Mã hóa thông tin

Mã hóa thông tin

- Giới thiệu mô hình mã hóa
  - Mã đối xứng
  - Mã hóa phi đối xứng
- Giới thiệu hàm băm
- Giới thiệu mô hình truyền khóa
- Ứng dụng mã hóa, hàm băm trong bảo vệ và kiểm tra dữ liệu

Mô hình hệ thống

- Hệ thống mã hóa (cryptosystem) là một bộ năm (P, C, K, E, D) thỏa mãn các điều kiện sau:
  1. *Tập nguồn P là tập hữu hạn tất cả các bản tin nguồn cần mã hóa có thể có*
  2. *Tập đích C là tập hữu hạn tất cả các bản tin có thể có sau khi mã hóa*
  3. *Tập khóa K là tập hữu hạn các khóa có thể được sử dụng*

Mô hình hệ thống (t)

- (P, C, K, E, D) :

  4. E, D là tập luật mã hóa và giải mã. Với mỗi khóa k tồn tại một luật mã hóa $e_k \in E$ và luật giải mã tương ứng $d_k \in D$. Luật mã hóa e_k: $P \rightarrow C$ và d_k: $C \rightarrow D$ thỏa mãn. $d_k(e_k(x))=x, \forall x \in P$.

Phân loại mã hóa

- Mã đối xứng – mật – quy ước
  - Từ e_k có thể suy ra d_k và ngược lại
- Mã phi đối xứng – công khai
  - Từ e_k không thể suy ra được d_k và ngược lại

Một số mã hóa kinh điển

- Mã hóa dịch vòng
- Phương pháp thay thế
- Phương pháp Affine
- Phương pháp Vigenere
- Phương pháp Hill
- Phương pháp hoán vị

Mã hóa dịch vòng

- $P=C=K=Z_n$
- Khóa k định nghĩa $k \in K$ định nghĩa
- $e_k(x)=(x+k) \bmod n$
- $d_k(y)=(y-k) \bmod n$
- $x, y \in Z_n$
- $E=\{e_k, k \in K\}$
- $D=\{d_k, k \in K\}$

Mã hóa dịch vòng (t)

- Ví dụ: trong tiếng anh có a->z vậy n=26
- Chọn k=12 vậy
- NOTHINGIMPOSSIBLE
- Thứ tự là:
- 13,14,19,7,8,13,6,8,12,15,14,18,18,8,1,11,4
- Sau khi mã hóa là:
- 25,0,5,19,20,25,18,20,24,1,0,4,4,20,13,23,16
- ZAFTUZSUYBAEEUNXQ

Mã hóa dịch vòng (t)

- Thực hiện đơn giản
- Không gian khóa bé (26), dễ tấn công:
  - Vét cạn
  - Thống kê ký tự

Mã hóa thay thế

- $P=C=Z_n$
- K tập tất cả hoán vị của n phần tử
- k: là một hoán vị π
- $e_k(x)= \pi(x)$
- $d_k(y)= \pi^{-1}(y)$

Mã hóa thay thế (t)

- NOTHINGIMPOSSIBLE
- Thành
- WZCILWMLNTZXXLUPK
- Tra bảng ngược lại khi nhận
- NOTHINGIMPOSSIBLE

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

Mã hóa thay thế (t)

- Thời gian thực hiện ngắn
- Không gian khóa là n! khá lớn
- Tấn công theo phương pháp thống kê

Phương pháp Affine

- $P=C=Z_n$
- $K=\{(a,b) \in Z_n x Z_n : \gcd(a,n)=1\}$
- $e_k(x) = (ax + b) \bmod n$
- $d_k(x) = (a^{-1}(y-b)) \bmod n$
- $x, y \in Z_n$
- $E=\{e_k, k \in K\}$
- $D=\{d_k, k \in K\}$

Phương pháp Affine (t)

- Trường hợp riêng của thay thế
- Tính toán đơn giản
- Số lượng khóa không lớn

Phương pháp Vigenere

- $P=C=K=(Z_n)^m$
- $K=\{(k_1, k_2,\dots ,k_r) \in (Z_n)^r\}$
- $e_k(x_1, x_2, ..x_r)=((x_1+k_1) \bmod n, \dots,(x_r+k_r) \bmod n)$
- $d_k(y_1, \dots, y_r)=((y_1-k_1) \bmod n), \dots,(y_r-k_r) \bmod n)$

