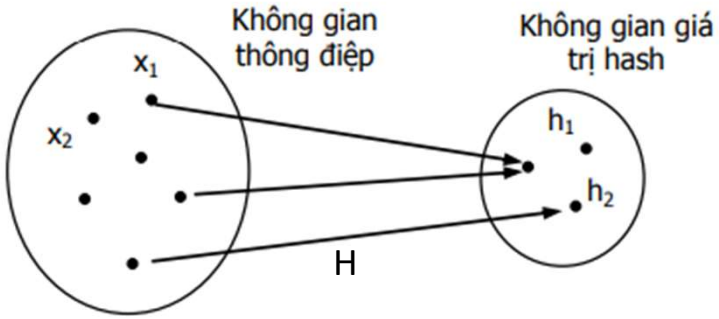
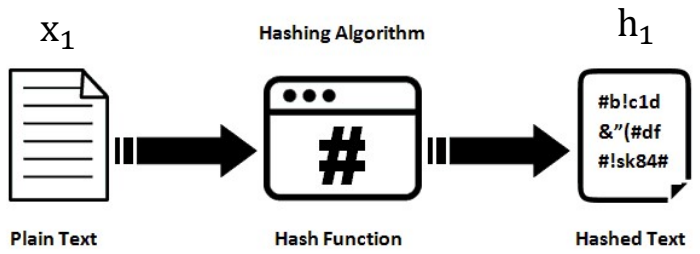
# CHƯƠNG 5: Hàm băm và ứng dụng

## Hàm băm (hash function)

- Hàm băm là gì? Tính chất của hàm băm?
- Một số hàm băm thông dụng: MD5, SHA-1, SHA-2, SHA-3
- Ứng dụng của hàm băm

**Hàm băm (hash function):** biến một dữ liệu đầu vào có độ dài bất kỳ thành một chuỗi đầu ra đặc trưng có độ dài cố định



