BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC CẦN THƠ
TRƯỜNG CÔNG NGHỆ THÔNG TIN & TRUYỀN THÔNG

BÁO CÁO TỔNG HỢP
An Toàn Bảo Mật Thông Tin

Sinh viên thực hiện:
Lê Ngọc Phương Đình B2017032

Cần Thơ, 12/2023

Mục Lục

1) Thuật toán DES
   a) Cơ sở lý thuyết

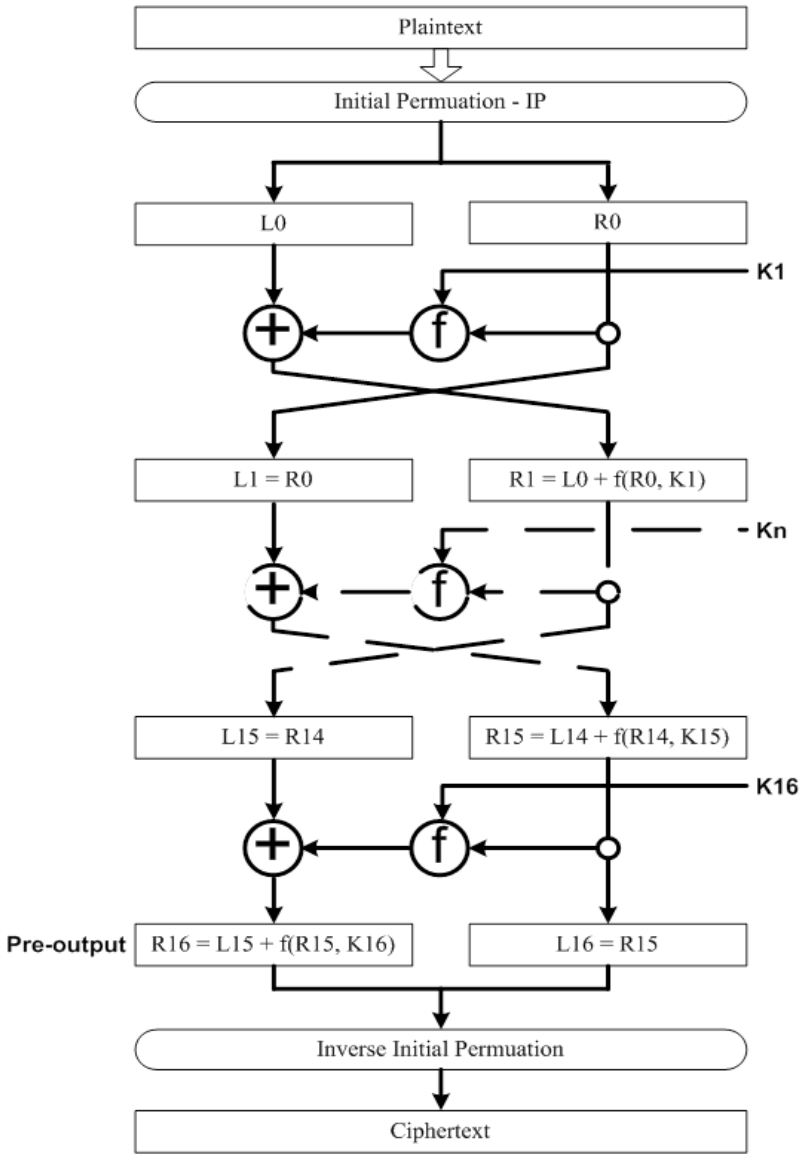
      Thuật toán DES (Data Encryption Standard) là một thuật toán mã hóa khối đối xứng được phát triển bởi IBM và được Cơ quan Tiêu chuẩn Quốc gia Hoa Kỳ (NIST) chấp nhận làm tiêu chuẩn mã hóa vào năm 1976. DES sử dụng khóa 64 bit để mã hóa và giải mã các khối dữ liệu 64 bit. DES là thuật toán mã hóa khối: nó xử lý từng khối thông tin của bản rõ có độ dài xác định và biến đổi theo những quá trình phức tạp để trở thành khối thông tin của bản mã có độ dài không thay đổi. Trong trường hợp của DES, độ dài mỗi khối là 64 bit. DES cũng sử dụng khóa để cá biệt hóa quá trình chuyển đổi. Nhờ vậy, chỉ khi biết khóa mới có thể giải mã được văn bản mã.

