COURSE TECHNOLOGY
CENGAGE Learning™

AN TOÀN VÀ BẢO MẬT THÔNG TIN

Đàm Quang Viễn

Chương 2

MÃ HÓA ĐỐI XỨNG CỔ ĐIỂN

NỘI DUNG

- Mô hình mã hóa đối xứng
- Các hệ mã hóa thay thế cổ điển
- Các kỹ thuật mã hóa hoán vị cổ điển

Mô hình mã hóa đối xứng

- Mật mã đối xứng sử dụng cùng một khóa cho việc mã hóa và giải mã. Có thể nói mã đối xứng là mã một khóa hay mã khóa riêng hay mã khoá thỏa thuận.
- E là hàm biến đổi bản rõ thành bản mã và D là hàm biến đổi bản mã trở về bản rõ
- Giả sử X là văn bản cần mã hóa và Y là dạng văn bản đã được thay đổi qua việc mã hóa

Y = EK(X) (*Encryption)*

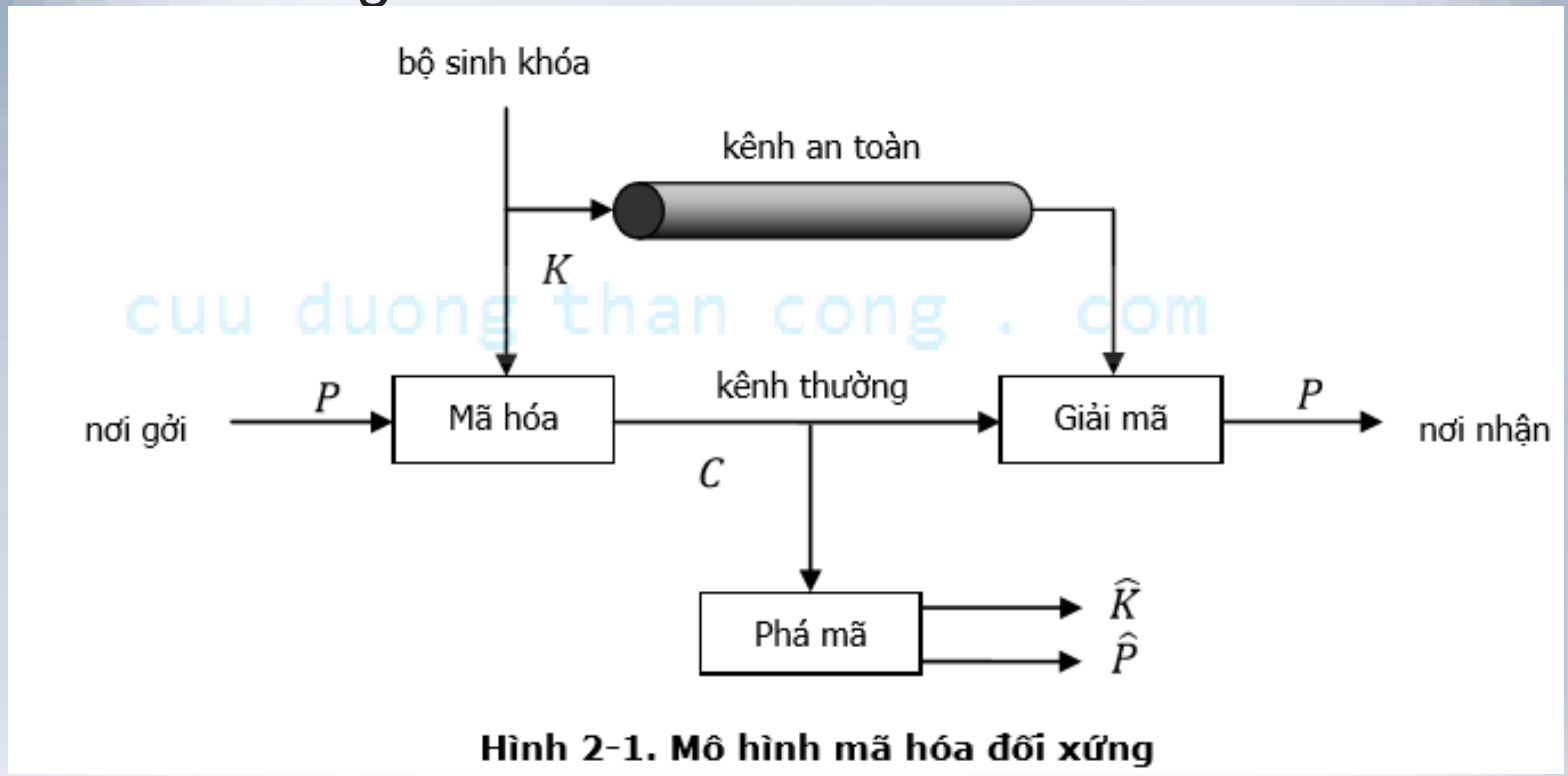
X = DK(Y) (*Decryption)*

Mô hình mã hóa đối xứng

- **Bản rõ X** được gọi là là bản tin gốc.
- **Bản mã Y** là bản tin gốc đã được mã hoá. thường xét mã hóa mà không làm thay đổi kích thước của bản rõ
- **Mã là thuật toán E** chuyển bản rõ thành bản mã. Thông thường chúng ta cần thuật toán mã hóa mạnh,
