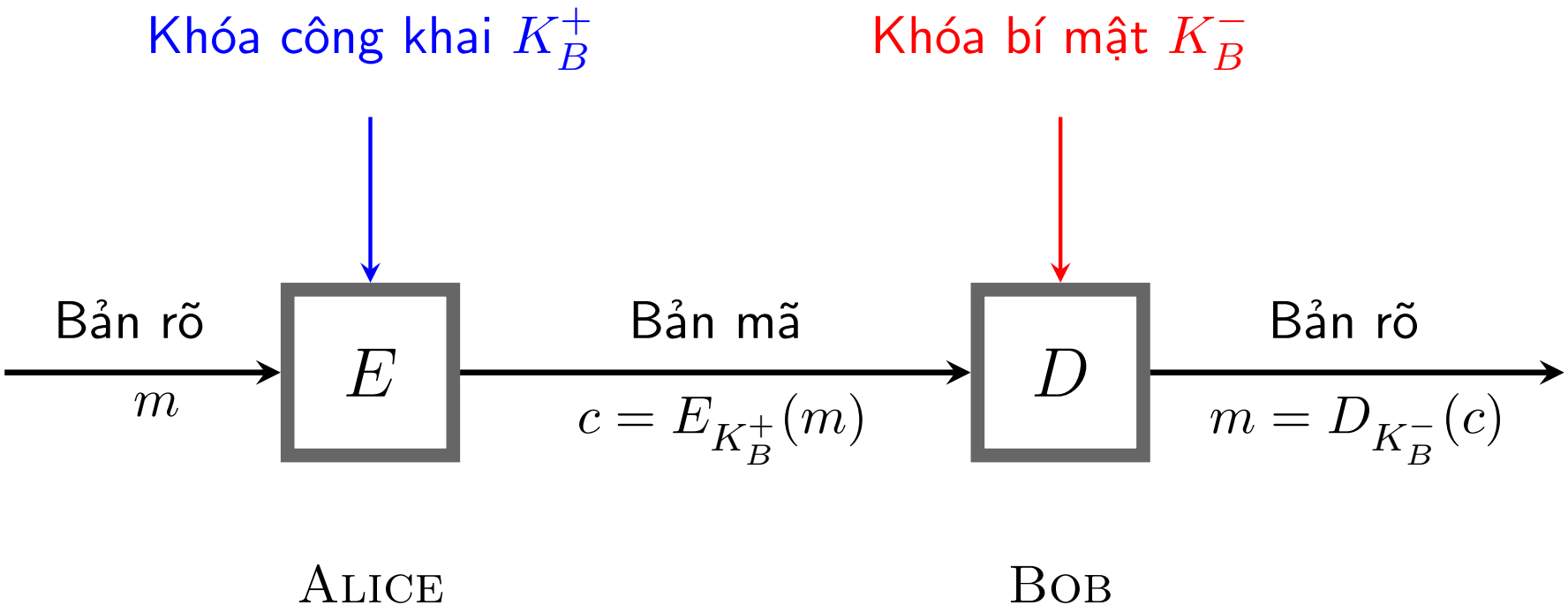
Mật mã và độ phức tạp thuật toán (Complexity and Cryptography) Chủ đề 6: Mật mã khóa công khai

PGS.TS. Nguyễn Đình Hân
(Mobile: 0915.046.320; Email: han.nguyendinh@hust.edu.vn)

Viện Toán ứng dụng và Tin học
Trường Đại học Bách khoa Hà Nội

Mật mã khóa công khai

Mật mã khóa công khai



Hình 6.1 Mô hình hệ mật khóa công khai tổng quát

Tổng quan về mật mã khóa công khai

- **Mật mã khoá công khai** (public-key cryptography) là một bước tiến lớn nhất và là cuộc cách mạng thực sự trong lĩnh vực mật mã.
- Hệ mật khoá công khai triệt để thay đổi những gì đã tồn tại trước đây. Ví dụ như các giải thuật mã công khai dựa trên các hàm toán học chứ không dựa trên kỹ thuật thay thế và dịch chuyển.

⋆ Quan trọng hơn nữa, hệ mã công khai là hệ mã không đối xứng, nghĩa là sử dụng hai khoá liên đới cho việc mã hoá và giải mã thay vì một khoá duy nhất như trong các hệ mã cổ điển (hay còn gọi là hệ mã đối xứng). Việc này đáp ứng được các yêu cầu trong các ứng dụng về *bảo mật riêng tư, phân phối khoá*, và *xác thực điện tử*.