   b) Nguyên lý hoạt động

      i. Các bước thực hiện:

- Phân chia dữ liệu thành các khối 64 bit.
- Thêm một số bí mật vào đầu mỗi khối dữ liệu.
- Tiến hành 16 vòng mã hóa.
- Thêm một số bí mật khác vào cuối mỗi khối dữ liệu.
- Tái hợp các khối dữ liệu đã được mã hóa.

ii. Lưu đồ mã hoá



2) Thuật toán Affine

a) Cơ sở lý thuyết

Mật mã Affine là một dạng mật mã thay thế dùng một bảng chữ cái, trong đó mỗi chữ cái được ánh xạ tới một số sau đó mã hóa qua một hàm số toán học đơn giản. Trong mật mã Affine, đầu tiên bảng chữ cái của thông điệp cần mã hóa có kích thước m sẽ được chuyển thành các con số tự nhiên từ 0..m−1. Sau đó dùng một hàm mô đun để mã hóa và chuyển thành bản mã.

b) Nguyên lý hoạt động

- Phân chia dữ liệu thành các khối.
- Áp dụng hàm affine cho mỗi khối dữ liệu.
- Tái hợp các khối dữ liệu đã được mã hóa.

3) Thuật toán Caeser
   a) Cơ sở lý thuyết

      Mật mã Caesar còn gọi là mật mã dịch chuyển, là một trong những mật mã đơn giản và được biết đến nhiều nhất. Hệ mã Caesar là một hệ mã hóa thay thế đơn âm, làm việc trên bằng chữ cái tiếng Anh 26 ký tự. Đó là một dạng của mật mã thay thế, trong đó mỗi ký tự trong văn bản được thay thế bằng một ký tự cách nó một đoạn trong bảng chữ cái để tạo thành bản mã.

   b) Nguyên lý hoạt động
      i. Mã hoá
         EK(i) = (i+k) mod N
      ii. Giải mã
         DK(i) = (i-k) mod N

      Với N là 26
4) Thuật toán Hill
   a) Cơ sở lý thuyết

      Mật Mã Hill được đề xuất bởi Lester.S.Hill năm 1929. Mã cũng được thực hiện trên từng bộ m ký tự. mỗi ký tự trong bản mã một tổ hợp tuyến tính (trên vành Z26) của m ký tự trong bản rõ. Khoá sẽ được cho bởi một ma trận cấp m. Mã Hill là bộ (P, C, K, E, D), thỏa mãn: • P = C = Z26m, với m là một số nguyên dương. • K = {k Z26m x m | (|k|, m) = 1}.

