ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

Nhập môn An Toàn Thông Tin

Sử dụng mã khối

Mã khối lý tưởng

- Trên thực tế, người ta xem AES hoặc 3DES như một hệ ***mã khối lý tưởng***;
- Tức là, với mỗi khóa k, ánh xạ

$$F_k(x) = e(k, x)$$

là một hoán vị ngẫu nhiên độc lập từ $\{0,1\}^{128}$ lên chính nó.

Các chế độ và mode sử dụng

Câu hỏi: Làm thế nào để mã hóa thông điệp với độ dài bất kỳ? (dùng AES)

Trả lời: Dùng một trong các mode sau:

- "ECB" = "Electronic code book"
- "CTR" = "Counter mode"
- "CBC" = "Cipher Block Chaining"
- "OFB" = "Output Feedback"
- ...

Chế độ sử dụng: Khoá chỉ sử dụng một lần và khoá dùng nhiều lần.

Ứng dụng của khoá dùng nhiều lần

Mã hóa hệ thống file

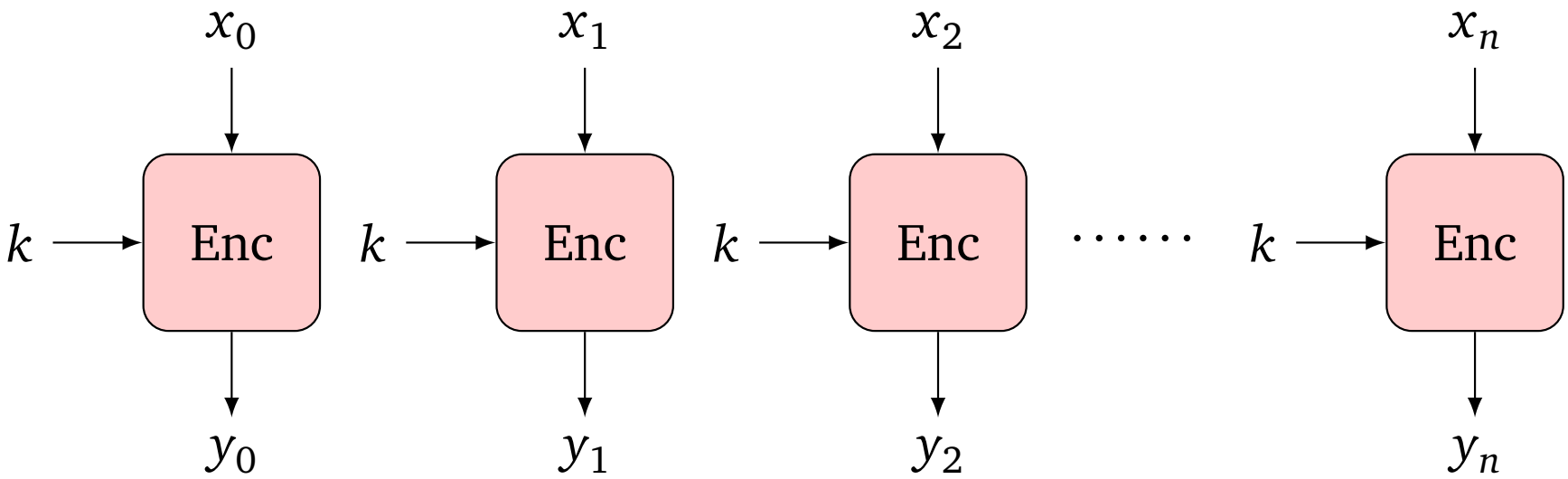
- Mã hóa nhiều file dùng AES với cùng khóa

IPSec

- Nhiều gói tin cùng được mã hóa bằng AES với cùng một khóa

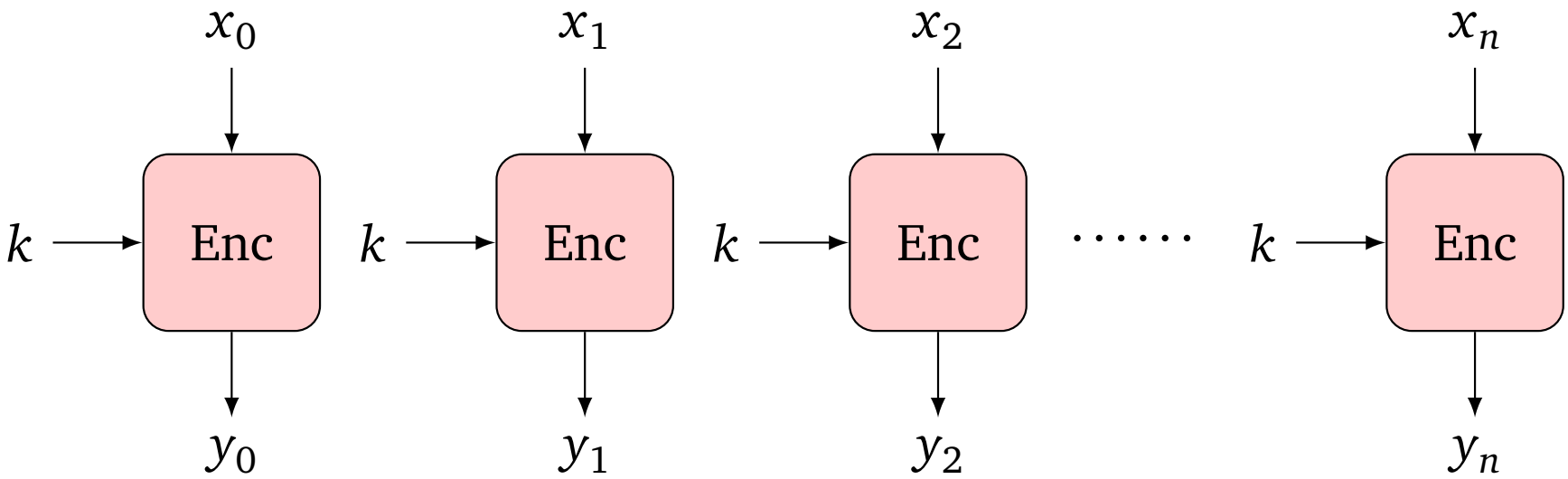
Nội dung

ECB (Electronic code book)



