BÀI 3.
XÁC THỰC THÔNG ĐIỆP

Bùi Trọng Tùng,
Viện Công nghệ thông tin và Truyền thông,
Đại học Bách khoa Hà Nội

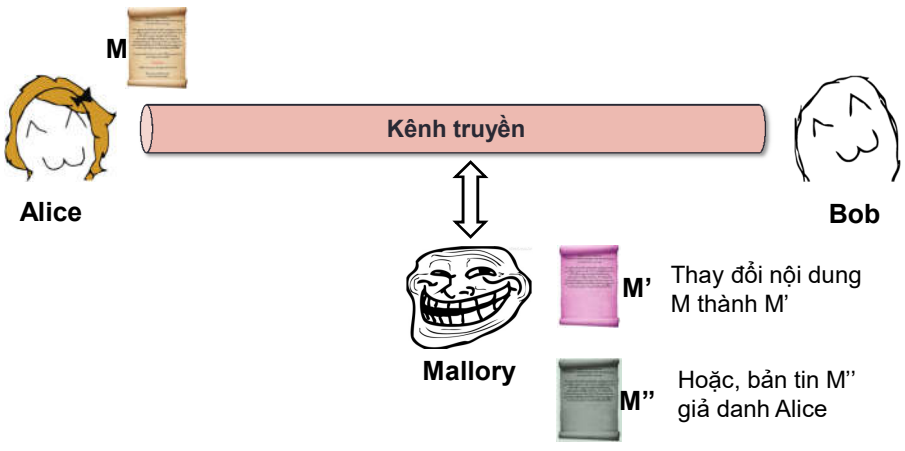
Nội dung

- Các vấn đề xác thực thông điệp
- Mã xác thực thông điệp (MAC)
- Hàm băm và hàm băm mật HMAC
- Chữ ký số

1. ĐẶT VẤN ĐỀ

3

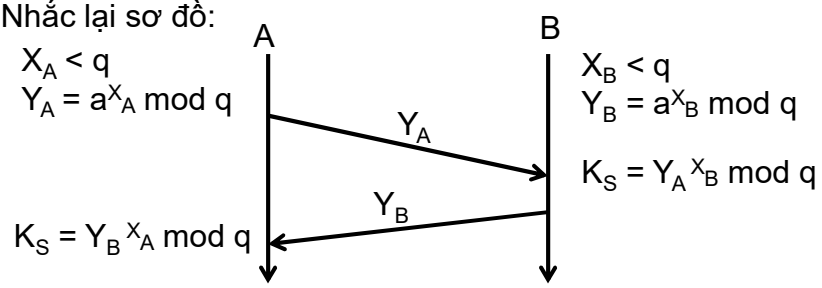
1. Đặt vấn đề



4

Một ví dụ - Tấn công vào sơ đồ trao đổi khóa Diffie-Hellman

- Nhắc lại sơ đồ:



- Kịch bản tấn công:
  - C sinh 2 cặp khóa (X'_A, Y'_A) và (X'_B, Y'_B)
  - Tráo khóa Y_A bằng Y'_A, Y_B bằng Y'_B
  - Hãy suy luận xem tại sao C có thể biết được mọi thông tin A và B trao đổi với nhau