Tổng quan về mật mã khóa công khai

Một số quan điểm không đúng về mật mã khóa công khai:

- Mật mã khoá công khai bảo mật hơn mật mã cổ điển.
- Mật mã khoá công khai là một kỹ thuật tổng quát và đã làm cho mật mã cổ điển trở nên lỗi thời (thực tế, áp dụng hệ mã công khai cho công tác quản lý khoá mã và các ứng dụng sử dụng chữ ký điện tử là thích hợp nhất).
- Phân phối khoá mã là công việc nặng nề đối với mật mã cổ điển (sử dụng các trung tâm phân phối) trong khi lại là đơn giản đối với mật mã khoá công khai.

Lưu ý $\rightsquigarrow$ Một hệ mật khóa công khai không bao giờ cung cấp độ mật vô điều kiện - thực tế, đó là *hàm cửa sập một chiều* (a trapdoor one-way function).

Những hệ mật khóa công khai quan trọng nhất

RSA dựa trên độ khó của phép phân tích các số nguyên lớn.

Merkle-Hellman Knapsack và các hệ liên quan dựa trên độ khó của bài toán subset sum (được biết là NP-complete). Tuy nhiên, có nhiều hệ mật dựa trên bài toán sắp ba lô đã được chứng minh là không bảo mật.

McEliece dựa trên bài toán giải mã của một mã tuyến tính (cũng được cho là NP-complete).

ElGamal dựa trên bài toán Logarit rời rạc trên trường hữu hạn.

Chor-Rivest là một hệ sắp ba lô nhưng được xem là bảo mật.

Elliptic Curve là sự cải tiến của các hệ mật khác, chẳng hạn tương tự ElGamal nhưng dựa trên các đường cong elíp thay vì trường hữu hạn. Ưu điểm của các hệ mật dạng này là có thể duy trì được độ bảo mật với khóa nhỏ hơn thông thường.

Phép chia (division)

Cho trước một số nguyên $b \neq 0$, ta nói số nguyên a **chia hết** cho b nếu $a = mb$ với m là số nguyên bất kỳ. Nghĩa là, phép chia này không có phần dư. Nếu a chia hết cho b, ta kí hiệu là $b|a$. Nếu $b|a$, ta nói b là ước số của a.

Ta có thể kiểm tra các quan hệ sau đây

1. Nếu $a|1$ thì $a = \pm 1$
2. Nếu $a|b$ và $b|a$ thì $a = \pm b$
3. Nếu $b \neq 0$ thì $b|0$
4. Nếu $b|g$ và $b|h$ thì $b|(mg + nh)$ với m, n là các số nguyên tùy ý.

Ước số chung lớn nhất (greatest common divisor)

Một số nguyên c được gọi là ước số chung lớn nhất của a và b nếu c là ước số của cả a và b. Hơn nữa, một ước số bất kì của a và b cũng là ước số của c.

Ta sẽ kí hiệu $\gcd(a, b)$ là ước số chung lớn nhất của hai số nguyên a và b. Khi đó, $\gcd(a, b) = \max[k,$ sao cho $k|a$ và $k|b]$.

Thuật toán Euclid //Tìm $\gcd(r_0, r_1)$ với $r_0 > r_1$

$$
\begin{aligned}
r_0 &= q_1 r_1 + r_2, \quad 0 < r_2 < r_1 \\
r_1 &= q_2 r_2 + r_3, \quad 0 < r_3 < r_2 \\
&\cdots \\
r_{m-2} &= q_{m-1} r_{m-1} + r_m, \quad 0 < r_m < r_{m-1} \\
r_{m-1} &= q_m r_m
\end{aligned}
$$

Ta có $\gcd(r_0, r_1) = r_m$.

Số nguyên tố (prime)

Một số nguyên $p > 1$ là **số nguyên tố** khi các ước số của nó là ± 1 và $\pm p$.

Ta lưu ý tính chất quan trọng sau đây:

Mỗi số nguyên $a > 1$ bất kì có thể được phân tích duy nhất thành dạng tích luỹ thừa của các số nguyên tố (*khai triển chính tắc*)

$$a = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_t^{\alpha_t}$$

Trong đó, $p_1 < p_2 < \dots < p_t$ là các số nguyên tố và $\alpha_i > 0$.

Số nguyên tố cùng nhau (relatively prime)

Các số nguyên a và b là **nguyên tố cùng nhau** nếu chúng không có thừa số nguyên tố phân tích chung khác 1, tức là $\gcd(a, b) = 1$.

