TRUYỀN VÀ BẢO MẬT THÔNG TIN

Bài 3:

Mã hóa đối xứng căn bản

VŨ THỊ TRÀ

Nội dung

- Mã hóa Ceasar
- Mô hình tổng quát mã hóa đối xứng
- Mã hóa thay thế đơn bảng
- Mã hóa thay thế đa ký tự (Playfair, Hill)
- Mã hóa thay thế đa bảng
- One-Time Pad & Mã hóa hoán vị +3 tình huống phá mã

Mã hóa Ceasar

Mã hóa Ceasar

- Gán cho mỗi chữ cái một con số nguyên từ 0 đến 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
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

- Phương pháp Ceasar được biểu diễn như sau:
  - **Lập mã:** $c = (p + k) \bmod 26$
  - **Giải mã:** $p = (c - k) \bmod 26$

  Trong đó, mod là phép chia module lấy số dư.

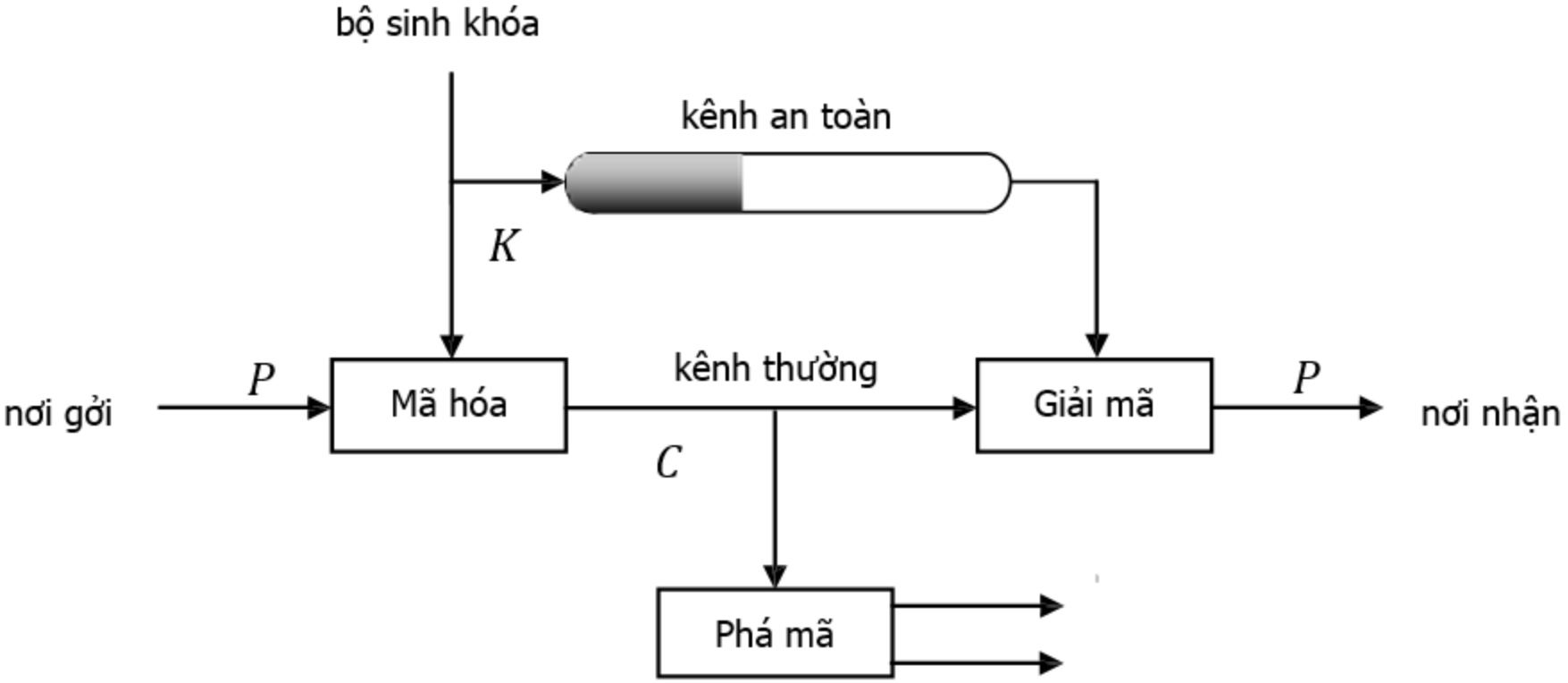
Mã thám mã Ceasar

B.mã:	PHHW PH DIWHU WKH WRJD SDUWB
1	oggv og chvgt vjg vqic rctva
2	nffu nf bgufs uif uphb qbsuz
3	**meet me after the toga party**
4	ldds ld zesdq sgd snfz ozqsx
5	kccr kc ydrcp rfc rmey nyprw
…	… … … …
25	qiix qi ejxiv xli xske tevxc

