

BÀI GIẢNG AN TOÀN & BẢO MẬT THÔNG TIN CHƯƠNG 1 - HỆ MÃ CỔ ĐIỂN

TS. NGUYỄN ĐÌNH DƯƠNG
BỘ MÔN KHMT - KHOA CÔNG NGHỆ THÔNG TIN

Email: duongnd@utc.edu.vn

Ngày 03/07/2022

Nội dung

Mã cổ điển

- 1.2 Một số thuật ngữ và kí hiệu
- 1.3 Mã thay thế
- 1.4 Mã hoán vị
- 1.5 Phụ lục: Phép toán modulo

Trao đổi

Nội dung

Mã cổ điển

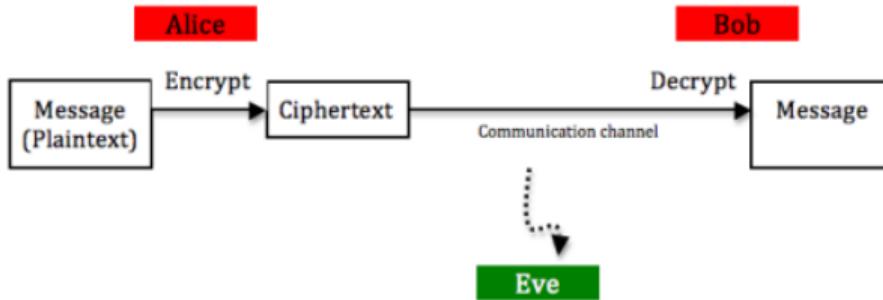
- 1.2 Một số thuật ngữ và kí hiệu
- 1.3 Mã thay thế
- 1.4 Mã hoán vị
- 1.5 Phụ lục: Phép toán modulo

Trao đổi



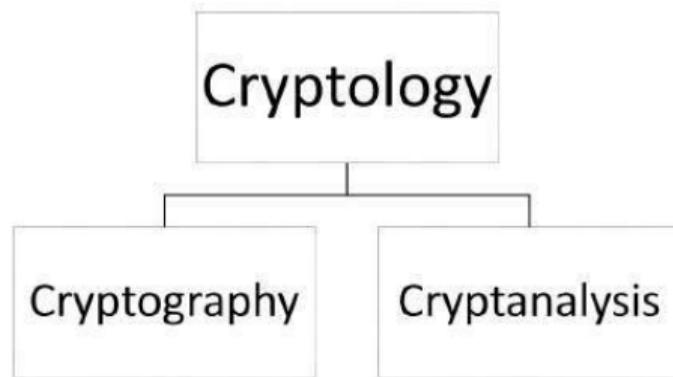
1. Mã cổ điển

1. 2. Một số thuật ngữ và kí hiệu



- Cryptography có nguồn gốc từ tiếng Hy Lạp
 - Crypto: Secret
 - Graphy: Writing
- Mục tiêu cơ bản nhất: gửi một tin nhắn mà không ai ngoài người nhận có thể đọc được
 - Làm được rất nhiều thứ khác trong ATBMTT: Bí mật, Nguyên vẹn, Xác thực, Không tì chối, Chống lặp lại
 - Encryption, Hash Function, MAC, Digital Signature

1. Mã cổ điển



1. 2. Một số thuật ngữ và kí hiệu

- Cryptography: method to send secret message using a code
- Cryptanalysis: trying to break the code and read those messages

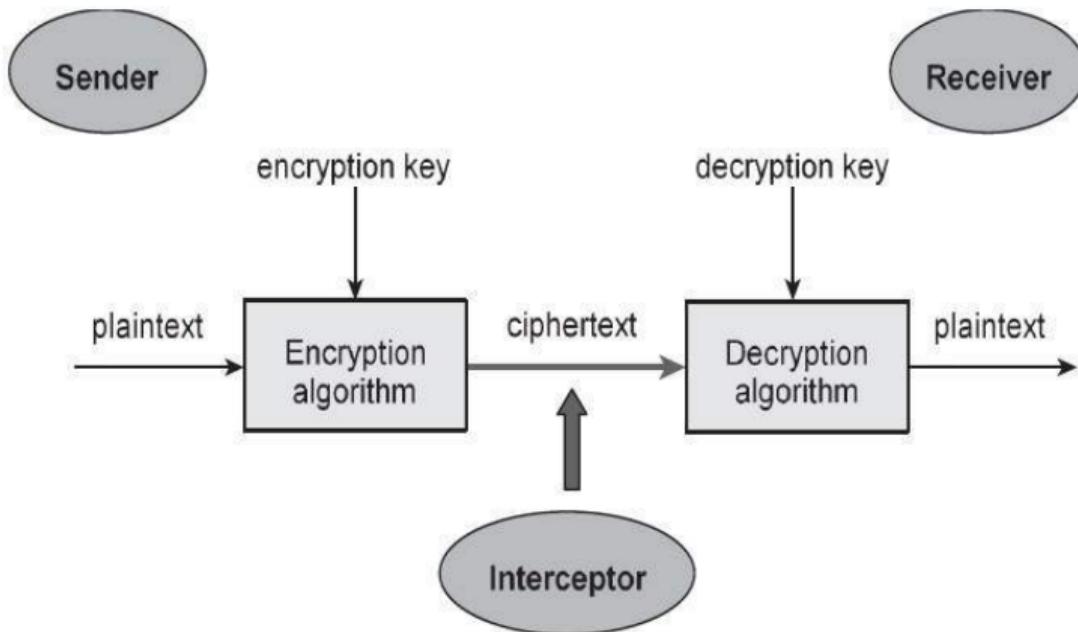
Primitives Service ↓	Encryption	Hash Function	MAC	Digital Signature
Confidentiality	Yes	No	No	No
Integrity	No	Sometimes	Yes	Yes
Authentication	No	No	Yes	Yes
Non Reputation	No	No	Sometimes	Yes



