Chủ đề 2:
Hệ thống Mã hóa đối xứng

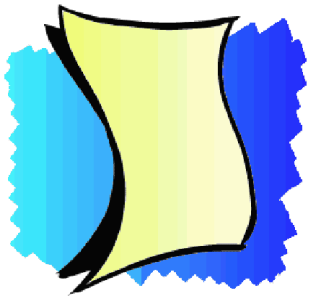
PGS.TS. Trần Minh Triết

KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

Mở đầu

- Hệ thống mã hóa đối xứng (symmetric cryptosystem)
  - Hệ thống mã hóa quy ước (conventional cryptosystem)
  - Hệ thống mã hóa trong đó quy trình mã hóa và giải mã đều sử dụng chung một khoá - *khóa bí mật*.
  - Việc bảo mật thông tin phụ thuộc vào việc bảo mật khóa.

Hệ thống mã hóa đối xứng



Các phương pháp truyền thống

- Các phương pháp truyền thống sử dụng:
  - Phép thay thế (substitution): thay thế 1 từ/ký tự bằng 1 từ/ký tự khác
  - Phép thay đổi vị trí (transposition): các ký tự được thay đổi vị trí
- Việc thay thế/thay đổi vị trí có thể được thực hiện:
  - Đơn ký tự (mono-alphabetic)
  - Đa ký tự (poly-alphabetic)

Phương pháp mã hóa dịch chuyển

Shift Cipher:

- Một trong những phương pháp lâu đời nhất được sử dụng để mã hóa
- Thông điệp được mã hóa bằng cách dịch chuyển xoay vòng từng ký tự đi *k* vị trí trong bảng chữ cái
- Trường hợp với *k*=3 gọi là phương pháp *mã hóa Caesar*.

Phương pháp mã hóa dịch chuyển

Cho $P = C = K = \mathbf{Z}_n$

Với mỗi khóa $k \in K$, định nghĩa:

$e_k(x) = (x + k) \bmod n$ và $d_k(y) = (y - k) \bmod n$ với $x, y \in \mathbf{Z}_n$

$E = \{e_k, k \in K\}$ và $D = \{d_k, k \in K\}$

- Phương pháp đơn giản,
- Thao tác xử lý mã hóa và giải mã được thực hiện nhanh chóng
- Không gian khóa $K = \{0, 1, 2, \ldots, n\text{-}1\} = Z_n$
- Dễ bị phá vỡ bằng cách thử mọi khả năng khóa *k*

Phương pháp mã hóa dịch chuyển

- Ví dụ:
  - Mã hóa một thông điệp được biểu diễn bằng các chữ cái từ A đến Z (26 chữ cái), ta sử dụng Z_{26}.
  - Thông điệp được mã hóa sẽ không an toàn và có thể dễ dàng bị giải mã bằng cách thử lần lượt 26 giá trị khóa *k*.
  - Tính trung bình, thông điệp đã được mã hóa có thể bị giải mã sau khoảng 26/2 = 13 lần thử khóa

- Cho bản mã

  JBCRCLQRWCRVNBJENBWRWN

- Lần lượt thử các khóa k = 0, 1, 2, … 25

  jbcrclqrwcrvnbjenbwrwn
  iabqbkpqvbqumaidmavqvm
  hzapajopuaptlzhclzupul
  gyzozinotzoskygbkytotk
  fxynyhmnsynrjxfajxsnsj
  ewxmxglmrxmqiweziwrmri
  dvwlwfklqwlphvdyhvqlqh
  cuvkvejkpvkogucxgupkpg
  btujudijoujnftbwftojof
  astitchintimesavesnine

- Cho bản mã

  JBCRCLQRWCRVNBJENBWRWN

- Lần lượt thử các khóa k = 0, 1, 2, … 25

  jbcrclqrwcrvnbjenbwrwn
  iabqbkpqvbqumaidmavqvm
  hzapajopuaptlzhclzupul
  gyzozinotzoskygbkytotk
  fxynyhmnsynrjxfajxsnsj
  ewxmxglmrxmqiweziwrmri
  dvwlwfklqwlphvdyhvqlqh
  cuvkvejkpvkogucxgupkpg
  btujudijoujnftbwftojof
  astitchintimesavesnine ← ***k*=9**

Phương pháp mã hóa thay thế

Substitution Cipher:

- Phương pháp mã hóa nổi tiếng
- Được sử dụng phổ biến hàng trăm năm nay
- Thực hiện việc mã hóa thông điệp bằng cách hoán vị các phần tử trong bảng chữ cái hay tổng quát hơn là hoán vị các phần tử trong tập nguồn *P*

Phương pháp mã hóa thay thế

Cho $P = C = Z_n$

K là tập hợp tất cả các hoán vị của n phần tử $0, 1, \ldots, n-1$. Như vậy, mỗi khóa $\pi \in K$ là một hoán vị của n phần tử $0, 1, \ldots, n-1$.