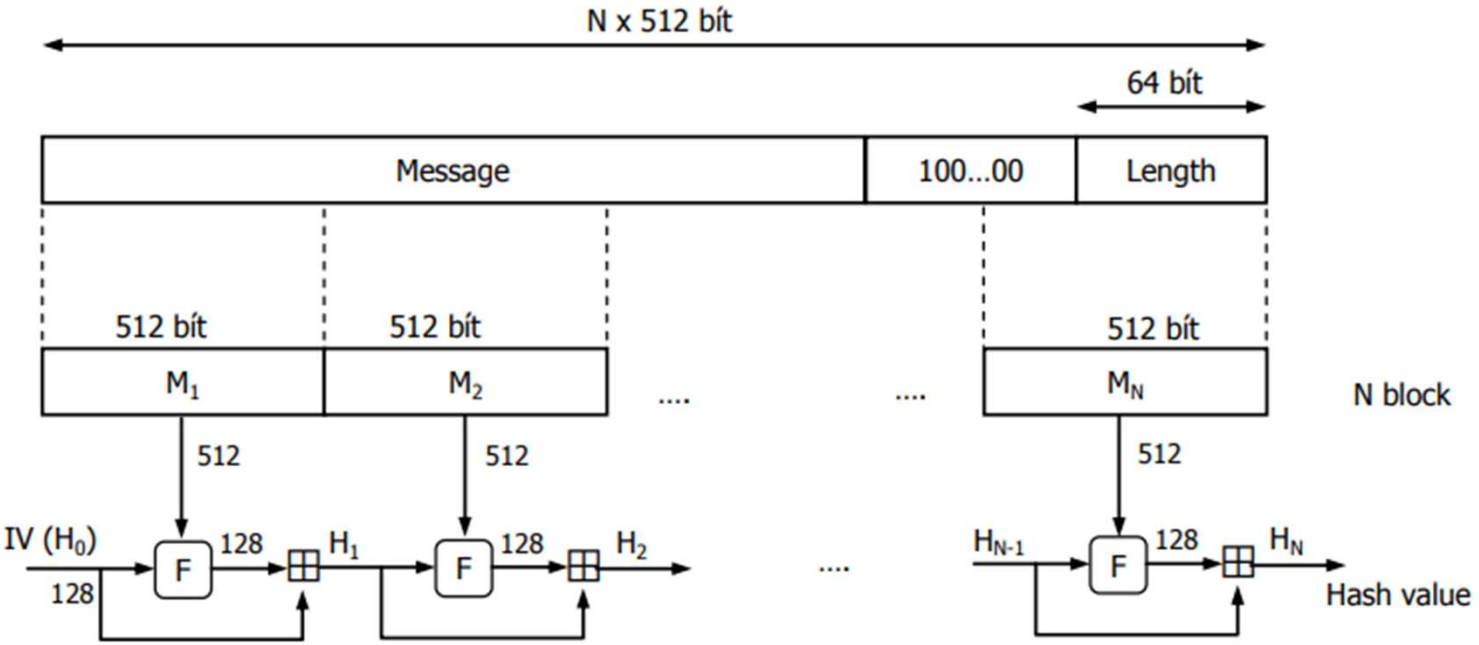




$H(x_1) = h_1$

- H có thể được áp dụng trên khối dữ liệu có độ dài bất kỳ
- H tạo đầu ra có độ dài cố định
- H(x) tính toán mọi x tương đối dễ dàng, tạo điều kiện cho việc cài đặt trên phần cứng lẫn phần mềm được thiết thực
- Với bất kỳ giá trị băm h, không thể tính được x sao cho H(x)=h. Hay H được gọi là **hàm một chiều**
- **Tính bền xung đột yếu (weak collision resistance):** với bất kỳ giá trị x, không thể tính được $y \neq x$ sao cho $H(y) = H(x)$.
- **Tính bền xung đột mạnh (strong collision resistance):** Không thể tính được một cặp (x, y) sao cho $H(x) = H(y)$

## Hàm băm MD5 (Message-Digest)

- Kích thước giá trị băm là **128** bít, thông điệp cần băm có kích thước tối đa là $2^{64}$ bít.