1. Mã cổ điển

1. 2. Một số thuật ngữ và kí hiệu

Cryptography



1. Mã cổ điển

1. 2. Một số thuật ngữ và kí hiệu

$$C = E(K, P); \quad P = D(K, C)$$

- **Plaintext (P)** = văn bản gốc (văn bản rõ)
- **Ciphertext (C)** = văn bản mã
- **Encryption** = chuyển plaintext thành ciphertext
- **Encryption algorithm** = phương pháp sử dụng để chuyển plaintext thành ciphertext.
- **Decryption** = chuyển ciphertext thành plaintext
- **Decryption algorithm** = phương pháp sử dụng để chuyển ciphertext thành plaintext.
- **Key (K)** = giá trị bí mật được sử dụng trong quá trình mã hoá và giải mã.

1. Mã cổ điển

1. 2. Một số thuật ngữ và kí hiệu

Liệu có an toàn?

- Làm thế nào để biết một hệ mã là "an toàn"?
 - đưa nhiều bản mã cho những người thật sự thông minh để giải mã (cryptanalysis)
 - nếu họ không giải mã → an toàn
- Điều này có thể sai !!!

1. Mã cổ điển

1. 2. Một số thuật ngữ và kí hiệu

Độ an toàn tính toán

Định nghĩa 1.1

Một hệ mã được gọi là an toàn tính toán nếu có một thuật toán tốt nhất để giải mã thì cần ít nhất N phép toán, với N là một số rất lớn nào đó.

- Thực tế, hệ mã là "an toàn tính toán" nếu có một thuật toán giải mã nhưng đòi hỏi nguồn lực máy tính và thời gian lớn đến mức không chấp nhận được
 - bài toán có **độ phức tạp hàm mũ**
 - bài toán thuộc lớp **bài toán có độ phức tạp NP**
- Một cách tiếp cận khác về "an toàn tính toán" là quy nó về bài toán được chứng minh là "rất khó"
 - bài toán "**phân tích ra thừa số nguyên tố của số n cho trước**"
 - bài toán "**logarit rời rạc**"

1. Mã cổ điển

1. 2. Một số thuật ngữ và kí hiệu

Phân loại hệ mật mã

Có **3** cách phân loại:

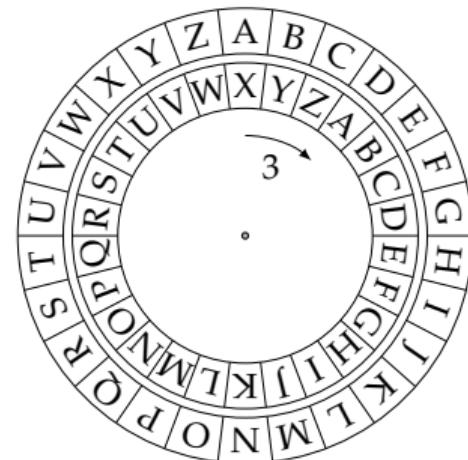
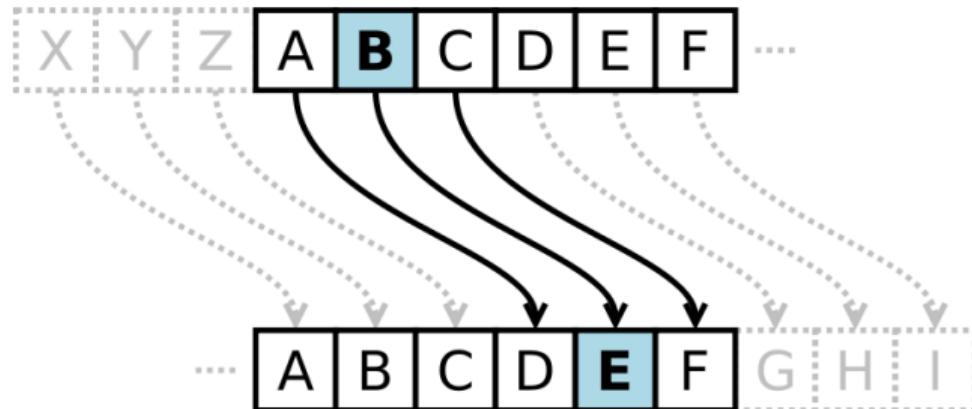
- **Kiểu của thao tác mã hóa được sử dụng trên bản rõ:**
 - Phép thế (Substitution): thay thế các ký tự trên bản rõ bằng các ký tự khác trên bản mã.
 - Hoán vị (Transposition): thay đổi vị trí các ký tự trong bản rõ, tức là thực hiện hoán vị các ký tự của bản rõ.
 - Tích (Product): kết hợp cả hai kiểu thay thế và hoán vị các ký tự của bản rõ.
- **Số khóa được sử dụng khi mã hóa và giải mã:**
 - Đôi xứng (Symmetric): một khóa duy nhất (khoá bí mật)
 - Không đối xứng (Asymmetric): hai khoá khác nhau (khoá công khai)
- **Cách mà bản rõ được xử lý:**
 - Khối (Block cipher): dữ liệu được chia thành từng khối có kích thước xác định và áp dụng thuật toán mã hóa với tham số khóa cho từng khối.
 - Dòng (Stream cipher): từng đơn vị thông tin đầu vào thường là bit hoặc byte được xử lý liên tục tạo phần tử đầu ra tương ứng.