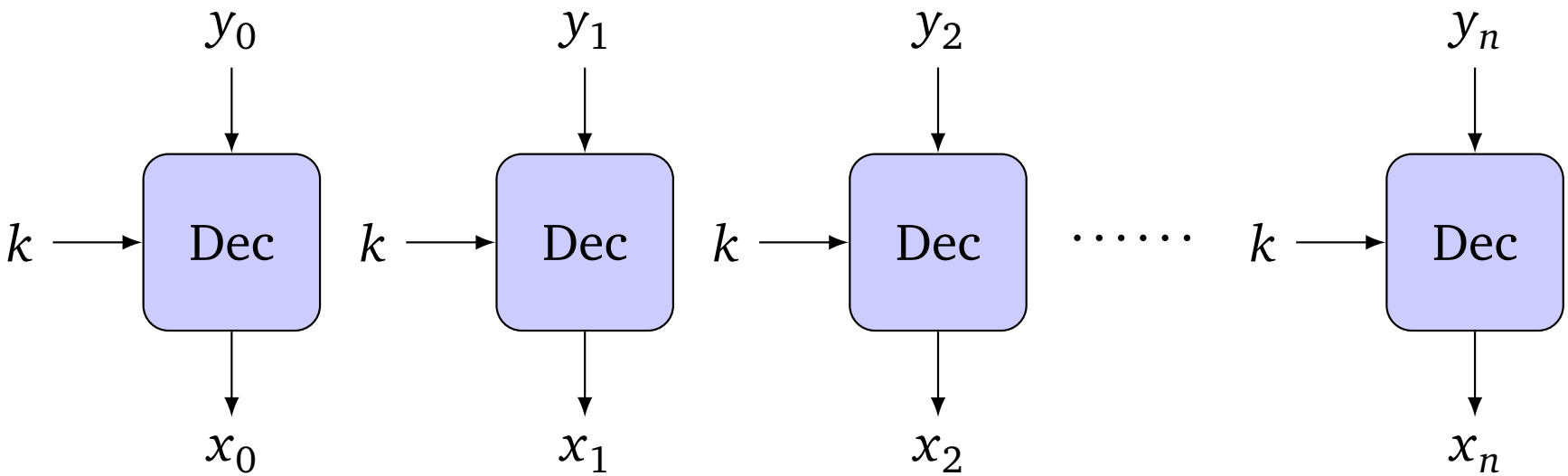
- Dữ liệu được chia thành các khối khối b bit, với $b =$ kích thước khối.
- Với dữ liệu không chia hết cho b bit: Thêm dãy "10..0" để độ dài thông điệp chia hết cho b.

ECB (Electronic code book)

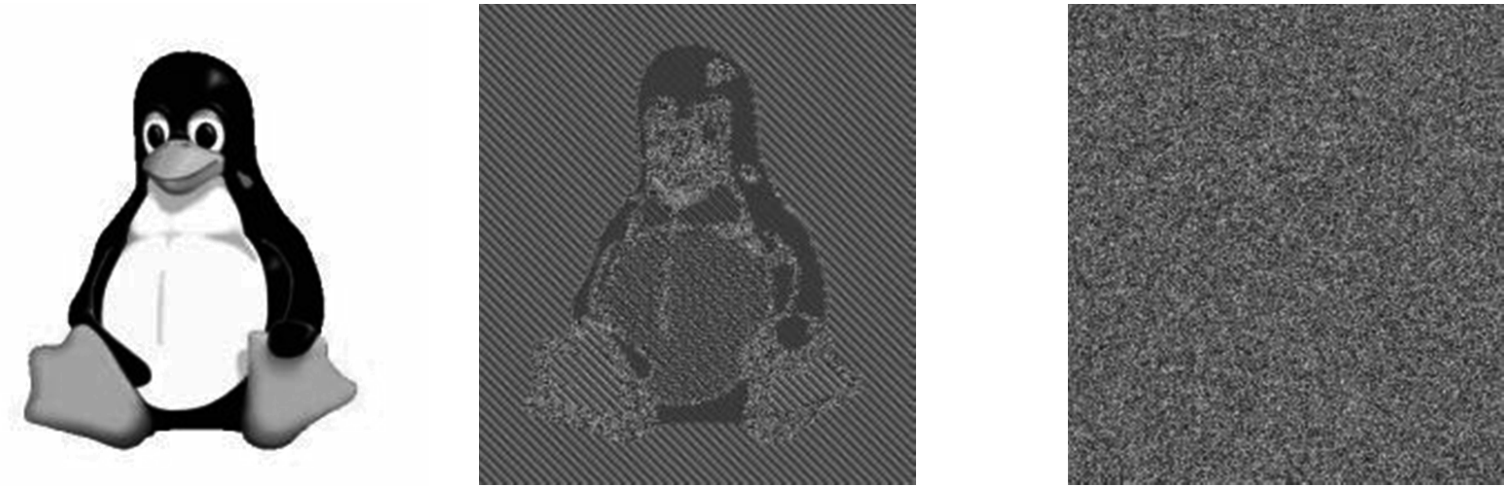


- Dữ liệu được chia thành các khối khối b bit, với $b =$ kích thước khối.
- Với dữ liệu không chia hết cho b bit: Thêm dãy "10..0" để độ dài thông điệp chia hết cho b.
- Phép toán padding này cho có tính khả nghịch. Nó cho phép giải mã.

ECB: Làm thế nào để giải mã?



ECB là không an toàn



Hình: Ảnh ở giữa là ECB mode, ảnh bên phải là mã hóa an toàn

- Vấn đề: Nếu $x_i = x_j$ thì $y_i = y_j$.
- ECB chỉ an toàn khi mã hóa dữ liệu ngẫu nhiên (ví dụ, các khóa).

Ví dụ: Chuyển tiền giữa hai ngân hàng dùng ECB

Block #	1	2	3	4	5
	Sending Bank A	Sending Account #	Receiving Bank B	Receiving Account #	Amount $

Hình: Giao thức trao đổi giữa các ngân hàng:

1. Giả sử: Mỗi trường đều là n-bit (ví dụ 128 bit)
2. Giả sử: Khoá k_{AB} để trao đổi thông tin giữa hai ngân hàng A và B là cố định.

Oscar tấn công

Block #	1	2	3	4	5
	Sending Bank A	Sending Account #	Receiving Bank B	Receiving Account #	Amount $

1. Oscar mở một tài khoản tại ngân hàng A và một tài khoản tại ngân hàng B;
2. Oscar chuyển nhiều lần 1$ từ tài khoản của anh ta ở ngân hàng A sang tài khoản ở ngân hàng B;
3. Oscar bắt gói tin trên đường truyền và nhận được các bản mã giống nhau

$$B_1 \| B_2 \| B_3 \| B_4 \| B_5$$

và anh ta giữ lại bản mã B_4.

4. Trong tương lai, mỗi khi thấy lệnh chuyển tiền từ B_1 tới B_3, thay block thứ 4 bởi B_4.

Nội dung

Bài toán

Ta cần giải quyết hai vấn đề:

- Làm cho hệ mã trở thành hệ mã xác suất;
- Ảnh hưởng của việc mã hoá trên mọi khối.