Phương pháp Vigenere (t)

- Thuật toán này là mở rộng thuật toán dịch vòng với khóa là bộ nhiều khóa dịch vòng
- Thực hiện đơn giản
- Không gian khóa lớn n^m

Phương pháp Hill

- $P=C=(Z_n)^m$
- K là tập hợp ma trận mxm khả nghịch

$$k=\begin{pmatrix} k_{1,1} & k_{1,2} & \cdots & k_{1,m} \\ k_{2,1} & \cdots & \cdots & k_{2,m} \\ \vdots & \vdots & & \vdots \\ k_{m,1} & k_{m,2} & \cdots & k_{m,m} \end{pmatrix} \in K$$

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

Phương pháp Hill

- Thực hiện đơn giản (dựa phép nhân ma trận)
- Không gian khóa lớn n^{mxm}

Phương pháp hoán vị

- $P=C=(Z_n)^m$
- $\pi \in K$ là một hoán vị
- $e_\pi(x_1, \ldots, x_m) = (x_{\pi(1)}, \ldots, x_{\pi(m)})$
- $d_\pi(y_1, \ldots, y_m)=(y_{\pi^{-1}(1)}, \ldots, y_{\pi^{-1}(m)})$

Phương pháp hoán vị (t)

- Trường hợp riêng của ma Hill
- Thực hiện đơn giản
- Không gian mã hóa là m!

Một số mã hóa tiêu chuẩn

- DES – Data Encryption Standard
- AES – Advanced Encryption Standard

DES

- $x \in$ P dãy 64 bit
- $k \in$ K dãy 56 bit
- Khóa thực tế sử dụng 48 bit
- Sử dụng tripple DES sử dụng 3 quá trình với 3 khóa khác nhau
- Có thể bị phá mã trong khoảng thời gian ngắn
- Tài liệu tham khảo

AES

- Sử dụng những khóa có độ dài là 128, 192 hoặc 256.
- Sử dụng cấu trúc toán đơn giản, thời gian thực hiện thuận tiện
- Đến hiện tại được xem là an toàn

Hàm băm

- Chuyển đổi một thông điệp có độ dài bất kỳ thành một độ dài cố định
- Hàm băm không là hàm song ánh

Hàm băm

- Sử dụng để kiểm tra tính toàn vẹn cho dữ liệu
- Sử dụng để đại diện cho phần chữ ký
- Sử dụng lưu trữ thông tin kiểm chứng (mật khẩu, …)
- Tấn công bằng phương pháp đụng độ

Hàm băm (t)

- Các hàm băm được sử dụng
  - MD4
  - MD5
  - SHA-1
  - SHA-256
  - SHA-384
  - SHA-512

Mã hóa công khai

- Dựa trên các hàm cửa sập một chiều
  - Phân tích thừa số nguyên tố (RSA)
  - Bài toán logarit rời rạc (ECC)
- Sử dụng hai khóa khác nhau, độc lập (không thể phân tích được khóa từ khóa còn lại)

