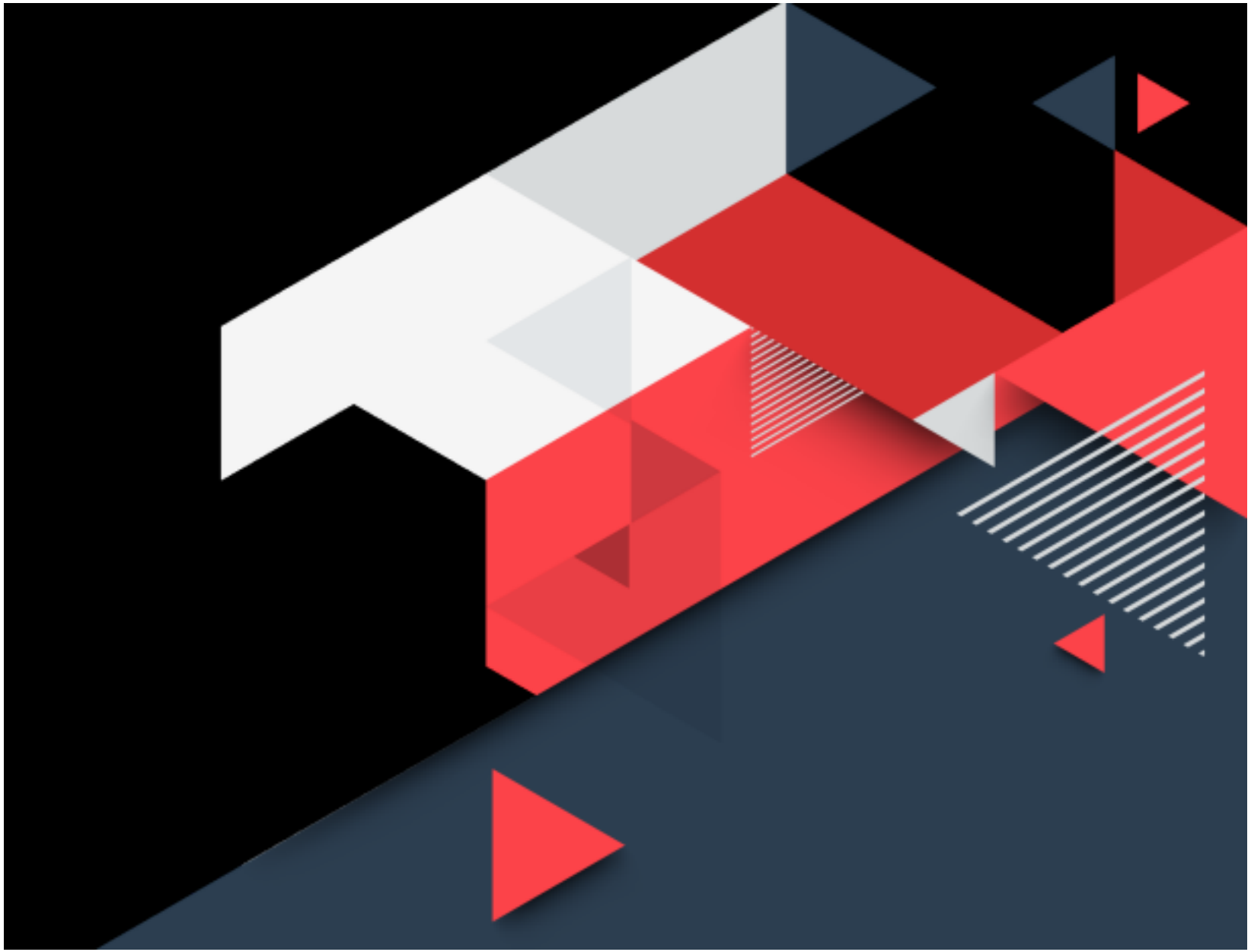


MỤC TIÊU

GIỚI THIỆU HỌC PHẦN

GIỚI THIỆU HỌC PHẦN

[1] Đặng Trường Sơn. (2012). *Giáo trình Bảo mật thông tin.* NXB Đại học quốc gia TP.HCM. [2] BehrouzA. Forouzan *Cryptography and Network Security 1st edition, MacGraw Hill* [3] William Stallings. (2010). *Cryptography and Network Security: Principles and Practice* (5th ed.). Prentice Hall.

PHƯƠNG PHÁP HỌC TẬP

QUY ĐỊNH HỌC TẬP

thoại

Lắng nghe

Đúng giờ

Tắt chuông điện

Hỏi lại những gì chưa hiểu

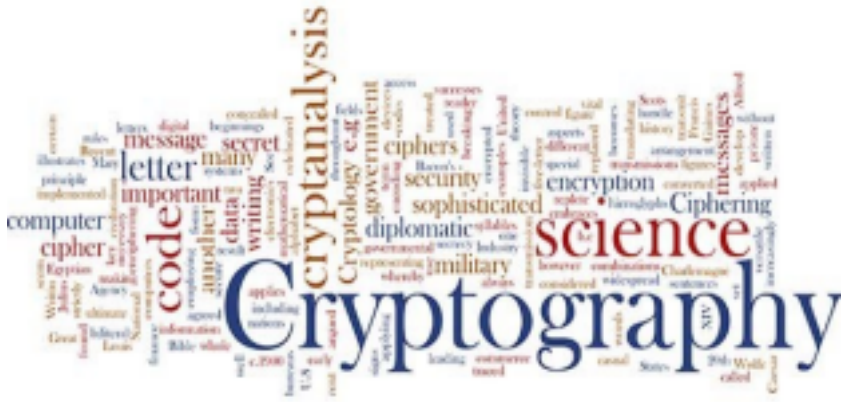
Đóng góp ý kiến và chia sẻ kinh

nghiệm **trong lớp học**

Không hút thuốc

BÀI 01 - MỤC TIÊU

TỔNG QUAN VỀ MẬT MÃ HỌC



Tránh bị truy nhập

sử dụng

Tránh bị phá hoại trái phép

Tránh bị tiết lộ

Tránh bị gián đoạn

Tránh bị sửa đổi

Các loại chung của mối đe dọa

Mối đe dọa bên trong

Mối đe dọa bên ngoài

Mối đe dọa kết hợp

Phân tích hệ quả của việc thực hiện các mối đe dọa

Vi phạm quy tắc kiểm soát truy cập đến các tài nguyên

Vi phạm tính bí mật

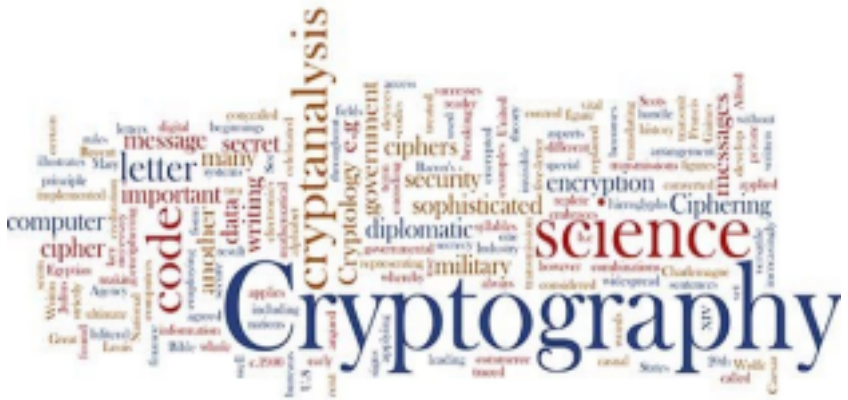
Vi phạm tính toàn vẹn

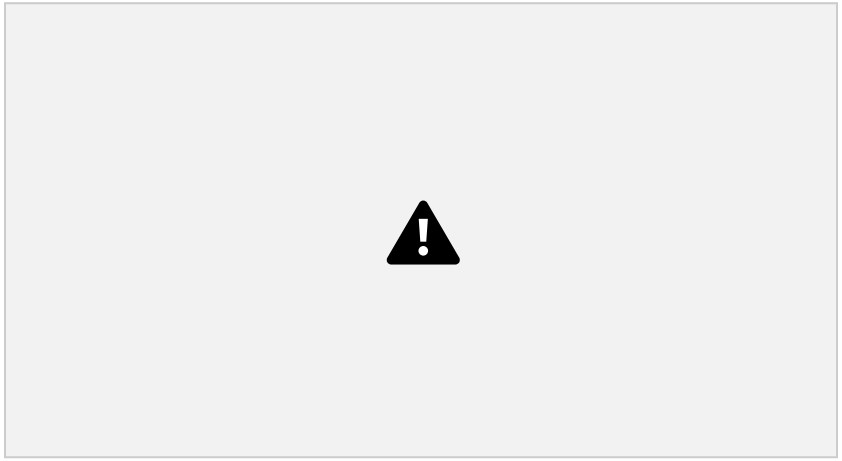
Vi phạm tính sẵn sàng phục vụ của hệ thống

