Thuật toán mã hóa RSA

Mã hóa (Cryptology)

- Mã hóa thông tin
  - Mã hóa Ceasar: Dịch các chữ cái 3 vị trí, ba chữ cái cuối thay thế bằng 3 chữ cái đầu.
  - Ví dụ: A thay bằng D, X thay bằng A
- Biểu diễn ý tưởng phép dịch 3 vị trí?
  - Có 26 chữ cái trong bảng chữ cái Tiếng Anh. Đánh số thứ tự cho chúng từ 0, 1, 2, 3, ..., 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
  - Mã hóa của chữ cái có số thứ tự p là: $f(p) = (p + 3) \bmod 26$

Mã hóa

- Mã hóa thông tin sử dụng phép dịch 3 vị trí
  - Mã hóa của chữ cái có số thứ tự p là: $f(p) = (p + 3) \bmod 26$
- Mã của các chữ cái

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

- Mã hóa thông tin: I LIKE DISCRETE MATH

Mã hóa

- Mã hóa thông tin sử dụng dịch chuyển 3 vị trí
  - Mã hóa của chữ cái có số thứ tự p là: $f(p) = (p + 3) \bmod 26$
- Mã của các chữ cái

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

- Mã hóa thông tin: I LIKE DISCRETE MATH
  - L OLNH GLYFUHVH PDVK

Mã hóa

- Làm thế nào để giải mã?
  - Mã hóa của chữ cái có số thứ tự p là: $f(p) = (p + 3) \bmod 26$
- Mã của các chữ cái

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 được sử dụng để giải mã thông tin:
  - $f^{-1}(p) = (p - 3) \bmod 26$

Hệ mật mã (Cryptosystem)

- Một hệ mật mã là một bộ-5 (P, C, K, E, D)
  - P là tập xâu rõ (gốc, plain text), C là tập xâu mã hóa, K là tập khóa (tập các khóa có thể có), E là tập các hàm mã hóa, D là tập các hàm giải mã.
- Ta kí hiệu E_k là một hàm trong E tương ứng với khóa k và D_k là một hàm trong D thực hiện giải mã xâu mã hóa được mã hóa bằng E_k, tức là $D_k(E_k(p)) = p$, với mọi xâu rõ p.

Mã hóa Khóa công khai

- Mã hóa dịch chuyển là ví dụ của hệ mật mã khóa bí mật
  - Với khóa mã hóa, bạn có thể dễ dàng tìm được khóa giải mã
    - Cần trao đổi khóa một cách bí mật
- Hệ mật mã khóa công khai là một cách khác
  - Được giới thiệu vào những năm 1970
  - Một khóa mã hóa công khai với các khóa giải mã bí mật
    - Cần rất nhiều thời gian (hàng tỉ năm tính toán) để có thể khôi phục được bản thông tin rõ khi không biết được khóa giải mã.

Hệ mật mã RSA

- Ronald **R**ivest, Adi **S**hamir và Leonard **A**dleman (1976) từ Massachusetts Institute of Technology.

- Mỗi người có một khóa mã hóa (n, e)
  - n (số mô đu lô): tích của hai số nguyên tố lớn p và q, (mỗi số có thể lên tới 200 chữ số hoặc hơn).
  - e: nguyên tố cùng nhau với $(p-1)(q-1)$

Mã hóa RSA

- Bản rõ M đầu tiên được chuyển thành chuỗi các cặp chữ số
  - A là 00, B là 01, …, và J là 09, …
  - Nối các cặp số này thành xâu của các số
  - Chia xâu này thành các khối (block) cùng cỡ gồm 2N số
    - 2N: số chẵn lớn nhất sao cho số 2525 . . . 25 với 2N số chữ số không vượt quá n
  - Từ đó, M được dịch thanh chuỗi của các số $m_1, m_2, \ldots, m_k$ với một số nguyên k nào đó.
  - Biến đổi mỗi block m_i thành block mã hóa c_i
    - $C = M^e \bmod n$ (xem Thuật toán 5 trong Phần 4.2)

Mã hóa RSA

- **Ví dụ**
  - Mã hóa "STOP" dùng RSA với khóa $(2537, 13)$. Lưu ý $2537 = 43 \times 49$, $p = 43$, $q = 59$ là các số nguyên tố, và $\gcd\big(e, (p-1)(q-1)\big) = \gcd(13, 42 \times 58) = 1$
- **Đáp án:**
  - Để mã hóa, ta dịch các chữ cái "STOP" sang dạng số tương đương của chúng. Nhóm chúng lại thành các block 4 chữ số (vì $2525 < 2537 < 252525$), ta có: 1819 1415
  - Ta mã hóa từng block sử dụng:
    $$C = M^{13} \textbf{ mod } 2537$$
    Tính toán sử dụng nhân mô đun nhanh ta có:
    $1819^{13} \textbf{ mod } 2537 = 2081$ và $1415^{13} \textbf{ mod } 2537 = 2182$.
  - Bản mã là 2081 2182.