→ ***k=3*** *: bản giải mã có nghĩa*

Mô hình mã hóa đối xứng (Symmetric Ciphers)

Mô hình mã hóa đối xứng (Symmetric Ciphers)



Mô hình mã hóa đối xứng (Symmetric Ciphers)

Trong đó:

- Bản rõ *P* (plaintext)
- Thuật toán mã hóa *E* (encrypt algorithm)
- Khóa bí mật *K* (secret key)
- Bản mã *C* (ciphertext)
- Thuật toán giải mã *D* (decrypt algorithm)

Và $C = E(P, K)$

$P = D(C, K)$

Thời gian phá mã trung bình tương ứng với kích thước của khóa

Kích thước khóa (bít)	Số lượng khóa	Thời gian thực hiện (tốc độ 10^6 khóa/giây)	Thời gian thực hiện (tốc độ 10^9 khóa/giây)
32	?	?	?
64	?	?	?
128	?	?	?
Hoán vị 26 ký tự	?	?	?

Thời gian phá mã trung bình tương ứng với kích thước của khóa

Kích thước khóa (bít)	Số lượng khóa	Thời gian thực hiện (tốc độ 10^6 khóa/giây)	Thời gian thực hiện (tốc độ 10^9 khóa/giây)
32	$2^{32} \approx 4,3 \times 10^9$	4295 giây	4,3 giây
64	$2^{64} \approx 18,45 \times 10^{18}$	584.942,4 năm	585 năm
128	$2^{128} \approx 3,4 \times 10^{38}$	$10,8 \times 10^{24}$ năm	$10,8 \times 10^{21}$ năm
Hoán vị 26 ký tự	$26! \approx 4,03 \times 10^{26}$	$12,8 \times 10^{12}$ năm	$12,8 \times 10^9$ năm

Mã hóa đối xứng được xem là an toàn

Phương pháp mã hóa đối xứng là an toàn nếu

1. Không tồn tại kỹ thuật tấn công tắt nào khác tốt hơn phương pháp vét cạn khóa.
2. Miền giá trị khóa đủ lớn để việc vét cạn khóa là bất khả thi.

Mã hóa thay thế đơn bảng

Mã hóa thay thế đơn bảng (Monoalphabetic Substitution Cipher)

Nguyên tắc: phép hoán vị của 26 chữ cái. Mỗi hoán vị được xem như là một khóa.

- Chữ 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
- Chữ thay thế: **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

VD: bản rõ: meet me after the toga party

bản mã: NJJU NJ ZRUJM UKJ UVSZ DZMUE

Bảng thống kê tần suất sử dụng của các chữ cái, cụm 2 chữ, cụm 3 chữ trong tiếng Anh

Chữ cái (%)		Cụm 2 chữ (%)		Cụm 3 chữ (%)		Từ (%)	
E	13.05	TH	3.16	THE	4.72	THE	6.42
T	9.02	IN	1.54	ING	1.42	OF	4.02
O	8.21	ER	1.33	AND	1.13	AND	3.15
A	7.81	RE	1.30	ION	1.00	TO	2.36
N	7.28	AN	1.08	ENT	0.98	A	2.09
I	6.77	HE	1.08	FOR	0.76	IN	1.77
R	6.64	AR	1.02	TIO	0.75	THAT	1.25
S	6.46	EN	1.02	ERE	0.69	IS	1.03
H	5.85	TI	1.02	HER	0.68	I	0.94
D	4.11	TE	0.98	ATE	0.66	IT	0.93
L	3.60	AT	0.88	VER	0.63	FOR	0.77
C	2.93	ON	0.84	TER	0.62	AS	0.76
F	2.88	HA	0.84	THA	0.62	WITH	0.76
U	2.77	OU	0.72	ATI	0.59	WAS	0.72
M	2.62	IT	0.71	HAT	0.55	HIS	0.71
P	2.15	ES	0.69	ERS	0.54	HE	0.71
Y	1.51	ST	0.68	HIS	0.52	BE	0.63
W	1.49	OR	0.68	RES	0.50	NOT	0.61
G	1.39	NT	0.67	ILL	0.47	BY	0.57
B	1.28	HI	0.66	ARE	0.46	BUT	0.56
V	1.00	EA	0.64	CON	0.45	HAVE	0.55
K	0.42	VE	0.64	NCE	0.45	YOU	0.55
X	0.30	CO	0.59	ALL	0.44	WHICH	0.53
J	0.23	DE	0.55	EVE	0.44	ARE	0.50
Q	0.14	RA	0.55	ITH	0.44	ON	0.47
Z	0.09	RO	0.55	TED	0.44	OR	0.45