- **Khoá K** là thông tin tham số dùng để mã hoá, chỉ có người gửi và người nhận biết.
- **Mã hoá** là quá trình chuyển bản rõ thành bản mã,
- **Giải mã** chuyển bản mã thành bản rõ, đây là quá trình ngược lại của mã hóa.
- **Mật mã** là nghiên cứu về các nguyên lý và pp mã hoá

Mô hình mã hóa đối xứng (Symmetric Ciphers)

Mã hoá đối xứng sử dụng cùng một khoá cho cả hai quá trình mã hoá và giải mã.



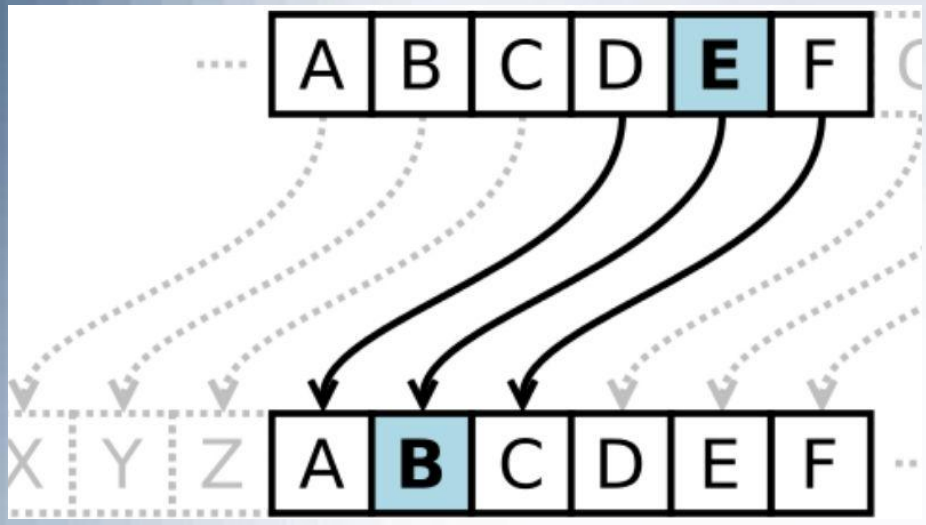
Hình 2-1. Mô hình mã hóa đối xứng

Mô hình mã hóa đối xứng (Symmetric Ciphers)

Quá trình thực hiện như sau:

Trong hệ thống mã hoá đối xứng, trước khi truyền dữ liệu, 2 bên gửi và nhận phải thoả thuận về khoá dùng chung cho quá trình mã hoá và giải mã. Sau đó, bên gửi sẽ mã hoá bản rõ (Plaintext) bằng cách sử dụng khoá bí mật này và gửi thông điệp đã mã hoá cho bên nhận. Bên nhận sau khi nhận được thông điệp đã mã hoá sẽ sử dụng chính khoá bí mật mà hai bên thoả thuận để giải mã và lấy lại bản rõ (Plaintext).