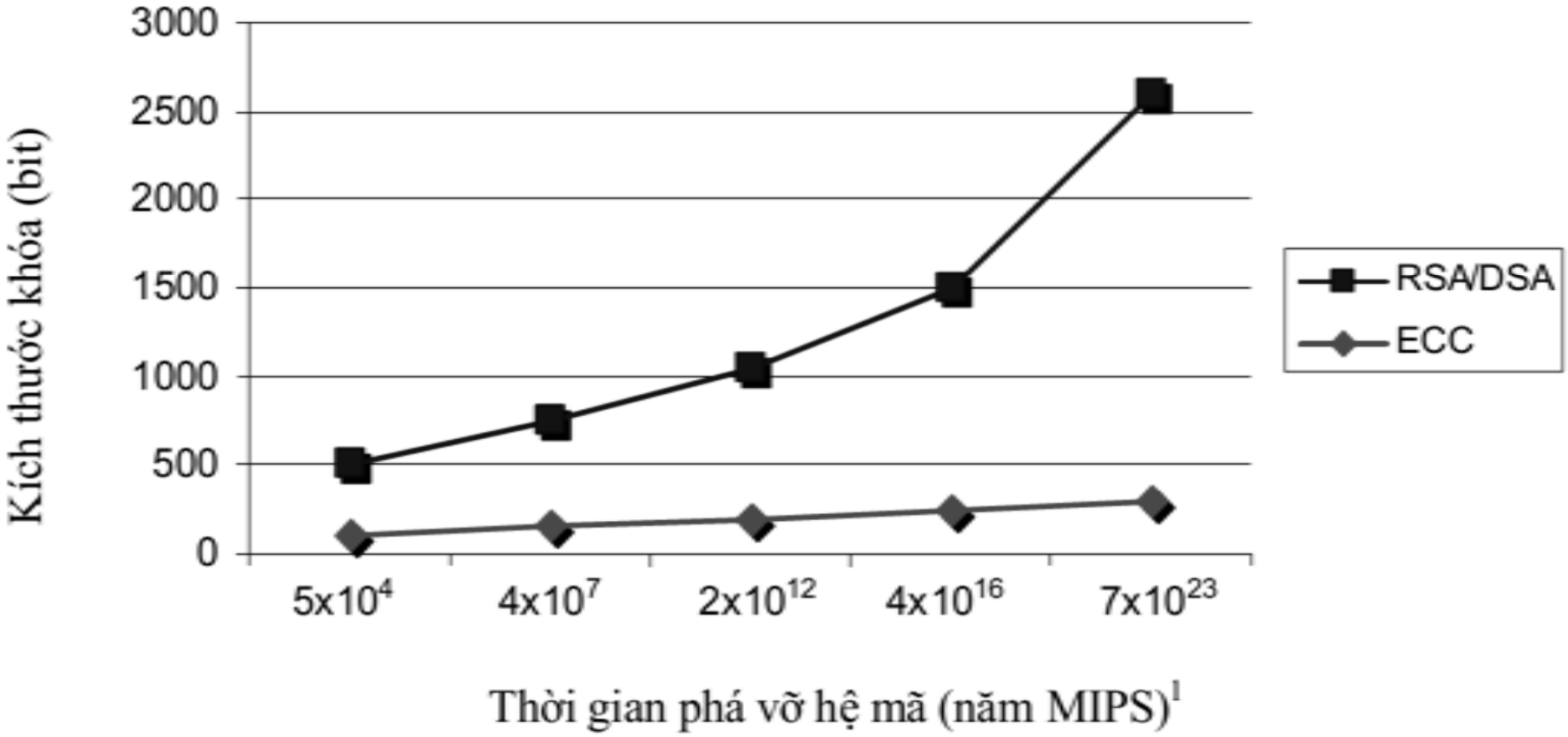
RSA

- Dựa trên phép tính số nguyên tố lớn
  - Khó khăn phân tích thừa số nguyên tố của n
  - Dựa trên phép mũ số nguyên tố lớn
- Thời gian thực hiện chậm
- Đảm bảo an toàn với 512 bit
- Tài liệu tham khảo

ECC

- Mã hóa dựa trên các đường cong eliptic
  - Tìm được đường cong
  - Tìm được nghiệm trên đường cong thỏa mãn số bậc của nghiệm lớn
- Thời gian tính toán nhanh hơn RSA
- Thời gian phá mã chậm hơn RSA
- Khó tìm được đường cong, nghiệm thỏa mãn điều kiện đã cho

So sánh


Kích thước khóa (bit)
3000
2500
2000
1500
1000
500
0
5x10^4
4x10^7
2x10^12
4x10^16
7x10^23
Thời gian phá vỡ hệ mã (năm MIPS)[1]
RSA/DSA
ECC

Mã hóa thông tin

- Giới thiệu mô hình mã hóa
  - Mã đối xứng
  - Mã hóa phi đối xứng
- Giới thiệu hàm băm
- Giới thiệu mô hình truyền khóa
- Ứng dụng mã hóa, hàm băm trong bảo vệ và kiểm tra dữ liệu

Bài tập

- Tìm hiểu thêm về mô hình RSA, ECC
- Thuật toán để thực hiện mã hóa DES, AES
- Thuật toán để tính MD5, SHA…