20/4

Nhập môn An toàn thông tin

PGS. Nguyễn Linh Giang

Bộ môn Truyền thông và Mạng máy tính

Nội dung

I. Nhập môn An toàn thông tin
II. Đảm bảo tính mật
    I. Các hệ mật khóa đối xứng (mã hóa đối xứng)
    II. Các hệ mật khóa công khai (mã hóa bất đối xứng)
III. Bài toán xác thực
    I. Cơ sở bài toán xác thực
    II. Xác thực thông điệp
    III. Chữ ký số và các giao thức xác thực
    IV. Các cơ chế xác thực trong các hệ phân tán
IV. An toàn an ninh hệ thống
    I. Phát hiện và ngăn chặn xâm nhập (IDS, IPS)
    II. Lỗ hổng hệ thống

Nội dung

- Tài liệu môn học:
  - W. Stallings “Networks and Internetwork security”
  - W. Stallings “Cryptography and network security”
  - Introduction to Cryptography – PGP
  - D. Stinson – Cryptography: Theory and Practice

Chương III. Các hệ mật khóa công khai

- Nguyên lý hệ mật khoá công khai
- Thuật toán RSA
- Sơ đồ trao đổi khoá Diffie-Hellman
- Một số hệ mật khóa công khai khác

Nguyên lý hệ mật khoá công khai

- Đặc điểm
  - Mật mã công khai dựa trên cơ sở của các hàm toán học.
  - Không dựa trên phép thay thế và đổi chỗ như trong phương pháp mã hoá đối xứng.
  - Mã mật công khai là bất đối xứng.
    - Trong cơ chế mã mật khoá công khai sử dụng hai khoá: khoá mật và khoá công khai.
    - Các hệ quả của việc sử dụng hai khoá bất đối xứng: tính toàn vẹn, tính xác thực, phân phối khoá.

Nguyên lý hệ mật khoá công khai

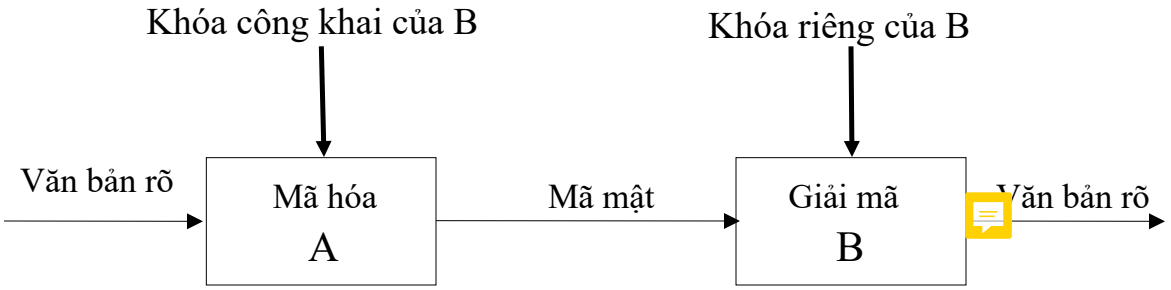
- Xuất xứ:
  - Hệ mã mật khoá công khai được phát triển nhằm giải quyết hai vấn đề phức tạp nảy sinh từ phương pháp mã hoá đối xứng:
    - Vấn đề thứ nhất: bài toán phân phối khoá;
    - Vấn đề thứ hai: chữ ký số.

Nguyên lý hệ mật khoá công khai

- Hệ mật khoá công khai.
  - Sơ đồ mã mật khoá công khai sử dụng một khoá để mã hoá và một khoá khác có liên quan để giải mã. Các thuật toán mã hoá và giải mã có một số đặc điểm quan trọng sau:
    - Không thể xác định được khoá giải mã nếu chỉ biết thuật toán mã hoá và khoá mã hoá.
    - Một số hệ mã mật khoá công khai (như RSA) còn cung cấp khả năng sử dụng bất kỳ một khoá trong cặp khoá làm khoá mã hoá thì khoá còn lại sẽ được dùng làm khoá giải mã.

