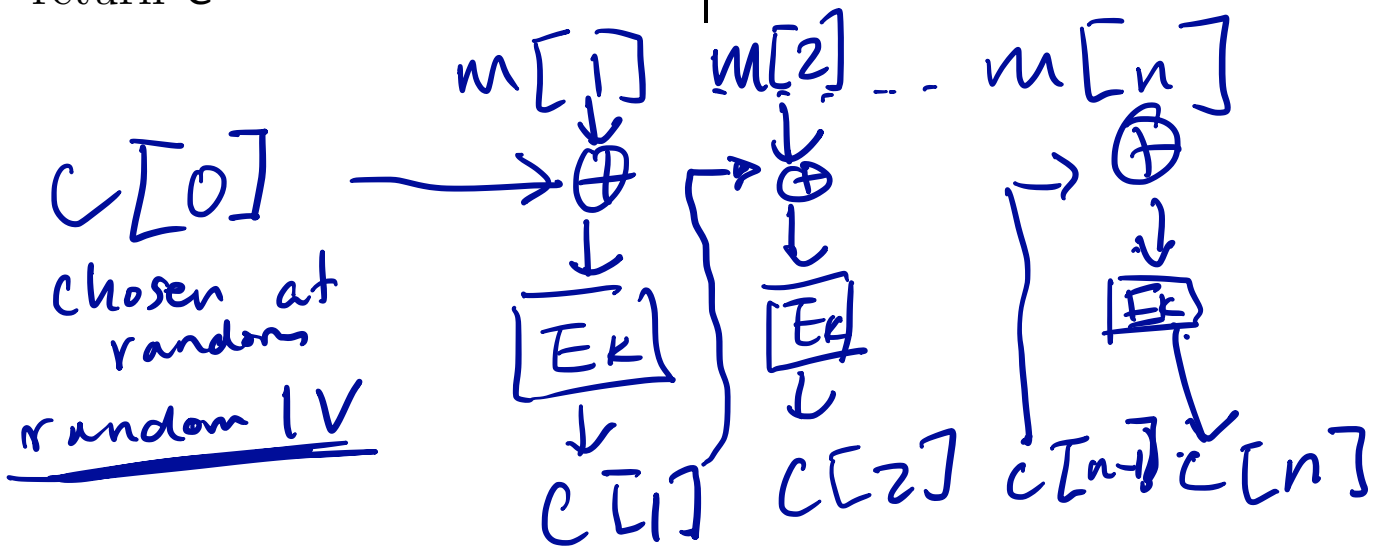
return C

Alg $\mathcal{D}_K(C)$

for $i = 1, \ldots, m$ do

$\quad M[i] \leftarrow E_K^{-1}(C[i]) \oplus C[i-1]$

return M



CTR-$ Mode

Counter-mode

If $X \in \{0,1\}^n$ and $i \in \mathbf{N}$ then $X + i$ denotes the n-bit string formed by converting X to an integer, adding i modulo 2^n, and converting the result back to an n-bit string.