Mã hoá xác suất

- Mã hóa hai lần của cùng một thông điệp sẽ cho hai bản mã khác nhau
- Bản mã phải dài hơn bản rõ
- Nói một cách nôm na:

  Kích thước bản mã = Kích thước bản rõ + "số bit ngẫu nhiên"

Bài tập

Hãy viết hàm giải mã D cho hàm mã hoá E được định nghĩa bởi:

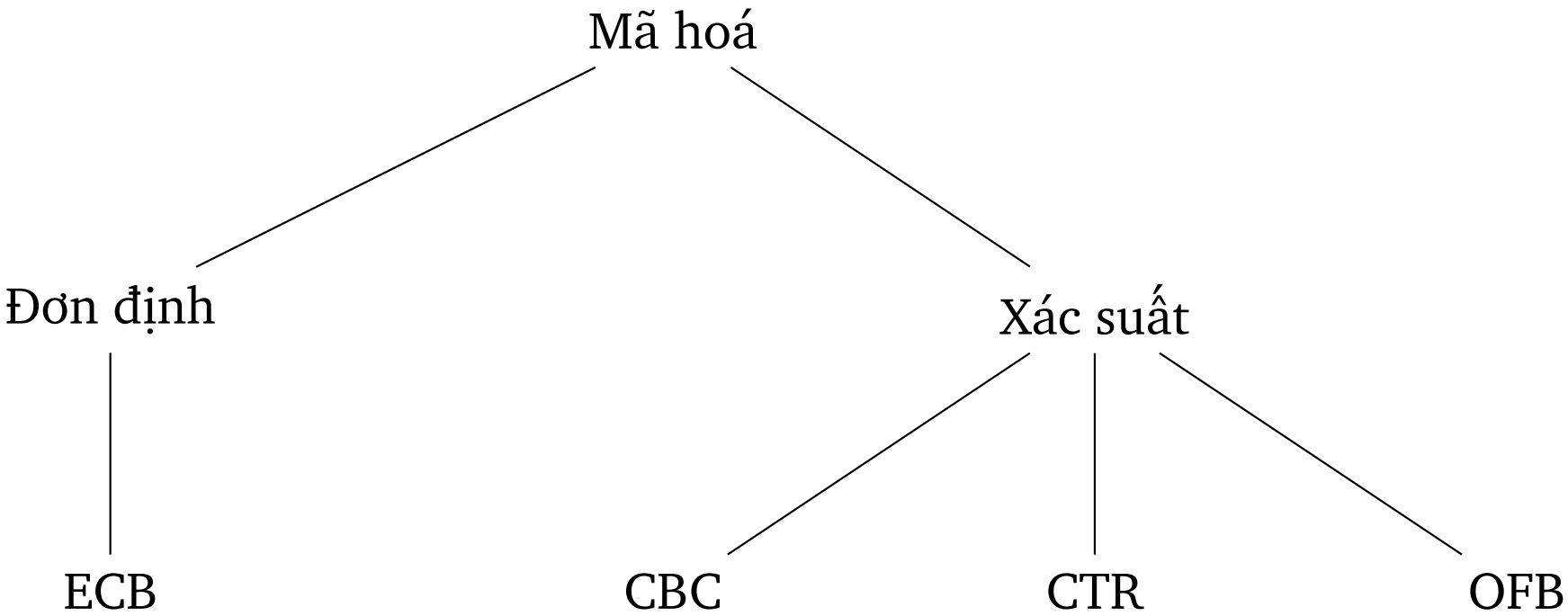
$E(k, m)$:

$r = \texttt{random}()$

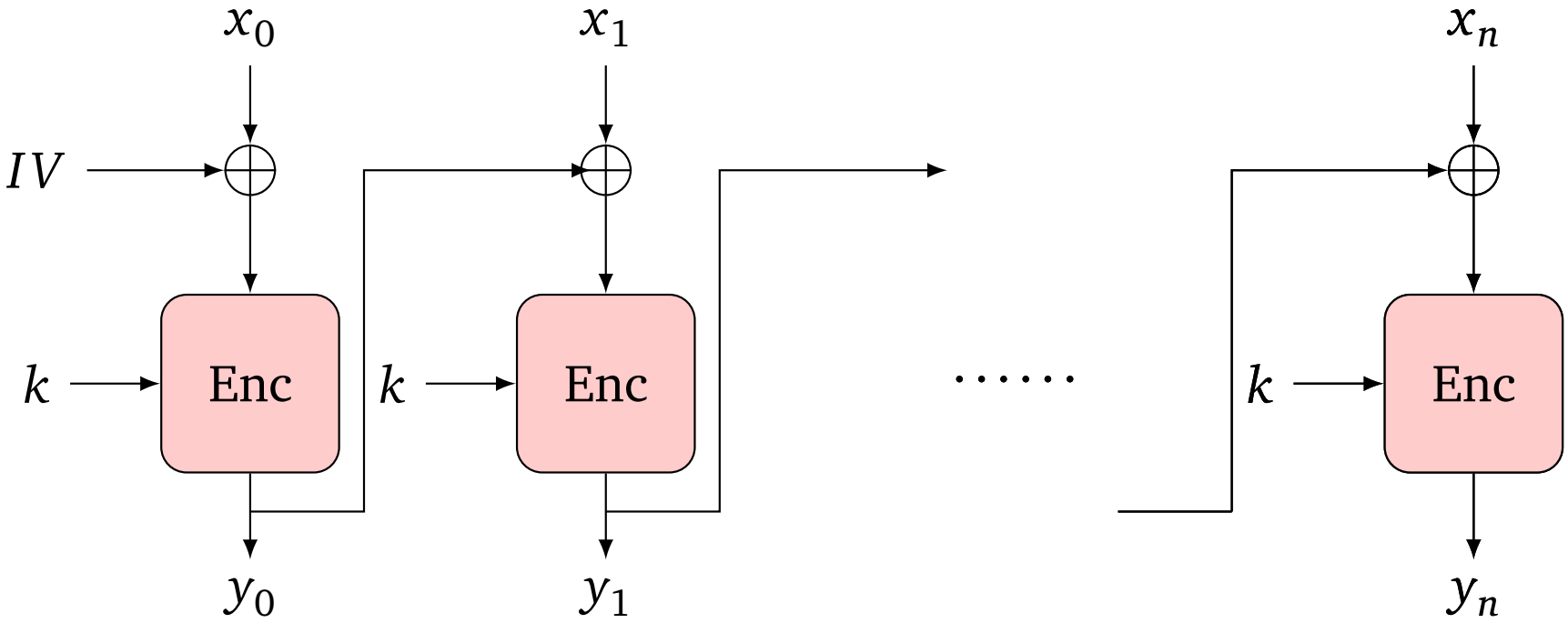
$c = \texttt{AES}(k, r) \oplus m$

$\texttt{return}\ (r, c)$

Dạng mã hoá



CBC (Cipher Block Chaining mode)



Thuật toán. Chọn IV ("initialization value") một cách ngẫu nhiên, sau đó dùng y_i như "IV" cho M_{i+1}. Gửi IV cùng với bản mã

$$IV \| y_0 \| y_1 \| \dots \| y_n$$

Sử dụng IV như thế nào?