5

Xác thực thông điệp

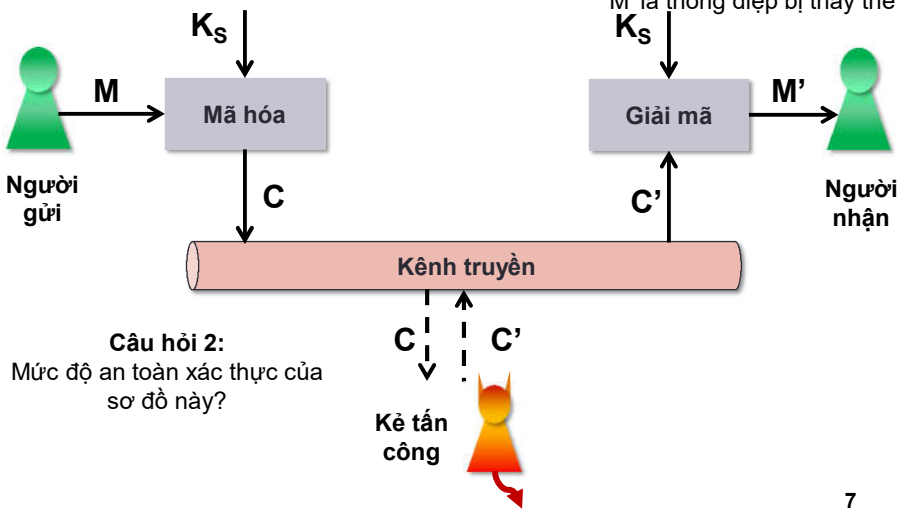
- Bản tin phải được xác minh:
  - Nội dung toàn vẹn: bản tin không bị sửa đổi
    - ✓ Bao hàm cả trường hợp Bob cố tình sửa đổi
  - Nguồn gốc tin cậy:
    - ✓ Bao hàm cả trường hợp Alice phủ nhận bản tin
    - ✓ Bao hàm cả trường hợp Bob tự tạo thông báo và "vu khống" Alice tạo ra thông báo này
  - Đúng thời điểm
  - → Các dạng tấn công điển hình vào tính xác thực: Thay thế (Substitution), Giả danh (Masquerade), tấn công phát lại (Reply attack), Phủ nhận (Repudiation)

6

Xác thực bằng mật mã khóa đối xứng

- Nhắc lại sơ đồ mật mã khóa đối xứng

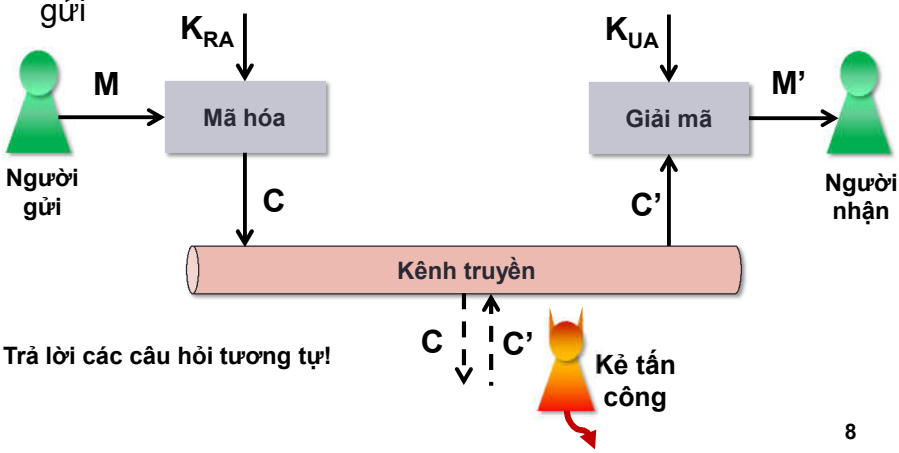
Câu hỏi 1:
Người nhận có nhận ra được M' là thông điệp bị thay thế?



Câu hỏi 2:
Mức độ an toàn xác thực của sơ đồ này?

Xác thực bằng mật mã khóa công khai

- Chúng ta đã biết sơ đồ bí mật: mã hóa bằng khóa công khai của người nhận
- Sơ đồ xác thực: mã hóa bằng khóa cá nhân của người gửi



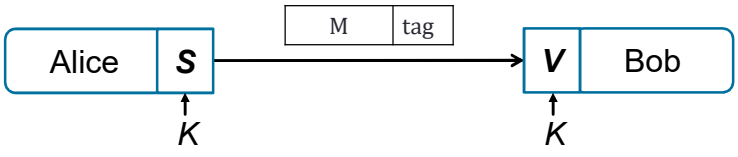
Trả lời các câu hỏi tương tự!