Một số kiến thức nền tảng

WITHNESS (a, n) //Kiểm tra n có là số nguyên tố không, $a < n$

1. Đặt $b_k b_{k-1} \cdots b_0$ là xâu nhị phân biểu diễn $(n-1)$
2. $d \leftarrow 1$
3. **for** $i \leftarrow k$ **downto** 0 **do**
4. $\quad x \leftarrow d$
5. $\quad d \leftarrow (d \times d) \mod n$
6. $\quad$ **if** $d = 1$ and $x \neq 1$ and $x \neq n-1$ **then**
7. $\quad\quad$ **return** TRUE
8. $\quad$ **if** $b_i = 1$ **then**
9. $\quad\quad d \leftarrow (d \times a) \mod n$
10. **if** $d \neq 1$ **then**
11. $\quad$ **return** TRUE
12. **return** FALSE

Định lý 6.1 (Định lý Fermat) Nếu p là số nguyên tố và a là số nguyên dương không chia hết cho p, thì

$$a^{p-1} \equiv 1 \pmod{p}.$$

Định lý 6.2 (Định lý Euler) Với mọi a và n nguyên tố cùng nhau, ta có

$$a^{\phi(n)} \equiv 1 \pmod{n}.$$

Lưu ý $\rightsquigarrow$ Phi hàm ơle $\phi(n)$ là số các số nguyên dương nhỏ hơn n và nguyên tố cùng nhau với n. Nếu p là số nguyên tố thì $\phi(p) = p - 1$.

Hệ mật RSA

Bài toán phân tích thừa số: Cho số nguyên dương $n \geq 2$. Hỏi n có là hợp số?

Thuật toán Pollard's $p-1$ (1974)

Input: n và β; Output: "thành công" hoặc "thất bại".

1. $a = 2$
2. **for** $j = 2$ **to** β **do**
3. $\quad a = a^j \mod n$
4. $d = \gcd(a-1, n)$
5. **if** $1 < d < n$ **then**
6. $\quad d$ là một thừa số của n (thành công)
7. **else**
8. $\quad$ không tìm thấy thừa số của n (thất bại)

Hệ mật RSA

- Hệ mật RSA, được phát triển bởi Ron Rivest, Adi Shamir và Leonard Adleman (1977), có thể được sử dụng trong bảo mật dữ liệu và công nghệ chữ ký điện tử.
- Bảo mật của RSA dựa trên giả thuyết không có các thuật toán đủ nhanh để khai triển luỹ thừa một số. Qui trình áp dụng RSA gồm hai bước

  1) Lựa chọn (sinh) cặp khoá công khai và khoá bí mật
  2) Thực hiện thuật toán mã hoá và thuật toán giải mã.

Lưu ý ⇝ Chọn kích cỡ khóa đủ lớn! Thế nào là đủ lớn? Nếu khóa càng lớn thì càng khó phá vỡ RSA, tuy nhiên thời gian để mã hoá và giải mã sẽ tăng lên đáng kể. Phòng nghiên cứu RSA đề xuất rằng kích cỡ khóa nên là **1024** bit, ngoài ra, kích cỡ có thể là **768** bit đối với các thông tin ít có giá trị hơn.

Sinh cặp khóa công khai-bí mật (Bob)

1. Chọn hai số nguyên tố đủ lớn, p và q
2. Tính toán $n = pq$ và $\phi(n) = (p-1)(q-1)$
3. Chọn một số, $e\,(1 < e < \phi(n))$ sao cho $\gcd(e, \phi(n)) = 1$. Giá trị e sẽ được sử dụng trong mã hoá
4. Tìm một số d sao cho $ed - 1$ chia hết cho $\phi(n)$ hay nói cách khác $d = e^{-1} \mod \phi(n)$. Giá trị d sẽ được sử dụng để giải mã
5. Công khai khoá $K_B^+ = (n, e)$ và giữ bí mật khoá $K_B^- = (n, d)$

Thuật toán mã hoá (Alice) và **thuật toán giải mã** (Bob)

1. Giả sử Alice muốn gửi cho Bob một mẫu bit, hoặc một số m sao cho $m < n$. Để mã hoá, Alice thực hiện tính luỹ thừa, m^e, sau đó tính toán số dư khi đem chia m^e cho n. Vì vậy, giá trị được mã hoá (c), của bản tin m là

$$c = m^e \mod n.$$

2. Để giải mã đoạn tin mã hoá nhận được (c), Bob tính toán

$$m = c^d \mod n.$$

Việc này đòi hỏi phải sử dụng khoá bí mật (n, d).