- IV không cần giữ bí mật
- Nhưng phải là "nonce" = "number used only once"

Ví dụ

1. Là ngẫu nhiên "thật"
2. Là bộ đếm "counter" (phải được lưu trữ bởi Alice)
3. $\mathrm{ID}_A \| \mathrm{ID}_B \| \mathrm{time}$

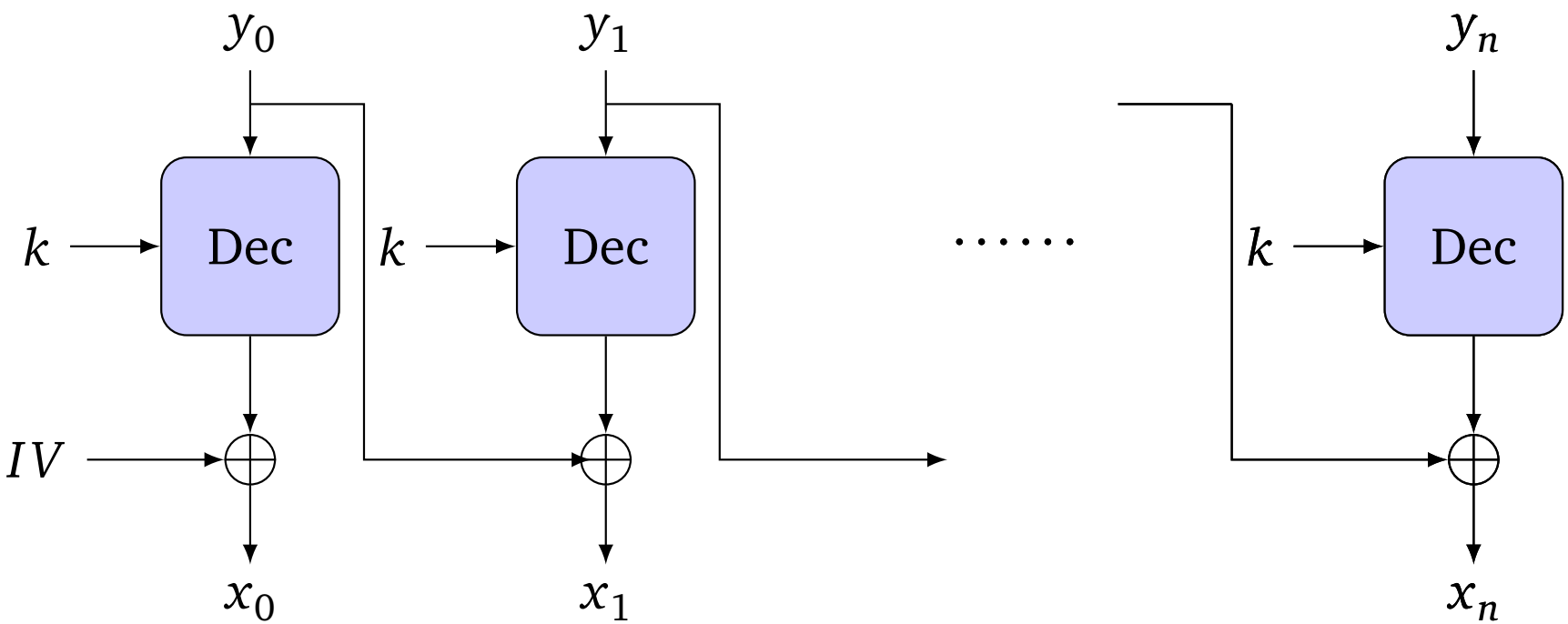
Một kỹ thuật padding cho CBC

- Padding theo từng byte;
- Giá trị mỗi byte được thêm là số byte cần được thêm.
  Ví dụ, nếu kích thước block là 8 và ta cần padding 4 byte:

  ```
  ... | DD DD DD DD 04 04 04 04 |
  ```

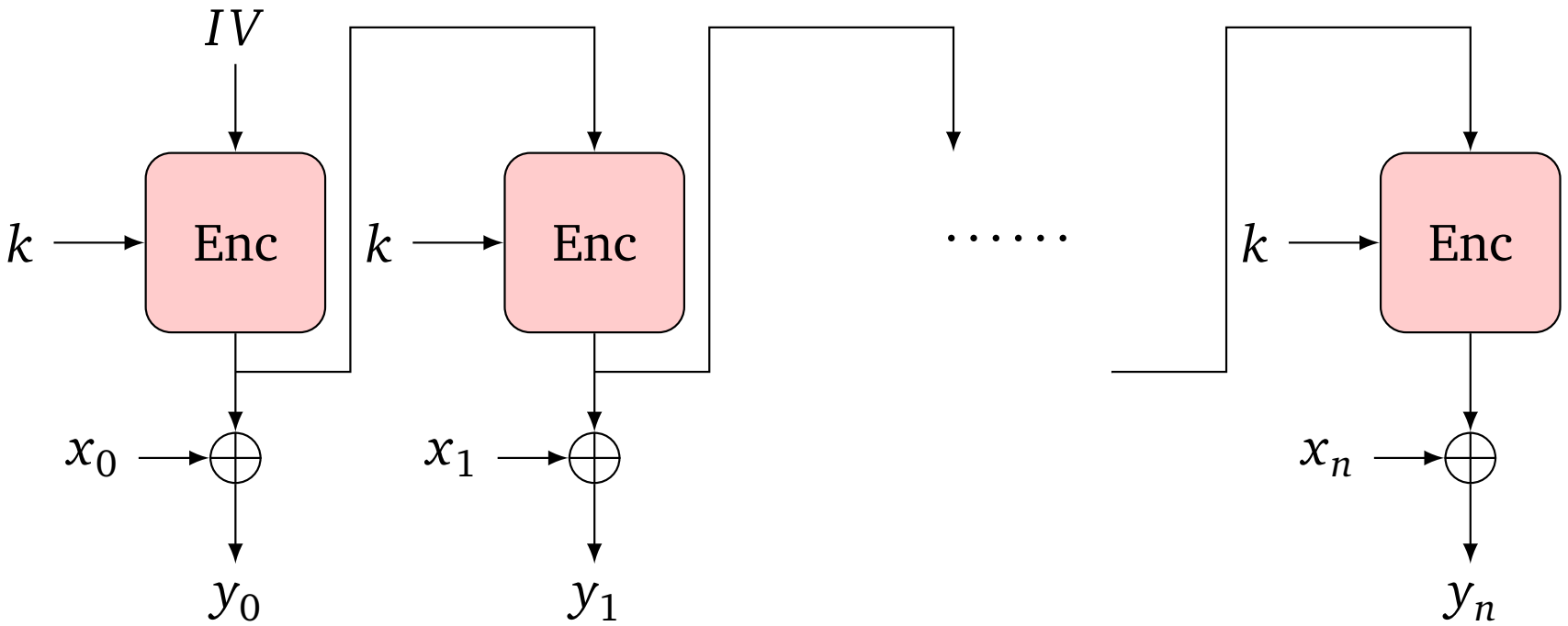
- nếu không cần padding, ta thêm một block giả.