1. Mã cổ điển

1. 3. Mã thay thế

Mã Caesar

- Tài liệu mã hoá sớm nhất được sử dụng bởi Ceasar trong quân sự
- Việc mã hoá đơn giản là **thay mỗi chữ trong bản rõ bằng chữ thứ ba tiếp theo** trong bảng chữ cái: A → D, P → S, ...



- “Meet me after the toga party” → “PHHW PH DIWHU WKH WRJD SDUWB”

1. Mã cổ điển

1. 3. Mã thay thế

Mã Caesar

- **Về toán học:** gán số thứ tự cho mỗi chữ trong bảng chữ cái (bắt đầu từ 0)

a	b	c	d	e	f	g	h	i	j	k	l	m
0	1	2	3	4	5	6	7	8	9	10	11	12
n	o	p	q	r	s	t	u	v	w	x	y	z
13	14	15	16	17	18	19	20	21	22	23	24	25

- **Mã dịch chuyển (Shift Cipher):**

$$\begin{cases} c = E(k, p) = (p + k) \mod 26 \\ p = D(k, c) = (c - k) \mod 26 \end{cases}, \quad \begin{cases} p \text{ là số thứ tự chữ cái trong bản rõ} \\ c \text{ là số thứ tự chữ t.ú trong bản mã} \\ k \text{ là khóa (số bước tịnh tiến các chữ).} \end{cases}$$

⊕ **Nhận xét:** $k = 3 \rightarrow$ mã Ceasar

1. Mã cổ điển

1. 3. Mã thay thế

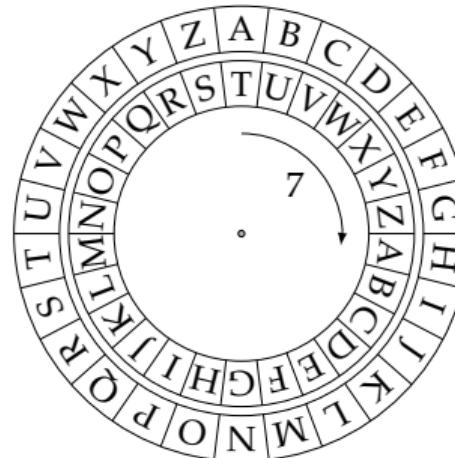
Mã dịch chuyển

$$\begin{cases} c = E(k, p) = (p + k) \mod 26 \\ p = D(k, c) = (c - k) \mod 26 \end{cases}, \quad \begin{cases} p \text{ là số thứ tự chữ cái trong bản rõ} \\ c \text{ là số thứ tự chữ t.ú trong bản mã} \\ k \text{ là khóa (số bước tịnh tiến các chữ.)} \end{cases}$$

- Giả sử chọn $k = 7$.
- Chữ A được mã thành H vì $0 + 7 \equiv 7 \pmod{26}$
- Chữ B được mã thành I vì $1 + 7 \equiv 8 \pmod{26}$
- Chữ H được mã thành O vì $7 + 7 \equiv 14 \pmod{26}$
- Chữ U được mã thành ???B

$\rightarrow 20 + 7 \equiv 1 \pmod{26}$

Ngày 03/07/2022



1. Mã cổ điển

1. 3. Mã thay thế

Giải mã Ceasar

- Làm thế nào để giải mã Ceasar?
- Vết cạn (brute-force): thử tất cả các giá trị của $k \in \{0, 1, 2, \dots, 25\} \rightarrow 26$ trường hợp
- Tính khả thi?
 - Easy by hand
 - Trivial by computer

1. Mã cổ điển

1. 3. Mã thay thế

Mã dịch chuyển

Luyện tập:

- ① Encode the message "MODULAR ARITHMETIC" using a Caesar Shift cipher with secret key $k = 7$.
- ② Decode the message "SLA AOLT LHA JHRL" using a Caesar Shift cipher with secret key $k = 7$.
- ③ Encode the message "ALL SQUARES ARE RECTANGLES" using a Caesar Shift cipher with secret key $k = 14$.
- ④ Decode the message "BCH OZZ FSQHOBUZSG OFS GEIOFSG" using a Caesar Shift cipher with secret key $k = 14$.
- ⑤ Find a partner to work with. Think of a secret message to send to them and encode it using a Caesar Shift cipher with a secret key of your choice. Give your partner the coded message and shift number. Decode your partner's message to you.