Mã hóa thay thế đơn bảng

- Xét bản mã sau:

  UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUD
  BMETSXAIZVUEPHZHMDZSHZOWSFPAPPDTSVPQUZ
  WYMXUZUHSXEPYEPOPDZSZUFPOMBZWPPDPTGU
  DTMOHMQ

- Số lần xuất hiện của các chữ cái được cho bởi bảng (1):

A 2	F 3	K 0	P 17	U 9
B 2	G 3	L 0	Q 3	V 5
C 0	H 6	M 7	R 0	W 4
D 6	I 1	N 0	S 10	X 5
E 6	J 0	O 9	T 4	Y 2
				Z 13

Mã hóa thay thế đơn bảng

- Số lần xuất hiện của các digram (xuất hiện từ 2 lần trở lên) được cho bởi bảng (2):

DT 2	HZ 2	PE 2	TS 2	XU 2
DZ 2	MO 2	PO 3	UD 2	ZO 2
EP 3	OH 2	PP 2	UZ 3	ZS 2
FP 3	OP 3	SX 3	VU 2	ZU 2
HM 2	PD 3	SZ 2	WS 2	ZW 3

- Theo bảng (1),(2), ta đoán P->e, Z->t, ZW->th, giả sử ZWSZ thuộc 1 từ thì từ đó có dạng th_t → S->a

Mã hóa thay thế đơn bảng

- Ta đã phá mã được như sau:

UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ

t a e e e a that e e a a t

VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX

e t ta t a e ee a e th t a

EPYEPOPDZSZUFPOMBZWPPDPTGUDTMOHMQ

e e tat e thee e

Mã hóa thay thế đơn bảng

- Cứ tiếp tục như vậy, dĩ nhiên việc thử không phải lúc nào cũng suôn sẻ, có những lúc phải thử và sai nhiều lần. Cuối cùng ta có được bản giải mã sau khi đã tách từ như sau:

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

→ *phá mã dựa trên tần suất chữ cái tốn thời gian ít hơn nhiều so với con số 12,8 tỷ năm*

Mã hóa thay thế đa ký tự – mã Playfair

Mã Playfair

- **Nguyên tắc:** xem hai ký tự đứng sát nhau là một đơn vị mã hóa, hai ký tự này được thay thế cùng lúc bằng hai ký tự khác. Playfair dùng một ma trận 5x5 các ký tự như sau:

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 Playfair

- Trong bảng trên, khóa là từ **MONARCHY** được điền vào các dòng đầu của bảng, các chữ cái còn lại được điền tiếp theo. Riêng hai chữ **I, J** được điền vào cùng một ô vì trong tiếng Anh, ít khi nhầm lẫn giữa chữ I và chữ J. Ví dụ, nếu gặp đoạn ký tự **CL_MATE**, ta sẽ biết đó là từ CLIMATE chứ không phải là từ CLJMATE.
- Trước khi mã hóa, bản rõ được tách ra thành các cặp ký tự. Nếu hai ký tự trong một cặp giống nhau thì sẽ được tách bằng chữ X (trong tiếng Anh ít khi có 2 ký tự X sát nhau). Ví dụ: từ **balloon** được tách thành **ba lx lo on** .