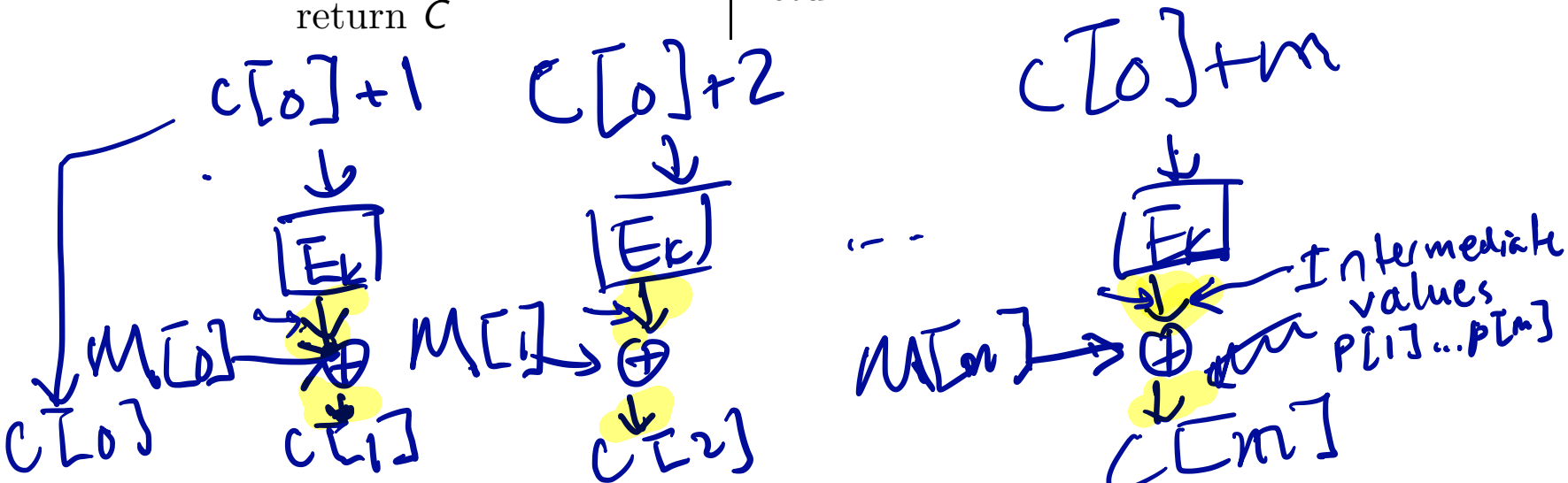
IV

Alg $\mathcal{E}_K(M)$

$C[0] \xleftarrow{\$} \{0,1\}^n$
for $i = 1, \ldots, m$ do
  $P[i] \leftarrow E_K(C[0] + i)$
  $C[i] \leftarrow P[i] \oplus M[i]$
return C

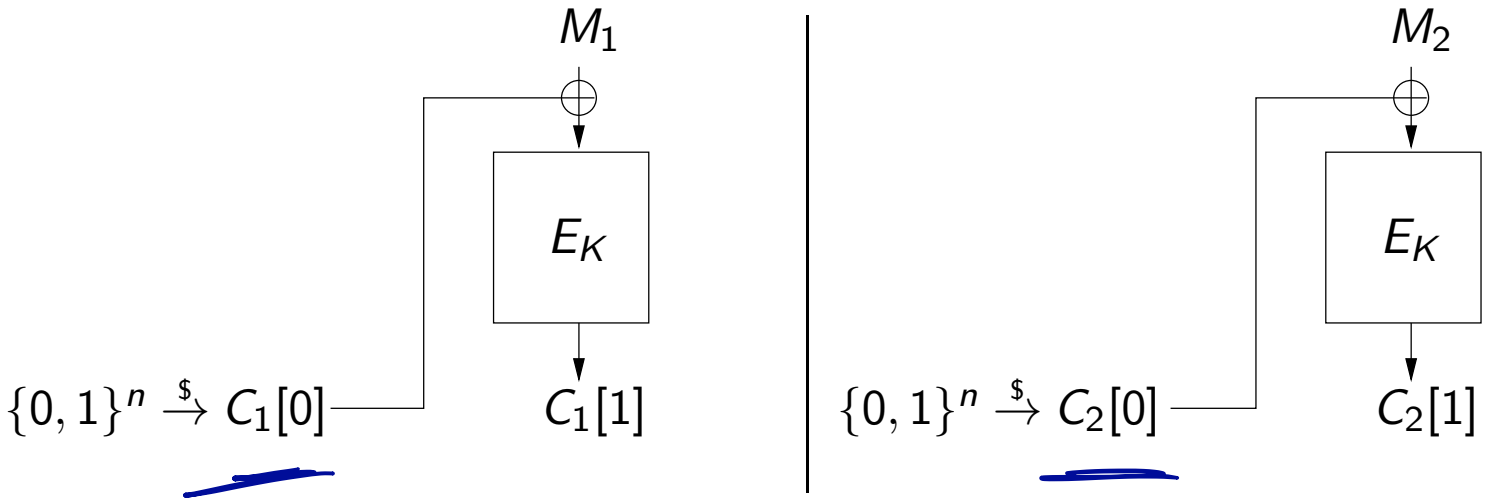
Alg $\mathcal{D}_K(C)$

for $i = 1, \ldots, m$ do
  $P[i] \leftarrow E_K(C[0] + i)$
  $M[i] \leftarrow P[i] \oplus C[i]$
return M



Voting with CBC-$

Suppose we encrypt $M_1, M_2 \in \{Y, N\}$ with CBC$.



Assessing Security

- How to determine which modes of operations are "good" ones?
- E.g., CBC-$ seems better than ECB. But is it secure? Or are there still attacks?
- Important since CBC-$ is widely used.

Towards a Master Property

- As before, one approach is to list requirements for a "good" encryption scheme.
  - Key recovery is hard.
  - Message recovery is hard
  - … Recovering the XOR of multiple messages.
- Better idea: Specify a master property that implies all the properties in such an (infinite) list.

Intuition

Encryption should hide all "partial information" !!!

In particular, if I encrypt a sequence of message, it shouldn't be any easier to guess the value of a function of those messages compared to when other messages are encrypted.

Indistinguishability under chosen-plaintext attack

IND-CPA

Let $\mathcal{SE} = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be an encryption scheme

Game $\mathrm{Left}_{\mathcal{SE}}$	Game $\mathrm{Right}_{\mathcal{SE}}$
procedure Initialize	**procedure Initialize**
$K \xleftarrow{\$} \mathcal{K}$	$K \xleftarrow{\$} \mathcal{K}$
procedure LR(M_0, M_1)	**procedure LR**(M_0, M_1)
Return $C \xleftarrow{\$} \mathcal{E}_K(M_0)$	Return $C \xleftarrow{\$} \mathcal{E}_K(M_1)$

Associated to $\mathcal{SE}, A$ are the probabilities

$$\Pr\left[\mathrm{Left}^A_{\mathcal{SE}} \Rightarrow 1\right] \qquad \Pr\left[\mathrm{Right}^A_{\mathcal{SE}} \Rightarrow 1\right]$$

that A outputs 1 in each world. The (ind-cpa) advantage of A is

$$\mathbf{Adv}^{\text{ind-cpa}}_{\mathcal{SE}}(A) = \Pr\left[\mathrm{Right}^A_{\mathcal{SE}} \Rightarrow 1\right] - \Pr\left[\mathrm{Left}^A_{\mathcal{SE}} \Rightarrow 1\right]$$

Advantage Interpretation

$\mathbf{Adv}^{\text{ind-cpa}}_{\mathcal{SE}}(A) \approx 1$ means A is doing well and $\mathcal{SE}$ is not ind-cpa-secure.