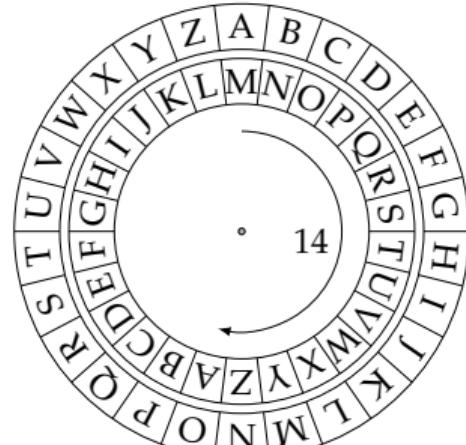
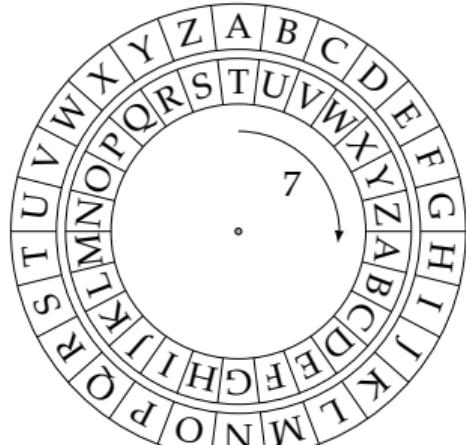
1. Mã cổ điển

1. 3. Mã thay thế

Mã dịch chuyển

Đáp số:

- ① TVKBSHY HYP AOT LAPJ
- ② LET THEM EAT CAKE
- ③ OZZ GEIOFSG OFS FSQHOBUZSG
- ④ BUT NOT ALL RECTANGLES ARE SQUARES



1. Mã cổ điển

1. 3. Mã thay thế

Bảng mã chữ đơn

- Mã Ceasar có **26** khoá khác nhau → việc thám mã khá đơn giản (sử dụng phương pháp Brute-Force).

Ví dụ 1.1

Cho bản mã "GCUA VQ DTGCM".

Bằng cách thử các phép tịnh tiến khác nhau, ta chọn được bước tịnh tiến thích hợp là 24 và cho bản rõ là "easy to break".

- Khắc phục:** mã hoá các chữ không chỉ là dịch chuyển bảng chữ, mà có thể tạo ra các **bước nhảy khác nhau** cho các chữ.

1. Mã cổ điển

1. 3. Mã thay thế

a	b	c	d	e	f	g	h	i	j	k	l	m
O	Y	C	P	K	G	V	W	B	Q	U	Z	J
n	o	p	q	r	s	t	u	v	w	x	y	z
X	E	N	M	H	T	D	S	I	L	F	R	A

- mỗi cách mã sẽ tương ứng với một hoán vị của bảng chữ và **hoán vị đó chính là khoá** của mã đã cho.
- Vết cạn: số khoá có thể có là $26!$ $\approx 4 \times 10^{26}$ → rất khó thực hiện bằng tay ...

Ví dụ 1.2

- Using the Substitution Cipher above, encrypt the message "I WANT COOKIES".
- Using the Substitution Cipher above, decrypt the message "BX DWK CEEUBK QOH".

1. Mã cổ điển

1. 3. Mã thay thế

Thám mã Bảng mã chữ đơn

- Liệu mã trên bảng chữ đơn sẽ an toàn? **Không, do đặc trưng ngôn ngữ!!!**
QYYQAV QY WQID
- Để ý rằng, trong mã thay thế, đặc trưng ngôn ngữ vẫn **giữ nguyên**
 - you can tell how often a letter occurs in a message
 - you can see when letters repeat
 - ...
- Sử dụng phương pháp **phân tích tần suất**

1. Mã cổ điển

1. 3. Mã thay thế

Phân tích tần suất

- Kỹ thuật thám mã này được đưa ra trong thế kỉ 9 bởi Al-Kindi ở Iraq
- Trong các ngôn ngữ, không phải tất cả các chữ trong một bảng chữ cái xuất hiện với **tần suất giống nhau**
- Mỗi ngôn ngữ đều có một số đặc trưng riêng
- Trong tiếng Anh
 - E is most common
 - Vowels are about 40%
 - Vowels tends to be separated by consonants
 - Q tends to be followed by U
 - ...

1. Mã cổ điển

1. 3. Mã thay thế

Phân tích tần suất

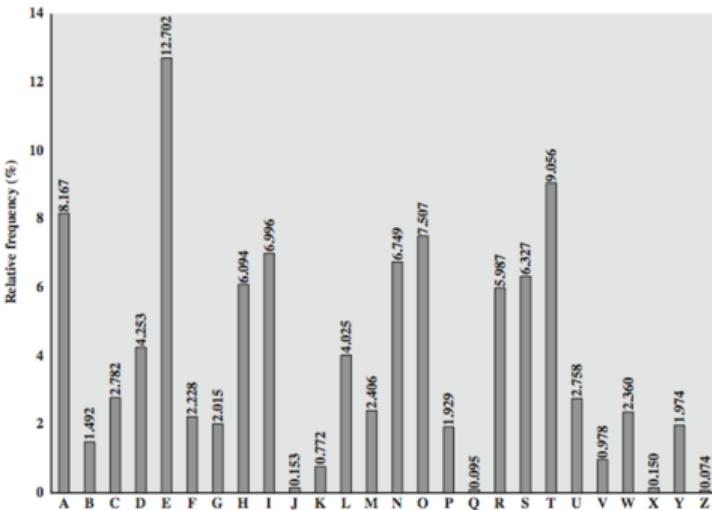
- Trong tiếng Anh, có một số chữ hoặc các cặp chữ hoặc bộ ba chữ được dùng thường xuyên hơn các bộ chữ cùng độ dài khác: "th lrd s m shphrd shll nt wnt"
- chữ E được sử dụng nhiều nhất, sau đó đến các chữ T, R, N, I, O, A, S; Một số chữ rất ít dùng như: Z, J, K, Q, X.
- Bằng phương pháp thống kê, ta có thể xây dựng các bảng các tần suất các chữ đơn, cặp chữ, bộ ba chữ.