2. MÃ XÁC THỰC THÔNG ĐIỆP (MAC)

9

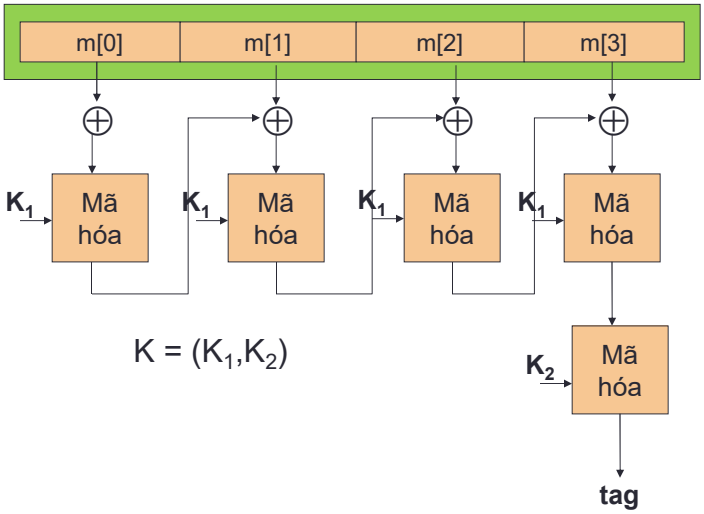
Message Authentication Code

- Xây dựng trên cơ sở hệ mật mã khóa đối xứng:
  - Hai bên đã trao đổi một cách an toàn khóa mật K
  - Sử dụng các thuật toán mã hóa khối ở chế độ CBC-MAC
- Bên gửi:
  - Tính toán tag t = MAC(K, M) : kích thước cố định, không phụ thuộc kích thước của M
  - Truyền (M||t)
- Bên nhận: xác minh Verify(K, M', t)
  - Tính t' = MAC(K,M')
  - So sánh: nếu t' = t thì Verify(K, M,t) = 1, ngược lại Verify(K, M,t) = 0



10

CBC-MAC



11

MAC – Ví dụ 1

Khách hàng chuyển khoản
1. K = KeyGen(l)
2. Xác thực thông tin CK:
   t = S(SoTK||money,K)

V **Ngân hàng**

Kẻ tấn công

t = A(SoTK'||money)

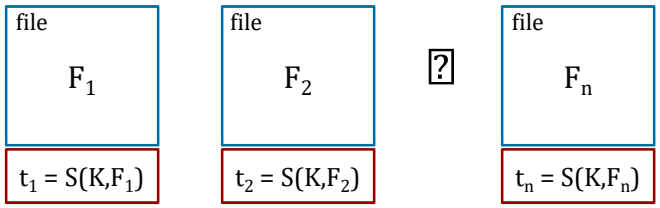
Mã MAC cho phép phát hiện thông tin bị sửa đổi

Thay đổi số tài khoản nhận tiền

12

MAC – Ví dụ 2: Phần mềm TripeWire

- Khi cài đặt, tính giá trị MAC của các file cần bảo vệ



- Khi máy tính khởi động khởi động, các file được kiểm tra mã MAC

→ Cho phép phát hiện các file bị sửa đổi (ví dụ do nhiễm virus)

13

Độ an toàn của MAC

- Giả sử M_1 và M_2 là hai bản tin có mã MAC giống nhau:

$MAC(M_1, K) = MAC(M_2, K)$

→$MAC(M_1||W, K) = MAC(M_2||W, K)$ với W bất kỳ

- Kịch bản tấn công:

1. Kẻ tấn công tính toán $t_x = MAC(M_i, K)$ với x = 1, …, N
2. Tìm cặp bản tin (M_i, M_j) có $t_i = t_j$. Nếu không tìm thấy thực hiện lại bước 1
3. Chọn bản tin W và tính $t = MAC(M_i||W, K)$
4. Thay $M_i || W$ bằng $M_j || W$ có lợi cho kẻ tấn công

14

Độ an toàn của MAC (tiếp)