$\mathbf{Adv}^{\text{ind-cpa}}_{\mathcal{SE}}(A) \approx 0$ (or ≤ 0) means A is doing poorly and $\mathcal{SE}$ resists the attack A is mounting.

Adversary resources are its running time t and the number q of its oracle queries, the latter representing the number of messages encrypted.

Security: $\mathcal{SE}$ is IND-CPA-secure if $\mathbf{Adv}^{\text{ind-cpa}}_{\mathcal{SE}}(A)$ is "small" for ALL A that use "practical" amounts of resources.

Insecurity: $\mathcal{SE}$ is not IND-CPA-secure if we can specify an explicit A that uses "few" resources yet achieves "high" ind-cpa-advantage.

$E: \{0,1\}^k \times \{0,1\}^l \to \{0,1\}^p$

Security Analysis of ECB

Claim: ECB is NOT IND-CPA secure. (In fact, no deterministic encryption scheme is IND-CPA)

Proof:

Adversary A

$c_1 \leftarrow LR(0^l, 1^l)$

$c_2 \leftarrow LR(0^l, 0^l)$

If $c_1 \neq c_2$

return 1

Else return 0

Claim 1:

$Pr[RIGHT_{ECB}^{A} \Rightarrow 1] = 1.$

proof.

Claim:

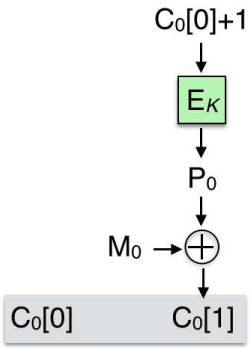
$Pr[LEFT_{ECB}^{A} \Rightarrow 1] = 0.$

proof.

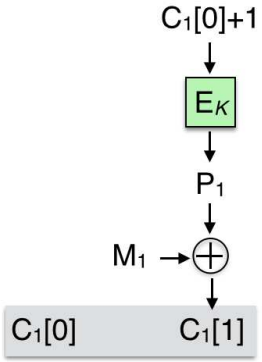
Security Analysis of CTR-$

Let $E : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^n$ be a blockcipher and $\mathcal{SE} = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ the corresponding CTR$ symmetric encryption scheme. Suppose 1-block messages M_0, M_1 are encrypted:

$C_0[0]C_0[1] \xleftarrow{\$} \mathcal{E}(K, M_0)$



$C_1[0]C_1[1] \xleftarrow{\$} \mathcal{E}(K, M_1)$



Let us say we are **lucky** If $C_0[0] = C_1[0]$. If so:

$$C_0[1] = C_1[1] \text{ if and only if } M_0 = M_1$$

So if we are lucky we can detect message equality and violate IND-CPA.

The Adversary

Advantage Analysis

Conclusion: CTR$ can be broken (in the IND-CPA sense) in about $2^{n/2}$ queries, where n is the block length of the underlying block cipher, regardless of the cryptanalytic strength of the block cipher.

So far: A q-query adversary can break CTR\$ with advantage $\approx \frac{q^2}{2^{n+1}}$

Question: Is there any better attack?

Answer: NO!

We can prove that the best q-query attack short of breaking the block cipher has advantage at most

$$\frac{\sigma^2}{2^n}$$

where σ is the total number of blocks encrypted.

Example: If q 1-block messages are encrypted then $\sigma = q$ so the adversary advantage is not more than $q^2/2^n$.

For $E = \text{AES}$ this means up to 2^{64} blocks may be securely encrypted, which is good.

Theorem Statement

Theorem: [BDJR98] Let $E : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^n$ be a block cipher and $\mathcal{SE} = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ the corresponding CTR$ symmetric encryption scheme. Let A be an ind-cpa adversary against $\mathcal{SE}$ that has running time t and makes at most q **LR** queries, these totalling at most σ blocks. Then there is a prf-adversary B against E such that

$$\mathbf{Adv}^{\text{ind-cpa}}_{\mathcal{SE}}(A) \leq 2 \cdot \mathbf{Adv}^{\text{prf}}_{E}(B) + \frac{\sigma^2}{2^n}$$

Furthermore, B makes at most σ oracle queries and has running time $t + \Theta(\sigma \cdot n)$.

- Analogous theorem holds for CBC-$.
- Provides a quantitative guarantee on how many blocks can be securely encrypted using these modes (assuming the underlying block cipher is good).