CBC: Giải mã



Nội dung

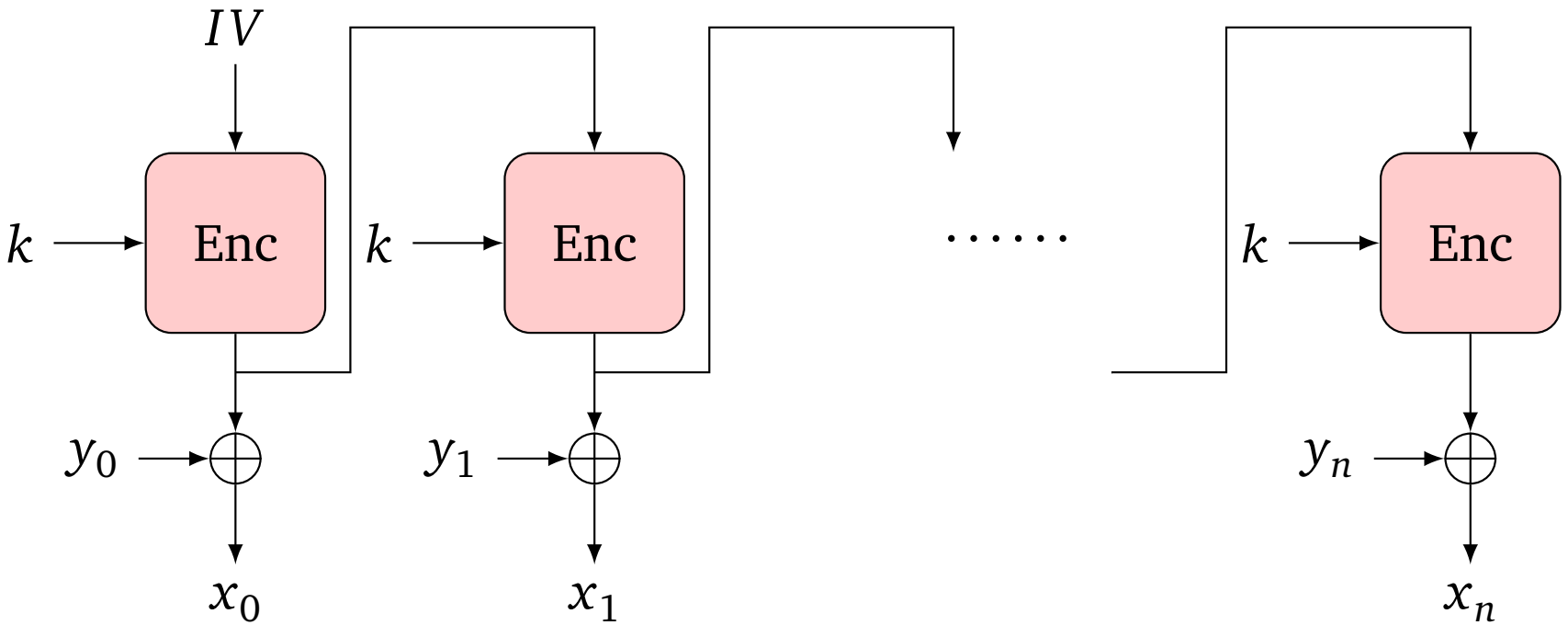
Output Feedback Mode (OFB)



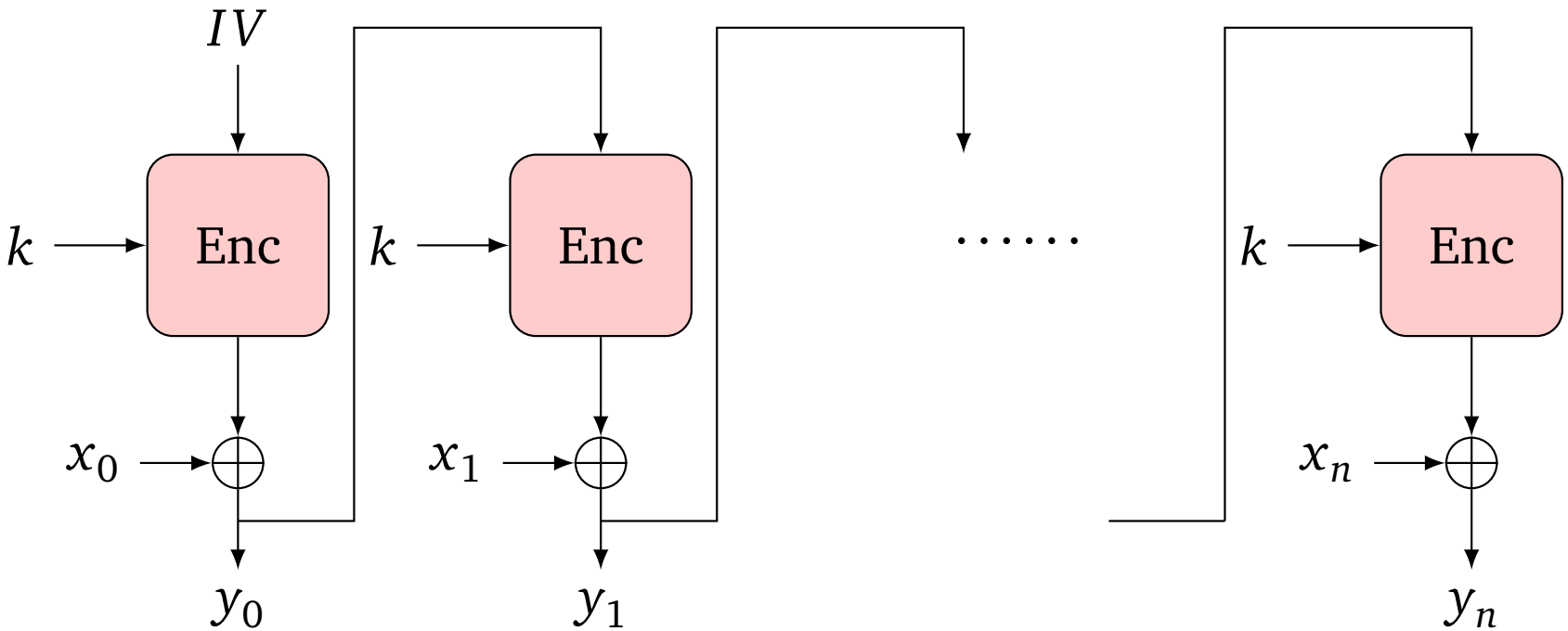
Thuật toán. Tương tự như CBC mode. Sử dụng IV ngẫu nhiên truyền cùng bản mã.