Mã Playfair

- Việc mã hóa từng cặp được thực hiện theo quy tắc:
  - Nếu hai ký tự trong cặp thuộc cùng một hàng, thì được thay bằng hai ký tự tiếp theo trong hàng. Nếu đến cuối hàng thì quay về đầu hàng. Ví dụ cặp ar được mã hóa thành RM.
  - Nếu hai ký tự trong cặp thuộc cùng một cột, thì được thay bằng hai ký tự tiếp theo trong cột. Nếu đến cuối cột thì quay về đầu cột. Ví dụ cặp ov được mã hóa thành HO.
  - Trong các trường hợp còn lại, hai ký tự được mã hóa sẽ tạo thành đường chéo của một hình chữ nhật và được thay bằng 2 ký tự trên đường chéo kia. Ví dụ: hs trở thành BP (B cùng dòng với H và P cùng dòng với S); ea trở thành IM (hoặc JM).

Mã thám mã Playfair

- Nếu chỉ xét trên 26 chữ cái thì mã hóa Playfair có 26x26=676 cặp chữ cái. Trong đó, các cặp chữ cái này ít bị chênh lệch về tần suất hơn so với sự chênh lệnh tần suất của từng chữ cái. Ngoài ra số lượng các cặp chữ cái nhiều hơn cũng làm cho việc phá mã tần suất khó khăn hơn. Đây chính là lý do mà người ta tin rằng mã hóa Playfair không thể bị phá và được quân đội Anh sử dụng trong chiến tranh thế giới lần thứ nhất.

Mã hóa thay thế đa ký tự – mã Hill

Mã Hill

- **Nguyên tắc:** mã hóa 1 lần *n* ký tự, trong đó, gán mỗi chữ cái cho 1 số nguyên, 26 ký tự tương ứng tập số nguyên [0..25], khóa *K là một ma trận cỡ nxn*

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

  - Lập mã: $\boldsymbol{C=PK \bmod 26}$ ***(****or* $C=KP \bmod 26$***)***
  - Giải mã: $\boldsymbol{P=CK^{-1} \bmod 26}$ ***(****or* $P=K^{-1}C \bmod 26$***)***

→ *Mã hóa Hill ẩn giấu các thông tin về tần suất nhiều hơn mã hóa Playfair do có thể mã hóa nhiều hơn 2 các ký tự cùng lúc.*

Ví dụ: Mã Hill

- Xét ví dụ bản rõ là **paymoremoney** cùng với khóa K là

$$\mathbf{K} = \begin{pmatrix} 17 & 17 & 5 \\ 21 & 18 & 21 \\ 2 & 2 & 19 \end{pmatrix}$$

- Ba chữ cái đầu tiên của bản rõ tương ứng với vector (15, 0, 24)

$$(15\ 0\ 24) \begin{pmatrix} 17 & 17 & 5 \\ 21 & 18 & 21 \\ 2 & 2 & 19 \end{pmatrix} \textit{mod}\ 26 = (303\ 303\ 531)\ \textit{mod}\ 26$$

$$= (17\ 17\ 11) = \text{RRL}$$

- Thực hiện tương tự cho 3 ký tự tiếp theo, cứ như vậy cho đến cuối bản ra ta sẽ có bản mã đầy đủ là RRLMWBKASPDH

Mã hóa thay thế đa bảng – mã Vigenere

Mã hóa thay thế đa bảng (Polyalphabetic Substitution Cipher)

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

Mã hóa thay thế đa bảng

Nguyên tắc:

1. Một tập hợp các qui tắc thay thế đơn bảng liên quan được sử dụng
2. Một chìa khóa xác định qui tắc cụ thể được chọn cho một chuyển đổi nhất định.

- **Trường hợp đặc biệt - mã Vigenere. :** cho dòng thứ k của bảng là một mã hóa Ceasar k-1 vị trí
- **Trường hợp tổng quát:** mỗi dòng của bảng là một mã hóa đơn bảng
- Để mã hóa một bản tin thì cần có một khóa có chiều dài bằng chiều dài bản tin. Thường thì khóa là một cụm từ nào đó và được viết lặp lại cho đến khi có chiều dài bằng chiều dài bản tin.

Mã hóa thay thế đa bảng

Ví dụ với bản tin là „**We are discovered, save yourself**“ và khóa là từ **DECEPTIVE**, chúng ta mã hóa như sau:

- Plaintext: **wearediscoveredsaveyourself**
- Key: **DECEPTIVEDECEPTIVEDECEPTIVE**
- Ciphertext: **ZICVTWQNGRZGVTWAVZHCQYGLMGJ**

Trong ví dụ trên, ứng với chữ w trong bản rõ là khóa D, nên dòng mã hóa thứ 4 ứng với khóa D trong bảng mã Vigenere được chọn. Do đó, chữ w được mã hóa thành chữ Z. Tương tự như vậy cho các chữ còn lại.