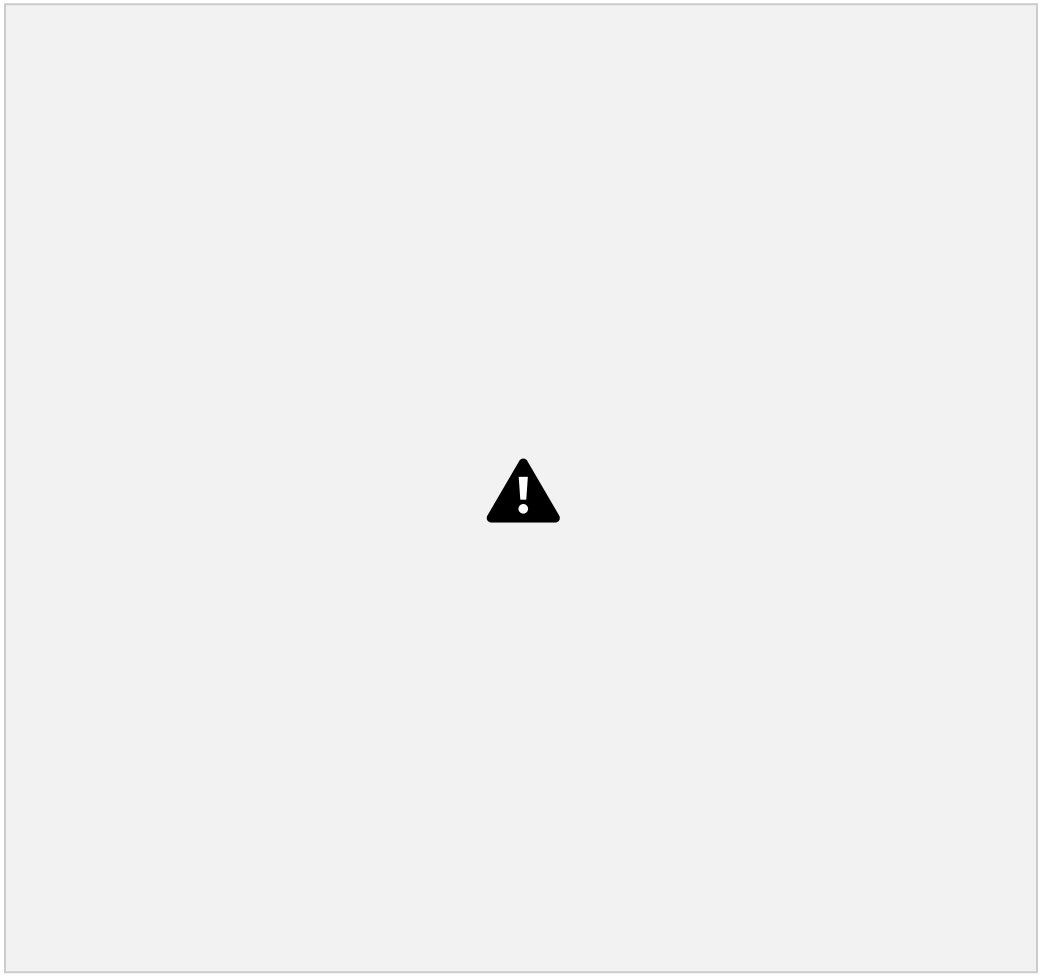
- Mã hóa
- Hàm băm, mã xác thực thông điệp (MAC)
- Chữ kí số

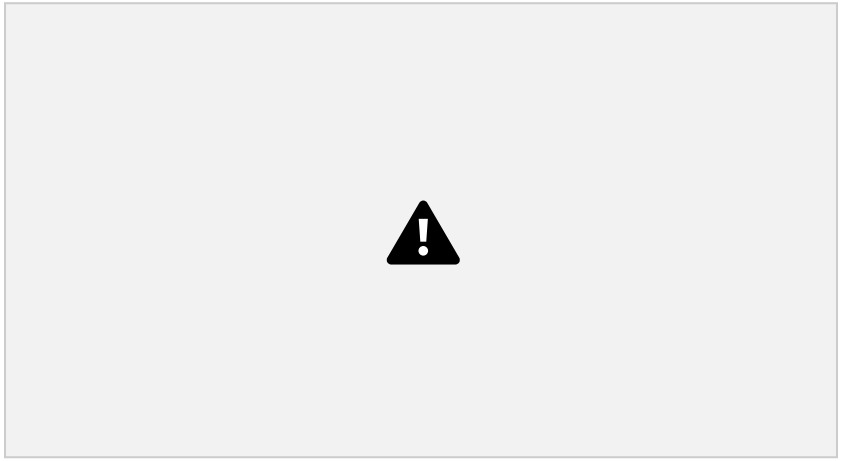
	Dịch vụ	Kỹ thuật
	Đảm bảo tính bí mật	Mã hóa
	Đảm bảo tính toàn vẹn	Chữ ký số, hàm băm, MAC

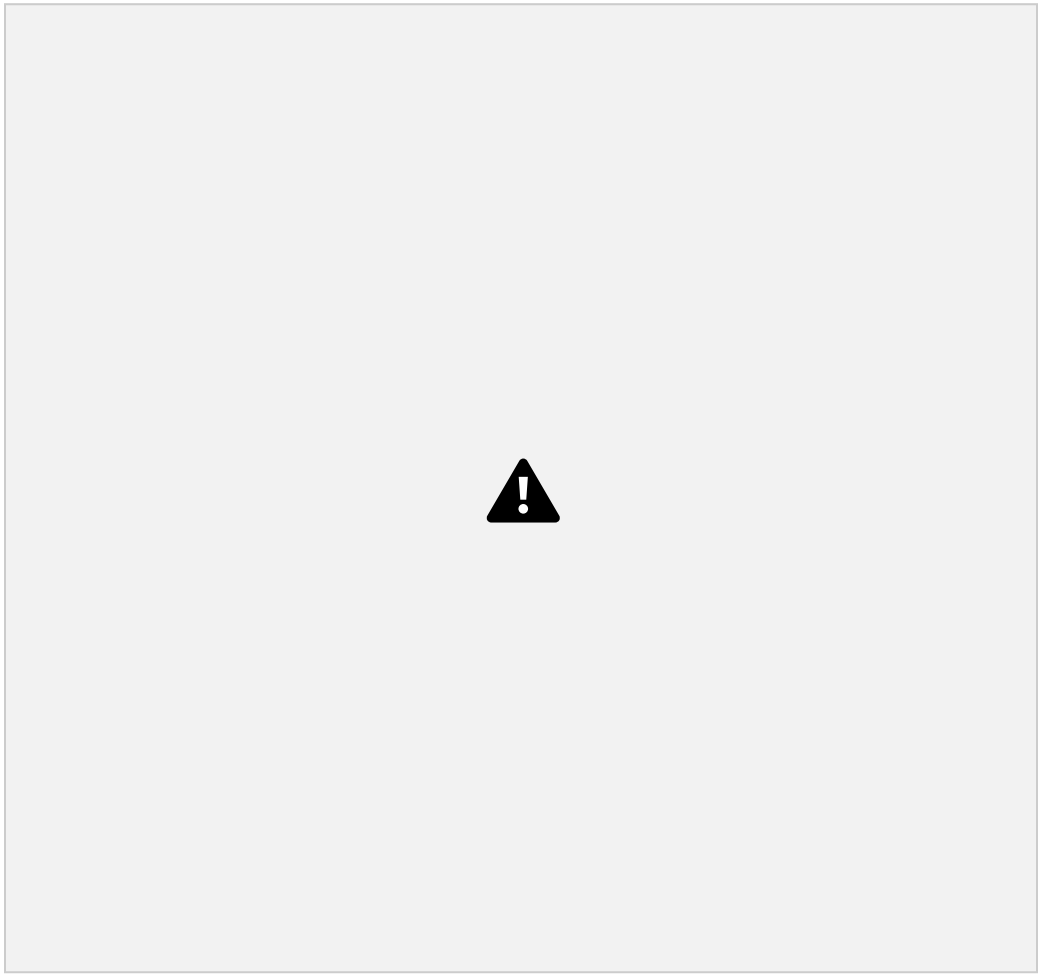
	Xác thực	**Chữ ký số, MAC**
	Chống chối bỏ	**Chữ ký số**


Cryptography
science
cryptanalysis
code
letter
message
secret
many
important
computer
cipher
data
writing
another
ciphers
security
encryption
sophisticated
Ciphering
diplomatic
military
messages
government
Cryptology
e.g.









- **P** là tập hữu hạn các bản rõ có thể
- **C** là tập hữu hạn các bản mã có thể
- **K** là tập hữu hạn các khoá có thể
- E_k: **P** → **C** - là quy tắc mã hóa với khóa $k \in$ **K**. Tập $\{E_k : k \in \mathbf{K}\}$

ký hiệu là **E**, còn tập {$e_K(x)$: $x \in$ **P**} ký hiệu là e_K(**P**).