Với mỗi khóa $\pi \in K$, định nghĩa:

$$e_\pi(x) = \pi(x) \quad \text{và} \quad d_\pi(y) = \pi^{-1}(y) \quad \text{với} \quad x, y \in \mathbf{Z}_n$$

$$E = \{e_\pi, \pi \in K\} \quad \text{và} \quad D = \{D_\pi, \pi \in K\}$$

Phương pháp mã hóa thay thế

- Đơn giản, thao tác mã hóa và giải mã được thực hiện nhanh chóng
- Không gian khóa K gồm $n!$ phần tử
- Khắc phục hạn chế của phương pháp Shift Cipher: việc tấn công bằng cách vét cạn các giá trị khóa $k \in K$ là không khả thi

Thật sự an toàn???

Phương pháp mã hóa thay thế

AO VCO JO IBU RIBU

AO VCO JO IBU

?A H?A ?A

MA HOA VA UNG DUNG

Tấn công dựa trên tần số xuất hiện của ký tự trong ngôn ngữ

Phương pháp mã hóa thay thế

L FDPH L VDZ L FRQTXHUHG

L FDPH L VDZ L FRQTXHUHG

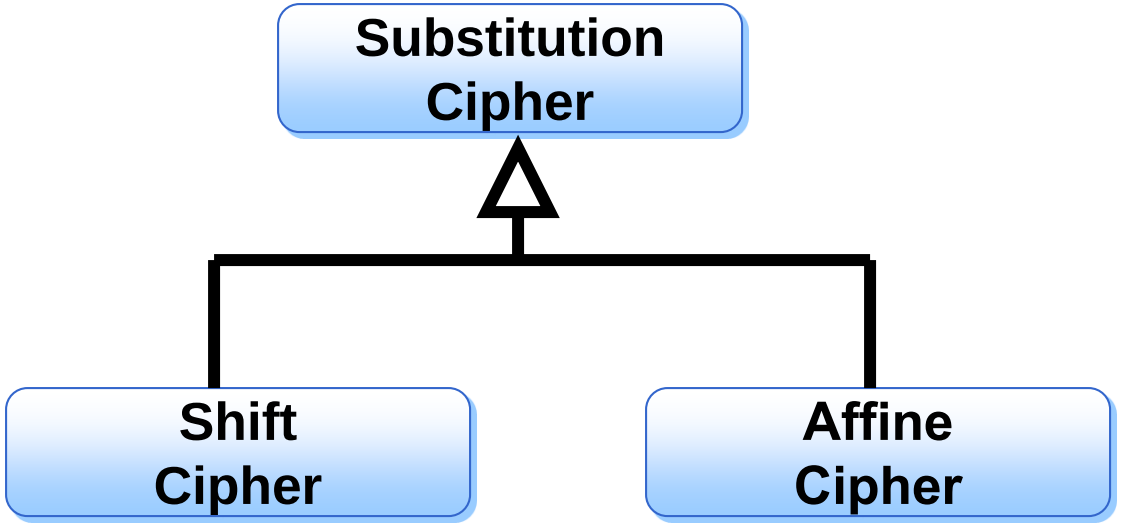
i ?a?e i ?a? i ?????e?e?

i came i saw i conquered

Phương pháp mã hóa thay thế

- Phân tích tần số
  - Ký tự: E > T > R > N > I > O > A > S
  - Nhóm 2 ký tự (digraph): TH > HE > IN > ER > RE > ON > AN > EN
  - Nhóm 3 ký tự (Trigraph): THE > AND > TIO > ATI > FOR > THA > TER > RES

Phương pháp Affine



Phương pháp Affine

Cho $P = C = \mathrm{Z}_n$

$K = \{(a,b) \in \mathbf{Z}_n \times \mathbf{Z}_n : \gcd(a,n) = 1\}$

Với mỗi khóa $k = (a,b) \in K$, định nghĩa:

$e_k(x) = (ax + b) \bmod n$ và $d_k(x) = (a^{-1}(y - b)) \bmod n$ với $x, y \in \mathbf{Z}_n$

$E = \{e_k, k \in K\}$ và $D = \{D_k, k \in K\}$

giải mã chính xác thông tin ???

e_k phải là song ánh

***a* và *n* nguyên tố cùng nhau: gcd(*a*,*n*)=1**

Phương pháp Affine

Ví dụ:

- Khóa
  - Plain: **abcdefghijklmnopqrstuvwxyz**
  - Cipher: **DKVQFIBJWPESCXHTMYAUOLRGZN**
- Mã hóa:
  - Plaintext: **ifwewishtoreplaceletters**
  - Ciphertext: **WIRFRWAJUHYFTSDVFSFUUFYA**

Phương pháp Affine

Gọi $\phi(n)$ là số lượng phần tử thuộc $\mathbf{Z}_n$ và nguyên tố cùng nhau với n.