   b) Nguyên lý hoạt động
      - Ma trận khóa (Key Matrix): Trong quá trình mã hóa và giải mã, một ma trận khóa vuông K được sử dụng. Kích thước của ma trận này phải là một số bình phương và nghịch đảo của nó trong modulo một số nguyên nào đó phải tồn tại.
      - Chuỗi văn bản và phân chia: Đầu tiên, văn bản cần mã hóa được chia thành các khối con với độ dài bằng với kích thước của ma trận khóa.
      - Ánh xạ các ký tự thành số: Mỗi ký tự trong khối văn bản được ánh xạ thành một số tương ứng, thường là dựa trên bảng mã ASCII.
      - Chuyển đổi khối thành vectơ: Mỗi khối số được chuyển đổi thành một vectơ cột, với mỗi phần tử của vectơ tương ứng với một số trong khối.
      - Mã hóa và giải mã: Quá trình mã hóa và giải mã được thực hiện bằng cách nhân ma trận khóa với vectơ đại diện cho mỗi khối. Công thức chung cho mã hóa một khối X là Y=K.XmodM, trong đó Y là vectơ kết quả và M là một số nguyên modulo.
      - Mã hóa: Y là vectơ kết quả, K là ma trận khóa, X là vectơ đại diện cho khối văn bản.
      - Giải mã: X là vectơ kết quả, K^(-1) là ma trận nghịch đảo của ma trận khóa, Y là vectơ đại diện cho khối mã hóa.
      - Chuyển đổi kết quả thành chuỗi ký tự: Vectơ kết quả sau cùng được chuyển đổi ngược lại thành chuỗi ký tự hoặc ký tự tương ứng dựa trên bảng mã ASCII.

5) Thuật toán Vigenère

a) Cơ sở lý thuyết

Mật mã Vigenère là một phương pháp mã hóa văn bản bằng cách sử dụng xen kẽ một số phép mã hóa Caesar khác nhau dựa trên các chữ cái của một từ khóa. Nó là một dạng đơn giản của mật mã thay thế dùng nhiều bảng chữ cái.

b) Nguyên lý hoạt động

- Đối với hệ mã hoá này không gian các bản mã và bản rõ là các thông điệp được tạo thành từ một bảng chữ cái A (giống hệ mã Caesar). Các chữ cái được đánh số từ 0 tới n - 1 trong đó n là số phần tử của bảng chữ cái.
- Không gian khoá K được xác định như sau:
- Với m là một số nguyên dương, khoá K là một xâu ký tự có độ dài
- Định nghĩa sơ đồ hệ mật mã:

$$P = C = K = (Z_n)^m$$

- Với khoá K = (k1, k2, k3, k4,…,km)
- Để mã hoá một bản rõ P, ta chia P thành các đoạn có độ dài và chuyển thứ tự tương ứng của chúng trong bảng không gian mã. Chẳng hạn X = (x1, x2, x3, x4,…,xm). Khi đó quá trình mã hoá và giải mã được thực hiện như sau:

$$e_K(x_1, x_2, x_3, \ldots, x_m) = \{(x_1 + k_1) \pmod n, \ldots, (x_m + k_m) \pmod n\}$$

$$d_K(y_1, y_2, y_3, \ldots, y_m) = \{(y_1 - k_1) \pmod n, \ldots, (y_m - k_m) \pmod n\}$$

Với n là không gian mã hoá, $Y = y_1y_2y_3 \ldots y_m$ là một đoạn bảng mã; $X, Y \in (Z_n)^m$