- d_K: d_K(**P**) → **P** - là quy tắc giải mã với khóa $K \in$ **K**. Tập {d_K: $K \in$ **K**} ký hiệu là **D**.
- Với mỗi $K \in$ **K** sẽ được mô tả dưới dạng $K = (e_K, d_K)$, trong đó: e_K- là khóa dùng cho mã hóa, d_K - là khóa dùng cho giải mã. Khi đó e_K được hiểu là hàm E_{xxxx}, d_K được hiểu là hàm D_{xxxx}

- **Định nghĩa hệ mật:** Một hệ mật là bộ 5 **(P, C, K, E, D)** thoả mãn các điều kiện sau:

  1) $\forall x \in$ **P**, $K \in$ **K** ta có:

$$D_K(E_K(x)) = x;$$

2)

$E_K(P)$

- **Ghi chú:**
- Mã hóa: $E_K(P) =$

$E_{K_{AB}}(P)$

- Giải mã: $D_K(C) =$

$D_{K_{AB}}(C)$

$C = E_K(P) \in K$

Mã hóa

Bản rõ Bản mã

Giải mã

Quá trình chung của hệ mật

Mã hóa đối xứng: biết được khóa mã hóa dễ dàng suy ra được khóa giải mã (thông thường $K_e = K_d$)

$$C = E_K(m)$$

$$D_K(c) = m$$

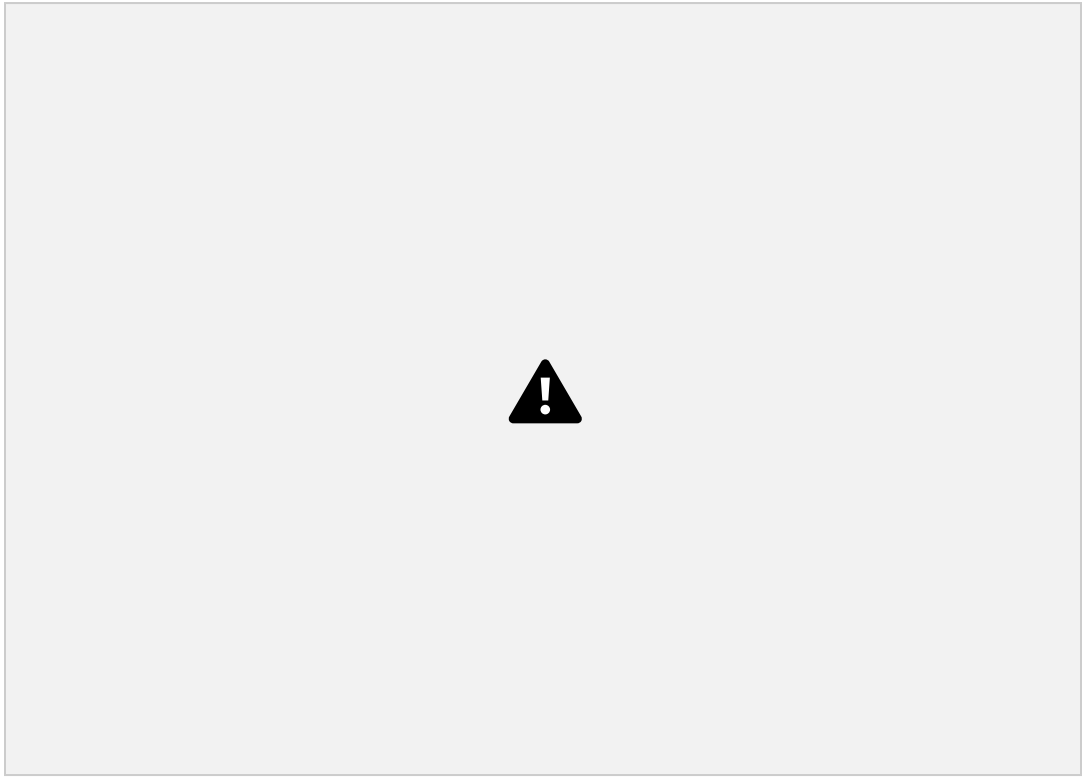

encryption key
eavesdroppin
g adversary
decryption key

Mã hóa bất đối xứng: Mỗi bên có một khoá công khai và một khoá bí mật. - Bên gửi dùng khoá công khai của bên nhận để mã hoá.
- Bên nhận dùng khoá bí mật của mình để giải mã.

$$c = E(m)$$

$$D(c) = m$$

- Mã khối:



Mã dòng:

(One-Time Pad)

Hệ mật Vernam

MÃ HÓA

Bộ kí tự: chữ cái latin

Khóa ngẫu nhiên:

PWKAX Thông điệp:

HELLO

Rõ	**H** (7)	**E** (4)	**L** (11)	**L** (11)	**O** (14)
Khóa	**P** (15)	**W** (22)	**K** (10)	**A** (0)	**X** (23)
Mã	**W** (22)	**A** (0)	**V** (21)	**?** (?)	**?** (?)

GIẢI MÃ

Khóa ngẫu nhiên:
PWKAX Bản mã: WAVLL

Bộ kí tự: chữ cái latin

Mã	**W** (22)	**A** (0)	**V** (21)	**L** (11)	**L** (11)
Khóa	**P** (15)	**W** (22)	**K** (10)	**A** (0)	**X** (23)
Rõ	**H** (7)	**E** (4)	**L** (11)	**L** (11)	**O** (14)

$$_{i}z$$

$$m\text{-}z\text{-}c_{\ i\ i\ i}\text{,-,-}0,1 \in \{\ \}$$

$$_{i}c$$

$$_{i}z$$

$$m_{i}$$

$$_{i}i\_z\ i\,c\_e\_m =$$

$$_{i}z$$

$m\text{-}d\text{-}c_{\ i\ z\ i}$

( )

,-, $m\text{-}c\text{-}z\text{-}A_{\ i\ i\ i} \in_i c$

$_i z$

( )