Giải mã RSA

- Bản rõ có thể được khôi phục nhanh chóng khi biết khóa giải mã d
  - Là nghịch đảo của e modulo $(p-1)(q-1)$
  - Nghĩa là $de \equiv 1 \left(\mathbf{mod}\ (p-1)(q-1)\right)$
  - $de = 1 + k(p-1)(q-1)$
  - $C^d \equiv (M^e)^d = M^{de} = M^{1+k(p-1)(q-1)} (\mathbf{mod}\ n)$
  - Đã được chứng minh rằng: $C^d \equiv M (\mathbf{mod}\ pq)$

Giải mã RSA

- **Ví dụ:**
  - Giả sử ta nhận được bản mã 0981 0461. Bản giải mã là gì nếu bản mã sử dụng mã hóa RSA từ **Ví dụ** trước?
- **Đáp án:**
  - Bản mã được mã hóa bằng RSA với $n = 43 \times 59$ và số mũ 13.
  - Như Bài tập 2 trong Phần 4.4, $d = 937$ là nghịch đảo của $13 \bmod 42 \times 58 = 2436$. Ta dùng 937 là số mũ giải mã. Do đó, để giải mã block C, ta tính $M = C^{937} \bmod 2537$
  - Để giải mã thông tin, ta dùng thuật toán lũy thừa mô đun nhanh để tính $0981^{937} \bmod 2537 = 0704$ và $0461^{937} \bmod 2537 = 1115$. Do đó, dạng số gốc của thông tin là 0704 1115.
  - Dịch ngược lại Tiếng Anh, ta thấy thông tin trong bản rõ là "HELP".

RSA: Hệ mật mã khóa công khai

- Hệ mật mã khóa công khai
  - Khi biết được 2 số nguyên tố lớn p và q, mỗi số có hơn 200 chữ số, có thể tìm nhanh được:
    - Số nguyên e nguyên tố cùng nhau với $(p-1)(q-1)$
    - Nghịch đảo d của e **mod** $(p-1)(q-1)$
  - **Tuy nhiên**, không có phương pháp nào giải mã được thông tin mà không dựa vào việc tìm 2 số nguyên tố lớn p và q khi biết n
    - Các phương pháp phân tích thừa số hiệu quả nhất đến nay (2010) cần hàng tỉ năm để phân tích số nguyên 400 chữ số.
    - Không có thuật toán thời gian đa thức nào thực hiện phân tích thừa số các số nguyên lớn cho đến thời điểm này.

Các giao thức mã hóa

- Trao đổi khóa (bí mật) theo giao thức Diffie-Hellman
  - Hai bên có thể sử dụng để trao đổi khóa bí mật qua kênh giao tiếp không an toàn mà không cần chia sẻ bất cứ thông tin gì trong quá khứ.
  - Giao thức trao đổi khóa Diffie-Hellman (1976)
    Giả sử Alice và Bob muốn chia sẻ một khóa chung. Giao thức gồm các bước sau, với các phép toán trên tập Z_p:
    1) Alice và Bob đồng ý sử dụng số nguyên tố p và căn nguyên thủy a của p.
    2) Alice chọn một số nguyên bí mật k_1 và gửi a^{k_1} **mod** p cho Bob.
    3) Bob chọn một số nguyên bí mật k_2 và gửi a^{k_2} **mod** p cho Alice.
    4) Alice tính $\left(a^{k_2}\right)^{k_1}$ **mod** p.
    5) Bob tính $\left(a^{k_1}\right)^{k_2}$ **mod** p.

    Khi kết thúc giao thức, Alice và Bob có khóa chia sẻ đã được tính ra, đó là: $\left(a^{k_1}\right)^{k_2}$ **mod** $p = \left(a^{k_2}\right)^{k_1}$ **mod** p.