1. Mã cổ điển

1. 3. Mã thay thế

① Chữ đơn:

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



② Cặp: Th, he, in, an, re, ed, on, es, st, en at, to

③ Bộ ba: The, ing, and, hex, ent, tha, nth, was eth, for, dth.

Ví dụ 1.3

Cho bản mã:

UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZVU
EPHZHMDZSHZOWSFAPPDTSVPQUZWYMXUZUHSXEPLYEPOPDZ
SZUFPOMBZWPFUPZHMDJUDTMOHMQ

- Bảng tần suất:

P 13.33	O 7.50	E 5.00	W 3.33	B 1.67	J 0.83	N 0.00
Z 11.67	M 6.67	V 4.17	Q 2.50	G 1.67	C 0.00	R 0.00
S 8.33	H 5.83	X 4.17	T 2.50	Y 1.67	K 0.00	
U 8.33	D 5.00	F 3.33	A 1.67	I 0.83	L 0.00	

- Đoán P và Z là e và t.
- Khi đó ZW là th và ZWP là the, ...

"it was disclosed yesterday that several informal but direct contacts have been made with political representatives in moscow"

1. Mã cổ điển

1. 3. Mã thay thế

Mã Playfair

⊕ **Nhận xét:** Mã bảng chữ đơn có số khoá lớn (26!) nhưng chưa đảm bảo an toàn.

 Mỗi chữ sẽ được mã bằng một số chữ khác nhau tùy thuộc vào các chữ mà nó đứng cạnh
→ **Playfair:**

- Là hệ mã đối xứng
- Được sáng tạo bởi Charles Wheatstone vào năm 1854 và mang tên người bạn là Baron Playfair.
- Mỗi chữ có thể được mã bằng một trong 7 chữ khác nhau tùy vào chữ cặp đôi cùng nó trong bản rõ.

1. Mã cổ điển

1. 3. Mã thay thế

Mã Playfair

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

①

Tạo ma trận khoá 5×5 : Chọn trước một từ làm khoá (không có chữ cái nào bị lặp), ví dụ MORNACHY

- Viết các chữ của từ khoá vào các hàng của ma trận bắt từ hàng thứ nhất.
- Viết các chữ khác trên bảng chữ cái chưa được sử dụng vào các ô còn lại.
- Vì có 26 chữ cái tiếng Anh, nên thiếu một ô → dồn hai chữ I và J vào 1 ô.

Mã Playfair

② Mã hoá:

- 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, "balloon" → "ba lx lo on".
- Nếu cả hai chữ trong cặp ở cùng một hàng, thì mã mỗi chữ bởi **chữ ở phía bên phải nó trong cùng hàng** của ma trận khóa (cuộn vòng), "st" → "TL".
- Nếu cả hai chữ trong cặp ở cùng một cột, thì mã mỗi chữ bởi **chữ ở phía bên dưới nó trong cùng cột** của ma trận khóa (cuộn vòng), "me" → "CL".
- Còn lại, 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, "nt" → "RQ", "ea" → "IM" hoặc "JM".

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

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

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

Ví dụ 1.4

Plaintext: "instrumentsz"

Plaintext =	in	st	ru	me	nt	sz
Ciphertext =	ga	tl	mz	cl	rq	tx

in:	M O N A R C H Y B D E F G I K L P Q S T U V W X Z	st:	M O N A R C H Y B D E F G I K L P Q S T U V W X Z	ru:	M O N A R C H Y B D E F G I K L P Q S T U V W X Z
me:	M O N A R C H Y B D E F G I K L P Q S T U V W X Z	nt:	M O N A R C H Y B D E F G I K L P Q S T U V W X Z	sz:	M O N A R C H Y B D E F G I K L P Q S T U V W X Z

Ví dụ 1.5

Ciphertext: "gatlmzclrqtx"

Ciphertext =	ga	tl	mz	cl	rq	tx
Plaintext =	in	st	ru	me	nt	sz

in:	M O N A R C H Y B D E F G I K L P Q S T U V W X Z	st:	M O N A R C H Y B D E F G I K L P Q S T U V W X Z	ru:	M O N A R C H Y B D E F G I K L P Q S T U V W X Z
me:	M O N A R C H Y B D E F G I K L P Q S T U V W X Z	nt:	M O N A R C H Y B D E F G I K L P Q S T U V W X Z	sz:	M O N A R C H Y B D E F G I K L P Q S T U V W X Z

1. Mã cổ điển

1. 3. Mã thay thế

Mã Playfair

• **Ưu điểm:**

- Khó thám mã hơn so với bảng đơn vì mỗi chữ có thể được mã bằng 7 chữ khác nhau → tần suất các chữ trên bản mã khác tần suất của các chữ cái trên văn bản tiếng Anh.
- Muốn sử dụng thống kê tần suất, cần phải có bảng tần suất của $25 \times 25 = 625$ cặp để thám mã (so với 26 của mã bảng đơn).