Nguyên lý hệ mật khoá công khai

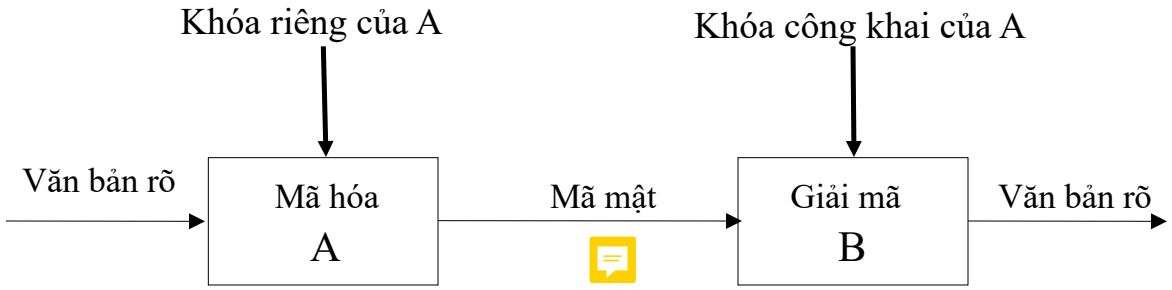
- Sơ đồ mã hoá công khai:
  - A và B có các cặp khóa (K_{RA}, K_{PA}), (K_{RB}, K_{PB}). Các khóa này dùng để mã hoá và giải mã các thông điệp.
  - A và B công bố khoá công khai K_{PA}, K_{PB} trong cặp khoá, khoá còn lại được giữ mật.
  - Khi gửi thông điệp cho B, A sẽ mã hoá văn bản bằng khoá công khai K_{PB} của B.
  - Khi nhận được thông điệp, B sẽ giải mã bằng khoá mật K_{RB}. Bên thứ ba không giải mã được thông điệp vì chỉ có B biết khoá mật K_{RB} của B.



Đảm bảo tính mật

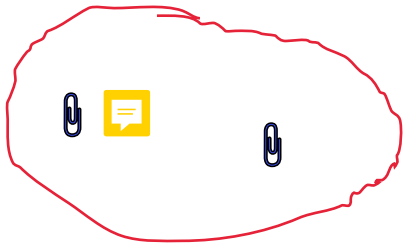
Nguyên lý hệ mật khoá công khai

- Sơ đồ xác thực:
  - Nếu A muốn gửi thông điệp được xác thực cho B, A sẽ mã hoá văn bản bằng khoá riêng của A.
  - Khi B nhận được thông điệp, B sẽ giải mã bằng khoá công khai của A. Không một bên thứ ba có thể giải mã được thông điệp vì chỉ có B biết khoá mật của B.



Đảm bảo tính xác thực

2 kịch bản:
- Ktr người gửi trc
- Ktra người nhận trc



Nguyên lý hệ mật khoá công khai

- Các yêu cầu đối với hệ mật khóa công khai
  - Quá trình sinh cặp khóa K_P, K_R là dễ trên phương diện tính toán;
  - Quá trình mã hóa bản tin bằng khóa công khai K_P ở bên gửi là dễ:

    $$Y = E_{KP}(M);$$
  - Quá trình giải mã ra văn bản rõ khi biết khóa riêng K_R và bản tin mật Y là dễ:

    $$M = D_{KR}(Y);$$
  - Đối với thám mã, nếu chỉ biết K_P sẽ rất khó trên phương diện tính toán để tính ra K_R;
  - Đối với thám mã, nếu chỉ biết K_P và bản tin mật Y sẽ rất khó trên phương diện tính toán để tính ra bản tin rõ M;
  - Nguyên lý đối xứng: quá trình mã hóa – giải mã có thể áp dụng theo hai chiều: $M = D_{KP}[E_{KR}(M)]$

Nguyên lý hệ mật khoá công khai

- Các ứng dụng của hệ mật khóa công khai
  - Ứng dụng trong mật mã – mã hóa, giải mã (RSA):
    - Bên gửi mã hóa bằng khóa công khai của bên nhận;
    - Bên nhận giải mã bằng khóa riêng.
  - Ứng dụng trong phân phối khóa(RSA, Diffie-Helman): duy trì kênh mật phân phối khóa đối xứng bằng cơ sở mã mật công khai;
  - Ứng dụng trong chữ ký số (RSA, DSS):
    - Bên gửi ký bằng khóa riêng.
    - Bên nhận xác thực chữ ký bằng khóa công khai của bên gửi.