6) Thuật toán RSA

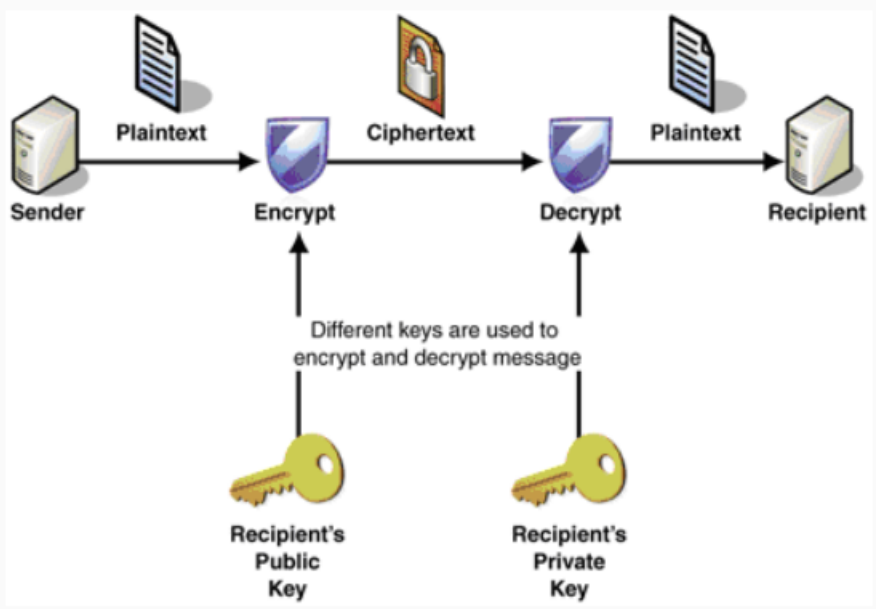
a) Cơ sở lý thuyết

Thuật toán RSA có hai khóa: khóa công khai (public key) và khóa bí mật (private key). Mỗi khóa là những số cố định sử dụng trong quá trình mã hóa và giải mã. Khóa công khai được công bố rộng rãi cho mọi người và được dùng để mã hóa. Những thông tin được mã hóa bằng khóa công khai chỉ có thể được giải mã bằng khóa bí mật tương ứng. Nói cách khác, mọi người đều có thể mã hóa nhưng chỉ có người biết khóa cá nhân (private) mới có thể giải mã được.

b) Nguyên lý hoạt động

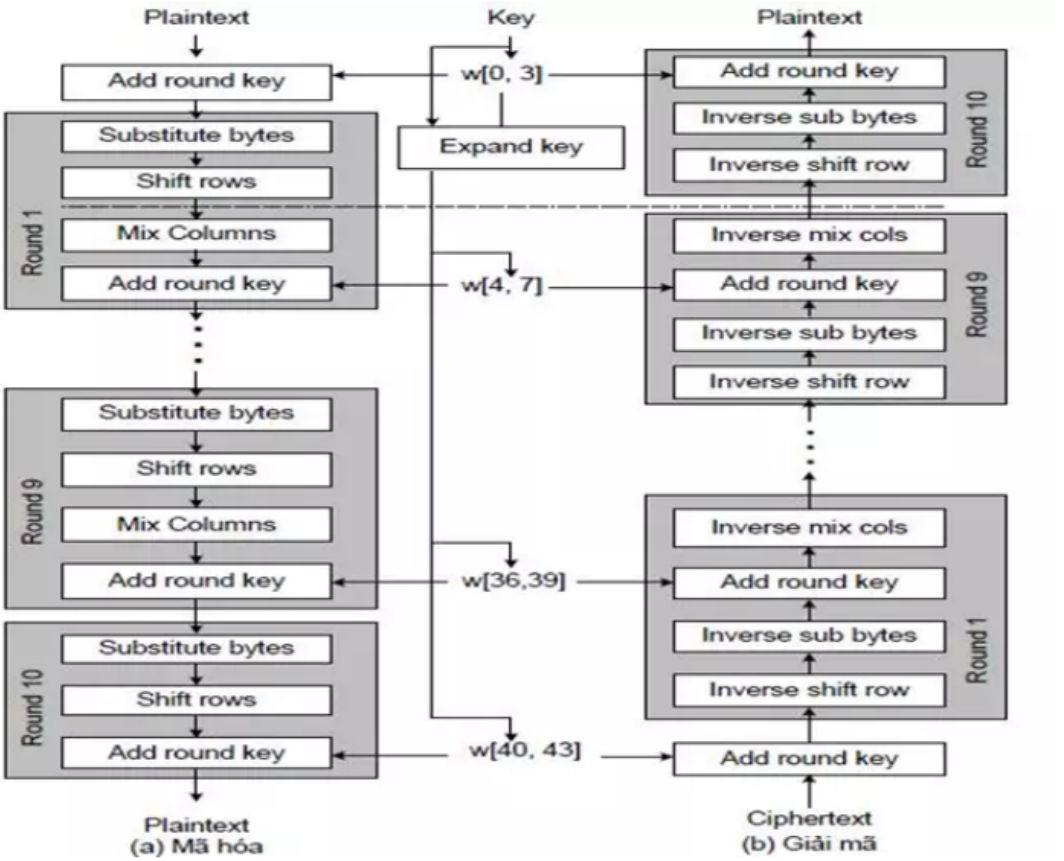
- Chọn hai số nguyên tố lớn ngẫu nhiên, gọi chúng là p và q. Độ dài của các số nguyên tố này thường được chọn sao cho kết hợp của chúng có độ dài mong muốn của khóa RSA.
- Tính n = p x q. Đây là một phần của khóa công khai và khóa bí mật.
- Tính φ(n)=(p−1)×(q−1), là số Euler của n.