Nếu kích thước của bản rõ M không chia hết cho b, ta chỉ cần truyền bản mã rút gọn (không cần padding).

OFB: Giải mã

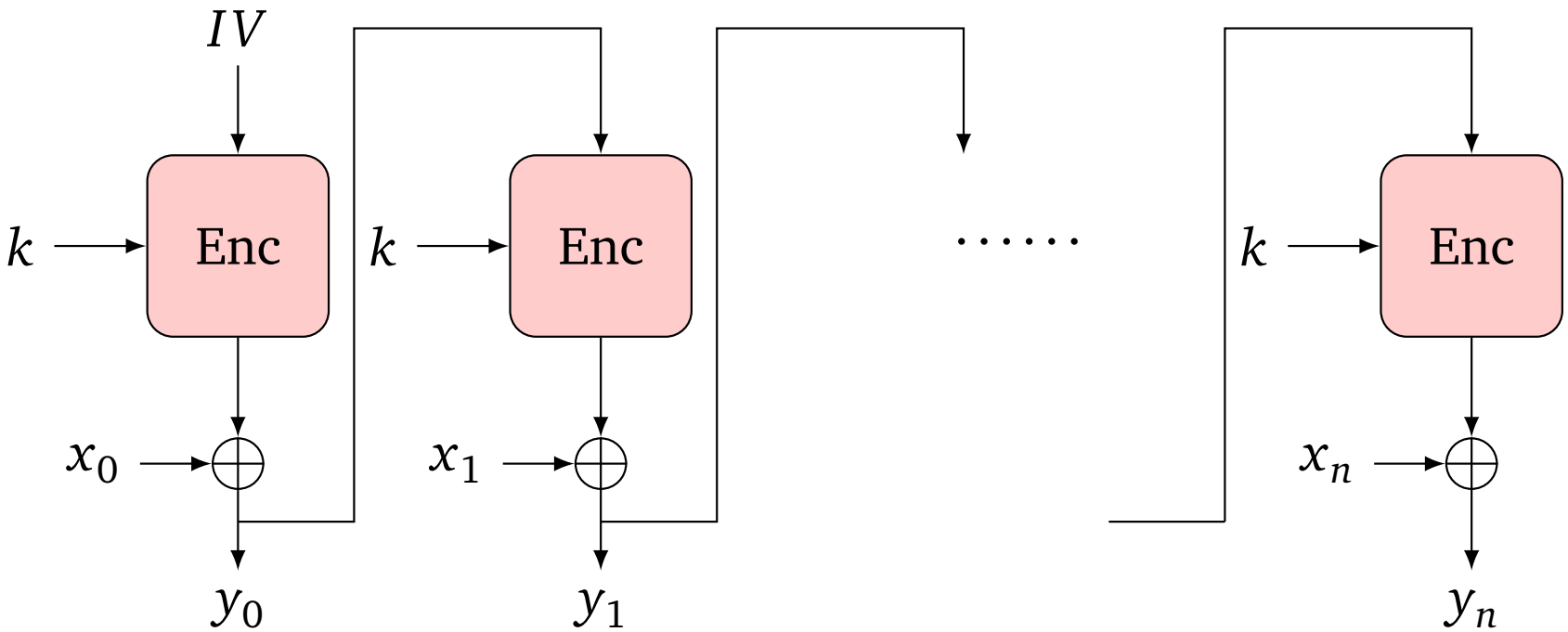


Cipher Feedback Mode (CFB)