Thuật toán mã hoá công khai RSA

- Cơ sở lý thuyết số
- Sơ đồ thuật toán
- Thám mã RSA

Sơ đồ thuật toán RSA

- Xuất xứ
  - RSA do Ron Rivest, Adi Shamir và Len Adlenman phát minh năm 1977;
  - Hệ thống mã khoá công khai phổ biến và đa năng:
    - Được sử dụng trong các ứng dụng mã hóa/giải mã;
    - Chứng thực;
    - Phân phối và trao đổi khoá.

Sơ đồ thuật toán RSA

- Thuật toán RSA:
  - Phương pháp mã hóa khối;
  - Văn bản rõ và văn bản mật là các số nguyên có giá trị từ 0 đến n-1, n – số nguyên lớn;
  - Mỗi khối có giá trị nhỏ hơn n.
  - Kích thước của khối (số bít) nhỏ hơn hoặc bằng $\log_2(n)$.
  - Thực tế, kích thước của khối là *k* bit với

$$2^k < n \leq 2^{k+1}.$$

Sơ đồ thuật toán RSA

- Cặp khóa: (e, d)
- Mã hoá

Bản rõ	M < n
Mã mật	$C = M^e \bmod n$

Bản tin tõ M được nâng lên luy thừa e mod n

- Giải mã

Mã mật	C
Bản rõ	$M = C^d \bmod n = (M^e)^d \bmod n$

Sơ đồ thuật toán RSA

- Bên gửi và bên nhận phải biết số *n*.
- Bên gửi biết khóa công khai là cặp (e, n).
- Bên nhận có khóa riêng là cặp (d, n).
- Các yêu cầu:
  - Có thể tìm được các số e, d, n sao cho:

    $M^{ed} = M \bmod n \ \forall \ M < n.$
  - Thực hiện tính M^e và C^d một cách đơn giản $\forall$ M < n.
  - Không thể xác định được d nếu biết e và n

Sơ đồ thuật toán RSA

- Tạo khoá
  - Tìm các số e, d sao cho:

$$M^{ed}=M \bmod n$$

  - Hệ quả của định lý Euler: cho p và q là số nguyên tố, n và m là hai số nguyên sao cho: n=pq và 0 < m < n, k là số nguyên bất kỳ. Đẳng thức sau nghiệm đúng:

$$m^{k\phi(n)+1}=m^{k(p-1)(q-1)+1}\equiv m \bmod n$$

  - Như vậy: $ed = k\phi(n)+1$, tức là:
  - $ed \equiv 1 \bmod \phi(n)$ hay $d \equiv e^{-1} \bmod \phi(n)$ có nghĩa là $\gcd(\phi(n), d) = 1$ và $\gcd(\phi(n), e) = 1$

Sơ đồ thuật toán RSA

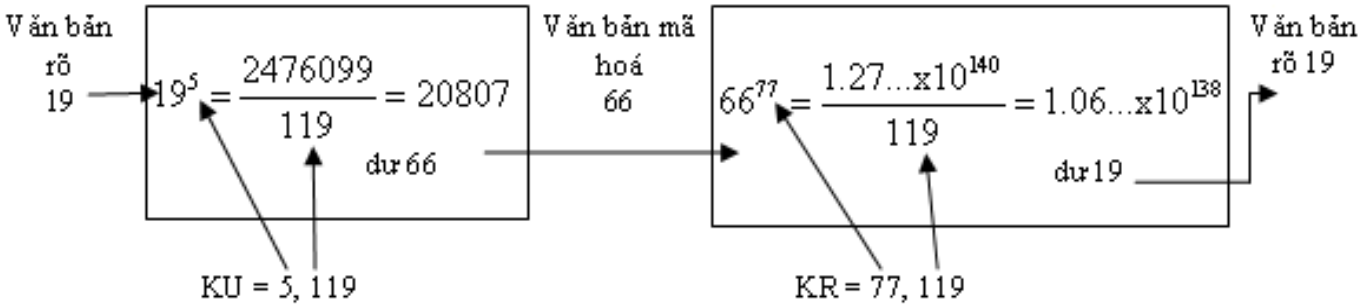
- Sơ đồ tạo khóa RSA

Chọn p, q	p và q là số nguyên tố
Tính $n = p \times q$	
Tính $\phi(n) = (p-1)(q-1)$	
Chọn số nguyên e	$\gcd(\phi(n), e) = 1;\ 1 < e < \phi(n)$
Tính d	$d \equiv e^{-1} \bmod \phi(n)$
Khoá công khai	$KU = [e, n]$
Khoá mật	$KR = [d, n]$