- Chọn một số nguyên e sao cho $1<e<\varphi(n)$ và e là số nguyên tố cùng nhau với $\varphi(n)$. e sẽ là một phần của khóa công khai.
- Tìm số nguyên d sao cho $d \times e \equiv 1 (\mod \varphi(n))$. d là nghịch đảo modular của e, và nó sẽ là một phần của khóa bí mật.
- Khóa công khai sẽ bao gồm cặp (e,n).
- Khóa bí mật sẽ bao gồm cặp (d,n).
- Mã hóa: Để mã hóa một thông điệp M, tính $C \equiv M^e (\mod n)$.
- Để giải mã C và nhận lại M, tính $M \equiv C^d (\mod n)$.



7) Thuật toán AES(Advanced Encryption Standard)
   a) Cơ sở lý thuyết
      - Thuật toán được xây dựng dựa trên Rijndael Cipher phát triển bởi 2 nhà mật mã học người Bỉ: Joan Daemen và Vincent Rijmen. AES làm việc với các khối dữ liệu 128bit và độ dài khóa 128bit, 192bit hoặc 256bit. Các khóa mở rộng sử dụng trong chu trình được tạo ra bởi thủ tục sinh khóa Rijndael.
      - Hầu hết các phép toán trong thuật toán AES đều thực hiện trong một trường hữu hạn của các byte. Mỗi khối dữ liệu đầu vào 128bit được chia thành 16byte, có thể xếp thành 4 cột, mỗi cột 4 phần tử hay một ma trận 4x4 của các byte, nó gọi là ma trận trạng thái.

b) Nguyên lý hoạt động



8) Tổ chức tin cậy (Trust Service Provider - TSP)

Tổ chức tin cậy là tổ chức cung cấp các dịch vụ nhằm thiết lập niềm tin trong các giao dịch điện tử và môi trường mạng. Một số dịch vụ chính của TSP bao gồm:

- Dịch vụ chứng thực chữ ký số: Cấp, quản lý và xác nhận chữ ký số điện tử.
- Dịch vụ chứng thực thời gian: Xác nhận thời điểm tạo, gửi, nhận thông tin điện tử.
- Dịch vụ chứng thực nguồn gốc: Xác nhận nguồn gốc của thông tin.
- Dịch vụ bảo mật, lưu trữ thông tin: Bảo vệ tính bảo mật và toàn vẹn dữ liệu.
- Dịch vụ xác thực website: Xác minh định danh và tính xác thực trang web.
- Dịch vụ kiểm toán, đánh giá an toàn thông tin.
- Để trở thành tổ chức tin cậy, doanh nghiệp cần đầu tư hạ tầng công nghệ hiện đại, tuân thủ các tiêu chuẩn và quy định nghiêm ngặt về bảo mật. TSP góp phần quan trọng trong việc xây dựng niềm tin số.

9) Chữ ký số

a) Cơ sở lý thuyết

Chữ ký số (digital signature) là một kỹ thuật toán học được dùng để cung cấp tính xác thực, tính yoàn vẹn và không thoái thác dưới dạng ký mã và chứng chỉ số.