Mô hình mã hóa đối xứng



Mã hóa Ceasar

Thế kỷ thứ 3 trước công nguyên, nhà quân sự người La Mã Julius Ceasar đã nghĩ ra phương pháp mã hóa một bản tin như sau: thay thế mỗi chữ trong bản tin bằng chữ đứng sau nó k vị trí trong bảng chữ cái. Giả sử chọn k = 3, ta có bảng chuyển đổi như sau

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Mã hóa Ceasar

Bản gốc (bản rõ): Meet me after the toga party

Bản mã hóa (bản mã): phhw ph diwhu wkh wrjd sduwb

Phương pháp Ceasar được biểu diễn như sau:

với mỗi chữ cái p thay bằng chữ mã hóa C, trong đó:

$$C = (p + k) \bmod 26$$

Và quá trình giải mã đơn giản là:

$$p = (C - k) \bmod 26$$

Ví dụ

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

Với bản mã: **PHHW PH DIWHU WKH WRJD SDUWB**

Hãy dung Excel giải mã là phép cộng trừ modulo 26 cho tất cả 25 trường hợp của k

Mã hóa thay thế đơn bảng (Monoalphabetic Substitution Cipher)

- Phương pháp đơn bảng tổng quát hóa phương pháp Ceasar bằng cách dòng mã hóa không phải là một dịch chuyển k vị trí của các chữ cái A, B, C, … nữa mà là một hoán vị của 26 chữ cái này. Lúc này mỗi hoán vị được xem như là một khóa. Giả sử có hoán vị sau:
- Ban đầu: a b c d e f g h i j k l m n o p q r s t u v w x y z
  Khóa : z p b y j r s k f l x q n w v d h m g u t o i a e c
- Như vậy bản rõ meet me after the toga party
- mã hóa thành: njju nj zrujm ukj uvsz dzmue

Mã hóa thay thế đơn bảng (Monoalphabetic Substitution Cipher)

- Việc mã hóa được tiến hành bằng cách thay thế một chữ cái trong bản rõ thành một chữ cái trong bản mã, (phương pháp thay thế).
- Số lượng hoán vị của 26 chữ cái là 26!, đây cũng chính là số lượng khóa của phương pháp này.
- Vì 26! là một con số khá lớn nên việc tấn công phá mã vét cạn khóa là bất khả thi.
- Vì vậy mã này được xem là một phương pháp mã hóa an toàn trong suốt 1000 năm sau công nguyên

Mã hóa thay thế đơn bảng (Monoalphabetic Substitution Cipher)

- **Ví dụ.** Ta có bản mã tương ứng với bản rõ trong mã bảng chữ đơn như sau:

  Plain: ABCDEFGHIJKLMNOPQRSTUVWXYZ

  Cipher: DKVQFIBJWPESCXHTMYAUOLRGZN

  Plaintext: ifwewishtoreplaceletters

  Tìm bản mã hóa?

Mã hóa thay thế đơn bảng (Monoalphabetic Substitution Cipher)

- Tuy nhiên vào thế kỷ thứ 9, một nhà hiền triết người Ả Rập tên là Al-Kindi đã phát hiện ra một phương pháp phá mã khả thi khác.
- Trong ngôn ngữ tiếng Anh, tần suất sử dụng của các chữ cái không đều nhau, chữ E được sử dụng nhiều nhất, còn các chữ ít được sử dụng thường là Z, Q, J.
- Đối với cụm 2 chữ cái (digram), cụm chữ TH được sử dụng nhiều nhất.