Sơ đồ thuật toán RSA

record 1h05: 4/5

- Ví dụ
  - p = 7, q = 17
  - n = pq = 119; $\phi(n)=(p-1)(q-1)=96$
  - Chọn e nguyên tố cùng nhau với $\phi(n)$, nhỏ hơn $\phi(n)$,
    - Chọn e = 5;
  - Tìm d: $d \equiv e^{-1} \bmod \phi(n)$
    - d=77 => cặp khóa: e=(5, 119); d=(77, 119)



Sơ đồ thuật toán RSA

- Mã hoá và giải mã
  - Vấn đề trong thuật toán mã hoá và giải mã RSA là việc thực hiện phép toán luỹ thừa và phép toán đồng dư với số nguyên lớn.
  - Giải quyết dựa trên tính chất của phép toán modun:

    [(a mod n) x (b mod n)] mod n = (a x b) mod n
  - Tính a^m với m lớn.
    - Biểu diễn nhị phân của m $=b_k b_{k-1} \ldots b_0 = \sum_{bi \neq 0} 2^i$
    - Do đó:

thay vì tính cả a^m mod n, tách a^m ra nhiều bi

$$a^m = a^{\left(\sum_{b_i \neq 0} 2^i\right)} = \prod_{b_i \neq 0} a^{2^i}$$

$$a^m \bmod n = \prod_{b_i \neq 0} a^{2^i} \bmod n = \prod_{b_i \neq 0} \left(a^{2^i} \bmod n\right)$$

Sơ đồ thuật toán RSA

- Sinh khoá
  - Các bước quan trọng trong tạo khóa:
    - Xác định 2 số nguyên tố p và q. Để tránh tấn công vét cạn, p và q phải lớn.
    - Xác định e và d

- Xác định số nguyên tố p, q (sử dụng thuật toán Miller – Rabin)

1. Chọn một số nguyên lẻ n ngẫu nhiên (sử dụng bộ sinh số giả ngẫu nhiên).
2. Chọn một số nguyên a < n ngẫu nhiên.
3. Thực hiện thuật toán xác suất để kiểm tra số nguyên tố. Nếu n test thành công thì loại bỏ giá trị n và quay lại bước 1.
4. Nếu n test thành công với số lượng test đủ, chấp nhận n; mặt khác, quay lại bước 2.

- Chọn e và tính d từ e và $\phi(n)$ (sử dụng thuật toán Euclid)

Thám mã RSA

- Tấn công vét cạn: thử vét cạn toàn bộ không gian khóa riêng.
- Tấn công toán học: thực hiện bài toán phân tích số nguyên thành tích hai số nguyên tố.
- Tấn công dựa vào thời gian: dựa vào thời gian để thực hiện thuật toán giải mã.

Thám mã RSA - tấn công vét cạn

- Phương pháp: thực hiện vét cạn toàn bộ không gian khoá.
- Biện pháp đối phó:
  - Sử dụng không gian có khoá kích thước lớn, tức là tăng số bít của d và e.
  - Không gian khóa có kích thước lớn sẽ làm quá trình sinh khoá, mã hoá, giải mã thực hiện chậm đi.

Thám mã RSA - Tấn công toán học

- Các phương pháp tấn công toán học vào RSA:
  - Phân tích n thành tích hai số nguyên tố p và q;
    - Sau đó cho phép tính $\phi(n)=(p-1)(q-1)$;
    - Từ $\phi(n)$ có thể tính $d=e^{-1} \bmod \phi(n)$.
  - Xác định $\phi(n)$ trực tiếp không qua p và q;
    - Cho phép từ $\phi(n)$ có thể tính $d=e^{-1} \bmod \phi(n)$.
  - Xác định trực tiếp d không qua tính $\phi(n)$.

Thám mã RSA - Tấn công toán học

- Trường hợp đơn giản nhất là người thám mã biết được $\phi(n)$
- Phân tích n thành tích của 2 thừa số nguyên tố: Có nhiều thuật toán phân tích n thành hai thừa số nguyên tố.
  - Có ba thuật toán hiệu quả trên các số rất lớn:
    - Thuật toán sàng bình phương (quadratic sieve),
    - Đường cong elip (elliptic curve) và
    - Sàng trường số.
  - Các thuật toán được biết đến nhiều trước đây:
    - Thuật toán p – 1 của Pollard,
    - Thuật toán p + 1 của William,
    - Thuật toán chia nhỏ liên tiếp
    - Thuật toán chia thử.