Có 3 thuật toán tiêu chuẩn chữ ký số(DSS) được sử dụng để tạo và xác minh chữ ký số:

- Thuật toán chữ ký số
- Thuật toán Rivest-Shamir Adelman (RSA)
- Thuật toán chữ ký số đường cong elip - Eliptic Curve Digital Sigature Algorithm (ECDSA)

Chữ ký số thường được sử dụng trong 2 trường hợp sau:

- Code signing(Ký mã): được sử dụng để xác minh tính toàn vẹn của tệp được thực thi được tải xuống từ trang web của nhà cung cấp.
- Digital certificates (chứng chỉ số): chúng được sử dụng để xác thực danh tính của hệ thống và trao đổi dữ liệu bí mật.

b) Nguyên lý hoạt động

- Người gửi sử dụng thuật toán băm (hash function) để tính toán ra một giá trị băm (hash value) từ nội dung thông điệp ban đầu.
- Giá trị băm này sau đó được mã hóa bằng khóa bí mật (private key) của người gửi tạo thành chữ ký số.
- Người gửi đính kèm chữ ký số vào thông điệp và gửi đi.
- Người nhận dùng khóa công khai (public key) của người gửi để giải mã chữ ký số về giá trị băm ban đầu.
- Người nhận dùng cùng thuật toán băm để tính toán giá trị băm từ nội dung thông điệp nhận được.
- Nếu 2 giá trị băm trên trùng khớp nhau thì thông điệp là chính xác và đến từ người gửi.

❖ Nhờ đó, chữ ký số đảm bảo tính xác thực, toàn vẹn và không thể phủ nhận thông điệp. Người nhận có thể xác minh chữ ký mà không cần biết khóa bí mật của người gửi.

10) Hash Fuction

a) Cơ sở lý thuyết

Hàm băm là giải thuật nhằm sinh ra các giá trị băm tương ứng với mỗi khối dữ liệu. Giá trị băm đóng vai gần như một khóa để phân biệt các khối dữ liệu.

b) Nguyên lý hoạt động

Dựa trên việc ánh xạ dữ liệu từ không gian lớn về không gian nhỏ hơn thông qua các phép toán nhất định. Cụ thể:

- Hàm băm nhận vào dữ liệu đầu vào (vd: chuỗi ký tự) có độ dài bất kỳ.
- Dữ liệu đầu vào sẽ được chia thành các khối có độ dài cố định (vd: 1024 bit). Mỗi khối sẽ được xử lý riêng biệt.
- Mỗi khối đầu vào sẽ đi qua một chuỗi các phép toán logic, như XOR, bit shift, modulo,... tạo ra một giá trị hash ngắn hơn.

- Các giá trị hash của từng khối sau đó được kết hợp lại thành một giá trị hash cuối cùng duy nhất đại diện cho toàn bộ dữ liệu đầu vào.
- Hàm băm được thiết kế để các giá trị hash sinh ra phân bố đồng đều trong không gian hash.
- Hàm băm là một hàm một chiều, không thể đảo ngược từ hash value để tìm ra input ban đầu.

 Nhờ cơ chế trên, hàm băm có thể ánh xạ dữ liệu lớn về không gian nhỏ hơn một cách hiệu quả, tạo điều kiện cho việc lưu trữ và tìm kiếm dữ liệu.

c) Hàm băm mật mã

Là một hàm băm với một số tính chất bảo mật nhất định để phù hợp việc sử dụng trong nhiều ứng dụng bảo mật thông tin đa dạng. Chẳng hạn như chứng thực và kiểm tra tính nguyên vẹn của thông điệp.