- Kích thước bản tin: ℓ_M
- Kích thước tag: ℓ_t
- Nếu $\ell_M \leq \ell_t$ và ℓ_M không đổi: Mã MAC an toàn
- Nếu ℓ_M thay đổi: $|M| > |t|$ nên tồn tại $M_2 \neq M_1$ sao cho $MAC(M_2) = MAC(M_1)$

→MAC bị giảm tính an toàn

- Yêu cầu với giải thuật tạo MAC:
  - Nếu biết trước (M_1,t_1), rất khó tìm M_2 sao cho $MAC(M_2) = t_1$
  - Xác suất tìm được cặp bản tin M_1 và M_2 sao cho $t_1 = t_2$ không lớn hơn 2^{-n}
  - Giả sử M' là một dạng biến đổi của M, xác suất để t' = t lớn nhất là 2^{-n}

15

Tấn công phát lại (Replay attack)

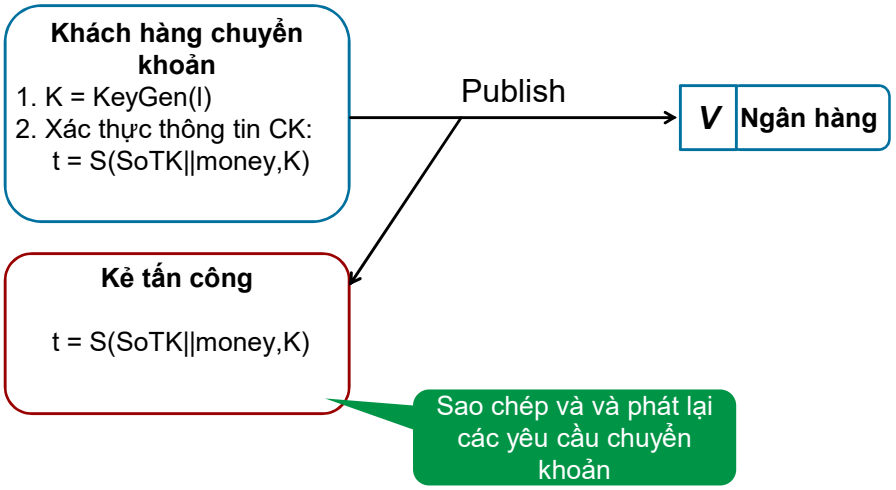
- Kẻ tấn công phát lại bản tin M đã được chứng thực trong phiên truyền thông trước đó
- Thiết kế MAC không chống được tấn công phát lại

→ cần thêm các yếu tố chống tấn công phát lại trong các giao thức truyền thông sử dụng MAC

- Một số kỹ thuật chống tấn công phát lại:
  - Giá trị ngẫu nhiên: MAC(M, Random, K)
  - Tem thời gian: MAC(M, TimeStamp, K)

16

Tấn công phát lại



17

Mật mã có xác thực

- Một hệ mật mã có xác thực (E, D) là một hệ mật mã mà

Hàm mã hóa E: K x M x N → C

Hàm giải mã D: K x C x N → $M \cup \{\perp\}$

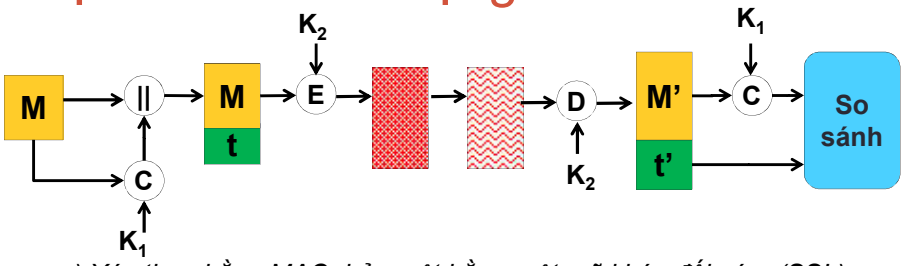
- Trong đó N là một dấu hiệu sử dụng để xác thực