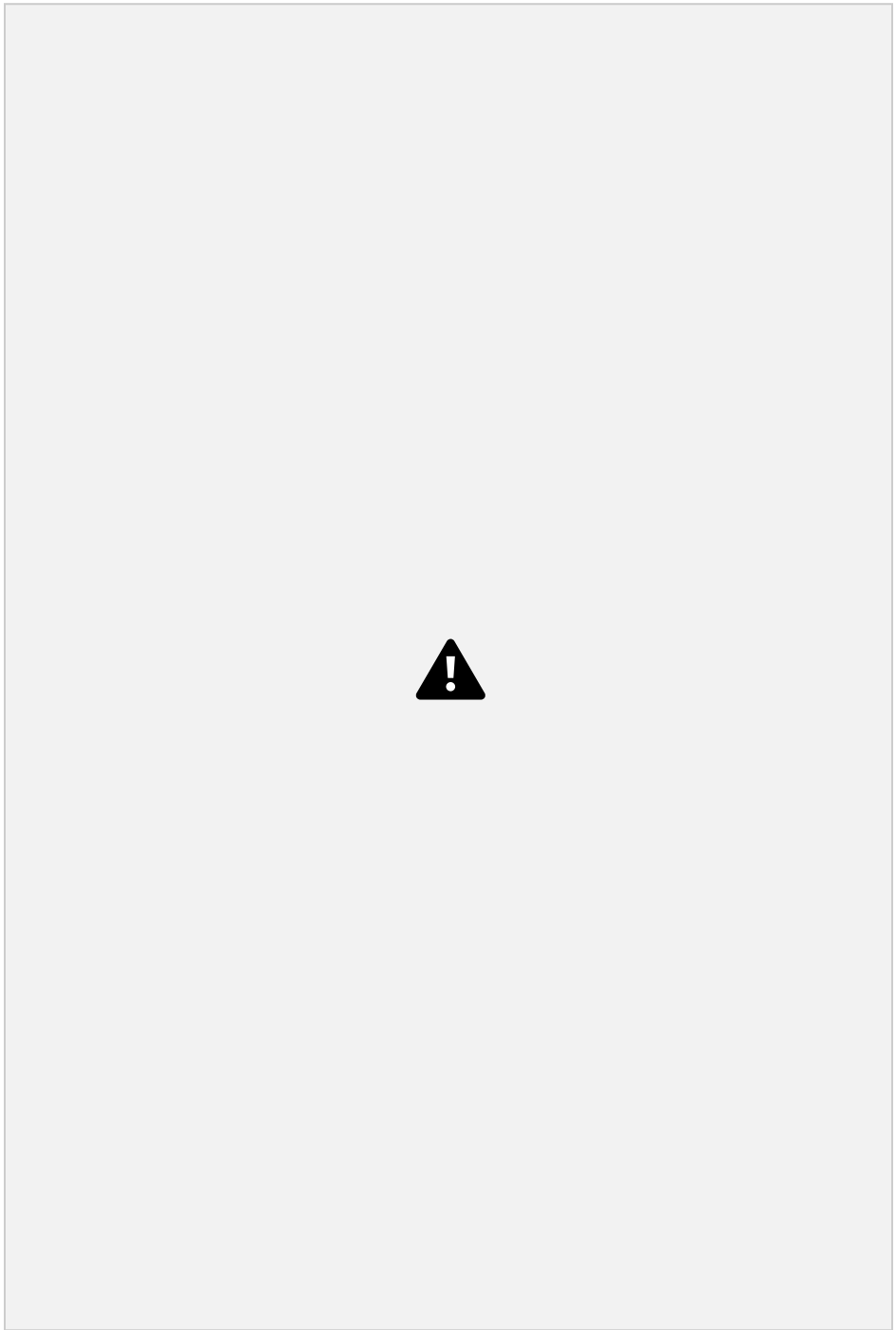
E D

m_i

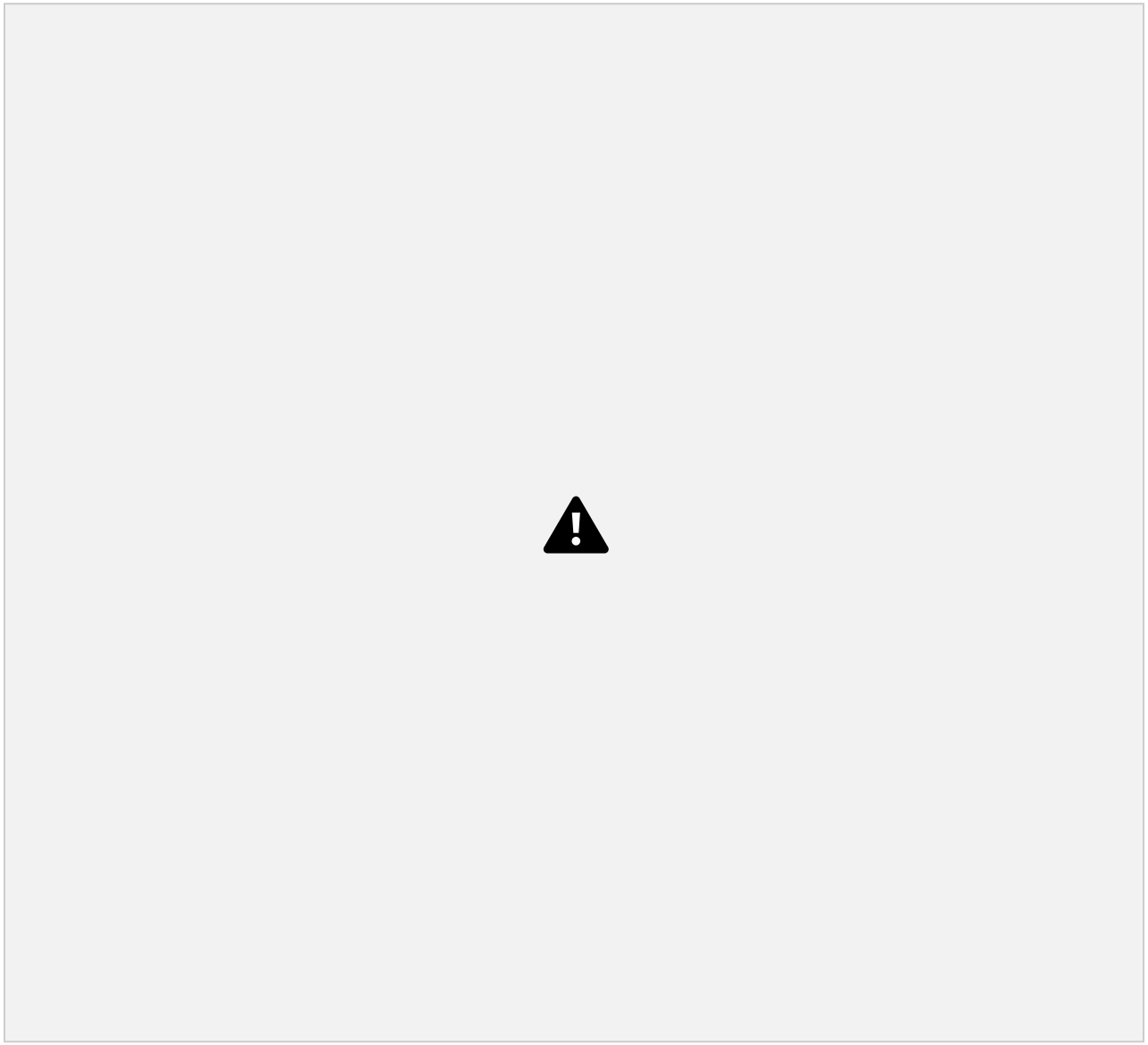
=

$_i$

- Khối «mở rộng khóa»
  - Là bộ sinh số giả ngẫu nhiên (PRNG)
  - Là quan trọng nhất
  - Quyết định độ an toàn của mã dòng
- Phân loại
  - z_i chỉ phụ thuộc K: «mã dòng đồng bộ» • z_i phụ thuộc $c_{i-n}, c_{i-n+1}, \ldots, c_{i-1}$: «mã dòng tự đồng bộ»



Mật mã học



- Sơ đồ khối của một hệ truyền tin mật sử dụng hệ mật

KBM? • Các phương pháp chính của hệ mật KBM?

- Yếu điểm của mã đơn biểu vs mã đa biểu?
- Hệ mật tích?
- Thám mã một số hệ mã cổ điển dựa trên phương pháp thống kê?

Sơ đồ khối của một hệ truyền tin

mật Oscar

Bản rõ Bản mã Kênh mở (không an toàn) Bản mã Bản rõ

Alice

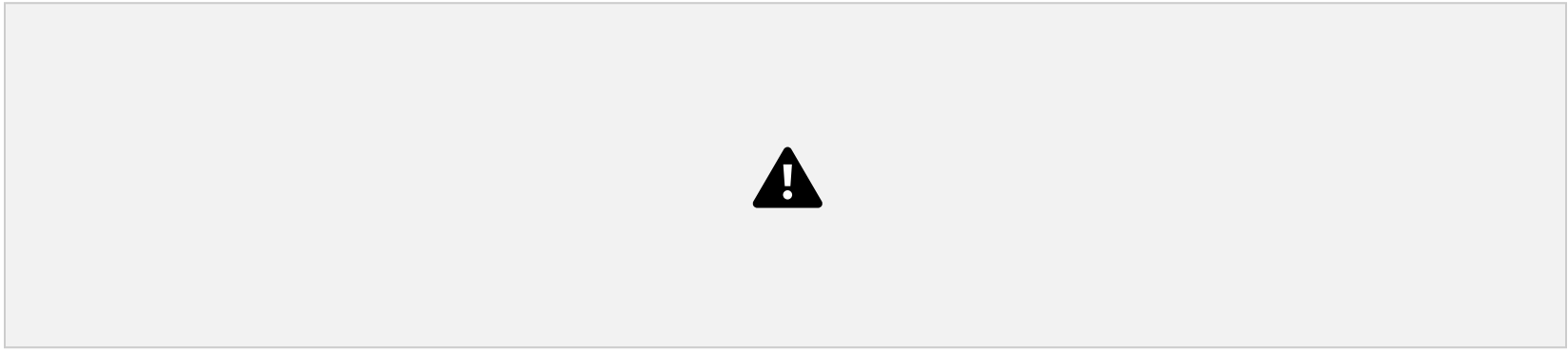
Bob

Kênh an toàn

- **Các hệ mật thay thế đơn biểu**
  - Khi khóa đã được chọn thì mỗi kí tự của bản rõ được ánh xạ đến một kí tự duy nhất của bản mã.
  - Như vậy độ dài của khóa ở đây là 26 và số khóa có thể có là 26!.

- **Mã dịch vòng (MDV – Shift cipher):**

- **Ví dụ:**
  - Bản rõ: HOC TAP TOT LAO DONG TOT
  - Khoá k = 5



  - Tìm bản mã
  - Từ bản mã thu được giải mã để thu bản rõ ban đầu.

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

Bản rõ																			
Mã tương ứng (x)					0				19				4	3	1				19
(x + 5) mod 26									24	5			19				8	11	
Bản mã		T		Y	F	U		T	5	Q	F	T		T	S	24	Y		
																			Y

Bản mã thu được: **MTHYFUYTYQFTITSLYTY**

Giải mã: SV tự làm!