Thám mã RSA - Tấn công toán học

- Những yêu cầu đối với p và q:
  - p và q chỉ nên khác nhau về độ dài khoảng vài hàng số nhị phân và trong khoảng từ 10^{75} đến 10^{100};
  - (p-1) và (q-1) phải có những thừa số nguyên tố lớn;
  - Gcd(p-1, q-1) phải nhỏ.
  - Thực tế cho thấy, nếu e < n và d < $n^{1/4}$ thì d có thể dễ dàng tính được!

d: phải là 1 khóa khó đoán

Thám mã RSA - Tấn công dựa thời gian

- Nội dung của phương pháp này dựa vào việc theo dõi thời gian thực hiện thuật toán giải mã;
  - Có thể áp dụng đối với những hệ mật khóa công khai khác!
  - Là dạng tấn công chỉ sử dụng mã mật (ciphertext only attack)
- Biện pháp đối phó:
  - Thời gian tính mũ là hằng: Làm cho thời gian tính mũ là như nhau trước khi trả về kết quả. Biện pháp này đơn giản nhưng làm giảm hiệu năng.
  - Thực hiện trễ ngẫu nhiên: Thêm các trễ thời gian ngẫu nhiên vào thuật toán mũ hoá.
  - Làm mù: Nhân văn bản mật với một số ngẫu nhiên trước khi thực hiện mũ hoá. Khi đó thám mã sẽ không biết bit nào của mã mật được xử lý và do đó ngăn chặn được quá trình phân tích mã mật.

Lý thuyết số

- Số học modun
- Định lý Euler và định lý Fermat
- Kiểm tra số nguyên tố
- Thuật toán Euclid
- Định lý số dư Trung Hoa
- Sinh giả ngẫu nhiên các số nguyên lớn

Số học modun

- ***Định lý về số dư.*** Cho một số nguyên dương n và một số nguyên a. Khi đó tồn tại duy nhất các số q và r với , sao cho a = qn + r.

  r gọi là số dư của phép chia a cho n.
- ***Định nghĩa số dư.*** Cho một số nguyên dương n và số nguyên a. Ký hiệu a mod n là số dư khi chia a cho n.

  a = x n + (a mod n)
- ***Định nghĩa 2.*** Hai số a và b được gọi là đồng dư theo mođun n nếu a mod n = b mod n, a $\equiv$ b (mod n)
- Ví dụ: 11 = 1 x 7 + 4 => 11 mod 7 = 4

  -11 = (-2) x 7 + 3 => -11 mod 7 = 3

  73 $\equiv$ 4 (mod 23)

Định lý Euler và định lý Fermat

- Định lý nhỏ Fermat: Nếu p là số nguyên tố và a là số nguyên dương không chia hết cho p thì

$$a^{p-1} \equiv 1 (\mathrm{mod}\ p)$$

- Hàm Euler
  - Hàm Euler được ký hiệu là $\phi(n)$ là số các số nguyên dương nhỏ hơn n và nguyên tố cùng nhau với n
  - Nếu p là số nguyên tố, $\phi(p) = p - 1$
  - Nếu $p,q \in P$, $n = pXq$, thì $\phi(n) = (p - 1)X(q - 1)$
- Định lý Euler: Nếu hai số nguyên a và n nguyên tố cùng nhau, ta có $a^{\phi(n)} \equiv 1 \pmod n$

Kiểm tra số nguyên tố

- Định lý

Nếu p là số lẻ thì $x^2 \equiv 1 \pmod{p} \Rightarrow x \equiv 1$ hoặc $x \equiv -1$

- Chứng minh
  - Xét trường hợp (x + 1), (x – 1) đồng thời chia hết cho p.
  - Xét trường hợp (x - 1) chia hết cho p.
  - Tương tự xét trường hợp (x + 1) chia hết cho p ta suy ra x -1 (mod p)
- Kết quả suy ra

Nếu tồn tại $x \mid x^2 \equiv 1 \pmod{n}, x \neq \pm 1$ thì n không phải là số nguyên tố.

1.3 Kiểm tra số nguyên tố (tiếp)

- Thuật toán Miller, Rabin: kiểm tra một số có phải là một số nguyên tố không dựa vào kết quả của định lý trên.