Từ chối giải mã các bản mã không hợp lệ

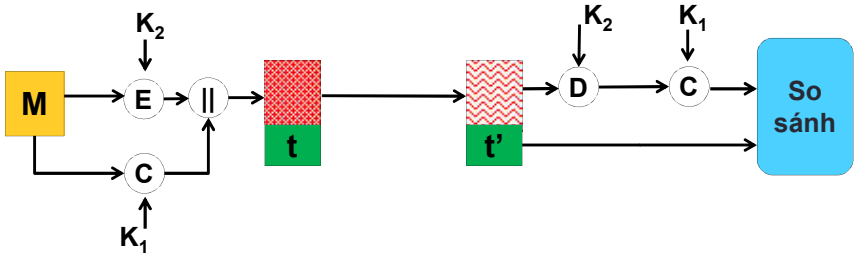
- Yêu cầu:
  - Chống tấn công chọn trước bản rõ, và
  - Kiểm tra được tính toàn vẹn của bản mật: xác suất kẻ tấn công tạo ra được một bản mật có thể giải mã là rất nhỏ
- Giải phát: Kết hợp mật mã và mã MAC

18

Một số sơ đồ sử dụng mã MAC



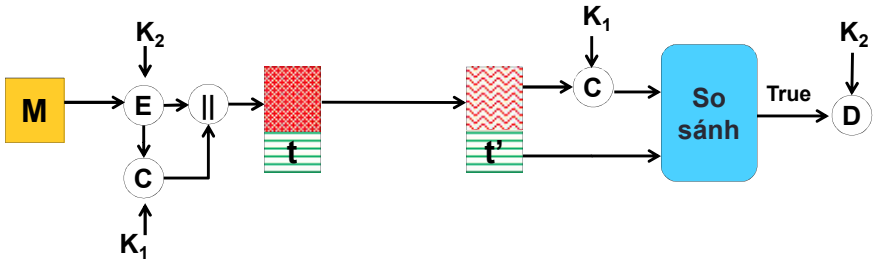
a) Xác thực bằng MAC, bảo mật bằng mật mã khóa đối xứng(SSL)



b) Xác thực bằng MAC, bảo mật bằng mật mã khóa đối xứng (SSH)

19

Một số sơ đồ sử dụng mã MAC(tiếp)



c) Xác thực bằng MAC, bảo mật bằng mật mã khóa đối xứng(IPSec)

- Một số chuẩn:

GCM: Mã hóa ở chế độ CTR sau đó tính CW-MAC

CCM: Tính CBC-MAC sau đó mã hóa ở chế độ CTR (802.11i)

EAX: Mã hóa ở chế độ CTR sau đó tính CMAC

- BTVN: Đánh giá độ an toàn của các sơ đồ

20

Nhận xét

Sơ đồ a

- Xác thực toàn vẹn bản rõ
- Không xác thực toàn vẹn bản mật(không phát hiện tấn công thay thế bản mật)
- MAC không cung cấp thông tin bản rõ

Sơ đồ b

- Xác thực toàn vẹn bản rõ
- Không xác thực toàn vẹn bản mật(có thể phát hiện bản mật bị thay thế)
- MAC chứa thông tin bản rõ

Sơ đồ c

- Xác thực toàn vẹn bản rõ
- Xác thực toàn vẹn bản mật(có thể phát hiện bản mật bị thay thế)
- MAC không chứa thông tin bản rõ

3.HÀM BĂM

22

Khái niệm

- *Hàm băm H*: thực hiện phép biến đổi:
  - Đầu vào: bản tin có kích thước bất kỳ
  - Đầu ra: giá trị *digest h = H(M)*có kích thước *n* bit cố định (thường nhỏ hơn rất nhiều so với kích thước bản tin đầu vào)
- Chỉ thay đổi 1 bit đầu vào, làm thay đổi hoàn toàn giá trị đầu ra
- Ví dụ:
  - Đầu vào: “The quick brown fox jumps over the lazy dog”
  - Mã băm: 2fd4e1c67a2d28fced849ee1bb76e7391b93eb12
  - Đầu vào: “The quick brown fox jumps over the lazy cog”
  - Đầu ra: de9f2c7fd25e1b3afad3e85a0bd17d9b100db4b3