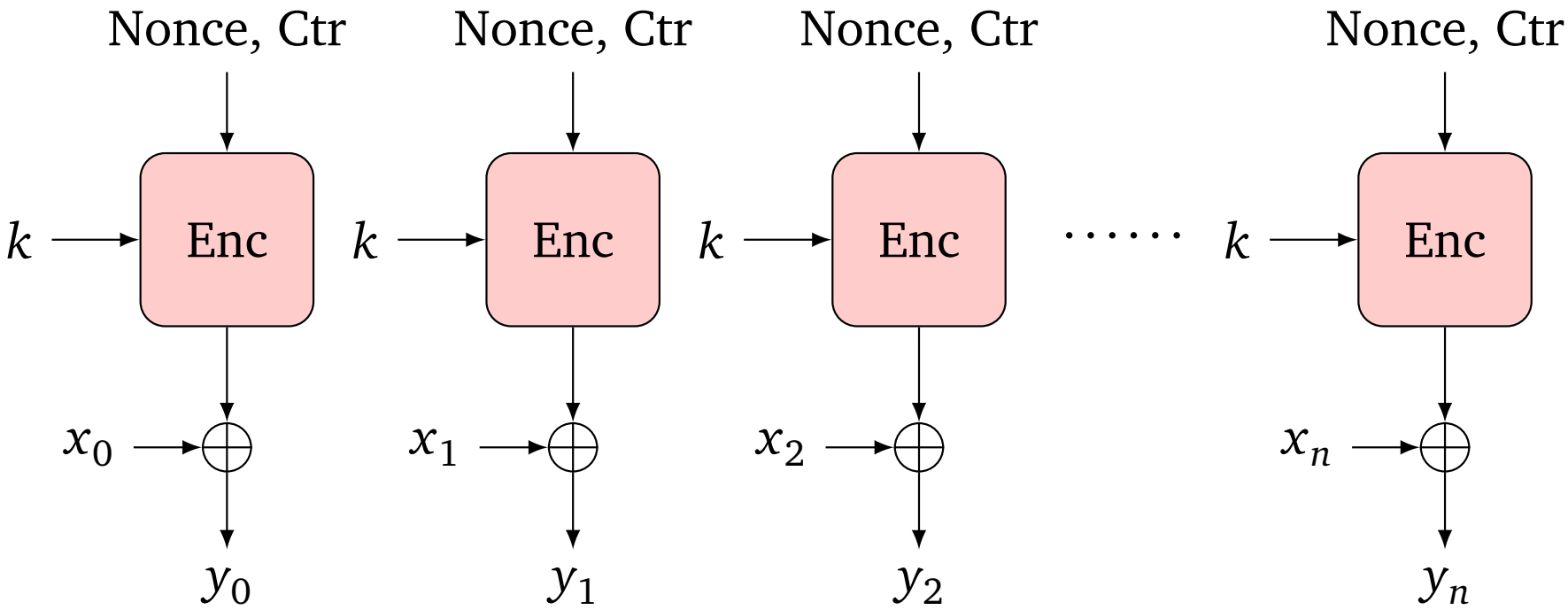


Bài tập

Hãy mô tả mạch giải mã CFB



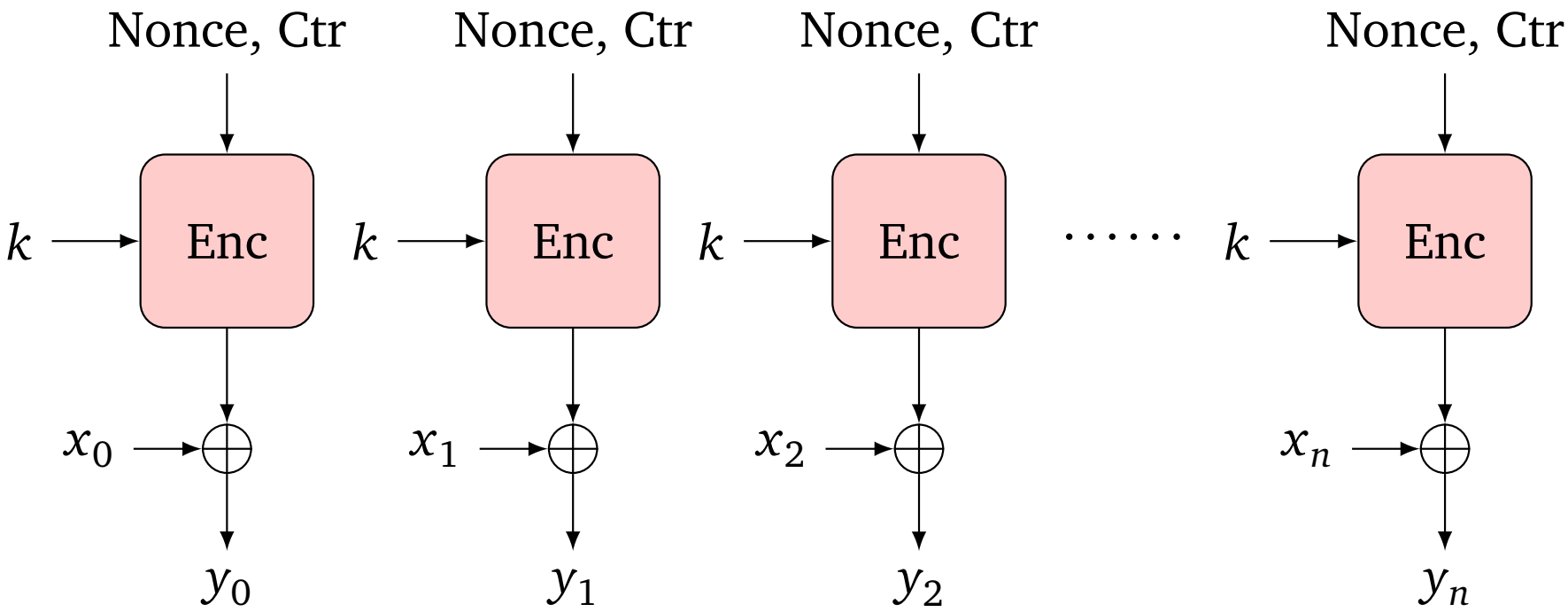
Counter Mode (CTR)



- Đảm bảo cặp Nonce||Ctr cặp không bao giờ lặp lại.
- Ctr được bắt đầu từ 0 cho mọi thông điệp.

Bài tập

Hãy mô tả mạch giải mã CTR



Bài tập

Xét thông điệp m gồm ℓ khối AES (ví dụ $\ell = 100$). Alice mã hóa m dùng CBC mode và truyền bản mã kết quả tới Bob. Do mạng lỗi, khối bản mã số $\ell/2$ bị mất trong khi truyền. Mọi bản mã khác được truyền và nhận đúng. Khi Bob giải mã bản mã nhận được, bao nhiêu khối bản rõ sẽ bị mất?

Bài tập

Xét thông điệp m bao gồm ℓ khối AES (ví dụ $\ell = 100$). Alice mã hóa m dùng randomized counter mode và truyền bản mã kết quả tới Bob. Do mạng lỗi, bản mã số $\ell/2$ bị mất trong khi truyền. Mọi khối bản mã khác được truyền và nhận đúng. Khi Bob giải mã bản mã nhận được, bao nhiêu khối bản rõ bị mất?

Nội dung

1. Electronic Codebook Mode (ECB)
2. Cipher Block Chaining Mode (CBC)
3. Mã dòng
4. Tính an toàn

Nên sử dụng mode nào?

Mục đích. Nếu hệ mã khối là không thể phân biệt với hệ mã khối lý tưởng, thì mode sử dụng nên đảm bảo tính không thể phân biệt dựa trên tấn công chọn bản mã:

- Định nghĩa trò chơi với kẻ tấn công.
- Mode là IND-CCA an toàn nếu kẻ tấn công có thể thắng trong trò chơi với xác suất nhiều nhất chỉ là $1/2 + \epsilon$ với ϵ là nhỏ "không đáng kể".

Thử nghiệm IND-CCA

Xét K là khóa được chọn ngẫu nhiên. E_K là hàm mã hóa với khóa K. D_K là hàm giải mã.

Pha I. ("Tìm kiếm")

- Kẻ tấn công có thể truy cập vào E_K, D_K như các hộp đen. (Có thể mã hóa/giải mã mọi thông điệp anh ta muốn)
- Kẻ tấn công đưa ra hai thông điệp M_0, M_1 cùng độ dài.

Pha II. ("Gợi ý")

- Ta bí mật chọn $d \leftarrow_{\$} \{0, 1\}$ và tính $Y = E_K(M_d)$.

Thử nghiệm IND-CCA

Xét K là khóa được chọn ngẫu nhiên. E_K là hàm mã hóa với khóa K. D_K là hàm giải mã.