Trong ví dụ trên, các chữ e trong bản rõ được mã hóa tương ứng thành I, T, G, T, H, M trong bản mã. Do đó phương pháp phá mã dựa trên thống kê tần suất chữ cái là không thực hiện được.

Mã hóa thay thế đa bảng

- Trong 3 thế kỷ sau đó mã hóa Vigenere được xem là mã hóa không thể bị phá và được biết dưới cái tên "le chipffre indechiffrable" (mật mã không thể phá nổi).
- Đến thế kỷ 19, nhà khoa học người Anh Charles Barbage, đã tìm ra cách phá mã Vigenere. Việc phá mã bằng cách thống kê sự lặp lại của các cụm từ để phỏng đoán chiều dài của khóa, trong ví dụ trên cụm từ VTW được lặp lại cách nhau 9 vị trí nên có thể đoán chiều dài của khóa là 9. Và từ đó có thể tách bản mã thành 9 phần, phần thứ nhất gồm các chữ 1, 10, 19, 28, … phần thứ hai gồm các chữ 2, 11, 20, 29….cho đến phần thứ chín. Mỗi phần coi như được mã hóa bằng phương pháp mã hóa đơn bảng. Từ đó áp dụng phương pháp phá mã dựa trên tần suất chữ cái cho từng phần một. Cuối cùng ráp lại sẽ tìm ra được bản rõ.

One-Time Pad

One-Time Pad

- Điểm yếu của mã hóa đa bảng là sự lặp lại các từ trong khóa, ví dụ từ DECEPTIVE được lặp đi lặp lại nhiều lần. Điều này làm cho vẫn tồn tại một mối liên quan giữa bản rõ và bản mã, ví dụ cụm từ red trong bản rõ được lặp lại thì cụm từ VTW cũng được lặp lại trong bản mã. Người phá mã tận dụng mối liên quan này để thực hiện phá mã. Do đó vấn đề ở đây là làm sao để giữa bản rõ và bản mã thật sự ngẫu nhiên, không tồn tại mối quan hệ nào. Để giải quyết vấn đề này, Joseph Mauborgne, giám đốc viện nghiên cứu mật mã của quân đội Mỹ, vào cuối chiến tranh thế giới lần thứ nhất, đã đề xuất phương án là dùng khóa ngẫu nhiên. Khóa ngẫu nhiên có chiều dài bằng chiều dài của bản rõ, mỗi khóa chỉ sử dụng một lần.

One-Time Pad

- **Ví dụ:**
  - Bản tin P: wearediscoveredsaveyourself
  - Khóa K1: FHWYKLVMKVKXCVKDJSFSAPXZCVP
  - Bản mã C: BLWPOODEMJFBTZNVJNJQOJORGGU
- **Trường hợp 1:**
  - Bản mã C: BLWPOODEMJFBTZNVJNJQOJORGGU
  - Khóa K2: IESRLKBWJFCIFZUCJLZXAXAAPSY
  - Bản giải mã: theydecidedtoattacktomorrow
    (they decided to attack tomorrow)

One-Time Pad

- **Trường hợp 2:**
  - Bản mã C: BLWPOODEMJFBTZNVJNJQOJORGGU
  - Khóa K2: FHAHDDRAIQFIASJGJWQSVVBJAZB
  - Bản giải mã: wewillmeetatthepartytonight

    (we will meet at the party tonight)
- Trong cả hai trường hợp trên thì bản giải mã đều có ý nghĩa. Điều này có nghĩa là nếu người phá mã thực hiện phá mã vét cạn thì sẽ tìm được nhiều khóa ứng với nhiều bản tin có ý nghĩa, do đó sẽ không biết được bản tin nào là bản rõ. Điều này minh chứng phương pháp **One-Time Pad** là phương pháp mã hóa an toàn tuyệt đối, và được xem là ly thánh của khoa học mật mã cổ điển.