23

Một hàm băm đơn giản

- Chia thông điệp thành các khối có kích thước n-bit
  - Padding nếu cần
- Thực hiện XOR tất cả các khối → mã băm có kích thước n bit
- Tất nhiên, hàm băm này không đủ an toàn để sử dụng trong bài toán xác thực thông điệp

$$m = \begin{bmatrix} m_1 \\ m_2 \\ \dots \\ m_l \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & m_{22} & \cdots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{l1} & m_{l2} & \cdots & m_{ln} \end{bmatrix}$$

$$\oplus \quad \oplus \quad \oplus \quad \oplus$$

$$\Downarrow \quad \Downarrow \quad \Downarrow \quad \Downarrow$$

$$\begin{bmatrix} c_1 & c_2 & \dots & c_n \end{bmatrix} = \text{H(m)}$$

24

Yêu cầu đối với hàm băm

1. Có thể áp dụng với thông điệp M với độ dài bất kỳ
2. Tạo ra giá trị băm h có độ dài cố định
3. H(M) dễ dàng tính được với bất kỳ M nào
4. Từ h rất khó tìm được M sao cho h = H(M): tính một chiều
5. Biết trước M_1 rất khó tìm được M_2 sao cho $H(M_1) = H(M_2)$
6. Rất khó tìm được cặp (M_1,M_2) sao cho $H(M_1)=H(M_2)$

25

Một số hàm băm phổ biến

- MD5
  - Kích thước digest: 128 bit
  - Công bố thuật toán tấn công đụng độ (collision attack) vào 1995
  - Năm 2005 tấn công thành công
- SHA-1
  - Kích thước digest: 160 bit
  - Đã có thuật toán tấn công đụng độ, nhưng chưa công bố tấn công thành công
- SHA-2: 256/512 bit

26

Tấn công ngày sinh

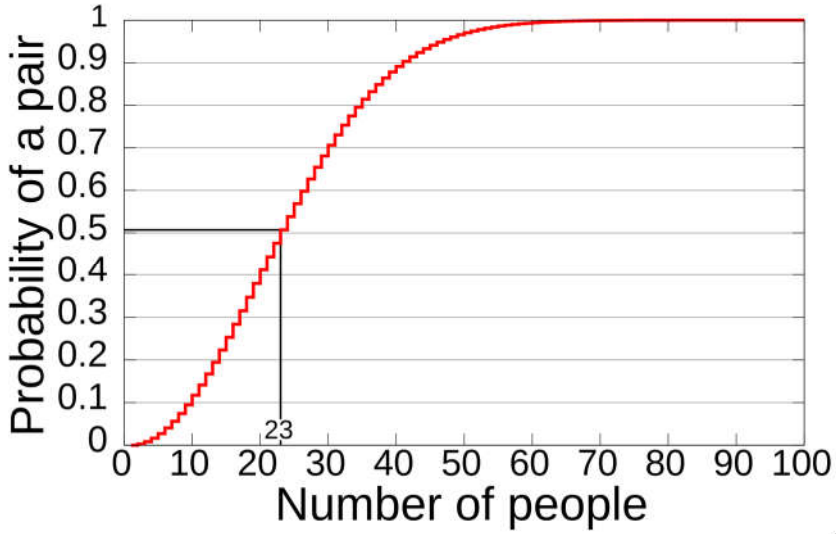
- Bài toán: Khi chọn n người bất kỳ, xác suất để có tối thiểu 2 người có trùng ngày sinh là bao nhiêu?
- Số cách chọn ra n người bất kỳ: 365^n
- Số cách chọn ra n người không có cặp nào trùng ngày sinh: 365 x 364 x … x (365-(n-1))
- Xác suất để chọn ra n người không có cặp nào trùng ngày sinh

$$Q = \frac{365 \times 364 \times \cdots \times (365 - (n-1))}{365^n}$$

- Xác suất cần tính: P = 1 – Q
- n = ? để P > 0.5 (cứ 2 lần chọn thì có 1 lần thỏa mãn)