  Input của thuật toán là số nguyên n và một số nguyên a nào đó nhỏ hơn n. Nếu WITNESS có giá trị trả về là TRUE thì n không phải là số nguyên tố, nếu WITNESS có giá trị trả về là FALSE thì n có thể là số nguyên tố
- Ví dụ
- Đánh giá độ phức tạp

```
WITNESS(a, n)
1.    b_k b_{k-1} ... b_0 là biểu diễn nhị phân của (n – 1)
2.    d ← 1
3.    for i ← k downto 0 do {
4.          x ← d
5.          d ← (d x d) mod n
6.          if d = 1 and x ≠ 1 and x ≠ n – 1 then
7.                return TRUE
8.          if b_i = 1 then
9.                d ← (d x a) mod n
      }
10.   if d ≠ 1 then
11.         return TRUE
12.   return FALSE
```

Thuật toán Euclid

- Tìm ước số chung lớn nhất

- Định lý

Với 2 số nguyên dương a và b bất kỳ chúng ta có

$$\gcd(a, b) = \gcd(b, a \bmod b)$$

- Chứng minh
- Ví dụ
- Đánh giá độ phức tạp

Thuật toán Euclid (tiếp)

- Tìm phần tử đối xứng

  Thuật toán Euclid mở rộng sẽ trả về phần tử đối xứng của d nếu gcd(d, f) = 1

EXTENDED EUCLID(d, f)

1. $(X1, X2, X3) \leftarrow (1, 0, f); (Y1, Y2, Y3) \leftarrow (0, 1, d)$
2. if $Y3 = 0$ return $X3 = \gcd(d, f)$;
3. if $Y3 = 1$ return $Y3 = \gcd(d, f)$; $Y2 = d^{-1} \bmod f$
4. $Q = \left\lfloor \frac{X3}{Y3} \right\rfloor$
6. $(T1, T2, T3) \leftarrow (X1 - QY1, X2 - QY2, X3 - QY3)$
7. $(Y1, Y2, Y3) \leftarrow (T1, T2, T3)$
8. goto 2

Định lý số dư Trung Hoa

- Định lý

$$M = \prod_{i=1}^{k} m_i$$

Trong đó m_i là nguyên tố cùng nhau từng đôi một, $\gcd(m_i, m_j) = 1$ với $1 \le i, j \le k$ và $i \ne j$. Chúng ta có thể biểu diễn bất kỳ số nguyên dương nào trong Z_M bởi k số trong các Z_{mi}:

$$A \leftrightarrow (a_1, a_2, \ldots, a_k)$$

$A \in Z_M$, $a_i \in Z_{mi}$ và $a_i = A \bmod m_i$ với $1 \le i \le k$.

  - Hai kết quả của định lý số dư Trung Hoa
  - Ứng dụng của định lý số dư Trung Hoa
  - Ví dụ

Sinh giả ngẫu nhiên các số nguyên lớn

- Bộ sinh số giả ngẫu nhiên

  Kỹ thuật được sử dụng rộng rãi trong việc sinh giả ngẫu nhiên là phương pháp đồng dư tuyến tính lần đầu tiên được đề xuất bởi Lehmer.
- Sinh số giả ngẫu nhiên dựa trên kỹ thuật mật mã
- Bộ sinh số giả ngẫu nhiên Blum Blum Shub

Để quản lí khóa

Hệ mật Diffie-Hellman

- Các sơ đồ quản lý khoá của hệ mật khoá công khai
  - Quản lý và chứng thực khoá công khai;
  - Cấp phát chứng thư số;
  - Trao đổi khoá mã hoá-giải mã của hệ mật khoá đối xứng:
    - Xây dựng kênh truyền bí mật để trao đổi phiên.
      - Dùng cơ chế bảo mật của hệ mật khoá công khai;

Nguyên lý trao đổi khóa Diffie-Hellman

- Được Diffie-Hellman đưa ra vào 1976
- Là sự kết hợp của hai mô hình xác thực và mật của hệ KCK
- Việc sinh ra các cặp khoá là hoàn toàn khác nhau đối với người sử dụng
- Sử dụng cơ chế trao đổi khoá trực tiếp không qua trung gian xác thực