One-Time Pad

- Phương pháp **One-Time Pad** là an toàn tuyệt đối khi mỗi khóa chỉ được sử dụng một lần. Nếu một khóa được sử dụng nhiều lần thì cũng không khác gì việc lặp lại một từ trong khóa (ví dụ khóa có từ DECEPTIVE được lặp lại). Ngoài ra các khóa phải thật sự ngẫu nhiên với nhau. Nếu các điều này bị vi phạm thì sẽ có một mối liên hệ giữa bản rõ và bản mã, mà người phá mã sẽ tận dụng mối quan hệ này.
- Phương pháp **One-Time Pad** không có ý nghĩa sử dụng thực tế. Vì chiều dài khóa bằng chiều dài bản tin, mỗi khóa chỉ sử dụng một lần, nên thay vì truyền khóa trên kênh an toàn thì có thể truyền trực tiếp bản rõ mà không cần quan tâm đến vấn đề mã hóa.

One-Time Pad

- Sau chiến tranh thế giới thứ nhất, người ta vẫn chưa thể tìm ra loại mật mã nào khác mà không bị phá mã. Mọi cố gắng vẫn là tìm cách thực hiện mã hóa thay thế đa bảng dùng một khóa dài, ít lập lại, để hạn chế phá mã. Máy **ENIGMA** được quân đội Đức sử dụng trong chiến tranh thế giới lần 2 là một máy như vậy.
- Sử dụng máy ENIGMA, Đức đã chiếm ưu thế trong giai đoạn đầu của cuộc chiến. Tuy nhiên trong giai đoạn sau, các nhà phá mã người Ba Lan và Anh (trong đó có Alan Turing, người phá minh ra máy tính có thể lập trình được) đã tìm ra cách **phá mã máy ENIGMA**. Việc phá mã thực hiện được dựa vào một số điểm yếu trong khâu phân phối khóa của quân Đức.

One-Time Pad

ENIGMA

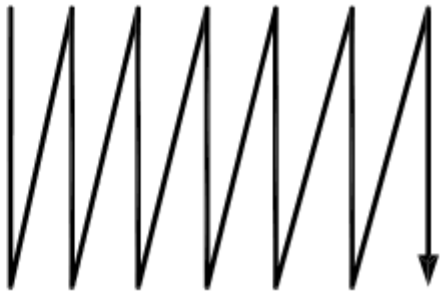
Mã hóa hoán vị

Mã hoán vị (Permutation Cipher)

Nguyên tắc: Xáo trộn thứ tự của các chữ cái trong bản rõ. Do thứ tự của các chữ cái bị thay đổi nên người đọc không thể hiểu được ý nghĩa của bản tin dù các chữ đó không thay đổi.

- **Phương pháp 1:** Ghi bản rõ theo từng hàng, sau đó kết xuất bản mã dựa trên các cột. Ví dụ bản rõ "**attackpostponeduntilthisnoon**" được viết lại thành bảng 4 x 7 như sau:

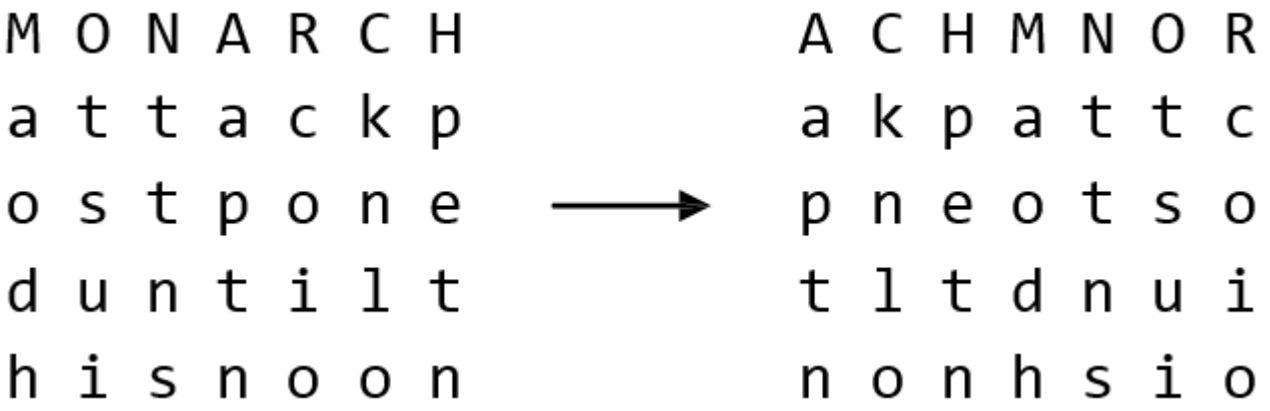
```
a t t a c k p
o s t p o n e
d u n t i l t
h i s n o o n
```