Mã hóa thay thế đơn bảng (Monoalphabetic Substitution Cipher)

Mã Playfair

- Ở đây mỗi chữ có thể được mã bằng một trong 7 chữ khác nhau tùy vào chữ cặp đôi cùng nó trong bản rõ
- Ma trận khoá Playfair. Cho trước một từ làm khoá, với điều kiện trong từ khoá đó không có chữ cái nào bị lặp. Ta lập ma trận Playfair là ma trận cỡ 5 x 5 dựa trên từ khoá đã cho và gồm các chữ trên bảng chữ cái, được sắp xếp theo thứ tự như sau:

Mã hóa thay thế đơn bảng (Monoalphabetic Substitution Cipher)

Mã Playfair

- Trước hết viết các chữ của từ khoá vào các hàng của ma trận bắt từ hàng thứ nhất.
- Nếu ma trận còn trống, viết các chữ khác trên bảng chữ cái chưa được sử dụng vào các ô còn lại. Có thể viết theo một trình tự qui ước trước, chẳng hạn từ đầu bảng chữ cái cho đến cuối
- Vì có 26 chữ cái tiếng Anh, nên thiếu một ô. Thông thường ta dồn hai chữ nào đó vào một ô chung, chẳng hạn I và J

Mã hóa thay thế đơn bảng (Monoalphabetic Substitution Cipher)

Mã Playfair

– Giả sử sử dụng từ khoá MONARCHY. Lập ma trận khoá Playfair tương ứng nhưsau:

M	O	N	A	R
C	H	Y	B	D
E	F	G	I/J	K
L	P	Q	S	T
U	V	W	X	Z

Mã hóa thay thế đơn bảng (Monoalphabetic Substitution Cipher)

Mã hoá và giải mã: bản rõ được mã hoá 2 chữ cùng một lúc theo qui tắc như sau:

- Chia bản rõ thành từng cặp chữ. Nếu một cặp nào đó có hai chữ như nhau, thì ta chèn thêm một chữ lọc chẳng hạn X. Ví dụ, trước khi mã **“balloon”** biến đổi thành **“ba lx lo on”.**
- Nếu cả hai chữ trong cặp đều rơi vào cùng một hàng, thì mã mỗi chữ bằng chữ ở phía bên phải nó trong cùng hàng của ma trận khóa (cuộn vòng quanh từ cuối về đầu), chẳng hạn “**ar”** biến đổi thành “RM

Mã hóa thay thế đơn bảng (Monoalphabetic Substitution Cipher)

- Nếu cả hai chữ trong cặp đều rơi vào cùng một cột, thì mã mỗi chữ bằng chữ ở phía bên dưới nó trong cùng cột của ma trận khóa (cuộn vòng quanh từ cuối về đầu), chẳng hạn "**mu**" biến đổi thành "CM"
  - Trong các trường hợp khác, mỗi chữ trong cặp được mã bởi chữ cùng hàng với nó và cùng cột với chữ cùng cặp với nó trong ma trận khóa. Chẳng hạn, "**hs**" mã thành "**BP**", và "**ea**" mã thành "**IM**" hoặc "**JM**" (tuỳ theo sở thích)

Mã hóa thay thế đơn bảng (Monoalphabetic Substitution Cipher)

- An toàn của mã Playfair:
  - An toàn được nâng cao so hơn với bảng đơn, vì ta có tổng cộng 26 x 26 = 676 cặp. Mỗi chữ có thể được mã bằng 7 chữ khác nhau, nên tần suất các chữ trên bản mã khác tần suất của các chữ cái trên văn bản tiếng Anh nói chung.
  - Muốn sử dụng thống kê tần suất, cần phải có bảng tần suất của 676 cặp để thám mã (so với 26 của mã bảng đơn). Như vậy phải xem xét nhiều trường hợp hơn và tương ứng sẽ có thể có nhiều bản mã hơn cần lựa chọn. Do đó khó thám mã hơn mã trên bảng chữ đơn.

Các mã đa bảng
(A polygram substitution cipher)

Một hướng khác làm tăng độ an toàn cho mã trên bảng chữ là sử dụng nhiều bảng chữ để mã. Ta sẽ gọi chúng là các mã thế đa bảng. Ở đây mỗi chữ có thể được mã bằng bất kỳ chữ nào trong bản mã tùy thuộc vào ngữ cảnh khi mã hoá. Làm như vậy để trải bằng tần suất các chữ xuất hiện trong bản mã. Do đó làm mất bớt cấu trúc của bản rõ được thể hiện trên bản mã và làm cho thám mã đa bảng khó hơn

Mã Vigenere

Không gian khóa *K* được xác định như sau:
Với mỗi số nguyên dương M, khóa có độ dài M là một xâu ký tự có độ dài M , $K = k_1 k_2 \ldots k_M$.
Để mã hóa một bản rõ P người ta chia P thành các đoạn độ dài M và chuyển thành số thứ tự tương ứng của chúng trong bảng chữ cái , chẳng hạn

$X = x_1 x_2 \ldots x_M$.