Nguyên lý trao đổi khóa Diffie-Helman

- Sử dụng trong các ứng dụng trao đổi khoá khi sử dụng hệ mật khoá đối xứng.
- Nguyên tắc: hai người sử dụng có thể trao đổi khoá phiên an toàn - được dùng để mã hoá và giải mã các thông điệp;
- Thuật toán tự giới hạn chỉ dùng cho các ứng dụng sử dụng kĩ thuật trao đổi khoá;

Cơ sở hình thành thuật toán

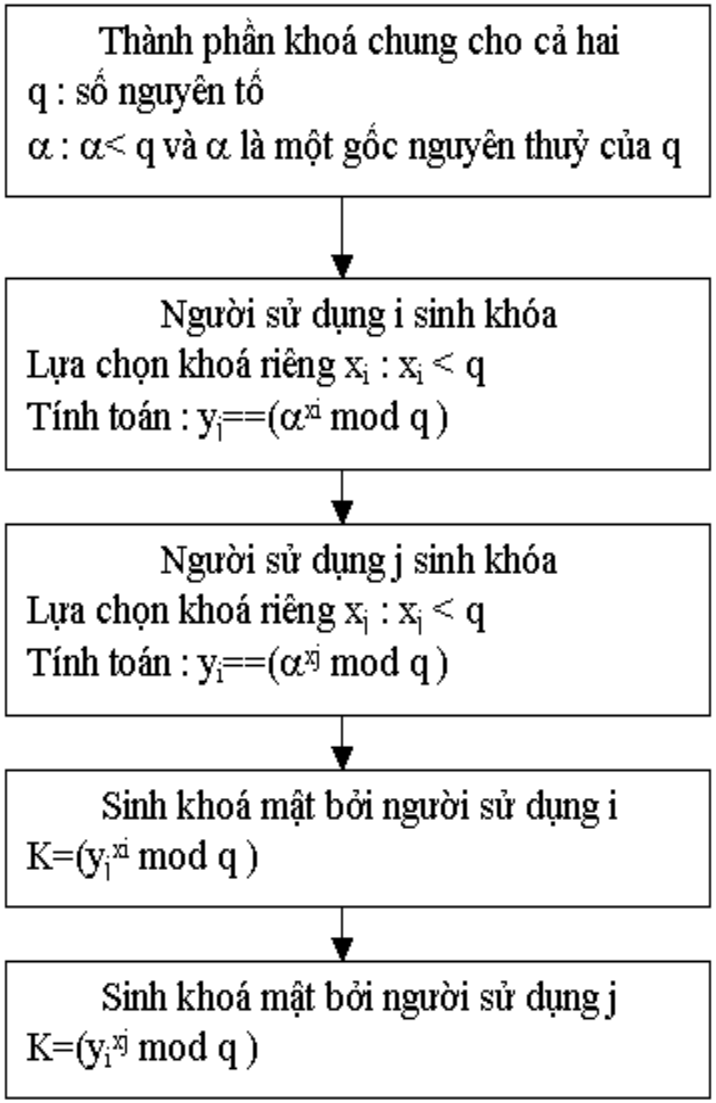
- Nguyên tắc toán học :
  - m là một số nguyên tố:
  - $y=a^i \bmod m$ là bài toán dễ;
  - Bài toán ngược là bài toán khó. Đặc biệt với m lớn.
- Dựa trên phép tính logarit rời rạc