• **Nhược điểm:**

- Mỗi cặp chữ trong bản mã (AB) và ngược lại (BA) sẽ có các bản rõ tương ứng như UR và RU (và ngược lại) → có thể bị khai thác nếu biết được ngôn ngữ của bản rõ và bảng phân tần số.
- Sử dụng cùng một khoá trong mã hoá và giải mã.

1. Mã cổ điển

1. 3. Mã thay thế

Mã Vigenere

- Để "trải bằng tần suất" các chữ xuất hiện trong bản mã và làm mất bớt cấu trúc của bản rõ được thể hiện trên bản mã → sử dụng nhiều bảng chữ để mã → **mã đa bảng Vigenere**.
- Mã hoá Vigenere là tiến hành nhiều mã Ceasar cùng một lúc trên bản rõ với nhiều khoá khác nhau.
- Giả sử khoá là một chữ có độ dài d được viết dạng $K = K_1K_2 \dots K_d$, trong đó K_i nhận giá trị nguyên từ 0 đến 25.
 - Chia bản rõ thành các khối gồm d chữ.
 - Mỗi chữ thứ i trong khối dùng bảng chữ với tịnh tiến là K_i giống như mã Ceasar.
 - Giải mã là quá trình làm ngược lại: mỗi chữ sử dụng bước nhảy lui lại về đầu.

1. Mã cổ điển

1. 3. Mã thay thế

Mã Vigenere

Ví dụ 1.6

- **Key:** "deceptive"
- **Plaintext:** we are discovered save yourself

key:	deceptive	deceptive	deceptive
plaintext:	wearedisc	overedsav	eyourself
ciphertext:	ZICVTWQNG	RZGVTWAVZ	???

HCQYGLMGJ

- Để mã chữ **w** đầu tiên ta tìm chữ đầu của khóa là **d** ⇒ w sẽ được mã trên bảng chữ tịnh tiến 3 (a tịnh tiến đến d): w → Z.
- Chữ thứ hai trong từ khóa là **e**, có nghĩa là chữ thứ hai trong bản rõ sẽ được tịnh tiến 4 (tịnh tiến 1) (tịnh tiến 2) (tịnh tiến 3) (tịnh tiến 4) ...

	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
0	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
1	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
2	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
3	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
4	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
5	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
6	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
7	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
8	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
9	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
10	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
11	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
12	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
13	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
14	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
15	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
16	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
17	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
18	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
19	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
20	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
21	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
22	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
23	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
24	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
25	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

1. Mã cổ điển

1. 3. Mã thay thế

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

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
26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51

Ví dụ 1.7

- Mã hoá Vigenere: ATTACK AT DAWN
- Key: MONKEY
- Kết quả: MHGKGI MH QKAL

1. Mã cổ điển

1. 3. Mã thay thế

An toàn của mã Vigenere

- Cùng một chữ của bản rõ có thể có nhiều mã khác nhau → tần suất của các chữ bị "là phẳng".
- Tuy nhiên do độ dài của khoá có hạn nên có thể tạo nên chu kỳ vòng lặp → cần tăng độ dài từ khoá để tăng số bảng chữ dùng khi mã.
- Lý tưởng nhất là khoá có độ dài bằng bản rõ → **mã khoá tự động**.

Ví dụ 1.8

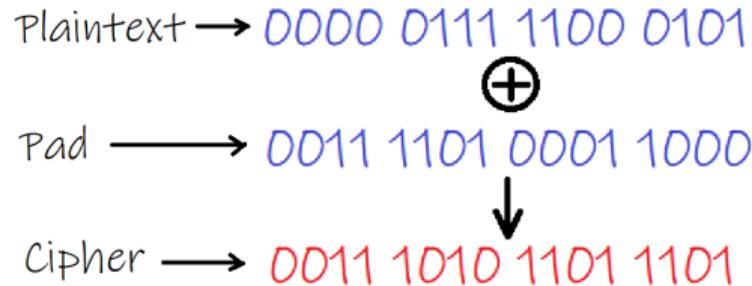
key:	deceptivewearediscoveredsav
plaintext:	wearediscoveredsaveyourself
ciphertext:	ZICVTWQNGKZEIIGASXSTSLVVWLA

1. Mã cổ điển

1. 3. Mã thay thế

Bộ đệm 1 lần (One-Time Pad)

- Với OTP, khóa:
 - *được chọn hoàn toàn ngẫu nhiên*
 - **có độ dài bằng bản rõ**
 - *chỉ được sử dụng 1 lần*
- Việc mã hóa (giải mã) được thực hiện bằng phép toán **XOR** từng bit giữa các bit có vị trí tương ứng ở bản rõ (bản mã) và khóa.
- **Ưu điểm:** mã OTP an toàn tuyệt đối (theo định lý Clause Shannon); tốc độ tính toán nhanh.
- **Nhược điểm:** việc sinh ngẫu nhiên khóa và phân phối khóa → OTP ít được sử dụng và chỉ dùng trong trường hợp đòi hỏi bảo mật rất cao.