Nếu $n = \prod_{i=1}^{m} p_i^{e_i}$ *với* p_i *là các số nguyên tố khác nhau và* $e_i \in \mathbf{Z}^+$, $1 \le i \le m$

thì $\phi(n) = \prod_{i=1}^{m} \left(p_i^{e_i} - p_i^{e_i - 1}\right)$.

- n khả năng chọn giá trị b
- $\phi(n)$ khả năng chọn giá trị a
- $n \times \phi(n)$ khả năng chọn lựa khóa $k = (a, b)$

Thuật toán Euclide mở rộng

$$r_0 = q_1 r_1 + r_2,\ 0 < r_2 < r_1$$

$$r_1 = q_2 r_2 + r_3,\ 0 < r_3 < r_2$$

$$\ldots$$

$$r_{m-2} = q_{m-1} r_{m-1} + r_m,\ 0 < r_m < r_{m-1}$$

$$r_{m-1} = q_m r_m$$

$$\gcd(r_0, r_1) = \gcd(r_1, r_2) = \ldots = \gcd(r_{m-1}, r_m) = r_m$$

Thuật toán Euclide

- Xây dựng dãy số:

$$t_0 = 0$$

$$t_1 = 1$$

$$t_j = (t_{j-2} - q_{j-1}t_{j-1}) \bmod r_0 \text{ với } j \geq 2$$

- Nhận xét:

Với mọi $j,\ 0 \leq j \leq m$ *, ta có* $r_j \equiv t_j r_1 (\bmod r_0)$

$$\gcd(r_0, r_1) = 1 \Rightarrow t_m = r_1^{-1} \bmod r_0$$

Phương pháp Vigenere

- Trong phương pháp mã hóa bằng thay thế: với một khóa k được chọn, mỗi phần tử $x \in P$ được ánh xạ vào duy nhất một phần tử $y \in C$.
- Phương pháp Vigenere sử dụng khóa có độ dài m.
- Được đặt tên theo nhà khoa học Blaise de Vigenere (thế kỷ 16)
- Có thể xem phương pháp mã hóa Vigenere bao gồm m phép mã hóa bằng dịch chuyển được áp dụng luân phiên nhau theo chu kỳ
- Không gian khóa K của phương pháp Vigenere có số phần tử là n^m
- Ví dụ: n=26, m=5 thì không gian khóa ~1.1 x 10^7

Phương pháp Vigenere

Chọn số nguyên dương m. Định nghĩa $P = C = K = (\mathbf{Z}_n)^m$

$K = \{(k_1, k_2, \ldots, k_m) \in (Z_n)^m\}$

Với mỗi khóa $k = (k_1, k_2, \ldots, k_m) \in K$, định nghĩa:

$e_k(x_1, x_2, \ldots, x_m) = ((x_1 + k_1) \bmod n, (x_2 + k_2) \bmod n, \ldots, (x_m + k_m) \bmod n)$

$d_k(y_1, y_2, \ldots, y_m) = ((y_1 - k_1) \bmod n, (y_2 - k_2) \bmod n, \ldots, (y_m - k_m) \bmod n)$

với $x, y \in (\mathbf{Z}_n)^m$.

Phương pháp Vigenere

- Ví dụ: m = 6 và keyword là CIPHER
- Suy ra, khóa k = (2, 8, 15, 7, 4, 17)
- Cho plaintext: **`thiscryptosystemisnotsecure`**

19	7	8	18	2	17
2	8	15	7	4	17
21	15	23	25	6	8

24	15	19	14	18	24
2	8	15	7	4	17
0	23	8	21	22	15

18	19	4	12	8	18
2	8	15	7	4	17
20	1	19	19	12	9

13	14	19	18	4	2
2	8	15	7	4	17
15	22	8	25	8	19

20	17	4
2	8	15
22	25	19

Phương pháp mã hóa Hill

- Phương pháp Hill (1929)
- Tác giả: Lester S. Hill
- Ý tưởng chính:
  - Sử dụng *m* tổ hợp tuyến tính của *m* ký tự trong plaintext để tạo ra *m* ký tự trong ciphertext
- Ví dụ:

$$y_1 = 11x_1 + 3x_2$$
$$y_2 = 8x_1 + 7x_2.$$

$$(y_1, y_2) = (x_1, x_2) \begin{pmatrix} 11 & 8 \\ 3 & 7 \end{pmatrix}$$

Phương pháp mã hóa Hill

Chọn số nguyên dương m. Định nghĩa:

$P = C = (\mathbf{Z}_n)^m$ và K là tập hợp các ma trận $m \times m$ khả nghịch

Với mỗi khóa $k = \begin{pmatrix} k_{1,1} & k_{1,2} & \cdots & k_{1,m} \\ k_{2,1} & \cdots & \cdots & k_{2,m} \\ \vdots & \vdots & & \vdots \\ k_{m,1} & k_{m,2} & \cdots & k_{m,m} \end{pmatrix} \in K$, định nghĩa:

$$e_k(x) = xk = (x_1, x_2, ..., x_m) \begin{pmatrix} k_{1,1} & k_{1,2} & \cdots & k_{1,m} \\ k_{2,1} & \cdots & \cdots & k_{2,m} \\ \vdots & \vdots & & \vdots \\ k_{m,1} & k_{m,2} & \cdots & k_{m,m} \end{pmatrix} \text{ với } x = (x_1, x_2, ..., x_m) \in P$$

và $d_k(y) = yk^{-1}$ với $y \in C$.

Mọi phép toán số học đều được thực hiện trên $\mathbf{Z}_n$.

Xác định ma trận nghịch đảo

- Cách xác định ma trận nghịch đảo
- Cho ma trận K khả nghịch, cần xác định K^{-1}
- Thực hiện:
  - Biến đổi sơ cấp từ ma trận $(K \mid I_n)$ thành $(I_n \mid K^{-1})$
  - Các phép biến đổi sơ cấp:
    - Nhân 1 dòng với 1 số khác 0
    - Thay 1 dòng bằng cách lấy dòng đó cộng hay trừ α lần dòng khác

Phương pháp mã hóa bằng hoán vị

- ý tưởng của các phương pháp đã trình bày: thay thế mỗi ký tự trong thông điệp nguồn bằng một ký tự khác để tạo thành thông điệp đã được mã hóa.
- Ý tưởng chính của phương pháp mã hóa hoán vị (Permutation Cipher) là vẫn giữ nguyên các ký tự trong thông điệp nguồn mà chỉ thay đổi vị trí các ký tự

Phương pháp mã hóa bằng hoán vị

Chọn số nguyên dương m. Định nghĩa:

$P = C = (\mathbf{Z}_n)^m$ và K là tập hợp các hoán vị của m phần tử $\{1, 2, \ldots, m\}$

Với mỗi khóa $\pi \in K$, định nghĩa:

$$e_\pi(x_1, x_2, \ldots, x_m) = \left(x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(m)}\right) \text{ và}$$

$$d_\pi(y_1, y_2, \ldots, y_m) = \left(y_{\pi^{-1}(1)}, y_{\pi^{-1}(2)}, \ldots, y_{\pi^{-1}(m)}\right)$$

với π^{-1} hoán vị ngược của π

Phương pháp mã hóa bằng hoán vị

$$\pi = \begin{array}{|c|c|c|c|c|c|} \hline 1 & 2 & 3 & 4 & 5 & 6 \\ \hline 3 & 5 & 1 & 6 & 4 & 2 \\ \hline \end{array}. \qquad \pi^{-1} = \begin{array}{|c|c|c|c|c|c|} \hline 1 & 2 & 3 & 4 & 5 & 6 \\ \hline 3 & 6 & 1 & 5 & 2 & 4 \\ \hline \end{array}.$$

shesellsseashellsbytheseashore

shesel | lsseas | hellsb | ythese | ashore

EESLSH | SALSES | LSHBLE | HSYEET | HRAEOS

EESLSHSALSESLSHBLEHSYEETHRAEOS

Xác định ma trận nghịch đảo

$$\left(\begin{array}{ccc|ccc} 2 & 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 3 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{array}\right)$$

(3) → (3) - (1)
(1) → (1) - (2)

$$\left(\begin{array}{ccc|ccc} 2 & 0 & -4 & 1 & -1 & 0 \\ 0 & 1 & 3 & 0 & 1 & 0 \\ 0 & 0 & 2 & -1 & 0 & 1 \end{array}\right)$$

(1) → (1) + 2(3)
(3) → ½ * (3)
(2) → (2) – 3(3)

$$\left(\begin{array}{ccc|ccc} 2 & 0 & 0 & -1 & -1 & 2 \\ 0 & 1 & 0 & 3/2 & 1 & -3/2 \\ 0 & 0 & 1 & -1/2 & 0 & 1/2 \end{array}\right)$$

(1) → ½ (1)

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1/2 & -1/2 & 1 \\ 0 & 1 & 0 & 3/2 & 1 & -3/2 \\ 0 & 0 & 1 & -1/2 & 0 & 1/2 \end{array}\right)$$