→ bản mã "**AODHTSUITTNSAPTNCOIOKNLOPETN**"

Mã hoán vị

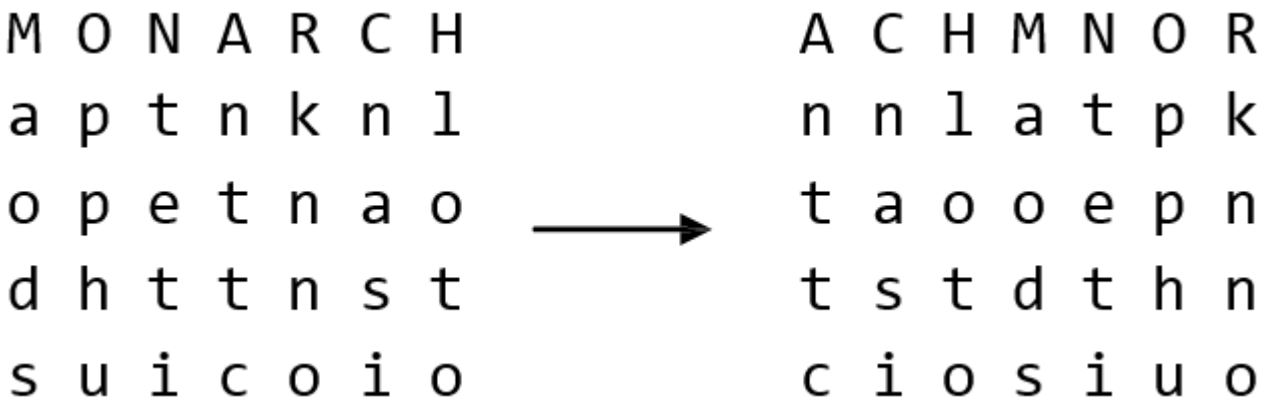
- **Phương pháp 2:** hoán vị các cột trước khi kết xuất bản mã. Ví dụ chọn một khóa là **MONARCH**, ta có thể hoán vị các cột:



→ bản mã: "**APTNKNLOPETNAODHTTNSTSUICOIO**"

Mã hoán vị

- **Phương pháp 3:** áp dụng phương pháp hoán vị 2 lần (double transposition), tức sau khi hoán vị lần 1, ta lại lấy kết quả đó hoán vị thêm một lần nữa.



→ bản mã: "**NTTCNASILOTOAODSTETIPPHUKNNO**"

Mã hoán vị: Nhận xét

- Phá mã phương pháp hoán vị 2 lần không phải là chuyện dễ dàng vì rất khó đoán ra được quy luật hoán vị.
- Không thể áp dụng được phương pháp phân tích tần suất chữ cái giống như phương pháp thay thế vì tần suất chữ cái của bản rõ và bản mã là giống nhau.