1. Mã cổ điển

1. 3. Mã thay thế

One-time Pad: Encryption

e=000 h=001 i=010 k=011 l=100 r=101 s=110 t=111

Encryption: Plaintext \oplus Key = Ciphertext

	h	e	i	l	h	i	t	l	e	r
Plaintext:	001	000	010	100	001	010	111	100	000	101
Key:	111	101	110	101	111	100	000	101	110	000
Ciphertext:	110	101	100	001	110	110	111	001	110	101

	s	r	l	h	s	s	t	h	s	r
	001	000	010	100	001	010	111	100	000	101

One-Time Pad: Decryption

e=000 h=001 i=010 k=011 l=100 r=101 s=110 t=111

Decryption: Ciphertext \oplus Key = Plaintext

	s	r	l	h	s	s	t	h	s	r
Ciphertext:	110	101	100	001	110	110	111	001	110	101
Key:	111	101	110	101	111	100	000	101	110	000
Plaintext:	001	000	010	100	001	010	111	100	000	101

	h	e	i	l	h	i	t	l	e	r
	001	000	010	100	001	010	111	100	000	101

- Chuyển dữ liệu sang dạng nhị phân (plaintext).
- Sinh ngẫu nhiên một mảng dữ liệu nhị phân với chiều dài bằng chiều dài của plaintext (đây là pad).
- XOR từng bit trong plaintext với bit ở vị trí tương ứng trong pad để được dữ liệu mã hóa (cipher).
- Để lấy plaintext từ cipher, thực hiện XOR cipher với pad.

1. Mã cổ điển

1. 4. Mã hoán vị

- Mã thay thế: các chữ trong bản rõ được thay thế bằng các chữ khác của bản mã, vd ABCU → XYZ
- Mã hoán vị: các chữ trong bản rõ không thay thế bằng các chữ khác mà chỉ thay đổi vị trí, vd ABCDE → BADEC
- Mã hoán vị giàu bản rõ bằng cách thay đổi thứ tự các chữ, không thay đổi các chữ thực tế được dùng → bản mã có **cùng bảng phân bố tần suất** như bản gốc → dễ bị khai thác trong thám mã.
- **Mã Rail Fence, Mã dịch chuyển dòng, Mã tích**

1. Mã cổ điển

1. 4. Mã hoán vị

Mã Rail Fence

- Viết các chữ của bản rõ **theo đường chéo** (hướng xuống dưới/hướng lên trên) trên một số dòng.
- Đọc các chữ theo **từng dòng** → bản mã.
- Số dòng chính là khoá.

Ví dụ 1.9

Plaintext: "This is a secret message"

- Key = 2

Rail Fence Encoding	T		I		I		A		E		R		T		E		S		G	
	H		S		S		S		C		E		M		S		A		E	
Ciphertext	T	I	I	A	E	R	T	E	S	G	H	S	S	S	C	E	M	S	A	E

1. Mã cổ điển

1. 4. Mã hoán vị

Plaintext

T H I S I S A S E C R E T M E S S S A G E

Rail Fence

Encoding

key = 3

T			I			E			T		S			
H		S	S	S	C	E	M	S	A		E			
I				A	R			E			G			

● Ciphertext

T I E T S H S S S S C E M S A E I A R E G

Plaintext

T H I S I S A S E C R E T M E S S S A G E

Rail Fence

Encoding

key = 4

T					A					T					G
H				S	S			E	M			A		E	
I		I				E	R			E	S				
S					C					S					

● Ciphertext

T A T G H S S S E M A E I I E R E S S S C S

1. Mã cổ điển

1. 4. Mã hoán vị

• Giải mã:

- Tạo ma trận gồm **số hàng = len(ciphertext)**; **số cột = key**.
- Điền các chữ của bản mã giống như quá trình mã hoá → bản rõ.

Ví dụ 1.10

- Ciphertext = "GSGSEKFREKEOE" và Key = 3.

*	-	-	-	*	-	-	-	*	-	-	-	*
-	*	*	*	-	*	-	*	*	-	-	*	*
-	-	*	-	-	*	-	-	-	*	*	-	-

- Plaintext = "GeeksforGeeks"

1. Mã cổ điển

1. 4. Mã hoán vị

Mã dịch chuyển cột (Columnar Transposition Ciphers)

- Giả sử lấy một số cột xác định và chọn một hoán vị chỉ số của các cột đó làm khóa.
- Viết các chữ của bản rõ lần lượt theo các dòng.
- Đọc lại chúng theo các cột với thứ tự chỉ số ở dòng khóa để nhận được bản mã.
- Quá trình giải mã được thực hiện ngược lại.

Key:	4	3	1	2	5	6	7
Plaintext:	a	t	t	a	c	k	p
	o	s	t	p	o	p	e
	d	u	n	t	i	l	t
	w	o	a	m	x	y	z

Ciphertext: TTNAAPMTSUOAODWCOIXKNLYPETZ

- Mỗi khóa trong ví dụ trên là một hoán vị của 7, nên số khóa có thể là $7! = 4032$.
- Do $2^{11} < 4032 < 2^{12} \rightarrow$ cần 12 bit để biểu diễn không gian khóa (độ dài khóa biểu diễn dạng bit là 12).

1. Mã cổ điển

1. 4. Mã hoán vị

Mã tích (Product Ciphers)

⊕ **Nhận xét:** Mã dùng hoán vị hoặc dịch chuyển không an toàn vì các đặc trưng tần suất của ngôn ngữ không thay đổi.