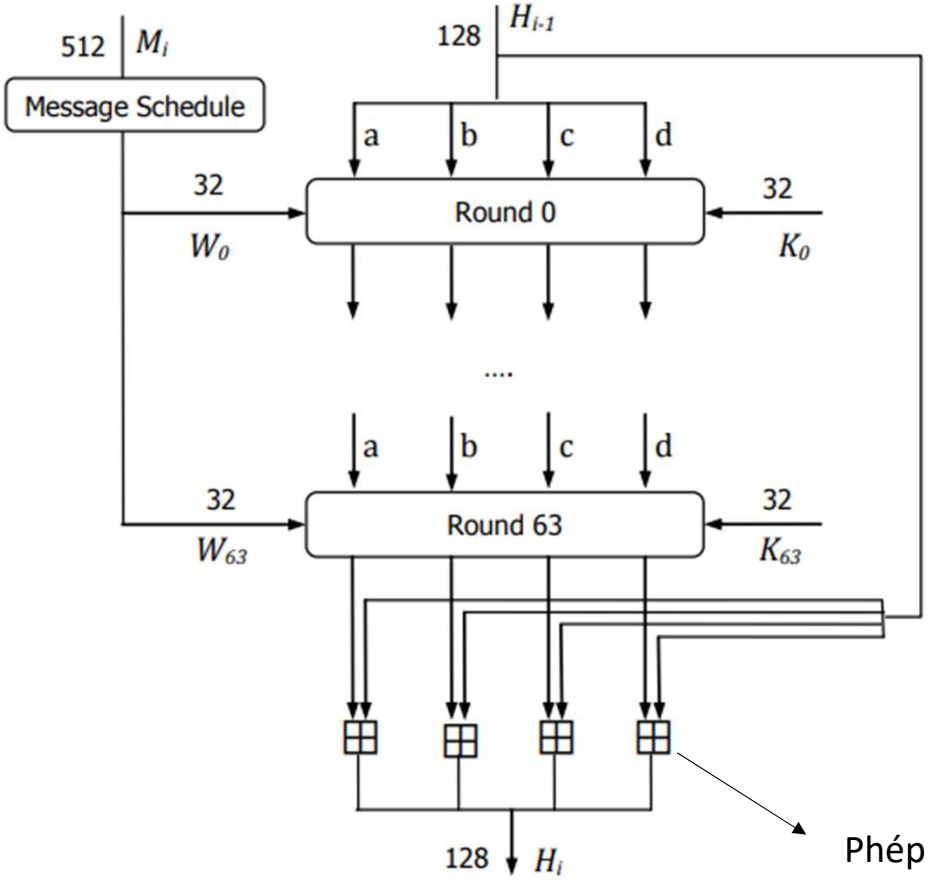



$H_0$ = abcd
(*hexadecimal*)

a = 01234567
b = 89abcdef
c = fedbca98
d = 76543210

| Hex | Binary |
|---|---|
| 0 | 0000 |
| 1 | 0001 |
| 2 | 0010 |
| 3 | 0011 |
| 4 | 0100 |
| 5 | 0101 |
| 6 | 0110 |
| 7 | 0111 |
| 8 | 1000 |
| 9 | 1001 |
| A | 1010 |
| B | 1011 |
| C | 1100 |
| D | 1101 |
| E | 1110 |
| F | 1111 |

## Function F



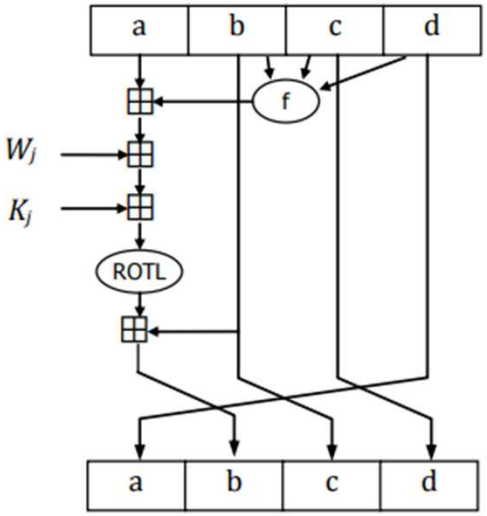

- $K_i$ và $W_i$ có độ dài 32 bit

$$K_i = \lfloor 2^{32}.|\sin(i+1)| \rfloor$$

$$M_i = W_0 W_1 \dots W_{15}$$

Phép cộng modulo $2^{32}$

## Round




| *i* | *s* |
|---|---|
| 0, 4, 8, 12 | 7 |
| 1, 5, 9, 13 | 12 |
| 2, 6,10, 14 | 17 |
| 3, 7, 11, 15 | 22 |
| 16, 20, 24, 28 | 5 |
| 17, 21, 25, 29 | 9 |
| 18, 22, 26, 30 | 14 |
| 19, 23, 27, 31 | 20 |
| 32, 36, 40, 44 | 4 |
| 33, 37, 41, 45 | 11 |
| 34, 38, 42, 46 | 16 |
| 35, 39, 43, 47 | 23 |
| 48, 52, 56, 60 | 6 |
| 49, 53, 57, 61 | 10 |
| 50, 54, 58, 62 | 15 |
| 51, 55, 59, 63 | 21 |

Ở đây $b \rightarrow c$, $c \rightarrow d$, $d \rightarrow a$. Giá trị $b$ được tính qua hàm:

$$t = a + f(b, c, d) + W_i + K_i$$

$$b = b + ROTL(t, \mathrm{s})$$

Trong đó:

- Hàm $f(x, y, z)$:

  $f(x,y,z) = (x \wedge y) \vee (\neg x \wedge z)$ nếu là vòng 0 đến 15

  $f(x,y,z) = (z \wedge x) \vee (\neg \mathrm{z} \wedge y)$ nếu là vòng 16 đến 31

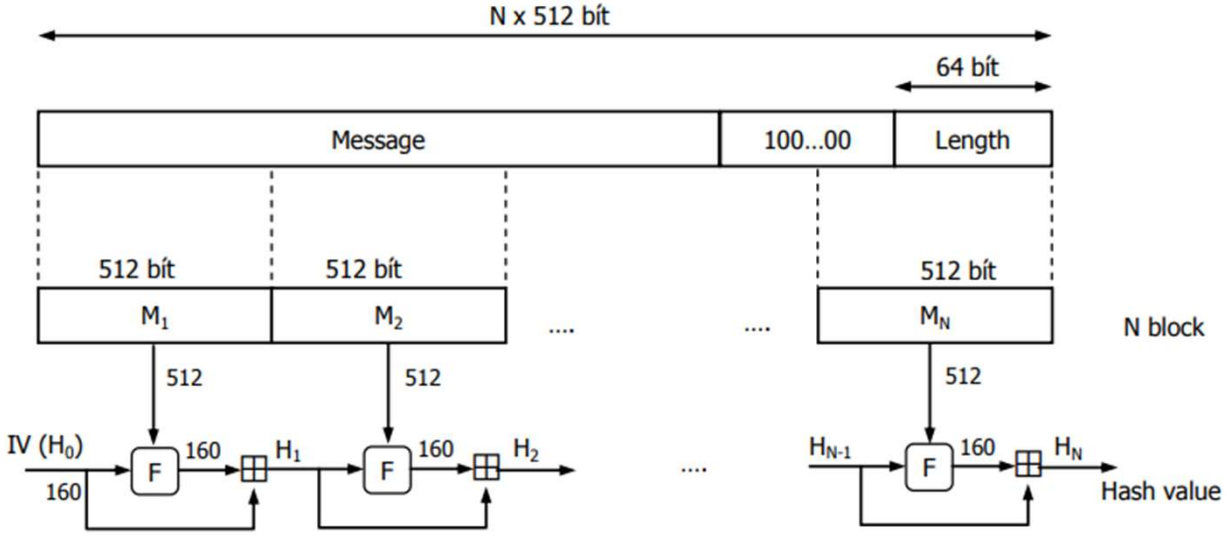
  $f(x,y,z) = x \oplus y \oplus z$ nếu là vòng 32 đến 48

  $f(x,y,z) = y \oplus (x \vee \neg z)$ nếu là vòng 49 đến 63

- Hàm $ROTL(\mathrm{t}, s)$: t được dịch vòng trái s bít, với s là các hằng số cho vòng thứ $i$