Pha I. ("Tìm kiếm")

- Kẻ tấn công có thể truy cập vào E_K, D_K như các hộp đen. (Có thể mã hóa/giải mã mọi thông điệp anh ta muốn)
- Kẻ tấn công đưa ra hai thông điệp M_0, M_1 cùng độ dài.

Pha II. ("Gợi ý")

- Ta bí mật chọn $d \leftarrow_{\$} \{0,1\}$ và tính $Y = E_K(M_d)$.
- Kẻ tấn công nhận Y, và có thể tiếp tục truy cập vào E_K và D_K (ngoại trừ trên Y).

Thử nghiệm IND-CCA

Xét K là khóa được chọn ngẫu nhiên. E_K là hàm mã hóa với khóa K. D_K là hàm giải mã.

Pha I. ("Tìm kiếm")

- Kẻ tấn công có thể truy cập vào E_K, D_K như các hộp đen. (Có thể mã hóa/giải mã mọi thông điệp anh ta muốn)
- Kẻ tấn công đưa ra hai thông điệp M_0, M_1 cùng độ dài.

Pha II. ("Gợi ý")

- Ta bí mật chọn $d \leftarrow_{\$} \{0,1\}$ và tính $Y = E_K(M_d)$.
- Kẻ tấn công nhận Y, và có thể tiếp tục truy cập vào E_K và D_K (ngoại trừ trên Y).
- Kẻ tấn công tính toán và đưa ra d' là gợi ý cho d.

IND-CCA an toàn

Hệ mã gọi là an toàn chống lại tấn công CCA (hay IND-CCA) nếu trong thử nghiệm IND-CCA, lợi thế của kẻ tấn công

$$\text{Adv} = |\Pr(d = d') - 1/2|$$

là nhỏ "không đáng kể".

INC-CCA an toàn

Sự kiện. Để là IND-CCA an toàn, phương pháp mã hóa phải ngẫu nhiên. !

Ngược lại, kẻ tấn công có thể mã hóa M_0 và M_1 rồi so sánh với y.

Các mode đã biết là không an toàn!

- ECB. Không ngẫu nhiên.
- CTR. Giá trị Ctr bắt đầu là ngẫu nhiên, nhưng nó được truyền dưới dạng bản rõ. Trong trường hợp này, kẻ tấn công có thể yêu cầu giải mã một khúc đầu của Y, và anh ta được khúc đầu của M_d.

Các mode đã biết là không an toàn!

- ECB. Không ngẫu nhiên.
- CTR. Giá trị Ctr bắt đầu là ngẫu nhiên, nhưng nó được truyền dưới dạng bản rõ. Trong trường hợp này, kẻ tấn công có thể yêu cầu giải mã một khúc đầu của Y, và anh ta được khúc đầu của M_d.
- CBC. Tương tự CTR: IV ngẫu nhiên nhưng được truyền dưới dạng bản rõ. Kẻ tấn công có thể dùng kỹ thuật giải mã khúc đầu.

Các mode đã biết là không an toàn!

- ECB. Không ngẫu nhiên.
- CTR. Giá trị Ctr bắt đầu là ngẫu nhiên, nhưng nó được truyền dưới dạng bản rõ. Trong trường hợp này, kẻ tấn công có thể yêu cầu giải mã một khúc đầu của Y, và anh ta được khúc đầu của M_d.
- CBC. Tương tự CTR: IV ngẫu nhiên nhưng được truyền dưới dạng bản rõ. Kẻ tấn công có thể dùng kỹ thuật giải mã khúc đầu.
- OFB. Tương tự. Kẻ tấn công có thể sử dụng kỹ thuật giải mã khúc đầu.

IND-CCA an toàn

Định lý. Các mode ECB, CTR, CBC, OFB không phải IND-CCA an toàn.

Chứng minh.

- Kẻ tấn công chọn $M_0 = 0^x$ và $M_1 = 1^x$ với x lớn.
- Khi đó $Y = E_K(M_d)$.

IND-CCA an toàn

Định lý. Các mode ECB, CTR, CBC, OFB không phải IND-CCA an toàn.

Chứng minh.

- Kẻ tấn công chọn $M_0 = 0^x$ và $M_1 = 1^x$ với x lớn.
- Khi đó $Y = E_K(M_d)$.
- Xét $Z =$ nửa đầu của Y.

IND-CCA an toàn

Định lý. Các mode ECB, CTR, CBC, OFB không phải IND-CCA an toàn.

Chứng minh.

- Kẻ tấn công chọn $M_0 = 0^x$ và $M_1 = 1^x$ với x lớn.
- Khi đó $Y = E_K(M_d)$.
- Xét $Z =$ nửa đầu của Y.
- Vì $Y \neq Z$ nên kẻ tấn công được phép yêu cầu tính $D_K(Z)$ trong Pha II.

IND-CCA an toàn

Định lý. Các mode ECB, CTR, CBC, OFB không phải IND-CCA an toàn.

Chứng minh.

- Kẻ tấn công chọn $M_0 = 0^x$ và $M_1 = 1^x$ với x lớn.
- Khi đó $Y = E_K(M_d)$.
- Xét $Z =$ nửa đầu của Y.
- Vì $Y \neq Z$ nên kẻ tấn công được phép yêu cầu tính $D_K(Z)$ trong Pha II.
- Vậy nó cho phép tính được một nửa đầu của M_d.

IND-CCA an toàn

Định lý. Các mode ECB, CTR, CBC, OFB không phải IND-CCA an toàn.

Chứng minh.

- Kẻ tấn công chọn $M_0 = 0^x$ và $M_1 = 1^x$ với x lớn.
- Khi đó $Y = E_K(M_d)$.
- Xét $Z =$ nửa đầu của Y.
- Vì $Y \neq Z$ nên kẻ tấn công được phép yêu cầu tính $D_K(Z)$ trong Pha II.
- Vậy nó cho phép tính được một nửa đầu của M_d.
- Vậy kẻ tấn công luôn thắng.

Có thể xây dựng IND-CCA an toàn?

Trả lời: Có.

Yêu cầu khi xây dựng. Đưa ra bản mã Y của thông điệp M, kẻ tấn công không thể tạo được bản mã Z cho một thông điệp có liên quan.

Kỹ thuật. Sử dụng kết hợp tính bí mật và tính toàn vẹn thông điệp.

Hệ mã IND-CCA an toàn. Hệ mã có xác thực.
Ví dụ: Sơ đồ UFE = "Unbalanced Feistel Encryption" của Desai (Crypto 2006).

ĐẠI HỌC BÁCH KHOA

25 YEARS ANNIVERSARY SOICT

VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGY

soict.hust.edu.vn/ fb.com/groups/soict