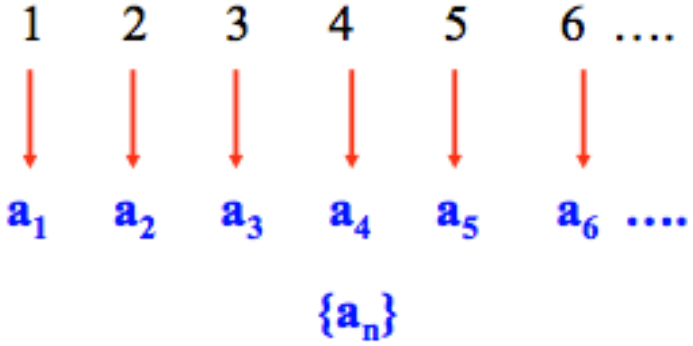
Giao thức trao đổi khóa Diffie-Hellman

- Không có phương pháp nào hiệu quả đến thời điểm hiện nay có thể tìm được khóa bí mật (được chia sẻ) nếu chỉ sử dụng thông tin được công khai.
  - Không khả thi về mặt tính toán khi p và a đủ lớn.
  - Với khả năng tính toán hiện thời, hệ thống được coi là không thể bị phá khi p có hơn 300 kí tự số và k_1 và k_2 có hơn 100 kí tự số.

Chuỗi [số] & Tổng

Chuỗi (sequence)

- Một chuỗi là một ánh xạ từ một tập con của một tập các số nguyên (thường là tập {0, 1, 2,…} hoặc tập {1, 2, 3,…}) đến một tập S.
  - Ta dùng a_n để kí hiệu ảnh của số nguyên n.
    - a_n là một phần tử (term) của chuỗi.
  - $\{a_n\}$ được dùng để biểu diễn một chuỗi $a_1, a_2, a_3,\ldots$
    - Chú ý: "{}" giống với kí hiệu dùng trong tập hợp.



Chuỗi

Ví dụ:

- $a_n = n^2$, với n = 1, 2, 3, …
  - Các phần tử của chuỗi là gì?
    - 1, 4, 9, 16, 25, …
- $a_n = (-1)^n$, với n = 0, 1, 2, 3, …
  - Các phần tử của chuỗi là gì?
    - 1, -1, 1, -1, 1, …
- $a_n = 2^n$, với n = 0, 1, 2, 3, …
  - Các phần tử của chuỗi là gì?
    - 1, 2, 4, 8, 16, 32, …

Chuỗi: Cấp số cộng

- Một cấp số cộng (arithmetic progression) là một chuỗi số ở dạng:

$$a, a+d, a+2d, \ldots, a+nd, \ldots$$

Với a là số hạng khởi đầu và d là công sai; $a, d \in R$.

- **Ví dụ:**
  - $s_n = -1 + 4n$ với $n = 0, 1, 2, 3, \ldots$
    - Các phần tử: -1, 3, 7, 11, ...

Chuỗi: Cấp số nhân

- Một cấp số nhân (geometric progression) là một chuỗi số ở dạng:

$$a, ar, ar^2, \ldots, ar^k, \ldots$$

Với a là số hạng khởi đầu và r là công bội; a, r ∈ R.

- **Ví dụ:**
  - $a_n = (1/2)^n$ với n = 0, 1, 2, 3, ...
    - Các phần tử: 1, ½, ¼, 1/8, ...

Chuỗi: tìm luật sinh

- Việc tìm ra luật (hàm) để sinh ra một chuỗi cho trước là không đơn giản.

- **Ví dụ 1:**
  - Giả sử chuỗi: 1, 3, 5, 7, 9, …
    - Công thức của chuỗi số là gì?
      - Từng số hạng được tạo ra nhờ cộng 2 vào số hạng trước nó.

        1, 1+2=3, 3+2=5, 5+2=7
      - Một **cấp số cộng**: a + nd, với a = 1 và d = 2: $a_n = 1+2n$

Chuỗi: tìm luật sinh

- Việc tìm ra luật (hàm) để sinh ra một chuỗi cho trước là không đơn giản.

- **Ví dụ 1:**
  - Giả sử chuỗi: 1, 1/3, 1/9, 1/27, …
    - Công thức của chuỗi số là gì?
      - Các mẫu số đều là bội của 3.