Hệ mật El Gamal

Bài toán logarit rời rạc trong $\mathbb{Z}_n$: Cho $I = (n, \alpha, \beta)$, trong đó n là số nguyên tố, $\alpha \in \mathbb{Z}_n$ là phần tử nguyên thủy và $\beta \in \mathbb{Z}_n^*$. Tìm một số nguyên $a, 0 \leq a \leq n-2$, sao cho

$$\alpha^a \equiv \beta \,(\mathrm{mod}\ n).$$

Nhắc lại: $\mathbb{Z}_n^* = \{a \in \mathbb{Z}_n | \gcd(a, n) = 1\}$.

1. Cấp của $\mathbb{Z}_n^*$ là số các phần tử của nó.
2. Phần tử $a \in \mathbb{Z}_n^*$ có bậc m nếu m là số nguyên dương bé nhất sao cho $a^m \equiv 1 \,(\mathrm{mod}\ n)$. Nếu $m = \phi(n)$ thì a được gọi là *phần tử nguyên thủy* (hay phần tử sinh) của $\mathbb{Z}_n^*$.

Lưu ý:

1. Nếu p là số nguyên tố thì $\mathbb{Z}_p^* = \{1, 2, \ldots, p-1\}$ là nhóm cyclic và có $\phi(p-1)$ phần tử nguyên thủy.

2. Nếu $p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_s^{\alpha_s}$ là khai triển chính tắc của $p-1$ và $a^{\frac{p-1}{p_1}} \equiv 1 \mod p, \ldots, a^{\frac{p-1}{p_s}} \equiv 1 \mod p$ thì a là phần tử nguyên thủy theo môđun p của $\mathbb{Z}_p^*$.

3. Nếu a là phần tử nguyên thủy của $\mathbb{Z}_p^*$ thì $b = a^i \mod p$ với $\gcd(i, p-1) = 1$ cũng là phần tử nguyên thủy của $\mathbb{Z}_p^*$.

Hệ mật El Gamal

Chọn n là số nguyên tố sao cho bài toán logarit rời rạc trong $\mathbb{Z}_n$ rất khó giải và chọn một phần tử nguyên thủy $\alpha \in \mathbb{Z}_n^*$. Đặt $\mathcal{P} = \mathbb{Z}_n^*, \mathcal{C} = \mathbb{Z}_n^* \times \mathbb{Z}_n^*$ và định nghĩa

$$\mathcal{K} = \{(n, \alpha, a, \beta) : \beta \equiv \alpha^a (\bmod\ n)\}.$$

Công khai giá trị n, α và β, giữ bí mật a.

Với $K = (n, \alpha, a, \beta)$ và một giá trị ngẫu nhiên (bí mật) $k \in \mathbb{Z}_{n-1}$, định nghĩa

$$e_K(x, k) = (y_1, y_2).$$

Trong đó

$$y_1 = \alpha^k \mod n, \quad y_2 = x\beta^k \mod n.$$

Với $y_1, y_2 \in \mathbb{Z}_n^*$, định nghĩa

$$d_K(y_1, y_2) = y_2(y_1^a)^{-1} \mod n.$$

Sơ đồ chữ kí điện tử

Định nghĩa 6.1 Một sơ đồ chữ kí (signature scheme) là một bộ năm $(\mathcal{P}, \mathcal{A}, \mathcal{K}, \mathcal{S}, \mathcal{V})$, thỏa mãn các điều kiện sau đây:

1. $\mathcal{P}$ là một tập hữu hạn các thông điệp
2. $\mathcal{A}$ là một tập hữu hạn các chữ kí
3. $\mathcal{K}$ là một tập hữu hạn các khóa
4. Với mỗi $K \in \mathcal{K}$, tồn tại một giải thuật ký $sig_K \in \mathcal{S}$ và một giải thuật kiểm chứng liên đới $ver_K \in \mathcal{V}$. Mỗi $sig_K : \mathcal{P} \to \mathcal{A}$ và $ver_K \in \mathcal{P} \times \mathcal{A} \to \{true, false\}$ là các hàm sao cho công thức sau thỏa mãn với mọi mọi $x \in \mathcal{P}$ và với mọi $y \in \mathcal{A}$:

$$ver(x, y) = \begin{cases} true, & \text{nếu } y = sig(x), \\ false, & \text{nếu } y \neq sig(x). \end{cases}$$