Khi đó việc mã hóa và giải mã được thực hiện như sau:
$EK(X) = (x_1 + k_1, x_2 + k_2, \ldots, x_M + k_M) \bmod N$
$DK(Y) = (y_1 - k_1, y_2 - k_2, \ldots, y_M - k_M) \bmod N$ với N là số phần tử của bả ng chữ cái và $Y = y_1 y_2 \ldots y_M$ là bản mã

Mã Vigenere

Không gian khóa *K* được xác định như sau:
Với mỗi số nguyên dương M, khóa có độ dài M là một xâu ký tự có độ dài M , $K = k_1 k_2 \ldots k_M$.
Để mã hóa một bản rõ P người ta chia P thành các đoạn độ dài M và chuyển thành số thứ tự tương ứng của chúng trong bảng chữ cái , chẳng hạn

$X = x_1 x_2 \ldots x_M$.

Khi đó việc mã hóa và giải mã được thực hiện như sau:
$EK(X) = (x_1 + k_1, x_2 + k_2, \ldots, x_M + k_M) \bmod N$
$DK(Y) = (y_1 - k_1, y_2 - k_2, \ldots, y_M - k_M) \bmod N$ với N là số phần tử của bả ng chữ cái và $Y = y_1 y_2 \ldots y_M$ là bản mã

Mã Vigenere

Ví dụ: xét *A* là bảng chữ cái tiếng Anh , ta có N = 26 giả sử khóa có độ dài 6 và K =“CIPHER”,

bản rõ

P = “THIS CRYPTOSYSTEM IS NOT SECURE” .

Ta có K = 2 8 15 7 4 17,

P = 19 7 8 18 2 17 | 24 15 19 14 18 23 | 18 19 4 12 8 18 | 13 14 19 18 4 2 | 20 17 4.

Mã Vigenere

Quá trình mã hóa thực hiện như sau:

P = 19 7 8 18 2 17 | 24 15 19 14 18 23 | 18 19 4 12 8 18 | 13 14 19 18 4 2 | 20 17 4

K = 2 8 15 7 4 17 | 2 8 15 7 4 17 | 2 8 15 7 4 17 | 2 8 15 7 4 17 | 2 8 15

C = 21 15 23 25 6 8 | 0 23 8 21 22 14 | 20 1 19 19 12 9 | 15 22 8 25 8 19 | 22 25 19

Vậy bả n mã là C = “VPXZGI AXIVWO UBTTMJ PWIZIT WZT”.

Mã hoán vị (Permutation Cipher)

- Các chữ trong bản rõ không được thay thế bằng các chữ khác mà chỉ thay đổi vị trí, tức là việc mã hoá chỉ dịch chuyển vị trí tương đối giữa các chữ trong bản rõ.
- Như vậy, nó dấu bản rõ bằng cách thay đổi thứ tự các chữ, nó không thay đổi các chữ thực tế được dùng.
- Do đó bản mã có cùng phân bố tần suất xuất hiện các chữ như bản gốc

Mã hoán vị (Permutation Cipher)

- Mã Rail Fence
  - Viết các chữ của bản rõ theo đường chéo trên một số dòng
  - Sau đó đọc các chữ theo theo từng dòng sẽ nhận được bản mã
  - Vì khi biết số dòng ta sẽ tính được số chữ trên mỗi dòng và lại viết bản mã theo các dòng sau đó lấy bản rõ bằng cách viết lại theo các cột

Mã hoán vị (Permutation Cipher)

- Ví dụ. Viết bản tin “meet me after the toga party” lần lượt trên hai dòng như sau

  m e m a t r h t g p r y

  e t e f e t e o a a t

- Sau đó ghép các chữ ở dòng thứ nhất với các chữ ở dòng thứ hai cho bản mã:

  MEMATRHTGPRYETEFETEOAAT

Mã hoán vị (Permutation Cipher)

- Mã dịch chuyển
  - Viết các chữ của bản tin theo các dòng với số cột xác định
  - Sau đó thay đổi thứ tự các cột theo một dãy số khoá cho trước
  - Rồi đọc lại chúng theo các cột để nhận được bản mã
  - Quá trình giải mã được thực hiện ngược lại

Mã hoán vị (Permutation Cipher)

- Mã dịch chuyển
- Ví dụ:

Key:

4	3	1	2	5	6	7
A	T	T	A	C	K	P
O	S	T	P	O	N	E
D	U	N	T	I	L	T
W	O	A	M	X	Y	z

Ta đọc theo thứ tự các cột từ 1 đến 7 để nhận được bản mã:

Ciphertext: TTNAAPTMTSUOAODWCOIXKNLYPETZ

Mã hoán vị (Permutation Cipher)

Mã tích

- Mã dùng hoán vị hoặc dịch chuyển không an toàn vì các đặc trưng tần xuất của ngôn ngữ không thay đổi
- Có thể sử dụng một số mã liên tiếp nhau sẽ làm cho mã khó hơn
- Mã cổ điển chỉ sử dụng một trong hai phương pháp thay thế hoặc hoán vị
- Người ta nghĩ đến việc kết hợp cả hai phương pháp này trong cùng một mã và có thể sử dụng đan xen hoặc lặp nhiều vòng

Điểm yếu của mã cổ điển:

- Phương pháp mã hoá cổ điển có thể dễ dàng bị giải mã bằng cách đoán chữ dựa trên phương pháp thống kê tần xuất xuất hiện các chữ cái trên mã và so sánh với bảng thống kê quan sát của bản rõ.
- Để dùng được mã hoá cổ điển thì bên mã hoá và bên giải mã phải thống nhất với nhau về cơ chế mã hoá cũng như giải mã. Nếu không thì hai bên sẽ không thể làm việc được với nhau.

Summary

- Mô hình mã hóa đối xứng
- Các hệ mã hóa thay thế cổ điển
  - Mã Ceasar
  - Các mã bảng chữ đơn
  - Mã Playfair
  - Các mã đa bảng
  - Mã Vigenere
- Các kỹ thuật mã hóa hoán vị cổ điển
  - Mã Rail Fence
  - Mã dịch chuyển dòng
  - Mã tích

Câu hỏi ôn tập

1. Có bao nhiêu khóa Playfair khác nhau.
2. Giả sử dùng mã dịch chuyển dòng với 8 cột. Hỏi có bao nhiêu khóa khác nhau. Nêu thuật toán giải mã với từ khóa cho trước.
3. Hãy nêu tóm tắt các kỹ thuật mã hóa hoán vị cổ điển

Bài tập sinh viên

Bài tập 1: Cho biến đọan mã sau dùng mã Ceasar "GCUA VQ DTGCM"
Suy luận tìm bản rõ (sử dụng bảng chữ cái tiếng Anh).
Bài tập 2: Tìm bản mã của bản rõ “We are studying cryptography this year” sử dụng mã
Playfair với từ khóa “information technology”.

Bài tập 3: Cho hệ mã Vigenere có M = 6. Giải mã xâu C = “RANJLV” người ta thu được
bản rõ là “CIPHER”.
a) Tìm khóa đã sử dụng của hệ mã trên.
b) Dùng khóa tìm được ở phần trên hãy hãy giải mã xâu M = “PLDKCI DUJQJO“

Bài tập sinh viên (VN)

Bài tập 4: Sử dụng kỹ thuật thám mã bảng chữ đơn, lập bảng tần suất các chữ, bộ chữ đôi, bộ chữ ba của đọan mã sau:

UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIVUEPHZHMDZSH
ZOWSFPAPPDTSVPQUZWYMXUZUHSXEPYEPOPDZSZUFPOUDTMOHMQ

Lập luận và cho biết ánh xạ của bảng chữ đơn và đưa ra bản rõ phù hợp