Tính chất cơ bản:

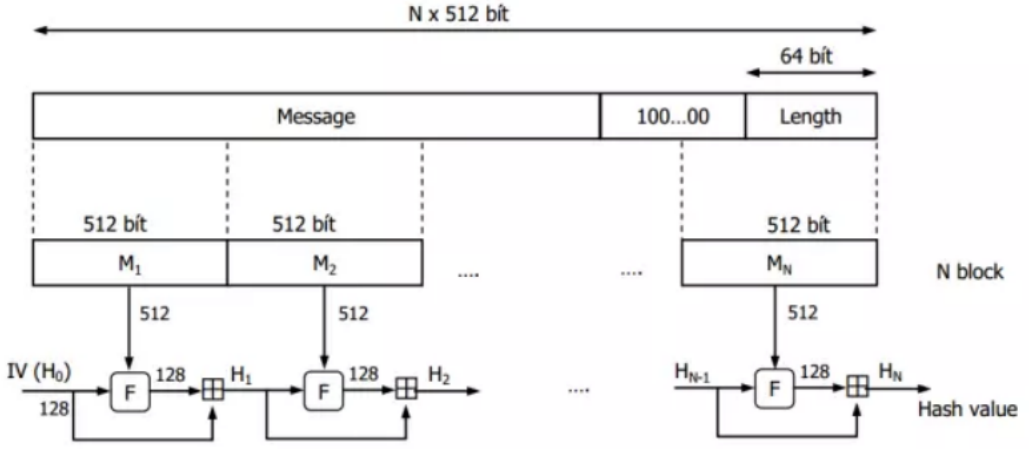
- Tính một chiều hay chống tiền ảnh
- Tính chống tiền ảnh thứ hai
- Tính chống xung đột

i. Giải thuật MD5

MD5 là viết tắt của Message-Digest algorithm 5 – Giải thuật Tiêu hóa tin 5, là một hàm băm được sử dụng phổ biến với giá trị Hash dài 128-bit. Giải thuật gồm 5 bước thao tác trên khối 512 bits.

Các bước thực hiện:

- Bước 1: nhồi dữ liệu
- Bước 2: thêm vào độ dài
- Bước 3: khởi tạo bộ đệm MD
- Bước 4: xử lý các khối dữ liệu 512 bits
- Bước 5: xuất kết quả



ii. Giải Thuật SHA

SHA (Secure Hash Algorithm hay Thuật toán Băm An toàn) là năm thuật toán được chấp nhận bởi FIPS dùng để chuyển một đoạn dữ liệu nhất định thành một đoạn dữ liệu có chiều dài không đổi với xác suất khác biệt cao.

Có 5 thuật toán SHA là SHA-1, SHA-224, SHA-256, SHA-384, SHA-512.

iii. Ứng dụng của hàm băm

- Key Stretching (tạo khoá bí mật từ mật khẩu)
- Integrity checking (kiểm tra tính toàn vẹn của dữ liệu)
- HMAC - Hashed Message Authentication Code (mã chứng thực thông điệp sử dụng hàm băm)
- Chữ ký điện tử

11) Tấn công mạng

a) Cơ sở lý thuyết

Tấn công là hoạt động có chủ ý của kẻ phạm tội lợi dụng các thương tổn của hệ thống thông tin và phá vỡ tính sẵn sàng, tính toàn vẹn, tính bí mật của hệ thống.

b) 5 loại xu hướng tấn công mạng:

- Khai thác lỗ hỏng mạng
- Từ chối dịch vụ : Truy cập máy chủ với số lượng lớn làm sập mạng
- Vi phạm chính sách
- Truy cập trái phép
- Tấn công mã độc : là phần mềm do các tin tặc hay các kẻ phá hoại tạo ra nhằm phá hoại hệ thống máy tính.

12) Các khái niệm về bản rõ, bản mật và giải mật.

- Bản rõ (Plaintext): Dữ liệu rõ ban đầu được gọi là Bản rõ.
- Dữ liệu đã được mật mã hóa còn được gọi là Bản mật (Ciphertext).
- Bản mật sẽ được giải mật (Decrypted) trở về bản rõ ban đầu.