## Hàm băm SHA-1 (Secure Hash Algorithm 1)

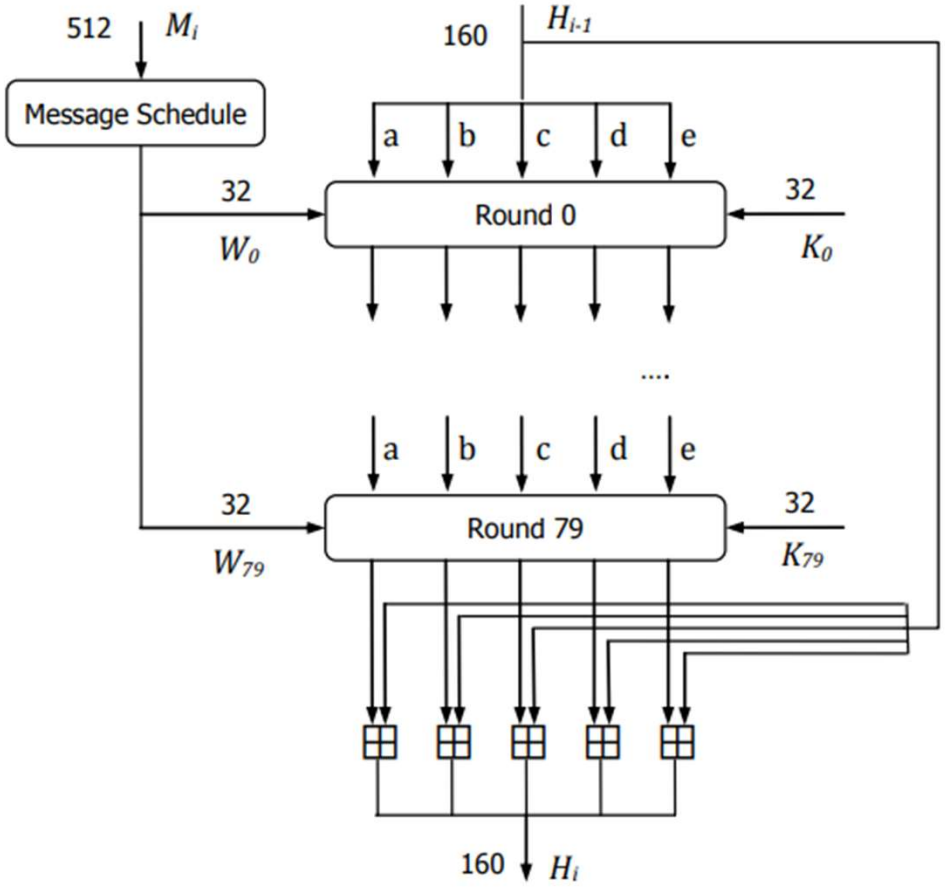
- Kích thước giá trị băm là **160** bít, thông điệp cần băm có kích thước tối đa là $2^{64}$ bít.




$H_0 =$ abcde
(*hexadecimal*)

a = 67452301
b = efcdab89
c = 98badcfe
d = 10325476
e = c3d2e1f0

## Function F




$K_i = 5A827999$ với $0 \le i \le 19$

$K_i = 6ED9EBA1$ với $20 \le i \le 39$

$K_i = 8F1BBCDC$ với $40 \le i \le 59$

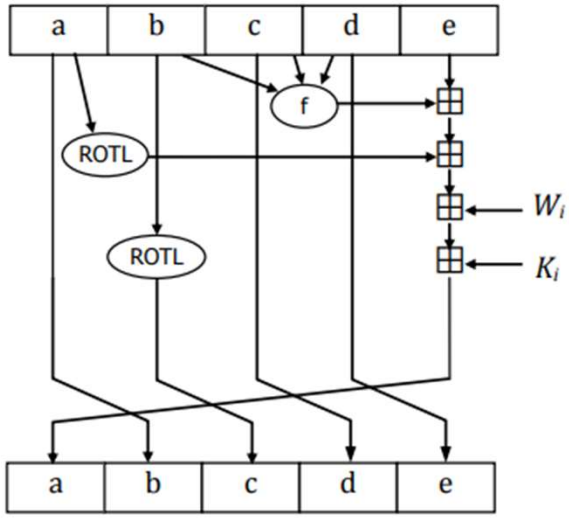
$K_i = CA62C1D6$ với $60 \le i \le 79$

$M_i = W_0 W_1 \ldots W_{15}$

Các giá trị $W_t$ $(16 \le t \le 79)$ được tính theo công thức:

$$W_t = ROTL(W_{t-3} + W_{t-8} + W_{t-14} + W_{t-16}, 1)$$

## Round




Ở đây $a \rightarrow b$, $c \rightarrow d$, $d \rightarrow e$. Giá trị $a$ và $c$ được tính qua các hàm:

$$a = ROTL(a, 5) + f(b, c, d) + e + W_i + K_i$$

$$c = ROTL(b, 30)$$

Trong đó, hàm $f(x, y, z)$:

$f(x, y, z) = Ch(x, y, z) = (x \wedge y) \oplus (\neg x \wedge z)$ nếu là vòng 0 đến 19

$f(x, y, z) = Parity(x, y, z) = x \oplus y \oplus z$ nếu là vòng 20 đến 39

$f(x, y, z) = Maj(x, y, z) = (x \wedge y) \oplus (y \wedge z) \oplus (z \wedge x)$ nếu là vòng 40 đến 59

$f(x, y, z) = Parity(x, y, z) = x \oplus y \oplus z$ nếu là vòng 60 đến 79

*Ý nghĩa của hàm Maj và hàm Ch:*

- Hàm *Maj:* giả sử $x_i$, $y_i$, $z_i$ là bít thứ $i$ của $x$, $y$, $z$, thì bít thứ $i$ của hàm *Maj* là giá trị nào chiếm đa số, 0 hay 1 (giống như hàm *maj* được định nghĩa trong phần thuật toán A5/1).
- Hàm *Ch*: bít thứ $i$ của hàm *Ch là phép chọn*: *if* $x_i$ *then* $y_i$ *else* $z_i$.

## Một số hàm băm SHA

- SHA-1 (trả lại kết quả dài 160 bit)
- SHA-224 (trả lại kết quả dài 224 bit)
- SHA-256 (trả lại kết quả dài 256 bit)
- SHA-384 (trả lại kết quả dài 384 bit)
- SHA-512 (trả lại kết quả dài 512 bit)

## Chữ ký số

Phân biệt giữa chữ ký thông thường và chữ ký số:

- Chữ ký thông thường là một phần vật lý của tài liệu. Chữ ký thông thường được kiểm tra bằng cách so sánh nó với các chữ ký xác thực khác, không phải là phương pháp an toàn vì nó dễ dàng giả mạo.
- Chữ ký số không gắn theo kiểu vật lý vào bức điện. Chữ ký số có thể được kiểm tra nhờ dùng một thuật toán kiểm tra công khai. Bất kỳ ai cũng có thể kiểm tra được chữ ký số. Việc dùng một sơ đồ chữ ký an toàn có thể sẽ ngăn chặn được khả năng giả mạo.
- Bản copy tài liệu được ký bằng chữ ký số đồng nhất với bản gốc còn bản copy tài liệu có chữ ký trên giấy thường có thể khác với bản gốc.

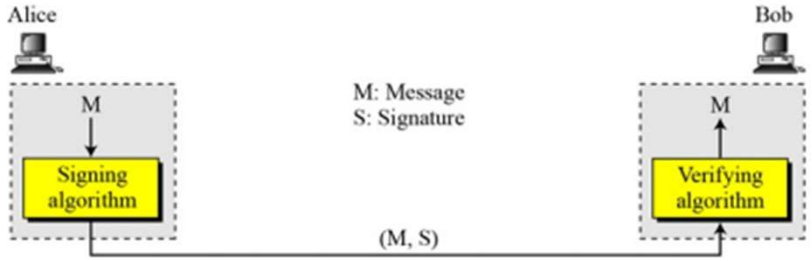
## Sơ đồ chữ ký số

Một sơ đồ chữ ký số là bộ 5 $(M, S, K, Sign, Verify)$ thoả mãn các điều kiện dưới đây:

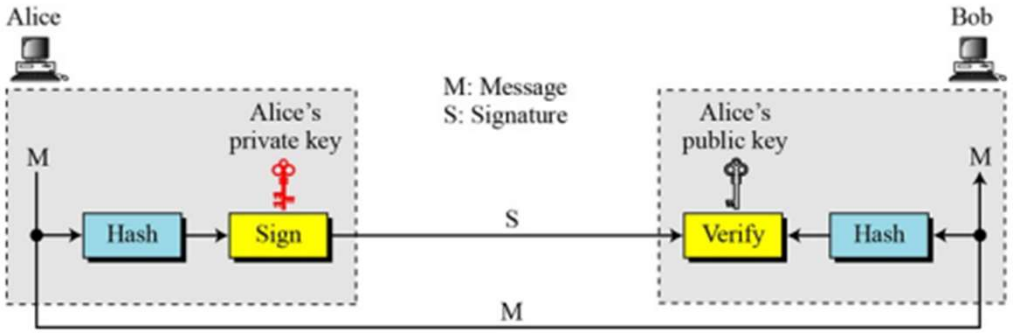
- M (*Message*) là tập hữu hạn các bức điện (thông điệp) có thể.
- S (*Signature*) là tập hữu hạn các chữ ký có thể.
- K không gian khoá là tập hữu hạn các khoá có thể. Khóa $K'$ và $K''$ tương ứng với không gian khóa mật và khóa công khai.
- Thuật toán tạo chữ ký Sign: $M \times K \mapsto S$).
- Thuật toán kiểm tra chữ ký Verify: $M \times K' \times K'' \mapsto \{True, False\}$

Với mỗi $k \in K$ tồn tại thuật toán kí $sig_k \in Sign$ và thuật toán xác minh $ver_k \in Verify$ thỏa mãn: $sig_k : M \to S$ và $ver_k : M \times K \to \{True, False\}$. Sao cho mỗi thông điệp $m \in M$ và mỗi chữ kí $s \in S$ thoả mãn phương trình: $ver_k = True$ nếu $S = sig(M)$ và ngược lại $ver_k = False$ nếu $S \neq sig(N)$. Hàm $ver_k$ là công khai, hàm $sig_k$ là bí mật.

# Sơ đồ chữ ký số



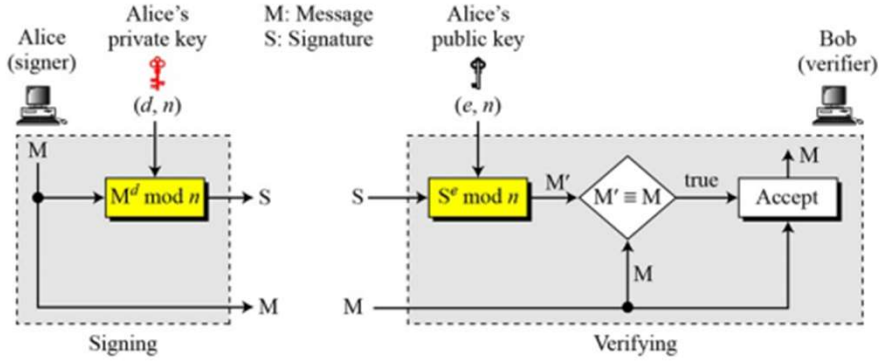

*Sơ đồ chữ ký tổng quát không dùng hàm băm*




*Sơ đồ chữ ký tổng quát dùng hàm băm*

## Sơ đồ chữ ký RSA

Sơ đồ chữ ký RSA Diffie-Hellman đề xuất năm 1973 và được Ronald Linn Rivest, Adi Shamir và Leonard Adleman thực hiện năm 1976. Sơ đồ chữ ký RSA được xác định như sau:



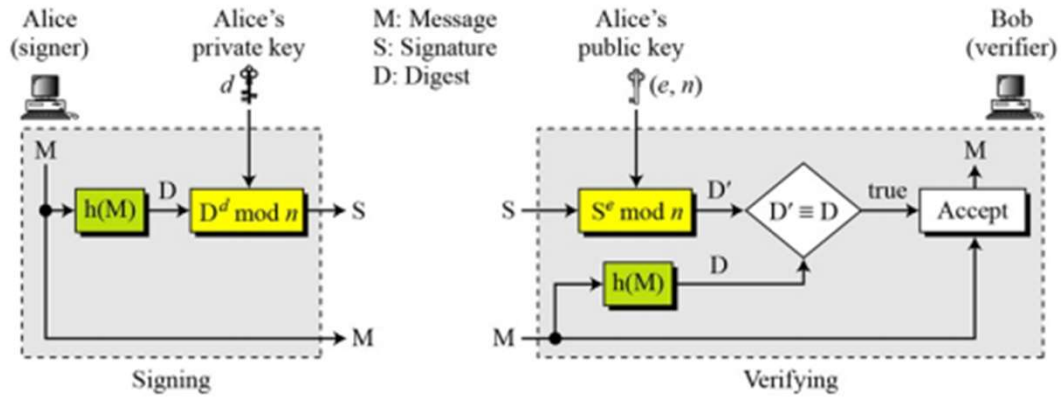

*Quá trình ký và xác thực của RSA (không dùng hàm băm h)*

- Thiết lập tham số:
  - Với $n = p \times q$ là tích của hai số nguyên tố lớn $p$, $q$.
  - $K$ là tập các cặp khoá $K = (K', K'')$, với $K' = a$ và $K'' = (n, b)$, $d$ và $e$ là hai số thuộc $Z_n^*$ thoả mãn $d.e \equiv 1 (\mathrm{mod}\,(n))$.
- Sinh chữ ký: $sig_{K'}(M) = M^d \mod n = S$
- Xác thực chữ ký: $ver_{K''}(M, S) = True \Leftrightarrow M \equiv S^e (\mathrm{mod}\, n)$

Dễ chứng minh được rằng sơ đồ được định nghĩa như vậy là hợp thức, tức là với mọi thông điệp $M \in P$ và mọi chữ ký $S \in A$: $ver_{K''}(M, S) = True \Leftrightarrow S = sig_{K'}(M)$.

## Sơ đồ chữ ký RSA

Nếu sử dụng hàm băm, trước khi ký thông điệp $M$ sẽ được băm $h(M)$ và quá trình ký thực hiện trên kết quả $D$ của hàm băm $h(M)$ tương ứng.




*Quá trình ký và xác thực của RSA (dùng hàm băm h)*

Ví dụ:

Alice chọn số nguyên tố $p = 113$ và $q = 171$. Alice tính $n = p \times q = 19323$ và $\phi(n) = 112 \times 170 = 19040$. Alice chọn $e = 311$ và tính $e \times d = 311 \times d \equiv 1 (\text{mod } 19040)$ được $d = 551$. Alice công khai khóa $(n = 19323, e = 311)$ và giữ bí mật khóa $d = 551$.

- Sinh chữ ký: Để ký một thông điệp M = 'LOVE'= "11 14 21 04", Alice tính:

$$S = M^d \text{ mod } n = 1114^{551} \text{ mod } 19323 = 14980 = 2104^{551} \text{ mod } 19323 = 3604$$

- Xác nhận chữ ký $S = 3604$: Bob tính:

$$M' = S^e \text{ mod } n = 14980^{311} \text{ mod } 19323 = 1114 = 3604^{311} \text{ mod } 19323 = 2104$$

- Cuối cùng Bob chấp nhận chữ ký vì $M' = M$.

Ví dụ:

Alice chọn số nguyên tố $p = 823$ và $q = 953$. Alice tính $n = p \times q = 784319$ và $\phi(n) = 822 \times 952 = 782544$. Alice chọn $e = 313$ và tính $e \times d = 313 \times d \equiv 1 (mod\, 782544)$ được $d = 160009$. Alice công khai khóa $(n = 784319, e = 313)$ và giữ bí mật khóa $d = 160009$.

- Sinh chữ ký: Để ký một thông điệp $M = 19070$, Alice tính:

$$S = M^d \bmod n = 19070^{160009} \bmod 784319 = 210625$$

- Xác nhận chữ ký $S = 210625$: Bob tính:

$$M' = S^e \bmod n = 210625^{313} \bmod 784319 = 19070$$

- Cuối cùng Bob chấp nhận chữ ký vì $M' = M$.