Kết luận

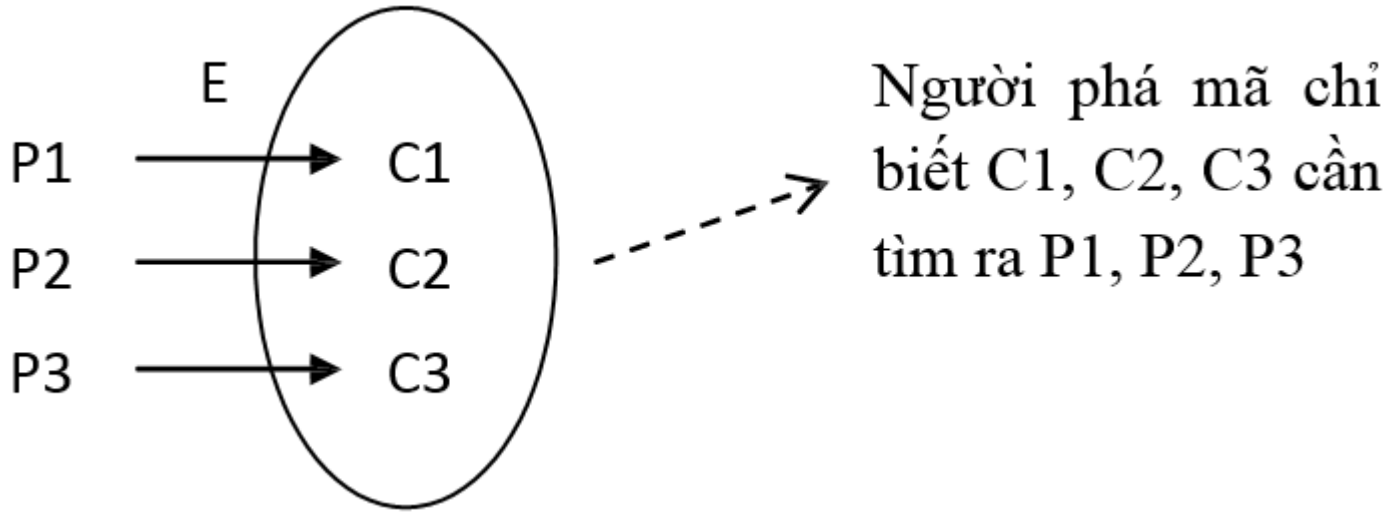
- Các phương pháp **mã hóa cổ điển** thường dựa trên hai phương thức:
  - **Phương thức thay thế** (substitution): biến một chữ cái trong bản rõ thành một chữ cái khác trong bản mã.
    Phương thức hoán vị (permutation): thay đổi thứ tự ban đầu của các chữ cái trong bản rõ.

Mục tiêu phá mã & 3 tình huống

- **Mục tiêu** của việc phá mã là từ bản mã đi tìm bản rõ, hoặc khóa, hoặc cả hai.
- **Giả định** rằng người phá mã biết rõ thuật toán mã hóa và giải mã (luật **Kerchoff**). Việc phá mã sẽ có 3 tình huống sau:
  - Chỉ biết bản mã (ciphertext–only)
  - Biết một số cặp bản rõ – bản mã (known–plaintext)
  - Một số cặp bản rõ – bản mã và bản rõ được lựa chọn (choosen–plaintext)

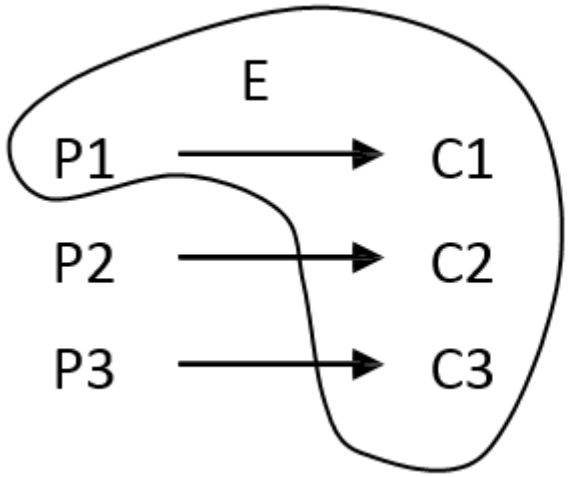
3 tình huống phá mã

1. **Chỉ biết bản mã (ciphertext–only):** đây là trường hợp gây khó khăn nhất cho người phá mã.



3 tình huống phá mã

2. **Biết một số cặp bản rõ – bản mã (known–plaintext):** trong trường hợp này, người phá mã có được một vài cặp bản rõ và bản mã tương ứng.



Người phá mã biết C1, C2, C3 và biết bản rõ tương ứng với C1 là P1. Cần tìm ra P2, P3.

3 tình huống phá mã

3. **Một số cặp bản rõ – bản mã và bản rõ được lựa chọn (choosen–plaintext):** trong trường hợp này, người phá mã có khả năng tự lựa một số bản rõ và quan sát được bản mã tương ứng.

→ *Thách thức cho các nhà nghiên cứu là phải tìm ra các thuật toán mã hóa sao cho không thể bị phá mã không chỉ trong trường hợp 1 mà còn ngay cả trong trường hợp 2 và 3*

Q&A

1. Hãy phân biệt các phương pháp mã hóa dựa trên cơ sở thay thế (substitution) và hoán vị (permutation)?
2. Phân loại các phương pháp mã hóa căn bản theo 2 nguyên lý trên?
3. Hãy đưa ra lộ trình phá khóa tương ứng với mỗi phương pháp mã hóa?