 Kết hợp hai phương pháp: mã thay thế và mã hoán vị theo kiểu đan xen hoặc lặp nhiều vòng.

- Trong thực tế, lặp nhiều lần cùng một loại mã không tạo nên mã phức tạp hơn mà về bản chất chúng cũng tương đương với một lần mã cùng loại: tích của hai phép thay thế sẽ là một phép thay thế; tích của hai phép hoán vị sẽ là một phép hoán vị.
- Tuy nhiên tích hai loại mã khác nhau sẽ tạo nên mã mới phức tạp hơn, ví dụ **phép thay thế được nối tiếp bằng phép dịch chuyển** tạo nên mã mới khó hơn rất nhiều → mã hiện đại.

1. Mã cổ điển

1. 4. Mã hoán vị

Quick Note

- Tất cả các hệ mã đều có 2 thành phần cơ bản:
 - Algorithm (What you do to the message)
 - Key (The secret that you need in order to encrypt/decrypt properly)
- Theo **luật Kirchoff** (1835-1903): "toute bộ thuật toán mã hoá/giải mã trừ khóa là không bí mật với kẻ địch"
 - Khoa cần bí mật
 - Thuật toán thì không

1. Mã cổ điển

1. 4. Mã hoán vị

Tổng kết

- Hệ mật mã là một lĩnh vực toán học nghiên cứu về bảo mật thông tin
- Chúng ta tin tưởng vào độ an toàn của một hệ mật mã nếu như nhiều người thông minh không giải mã được nó.
- Trong bài học, tìm hiểu một số loại hệ mã đơn giản:
 - Mã dịch chuyển
 - Mã thay thế
 - Mã Playfair
 - Mã Vigenere
 - One-Time Pad
 - Mã hoán vị: Rail Fence, Row Transposition
- Chúng đều có một thuật toán và một khoá

1. Mã cổ điển

1. 5. Phụ lục: Phép toán modulo

Phép toán modulo

- What is the remainder when 51 is divided by 7? Answer: 2

We write $51 \equiv 2 \pmod{7}$ and say "51 is congruent to 2 modulo 7"

This means that 51 and 2 have the same remainder when divided by 7.

In other words, 51 and 2 differ by a multiple of 7, or $51 - 2 = 49$ is a multiple of 7.

- In general, if a, b and m are integers,

$$a \equiv b \pmod{m}, \quad "a \text{ is congruent to } b \text{ modulo } m"$$

means that

a and b differ by a multiple of m , or $a - b = km$, where k is some integer.

- We will be interested in the smallest integer $b \geq 0$ such that $a - b = km$, where k is some integer.

1. Mã cổ điển

1. 5. Phụ lục: Phép toán modulo

Phép toán modulo

Ví dụ 1.11

Tính

- a) $52 \text{ mod } 8$
- b) $41 \text{ mod } 5$
- c) $84 \text{ mod } 4$
- d) $-17 \text{ mod } 4$
- e) $145672 \text{ mod } 13$

1. Mã cổ điển

1. 5. Phụ lục: Phép toán modulo

Phép toán modulo

- $28 \equiv 4 \pmod{8}$ and $11 \equiv 3 \pmod{8}$.

What happens when we add or subtract?

- $28 + 11 = 39 \equiv 7 \pmod{8}$.

Also notice: $28 \equiv 4 \pmod{8}$ and $11 \equiv 3 \pmod{8}$ and $3 + 4 = 7 \equiv 7 \pmod{8}$.

- $28 - 11 = 17 \equiv 1 \pmod{8}$.

Also notice: $28 \equiv 4 \pmod{8}$ and $11 \equiv 3 \pmod{8}$ and $4 - 3 = 1 \equiv 1 \pmod{8}$.

- **Tổng quát:** If $a \equiv A \pmod{m}$ and $b \equiv B \pmod{m}$, then

- $a + b \equiv A + B \pmod{m}$
- $a - b \equiv A - B \pmod{m}$

Phép toán modulo

Luyện tập

- ① Reduce 237288 modulo 5
- ② Reduce $192 + 118$ modulo 5
- ③ Reduce $192 - 118$ modulo 5
- ④ Reduce $118 - 192$ modulo 5
- ⑤ Today is a Wednesday. What day of the week will it be
 - a) 100 days from now?
 - b) 365 days from now?
 - c) 1000 days from now?
- ⑥ Emily celebrated her 13th birthday on Wednesday, February 19th, 2014. On what day of the week was she born? (Don't forget about the leap years in 2004, 2008 and 2012!)

1. Mã cổ điển

1. 5. Phụ lục: Phép toán modulo

Phép toán modulo

Đáp số:

- ① $237288 \equiv 3 \pmod{5}$
- ② $192 + 118 \equiv 0 \pmod{5}$
- ③ $192 - 118 \equiv 4 \pmod{5}$
- ④ $118 - 192 \equiv 1 \pmod{5}$
- ⑤
 - a) Friday
 - b) Thursday
 - c) Tuesday
- ⑥ Monday

Nội dung

Mã cổ điển

- 1.2 Một số thuật ngữ và kí hiệu
- 1.3 Mã thay thế
- 1.4 Mã hoán vị
- 1.5 Phụ lục: Phép toán modulo

Trao đổi

TRAO ĐỔI