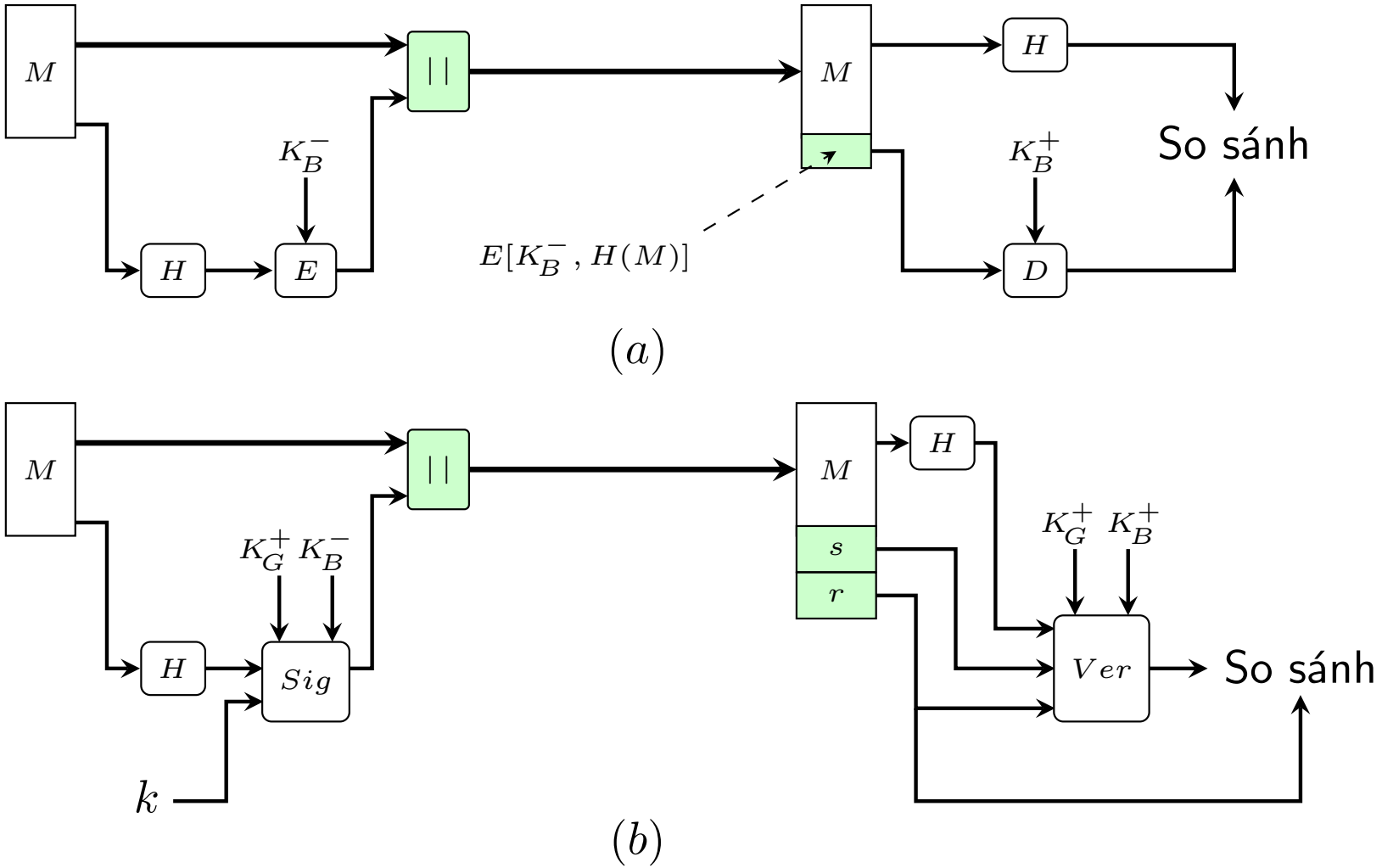
Hàm băm một chiều và ứng dụng

Định nghĩa 6.2 Một hàm băm h là hàm một chiều nếu với giá trị băm z, ta không có khả năng tìm ra thông điệp x sao cho $h(x) = z$.