        1, 1/3=1/3, (1/3)/3=1/(3*3)=1/9, (1/9)/3=1/27
      - Một **cấp số nhân**: ar^k, với a=1 và r=1/3: $(1/3)^n$

Chuỗi: Định nghĩa hồi quy (recursive)

- Phần tử thứ n của chuỗi $\{a_n\}$ được định nghĩa hồi quy dựa trên các phần tử trước đó của chuỗi và phần tử khởi đầu của chuỗi.
- **Ví dụ:**
  - $a_n = a_{n-1} + 2$ giả sử $a_0 = 1$;
    - $a_0 = 1$;
    - $a_1 = 3$;
    - $a_2 = 5$;
    - $a_3 = 7$;
  - Viết a_n dưới dạng tường minh?
    - $a_n = 1 + 2n$

Chuỗi Fibonacci

- Là một chuỗi được định nghĩa hồi quy, với
  - $f_0 = 0, f_1 = 1;$
  - $f_n = f_{n-1} + f_{n-2}$ với $n = 2, 3, \ldots$

- Các giá trị
  - $f_2 = 1;$
  - $f_3 = 2;$
  - $f_4 = 3;$
  - $f_5 = 5;$

Chuỗi: ví dụ

- **Lãi suất cộng dồn**: Giả sử một người gửi $10,000 vào tài khoản tiết kiệm tại một ngân hang với lãi suất cộng dồn là 11%/năm. Hỏi sau 30 năm thì số tiền trong tài khoản ngân hàng của người này là bao nhiêu?
- **Lời giải**:
  - Gọi P_n là số tiền trong tài khoản tiết kiệm sau n năm.
  - Chuỗi $\{P_n\}$ sẽ có định nghĩa hồi quy như sau

    $$P_n = P_{n-1} + 0.11\ P_{n-1} = (1.11)\ P_{n-1}.$$

  - Do đó, P_{30} = $228,922.97.

Tổng (summation)

- Tổng của các số hạng của một chuỗi:

$$\sum_{j=m}^{n} a_j = a_m + a_{m+1} + \cdots + a_n$$

Biến j được gọi là biến đếm (index) của tổng.

- m là giới hạn dưới và
- n là giới hạn trên của tổng

Tổng

Ví dụ:

- Tính tổng 7 số hạng đầu của $\{n^2\}$ với n = 1, 2, 3, …
  - $\sum_{j=1}^{7} a_j = \sum_{j=1}^{7} j^2 = 1 + 4 + 16 + 25 + 36 + 49 = 140$

- Giá trị của $\sum_{k=4}^{8} (-1)^k$
  - $\sum_{k=4}^{8} (-1)^j = 1 + (-1) + 1 + (-1) + 1 = 1$

Dãy cấp số cộng

- Tổng của các phần tử của một cấp số cộng a, a+d, a+2d, …, a+nd được gọi là một **chuỗi cấp số cộng (arithmetic series)**.
- **Định lý:** Tổng của các phần tử của một cấp số cộng a, a+d, a+2d, …, a+nd là

$$S = \sum_{j=1}^{n} (a + jd)$$

$$= na + d \sum_{j=1}^{n} j = na + d \frac{n(n+1)}{2}$$

Tại sao?

Chuỗi cấp số cộng

- **Định lý:** Tổng của các số hạng của một cấp số cộng a, a+d, a+2d, …, a+nd là

$$S = \sum_{j=1}^{n}(a + jd)$$

$$= na + d\sum_{j=1}^{n} j = na + d\frac{n(n+1)}{2}$$

- **Chứng minh:**

$$S = \sum_{j=1}^{n}(a + jd)$$

$$= \sum_{j=1}^{n} a + \sum_{j=1}^{n} jd = na + d\sum_{j=1}^{n} j$$

$\sum_{j=1}^{n} j = 1 + 2 + 3 + 4 + \cdots + (n-2) + (n-1) + n$

Chuỗi cấp số cộng

Ví dụ:

- $S = \sum_{j=1}^{5}(2 + j3) =$

$$= \sum_{j=1}^{5} 2 + \sum_{j=1}^{5} j3 =$$

$$= 2\sum_{j=1}^{5} 1 + 3\sum_{j=1}^{5} j =$$

$$= 2 * 5 + 3\sum_{j=1}^{5} j =$$

$$= 10 + 3\frac{(5+1)}{2} * 5 =$$

$$= 10 + 45 = 55$$

Chuỗi cấp số cộng

Ví dụ:

- $S = \sum_{i=1}^{4} \sum_{j=1}^{2} (2i - j) =$

$$= \sum_{i=1}^{4} \left[\sum_{j=1}^{2} 2i - \sum_{j=1}^{2} j \right] =$$

$$= \sum_{i=1}^{4} \left[2i \sum_{j=1}^{2} 1 - \sum_{j=1}^{2} j \right] =$$

$$= \sum_{i=1}^{4} \left[2i * 2 - \sum_{j=1}^{2} j \right] =$$

$$= \sum_{i=1}^{4} [2i * 2 - 3] =$$

$$= \sum_{i=1}^{4} 4i - \sum_{i=1}^{4} 3 =$$

$$= 4 \sum_{i=1}^{4} i - 3 \sum_{i=1}^{4} 1 = 4 * 10 - 3 * 4 = 28$$

Chuỗi cấp số nhân

- Tổng của các số hạng của một cấp số nhân a, ar, ar^2,…,ar^k được gọi là một **chuỗi cấp số nhân** (Geometric series).

- **Định lý:** Tổng của các số hạng của một cấp số nhân a, ar, ar^2,…,ar^n là

$$S = \sum_{j=0}^{n} (ar^j) = a \sum_{j=0}^{n} r^j = a\left[\frac{r^{n+1} - 1}{r - 1}\right]$$

Chuỗi nhân (Geometric series)

- **Định lý:** Tổng của các số hạng của một cấp số nhân a, ar, ar^2,…,ar^n là

$$S = \sum_{j=0}^{n} (ar^j) = a \sum_{j=0}^{n} r^j = a\left[\frac{r^{n+1} - 1}{r - 1}\right]$$

- **Chứng minh:**

$$S = \sum_{j=0}^{n} ar^j = a + ar + ar^2 + ar^3 + \cdots + ar^n$$

  - Nhân S với r

$$rS = r \sum_{j=0}^{n} ar^j = ar + ar^2 + ar^3 + \cdots + ar^{n+1}$$

  - Trừ $rS - S = [ar + ar^2 + ar^3 + \cdots + ar^{n+1}] - [a + ar + ar^2 + \cdots + ar^n]$
    $= ar^{n+1} - a$

$$S = \frac{ar^{n+1} - a}{r - 1} = a\left[\frac{r^{n+1} - 1}{r - 1}\right]$$

Chuỗi cấp số nhân

- **Ví dụ:**

$$S = \sum_{j=0}^{3} 2(5)^j =$$

- **Công thức tổng quát:**

$$S = \sum_{j=0}^{n} (ar^j) = a\sum_{j=0}^{n} r^j = a\left[\frac{r^{n+1} - 1}{r - 1}\right]$$

$$S = \sum_{j=0}^{3} 2(5)^j = 2 * \frac{5^4 - 1}{5 - 1} =$$

$$= 2 * \frac{625 - 1}{4} = 2 * \frac{624}{4} = 2 * 156 = 312$$

Chuỗi nhân vô hạn

- Chuỗi nhân vô hạn (Infinite geometric series) có thể được tính dưới dạng đóng với x < 1
- Bằng cách nào?

$$\sum_{n=0}^{\infty} x^n = \lim_{k\to\infty} \sum_{n=0}^{k} x^n = \lim_{k\to\infty} \frac{x^{k+1}-1}{x-1} = -\frac{1}{x-1}$$

$$= \frac{1}{1-x}$$

- Do đó:

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$$

Chuỗi: ví dụ

- Một số chuỗi hữu ích cùng công thức

Chuỗi	Tổng
$\sum_{k=0}^{n} ar^k \ (r \neq 0)$	$\frac{ar^{n+1} - a}{r - 1}, r \neq 1$
$\sum_{k=1}^{n} k$	$\frac{n(n+1)}{2}$
$\sum_{k=1}^{n} k^2$	$\frac{n(n+1)(2n+1)}{6}$
$\sum_{k=1}^{n} k^3$	$\frac{n^2(n+1)^2}{4}$
$\sum_{k=0}^{\infty} x^k, \lvert x \rvert < 1$	$\frac{1}{1-x}$
$\sum_{k=1}^{\infty} kx^{k-1}, \lvert x \rvert < 1$	$\frac{1}{(1-x)^2}$

- **Ví dụ**:
  - $\sum_{k=50}^{100} k^2 = ?$