27

Xác suất trong tấn công ngày sinh



28

Tấn công ngày sinh (Birthday paradox attack)

- h = H(M): kích thước n bit
  - n << $\mathscr{C}_M$ → luôn tồn tại $M_2 \neq M_1$ sao cho $H(M_2) = H(M_1)$

  → kẻ tấn công muốn được bản tin M_2 có lợi cho anh ta để thay thế M_1 đã được xác thực
- Phương pháp: vét cạn → số bản tin cần tính tối thiểu là bao nhiêu sẽ chắc chắn thành công?
- Cải tiến bằng tấn công ngày sinh: cho phép giảm số bản tin xuống chỉ còn $2^{n/2}$ với xác suất thành công là ≥ 0.5:
  - Công thức gần đúng tính xác suất thành công:

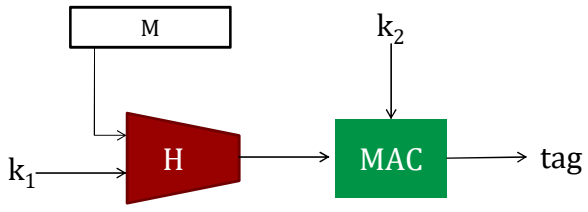
$$P(N,k) > 1 - e^{\frac{-k(k-1)}{2N}}$$

N: số giá trị *h*
k: số bản tin cần kiểm tra

29

HMAC

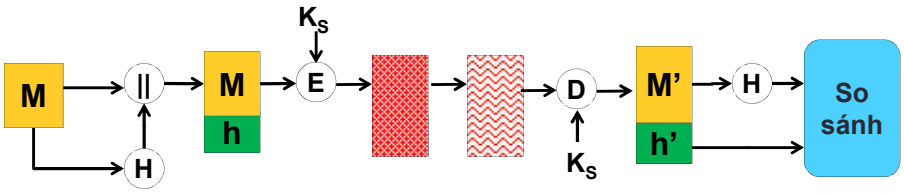
- Tương tự như CRC, hàm băm chỉ có thể phát hiện các lỗi ngẫu nhiên do nhiễu trong quá trình truyền
- Hashed MAC: kết hợp MAC và hàm băm để tăng cường an toàn cho hàm băm



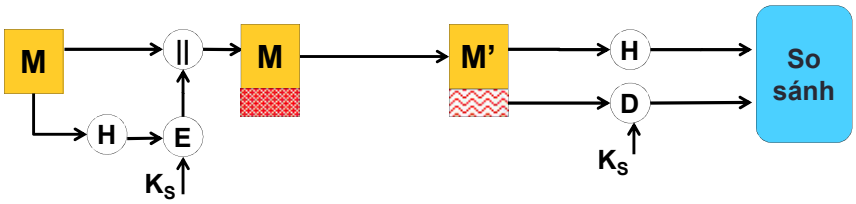
- Chữ ký số: kết hợp hàm băm với các phương pháp mật mã khóa công khai

30

Một số sơ đồ sử dụng hàm băm để xác thực



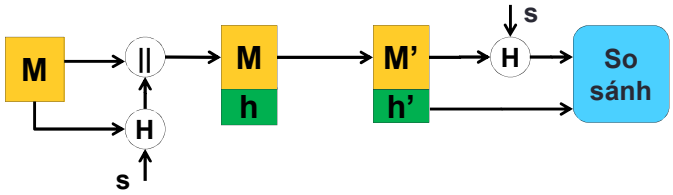
a) Xác thực thông điệp và bảo mật bằng mật mã khóa đối xứng



b) Xác thực thông điệp, mã băm được bảo vệ bằng mật mã khóa đối xứng

31

Một số sơ đồ sử dụng hàm băm để xác thực



c) Xác thực thông điệp sử dụng HMAC

Bài tập:

1. Kiểm tra những sơ đồ trên đáp ứng được yêu cầu nào về xác thực
2. Kết hợp sử dụng hệ mật mã khóa công khai để tạo ra một số sơ đồ mới

32

4. CHỮ KÝ ĐIỆN TỬ

Khái niệm – Digital Signature

- Chữ kí điện tử (chữ ký số) là đoạn dữ liệu được bên gửi gắn vào văn bản gốc trước khi truyền đi để chứng thực tác giả của văn bản và giúp người nhận kiểm tra tính toàn vẹn của dữ liệu mà mình thu được.
- Một số yêu cầu:
  - Chữ ký phải mang đặc trưng của người tạo văn bản
  - Chữ ký không thể sử dụng lại
  - Văn bản đã ký không được sửa đổi. Nếu có thì cần phải thực hiện ký lại trên văn bản mới.
- Đề xuất của Diffie-Hellman: sử dụng khóa cá nhân của người gửi để mã hóa bản tin
  - Hạn chế?

Chữ ký số

Sơ đồ chữ ký số gồm 2 hàm

- Hàm ký Sign(SK, M)
  - Đầu vào:
    - SK: Khóa cá nhân
    - M: Văn bản cần ký
  - Đầu ra: chữ ký số S
- Hàm kiểm tra: Vfy(PK, M, S)
  - Đầu vào:
    - PK: Khóa công khai
    - M, S
  - Đầu ra: True/False

- Hàm ký phải có tính ngẫu nhiên
- Bất kỳ ai có khóa SK đều có thể tạo chữ ký

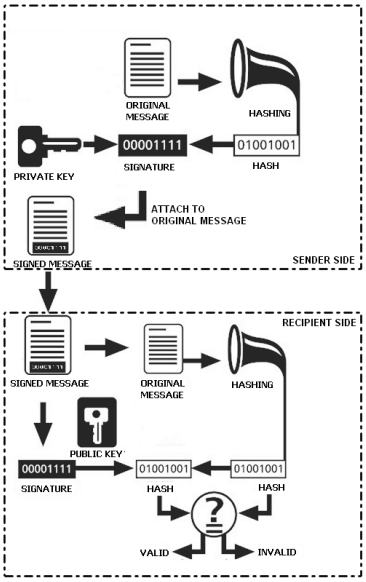
- Bất kỳ ai có khóa PK đều có thể kiểm tra chữ ký

35

Sơ đồ chung

- **Phía gửi :** hàm ký
  1. Băm bản tin gốc, thu được giá trị băm H
  2. Mã hóa giá trị băm bằng khóa riêng → chữ kí số S
  3. Gắn chữ kí số lên bản tin gốc (M || S)

- **Phía nhận :** hàm xác thực
  1. Tách chữ kí số S khỏi bản tin.
  2. Băm bản tin M, thu được giá trị băm H
  3. Giải mã S với khóa công khai của người gửi, thu được H'
  4. So sánh : H' và H''. Kết luận.



Một số dạng chữ ký số khác

- Chữ ký chống từ chối: quá trình kiểm tra chữ ký phải có sự tham gia của người ký
- Chữ ký mù
- Chữ ký nhóm
- Chữ ký ủy quyền
- Chữ ký chống sao chép, vi phạm bản quyền

37

An toàn cho chữ ký số

- Tính tin cậy của khóa công khai.
  - Vấn đề: kẻ tấn công làm sử dụng khóa công khai giả mạo. Nếu người dùng bị đánh lừa, họ sẽ tin cậy vào chữ ký giả mạo
  - Giải pháp: sử dụng hệ thống PKI để phát hành khóa công khai dưới dạng chứng thư số
- Tính an toàn của khóa cá nhân
  - Vấn đề: nếu khóa cá nhân bị kẻ tấn công đánh cắp, hắn có thể giả mạo chữ ký của người sở hữu khóa.
  - Giải pháp:
    - Bảo vệ bằng mật khẩu
    - Sử dụng thẻ thông minh(Smart Card)
    - Sử dụng thiết bị lưu trữ an toàn (USB Token)

38