❖ Mã Affine:

❖ Ví dụ:

- Cho k = (7, 3). Bản rõ: It is nice today
  - Tìm bản mã.
  - Giải mã bản mã thu được

❖ Giải:

- Tìm bản mã của bản rõ: It is nice today
  - Ta có hàm mã: $e_k(x) = 7x + 3 \bmod 26$

Ký tự	I	T	I	S	N	I	C	E	T	O	D	A	Y
	8	9	8	18	13	8	2	4	19	14	3	0	
	7	6	7	25	16	7	17	5	6	23	24	3	

Mã (x) 24 7x + 3 mod 26 15

Bản mã H G H Z Q H R F G X Y D P❑ Bản mã thu được là:

HGHZQHRFGXYDP

❖ Giải:

❑ Tìm bản rõ của bản mã: **HGHZQHRFGXYDP**

■ Ta có hàm mã:

$$d_k(y) = 7^{-1}.(y - 3) \bmod 26 = 15.(y - 3) \bmod 26$$

Bản mã H G H Z Q H R F G X Y D P

7	6	7	25	16	7	17	5	6	23	24	3
8	19	8	18	13	8	2	4	19	14	3	0

Mã (y) 15 15(y – 3) mod 26 24

Bản rõ I T I S N I C E T O D A Y❑ Bản mã thu được là: **It is nice**

today

- Nhận xét về các hệ mật thay thế đơn biểu:
  - Mỗi ký tự của bản rõ được ánh xạ đến một ký tự duy nhất của bản mã.
  - Các đặc trưng về ngôn ngữ, tần suất xuất hiện của các chữ trong bản rõ và chữ tương ứng trong bản mã là như

nhau ⇒ **Phương pháp thám mã bằng thống kê tần suất!**

- **Các hệ mật thay thế đa biểu**
  - Yếu điểm của các mã pháp đơn biểu
  - Một hướng khác làm tăng độ an toàn cho mã trên bảng chữ là sử dụng nhiều bảng chữ để mã.
  - Mỗi chữ sẽ được mã bằng bất kì chữ nào trong bản mã tùy

thuộc vào ngữ cảnh khi mã hóa. 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 mã thám đa bảng khó hơn.

Hệ mật thay thế đa biểu:

⚠

- **Hệ Vigenere:**
  - Blaise de Vigènere (1523 – 1596)
  - **Ý tưởng:**
    - Sử dụng một loạt mã Caesar khác nhau dựa trên các kí tự của một từ khóa

- Dễ hiểu và dễ thực hiện

- **Mô tả hệ mã Vigenere:**

- **Nhận xét:**
  - Số khoá: $\mathbf{26^m}$

  $\Rightarrow$ Tấn công tìm khoá vét cạn là không khả thi

- **Ví dụ minh hoạ:**
  - m = 6, k = cipher
    - Bản rõ: Information security
    - Hãy mã hoá bản rõ trên
    - Giải mã bản mã vừa thu được

- **Giải:**
  - Ta có từ khoá CIPHER, tương ứng với dãy số: k = (2, 8, 15, 7, 4, 17) • Chuyển các ký tự rõ thành mã trên Z_{26} rồi cộng với từ khoá

Bản rõ I N F O R M A T I O N S E C U R I T Y

8	**13**	**5**	**14**	**17**	**12**	**0**	**19**	**8**	**14**	**13**	**18**	**4**	**2**	**20**	**17**	**8**	**19**
2	**8**	**15**	**7**	**4**	**17**	**2**	**8**	**15**	**7**	**4**	**17**	**2**	**8**	**15**	**7**	**4**	**17**

Mã 24 Khoá 2 Bản mã 10 21 20 21 21 3 2 1 23 21 17 9 6 10 9 24 12 10 0

- Chuyển các ký tự số thành chữ cái tương ứng. Ta có bản mã: **KVUVVDCBXVRJGKJYMKA**

Mật mã học 24 February 2021 | Page 51

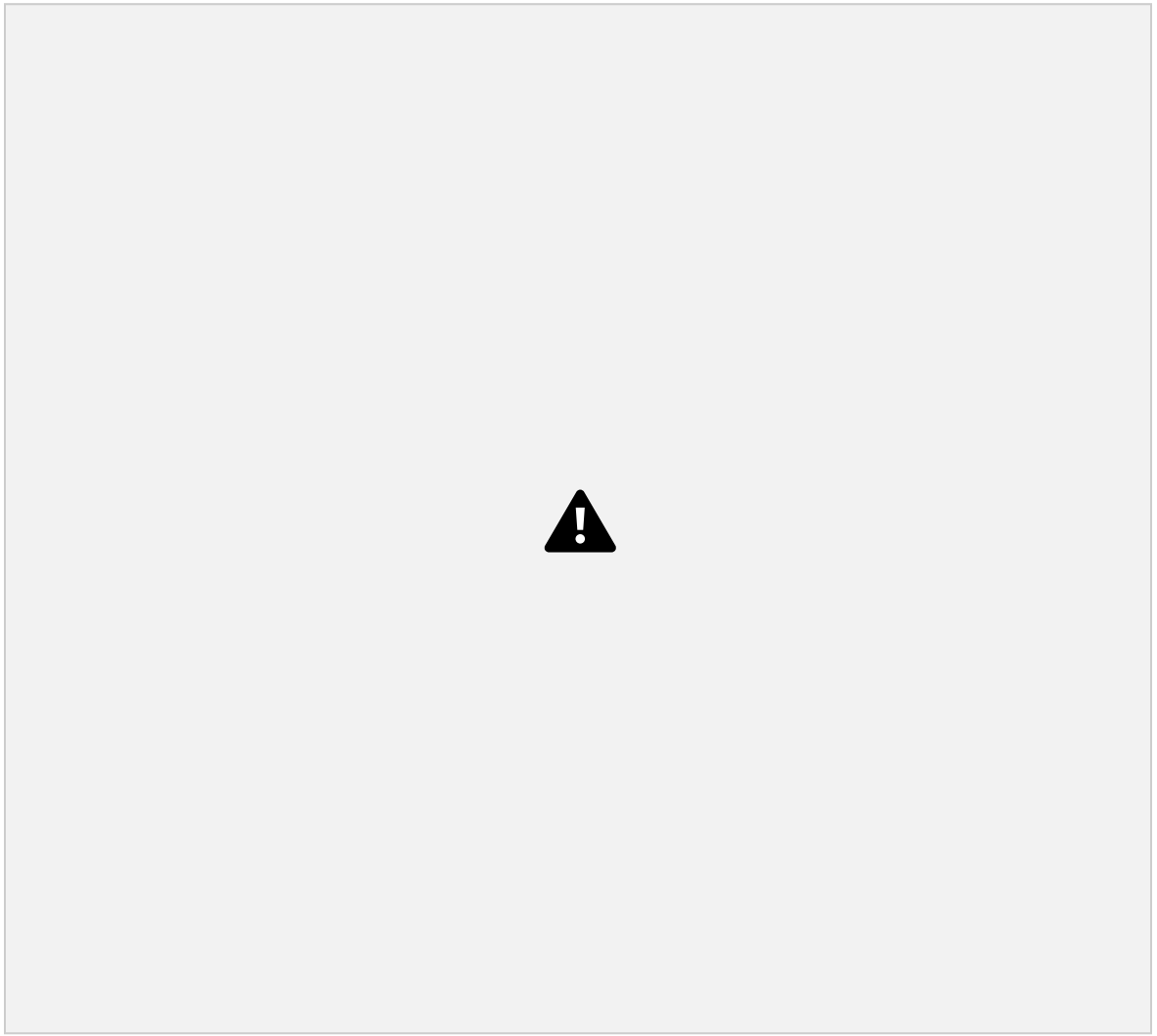
Bản rõ I N F O R M A T I O N S E C U R I T Y Bản mã K V U V V D C B X V R

J G K J Y M K A

- **Nhận xét?**
  - Như vậy có chữ mã khác nhau cho cùng một chữ của bản rõ. • Ví dụ: Chữ I được mã bởi các chữ: K, X, M; chữ N được mã bởi các chữ V, R; ...