Hai điều kiện sau đây thường được xem là hai điều kiện chủ yếu cho một hàm băm

1. Hàm băm phải là hàm một phía, nghĩa là cho x tính $z = h(x)$ là việc dễ, nhưng ngược lại, biết z tính x là việc cực khó (có thể qui ước dễ hay khó theo nghĩa tính được trong thời gian đa thức hay không).
2. Hàm băm phải là hàm không va chạm mạnh theo nghĩa sau đây: không có thuật toán tính được trong thời gian đa thức giải bài toán "tìm x_1 và x_2 thuộc Σ^* sao cho $x_1 \neq x_2$ và $h(x_1) = h(x_2)$"; nói cách khác, tìm hai văn bản khác nhau có cùng một đại diện là cực kỳ khó.

Thiết lập chữ kí điện tử



Hình 6.2 Hai cách tiếp cận chữ kí điện tử: (a) RSA và (b) DSS

Sơ đồ chữ ký RSA

Đặt $n = pq$ với p, q là các số nguyên tố. Đặt $\mathcal{P} = \mathcal{A} = \mathbb{Z}_n$ và định nghĩa

$$\mathcal{K} = \{(n, p, q, a, b) : n = pq,\ ab \equiv 1 (\mod \phi(n))\}.$$

Công khai giá trị n và b, giữ bí mật p, q và a.

Với $K = (n, p, q, a, b)$, định nghĩa

$$sig_K(x) = x^a \mod n$$

và

$$ver_K(x, y) = true \quad \Leftrightarrow \quad x \equiv y^b (\mod n)$$

với $(x, y \in \mathbb{Z}_n)$.

Sơ đồ chữ ký El Gamal

Chọn n là số nguyên tố sao cho bài toán logarit rời rạc trong $\mathbb{Z}_n$ rất khó giải và chọn một phần tử nguyên thủy $\alpha \in \mathbb{Z}_n^*$. Đặt $\mathcal{P} = \mathbb{Z}_n^*$, $\mathcal{A} = \mathbb{Z}_n^* \times \mathbb{Z}_{n-1}$ và định nghĩa

$$\mathcal{K} = \{(n, \alpha, a, \beta) : \beta \equiv \alpha^a (\bmod\ n)\}.$$

Công khai giá trị n, α và β, giữ bí mật a.

Với $K = (n, \alpha, a, \beta)$ và một giá trị ngẫu nhiên (bí mật) $k \in \mathbb{Z}_{n-1}^*$, định nghĩa

$$sig_K(x, k) = (\gamma, \delta).$$

Trong đó $\gamma = \alpha^k \pmod n$, $\delta = (x - a\gamma)\, k^{-1} \pmod{(n-1)}$.

Với $x, \gamma \in \mathbb{Z}_n^*$ và $\delta \in \mathbb{Z}_{n-1}$, định nghĩa

$$ver(x, \gamma, \delta) = true \quad \Leftrightarrow \quad \beta^\gamma \gamma^\delta \equiv \alpha^x (\bmod\ n).$$

Chuẩn chữ kí số (Digital Signature Standard)

Chọn p là số nguyên tố 512-bit sao cho bài toán logarit rời rạc trong $\mathbb{Z}_p$ rất khó giải và chọn q là một số nguyên tố 160-bit $(q \mid (p-1))$. Chọn $\alpha \in \mathbb{Z}_p^*$ là một căn bậc q của $1 \bmod p$. Đặt $\mathcal{P} = \mathbb{Z}_p^*, \mathcal{A} = \mathbb{Z}_q \times \mathbb{Z}_q$ và định nghĩa

$$\mathcal{K} = \{(p, q, \alpha, a, \beta) : \beta \equiv \alpha^a (\bmod\ p)\}.$$

Công khai giá trị p, q, α và β, giữ bí mật a. Với $K = (p, q, \alpha, a, \beta)$ và một giá trị ngẫu nhiên (bí mật) $1 \leq k \leq q-1$, định nghĩa

$$sig_K(x, k) = (\gamma, \delta).$$

Trong đó $\gamma = (\alpha^k \bmod p) \bmod q, \quad \delta = (x + a\gamma)\, k^{-1} \bmod q.$

Với $x \in \mathbb{Z}_p^*$ và $\gamma, \delta \in \mathbb{Z}_q$, định nghĩa

$$ver(x, \gamma, \delta) = true \quad \Leftrightarrow \quad (\alpha^{e_1} \beta^{e_2} \bmod p) \bmod q = \gamma.$$

Trong đó $e_1 = x\delta^{-1} \bmod q, \quad e_2 = \gamma\delta^{-1} \bmod q.$

TRÂN TRỌNG CẢM ƠN!