Cho a^i hỏi nó là mod của m nào

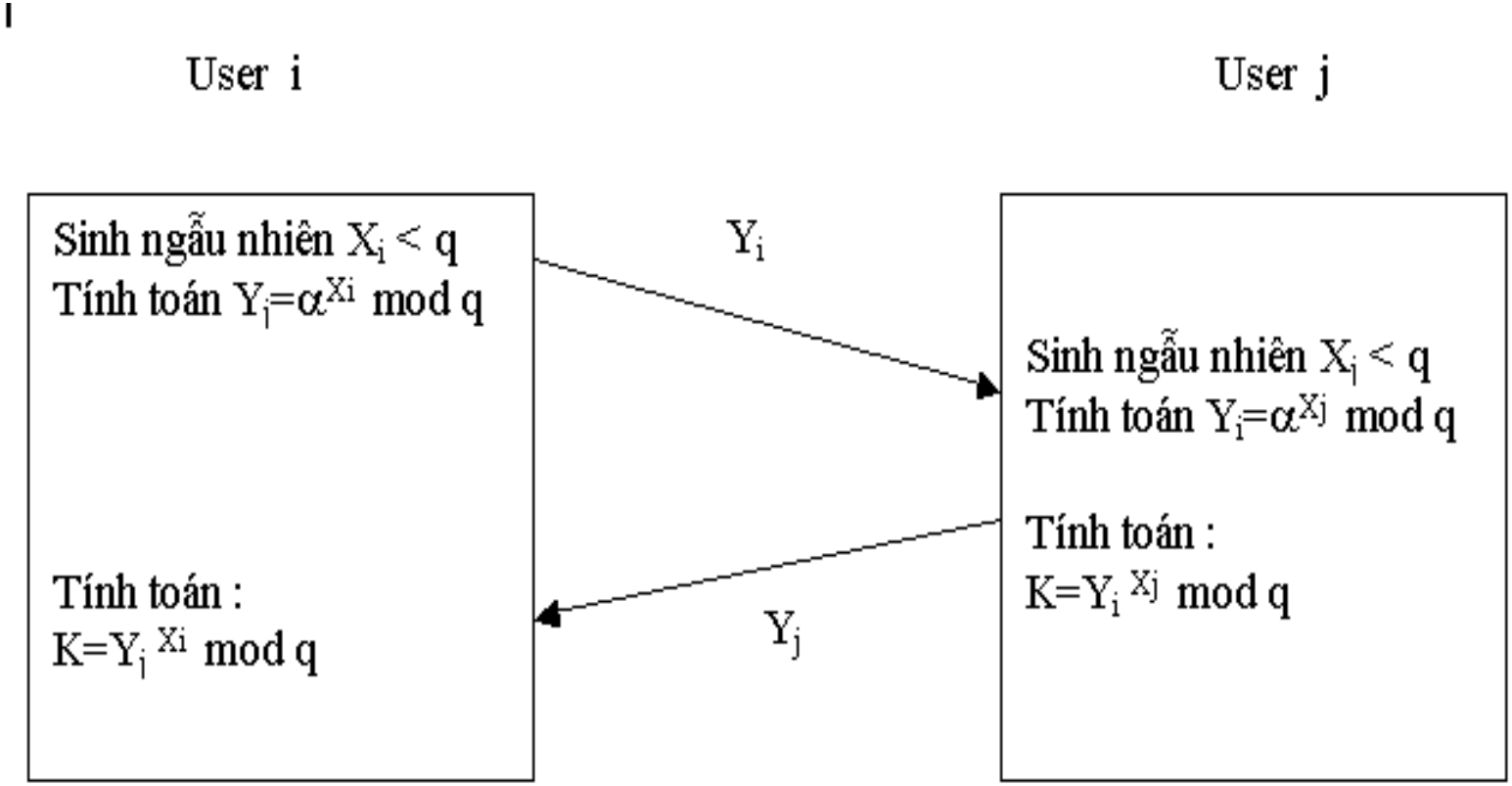
Thuật toán logarit rời rạc

- Thuật toán logarit rời rạc:
  - Một số nguyên tố p;
  - Một gốc nguyên thuỷ a của p : là các số mà luỹ thừa của a theo modul p thuộc (1,p-1)
  - Với b bất kì nguyên sẽ luôn $\exists$i sao cho b= a^i mod p.

Thuật toán Diffie-Hellman



Thuật toán trao đổi khoá



Tính bảo mật của hệ mật

- Thám mã có sẵn các thông tin :p,a,Y_i,Y_j
- Để có thể giải được K ,X bắt buộc thám mã phải sử dụng thuật toán logarit rời rạc : rất khó nếu p lớn
- Vì thế nên chọn p càng lớn càng tốt : như thế thì việc tính toán ra X coi như không thể

Hệ mật và thám mã

- Thám mã có thể tấn công vào các thông tin : p ,a,Y_j,Y_j
- Và sử dụng thuật toán rời rạc để tính ra X, sau đó tính ra K
- Quan trọng nhất là độ phức tạp của thuật toán logarit phụ thuộc vào chọn số nguyên tố p
- Tấn công man in the middle

Lĩnh vực ứng dụng

- Tự quá trình thuật toán đã hạn chế ứng dụng chỉ sử dụng cho quá trình trao đổi khoá mật là chủ yếu
- Sử dụng trong chữ kí điện tử.
- Các ứng dụng đòi hỏi xác thực người sử dụng.