  ⇒ Tần suất của các chữ bị là phẳng, nghĩa là tần suất xuất hiện các chữ trên bản mã tương đối đều nhau.
  - Tuy nhiên chưa mất hoàn toàn, do độ dài của khoá có hạn, nên có thể

tạo nên chu kỳ vòng lặp, hoặc việc lặp do ngẫu nhiên



Hệ mật Hill:

Lester S.Hill đưa ra năm 1929

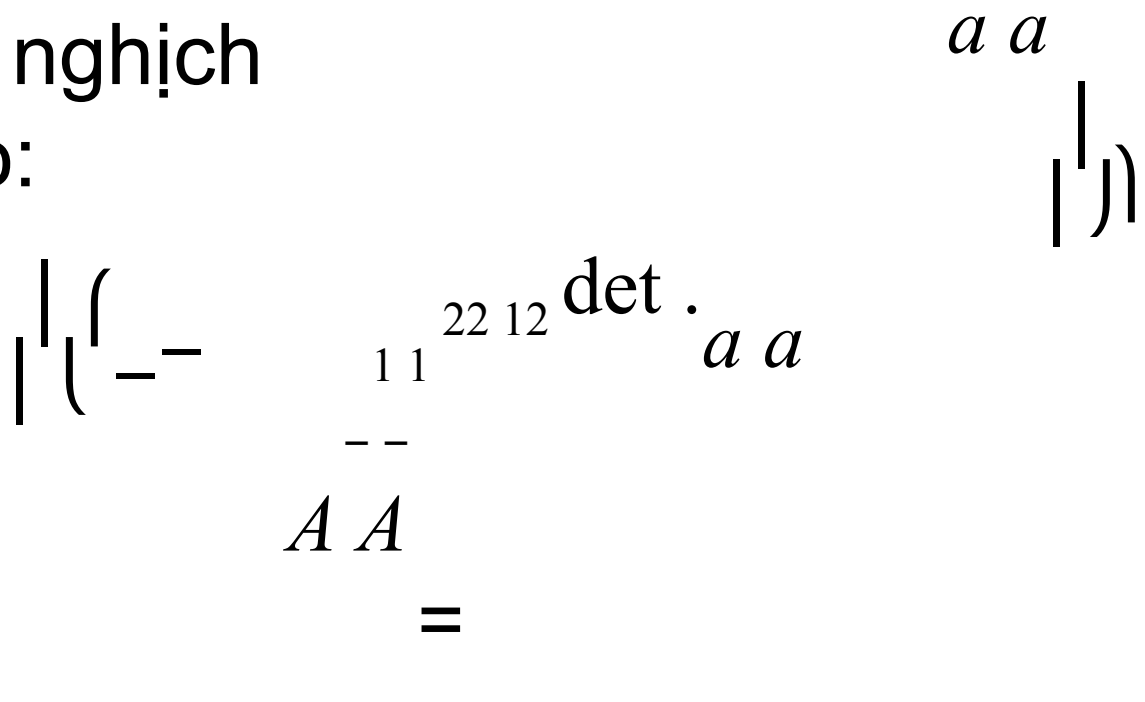
Ý tưởng: lấy *m* tổ hợp tuyến tính của *m* kí tự trong một phần tử bản rõ để tạo ra một phần tử *m* kí tự trong một phần tử của bản mã.

Mô tả:

- Cho ma trận:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

- Định thức detA = $a_{11}.a_{22} - a_{12}.a_{21}$
- Ma trận là khả nghịch $\Leftrightarrow$ det A ≠ 0
- Vì các phép toán tính theo modulo 26 nên phải có điều kiện: UCLN(detA, 26) = 1.
  - Ma trận nghịch đảo:



• **Ví dụ minh hoạ:** • Bản rõ: “july” • Ma trận khoá:

$$k(\ ) = \begin{vmatrix} \\ (\ J3\ 7 \end{vmatrix}$$

11 8

- Tìm bản mã của bản rõ trên
- Từ bản mã thu được tìm bản rõ ban đầu.

- **Các hệ mật thay thế không tuần hoàn**
  - Phép thế lý tưởng là dùng nhiều bảng chữ cái để không nhận diện được phân bố tần suất
  - Điều gì xảy ra nếu văn bản được mã bằng số bảng chữ cái không hạn chế?
  - Vigenere đề xuất hệ mật khoá tự sinh (hệ mật khoá chạy)

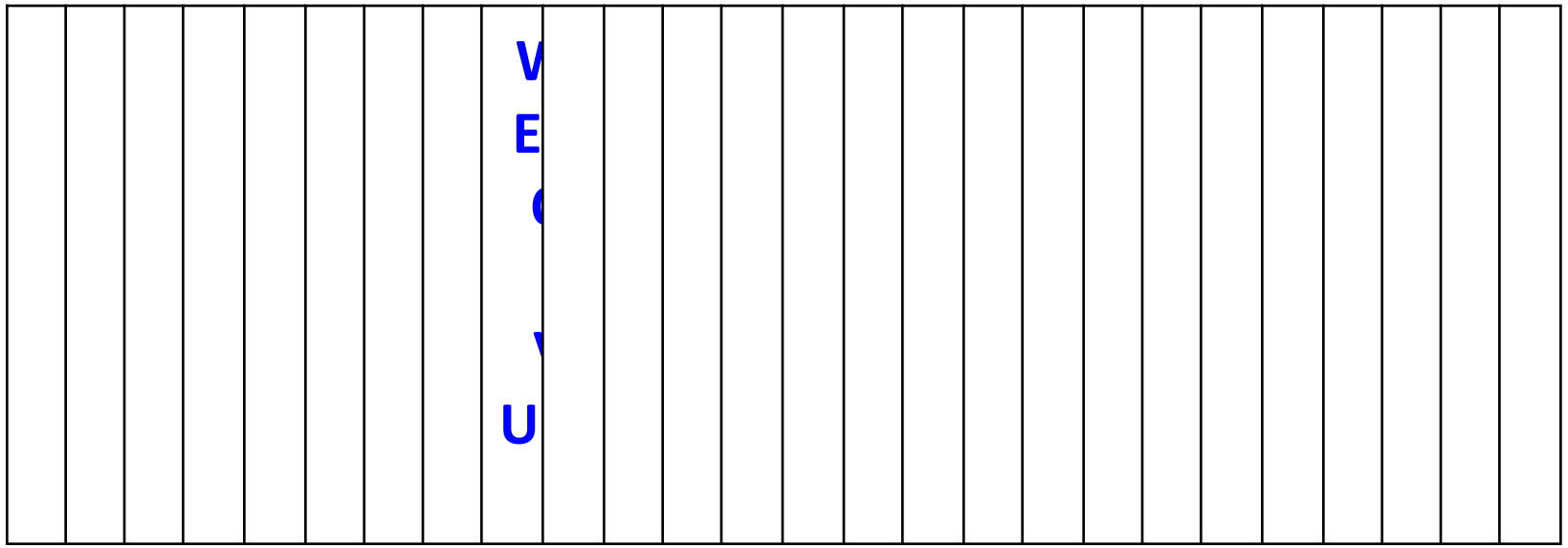
- **Hệ mật khoá chạy:**
  - **Ý tưởng:**
    - Từ khoá được nối tiếp bằng chính bản rõ, sau đó sử dụng mã Vigenere để mã • Khi biết từ khoá, giải được một số chữ của bản rõ rồi dùng chúng giải nốt phần còn lại
    - Sự cải tiến này gây mất khái niệm chu kỳ.
  - **Ví dụ:**
    - Key = deceptive
    - Bản rõ: we are discovered save yourself

- Hãy mã hoá bản rõ trên.
- Giải mã bản mã thu được

Mã hoá bản rõ

Khoá D E C E P T I V E

W E A R E D I S C O V E R E D S A V



Bản rõ

S L V V W L A

Bản mã

Z I C V T W Q N G K Z E I I G A S X S T

O V E R E D S A V

Khoá D E C E P T I V E

Bản mã

Z	I							C	C	Z	

								A							

- **Hệ mật OTP**
  - Do Gillbert Vernam đưa ra 1971
  - Mô tả:

$P = C = K = (Z_2)^n$, $n \geq 1$ là số nguyên, $K \in (Z_2)^n$,

với $x = (x_1, \ldots, x_n)$ và $K = (K_1, \ldots, K_n)$, ta có hàm mã

hoá: $e_K(x) = (x_1 \oplus K_1, \ldots, x_n \oplus K_n)$

Phép mã đồng nhất với phép giải. Nếu $y = (y_1, \ldots, y_n)$ ta có:

$$d_K(y) = (y_1 \oplus K_1, \ldots, y_n \oplus K_n)$$

- **Hệ mật hoán vị (MHV):**
  - **Ý tưởng:**
    - Các chữ trong bản rõ không được thay thế bằng các chữ khác mà chỉ thay đổi vị trí giữa các chữ trong bản rõ.
  - Nhận xét?
    - Bản mã có cùng tần suất xuất hiện các chữ như trong bản gốc • Dễ thám mã

- **MHV:**

- **Ví dụ 1:**
  - m =6; khóa là phép hoán vị π sau: **1 2**

    3 4 5 6

    3 5 1 6 4 2

- Khi đó phép HV ngược π^{-1}:

1 2 3 4 5 6

3 6 1 5 2 4

- Bản rõ: **asecondclasscarriageonthetrain**

- **Mã hóa:**
  - **B1:** Nhóm bản rõ thành các nhóm 6 kí tự

  - **B2:** Mỗi nhóm 6 kí tự sẽ được sắp xếp lại theo theo phép HV π (3, 5, 1, 6, 4, 2), ta có:

- Khi đó ta có bản mã:

EOANCSLSDSACRICARAOTGHNERIENAT

- **Giải mã:**
  - **B1:** Nhóm bản mã thành các nhóm 6 kí tự

  - **B2:** Mỗi nhóm 6 kí tự sẽ được sắp xếp lại theo theo phép HV π^{-1}(3, 6, 1, 5, 2, 4), ta có:

- Khi đó ta có bản rõ tương ứng:

asecondclasscarriageonthetrain

- **Hệ mật tích:**
  - Ý tưởng: kết hợp các hệ mật bằng cách tạo tích của chúng. • Xét các hệ mật có C = P (các hệ mật loại này được gọi là tự đồng cấu)

- **Thám mã một số hệ mật cổ điển:**
  - Nhận xét về các hệ mật thay thế đơn biểu:
    - Mỗi ký tự của bản rõ được ánh xạ đến một ký tự duy nhất

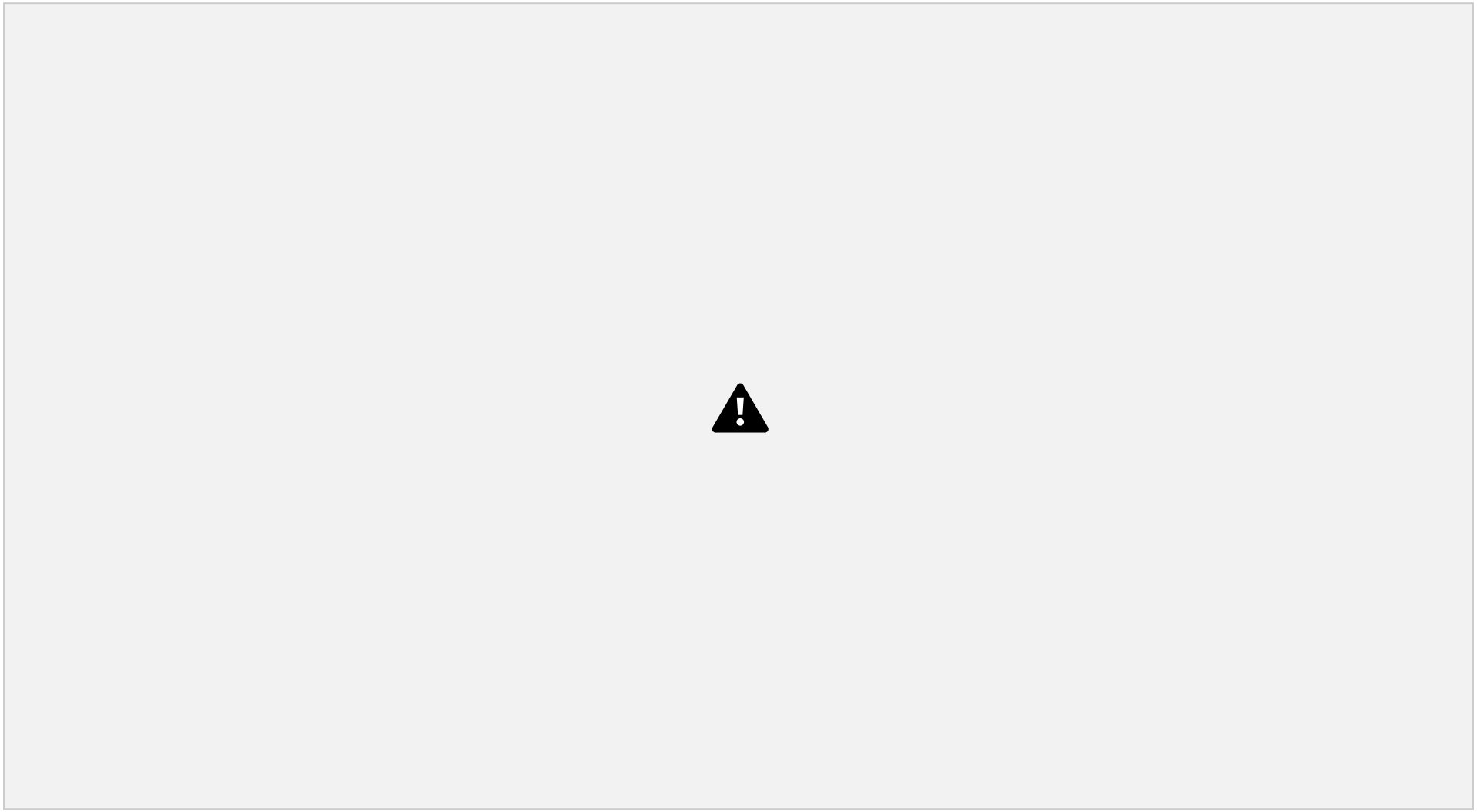
của bản mã.

- Các đặc trưng về ngôn ngữ, tần suất xuất hiện của các chữ trong bản rõ và chữ tương ứng trong bản mã là như nhau

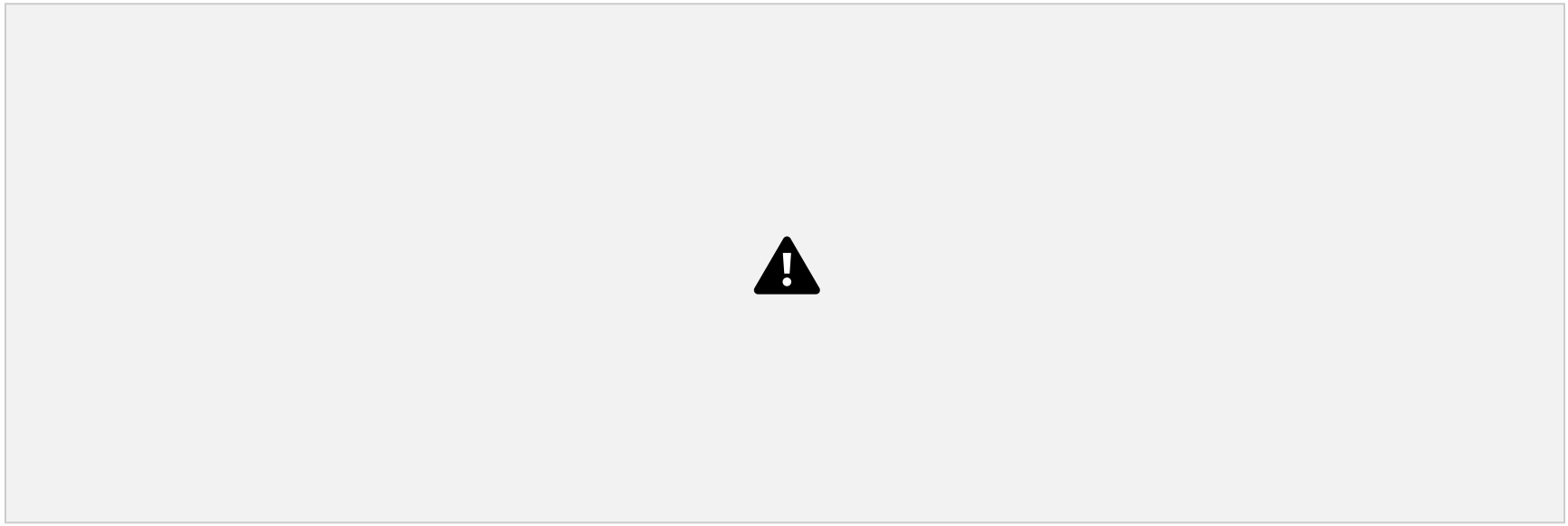
⇒ **Phương pháp thám mã bằng thống kê tần suất!**

- Phương pháp thám mã bằng thống kê tần suất:
  - Ngôn ngữ có tính dư thừa
    - *Ví dụ:* tiếng Anh, 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
    - Các bộ chữ thường xuất hiện: th, nt, lrd, shll, …

- Bằng phương pháp thống kê, ta có thể xây dựng các bảng các tần suất các chữ đơn, cặp hay bộ ba các chữ

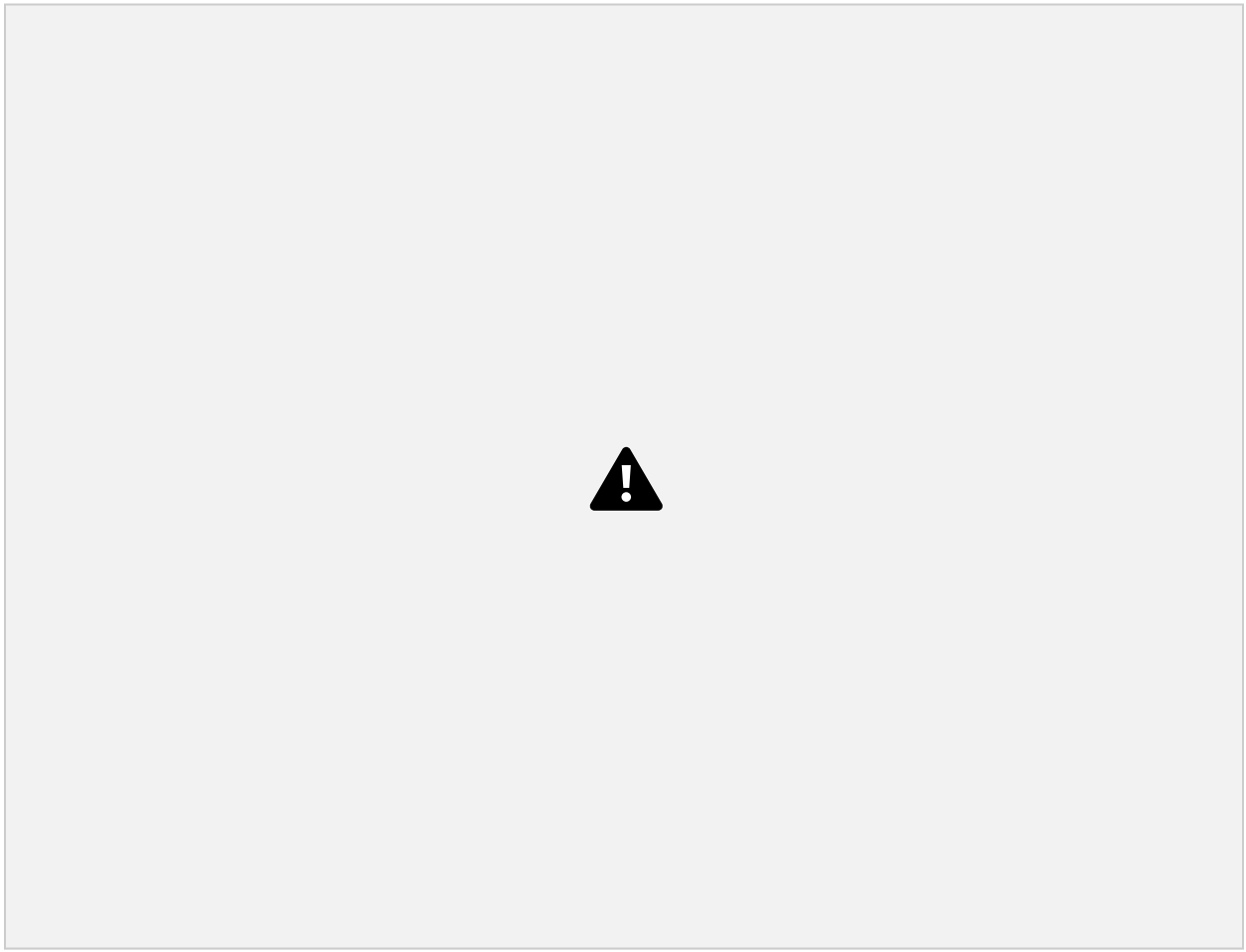


- **Ví dụ:** Giả sử ta có bản mã



- Thám mã Affine bằng phương pháp thống kê tần suất.

- **Bước 1:** xác định tần suất



G	22	I	
N	21	K	5
R	21	L	4
B	16	Q	4
O	16	E	4
W	13	P	3
Z	11	F	3
A	9	T	2
D	9	M	2
H	7	U	1
J	6	V	1
			1

- **Bước 2:** Tách nhóm

- Từ bảng tần suất, giả sử
  - G (6) là mã hóa của E (4)
  - N (13) là mã hóa của T (19)
  - Thiết lập hệ phương trình

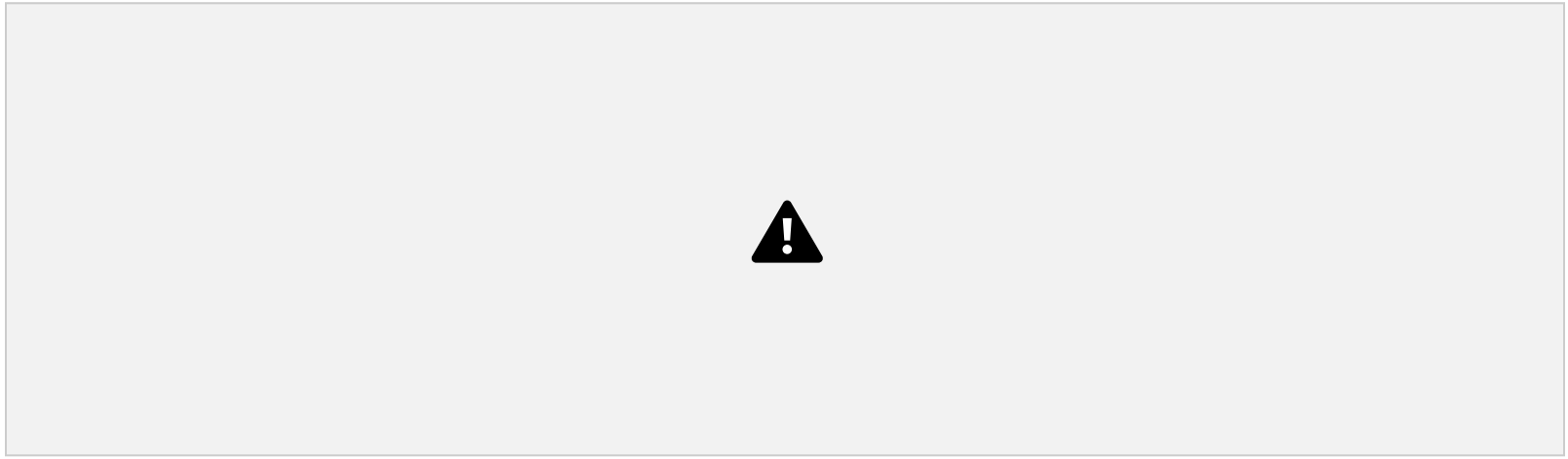
$$4a + b = 6$$

$$19a + b = 13$$

  - Giải hệ được **a = 23, b = - 8 = 18.** Ta có hàm mã: $\mathbf{e_k(x) = 23x + 18 \bmod 26}$
  - Hàm giải mã tương ứng:

$$\mathbf{d_k(y) = 17(y - 18) = (17y + 6) \bmod 26}$$

- **Bước 3:** Tìm bản rõ
  - Từ hàm giải mã, ta có bản rõ tương ứng:



- Nhận xét gì về bản rõ?

- **Quay lại bước 2**
  - Giả sử G là mã hóa của T
  - N là mã hóa của E
  - Ta lại có hệ:

$19a + b = 6$

$4a + b = 13$

- Giải hệ được **a = 3, b = 1.** Ta có hàm mã:

$$\mathbf{e_k(x) = 3x + 1 \bmod 26}$$

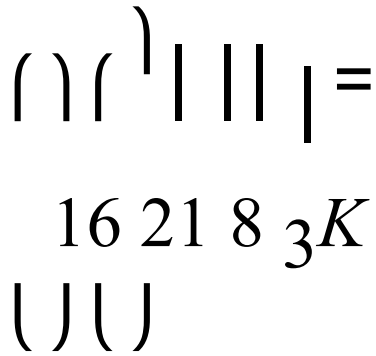
- Hàm giải mã tương ứng:

$$\mathbf{d_k(y) = 9(y - 1) = 9y + 17 \bmod 26}$$

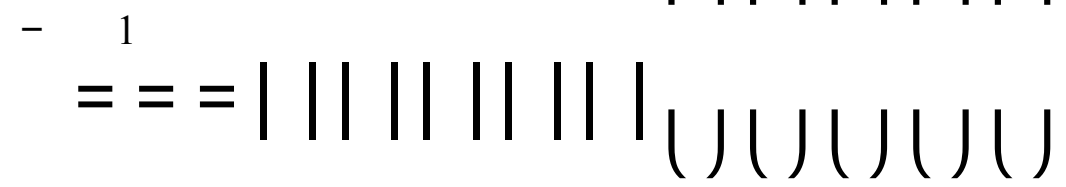
- **Bước 3:** Bản rõ tương ứng

- **Ví dụ minh hoạ:**
  - Giả sử ta dùng hệ mật Hill với bản rõ: "Friday", bản mã tương ứng "ZIQVSO"
  - Hãy tìm ma trận khoá.
  - Giải:
    - Ta có FR (5; 17), ID (8;3), AY (0; 24)
      ZI (25; 8) QV(16; 21); SO (18; 14)

- PT: 25 8 5 17 .



5 17 25 8 9 1 25 8 7 15



8 3 16 21 2 15 16 21 4 19

- Thám mã hệ mã Vigerner: đọc thêm tài liệu [4, 1.2 chapter 1]

Bài 02. Các hệ mật KBM

❖ Mục tiêu:

- Thực hiện làm thành thạo các bài tập minh họa phần lí thuyết về các hệ mật cổ điển học trong Bài 01

- **Bài 1.**
  - Bản rõ P được mã bằng mã Affine với k = (7, 11), đầu ra tiếp tục được mã bằng hệ mã Vigenere với từ khóa là CIPHER thu bản mã

BNNIECQDZSSGHG. Hãy tìm P ban đầu

- **Bài 2.**
  - Giải mã bản mã QQCD thu được khi mã bản rõ bằng hệ mã Hill.

Cho khóa $k = \begin{matrix} 5 & 8 \\ 12 & 7 \end{matrix}$