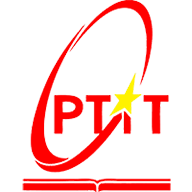
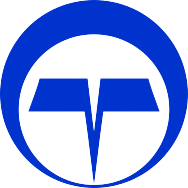


**HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG**

**VIỆN KHOA HỌC KỸ THUẬT BƯU ĐIỆN**



**BÁO CÁO MÔN HỌC**

**AN TOÀN VÀ BẢO MẬT HỆ THỐNG THÔNG TIN**

**CHỦ ĐỀ 9: CÁC HÀM BĂM, CHỮ KÝ SỐ, CHỨNG CHỈ SỐ VÀ KPI**

**Lớp:** INT1303-20242-13

**Giảng viên hướng dẫn:** Trần Đức Sự

**Nhóm thực hiện:**

Nguyễn Khả Phong - B23DCCC129

Phạm Tiến Công - B23DCCC025

Đinh Hoàng Long - B23DCCC102

***Hà Nội – 2/2025***

# Các hàm băm

## Khái niệm

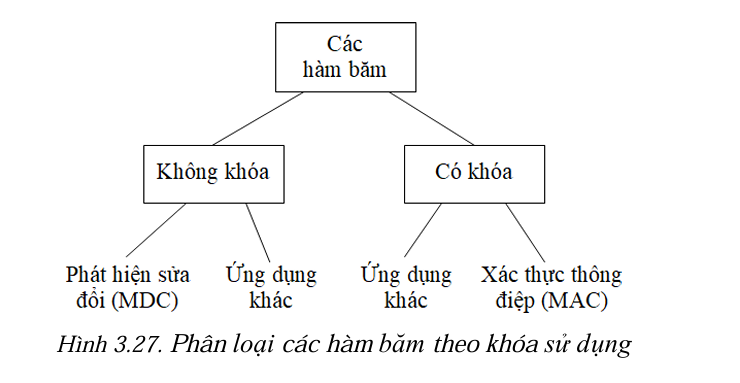
* + - Hàm băm (Hash function) là một ánh xạ chuyển đổi dữ liệu đầu vào có kích thước thay đổi sang dữ liệu đầu ra có kích thước cố định để tạo các bản tóm lược/ chuỗi đại diện (digest) của thông điệp, thường được sử dụng để nhận dạng và đảm bảo tính toàn vẹn của thông điệp.
    - Là một hàm toán học h có tối thiểu 2 thuộc tính: Nén và Dễ tính toán   
      VD : Khi lưu mật khẩu người dùng thì mật khẩu sẽ không được lưu trực tiếp trong database mà sẽ được thông qua một hàm băm. Kết quả output của hàm băm lúc này mới được lưu trong database

## Tính chất cơ bản

* + - Dễ tính toán: Trả về giá trị hash đầu vào một cách nhanh chóng
    - Tính xác định : Với một đầu vào chỉ cho một đầu ra duy nhất.Bất kể bạn nhập một kết quả bao nhiều lần đều cho ra một kết quả
    - Không thể đảo ngược ( Kháng tiền ảnh): Không thể xác định dữ liệu đầu vào dựa trên hash đầu ra ( không khả thi )
    - Chống va chạm : Mỗi đầu vào sẽ có hàm băm độc đáo của riêng mình
    - Thay đổi nhỏ trong đầu vào làm thay đổi đầu ra
    - Chấp nhận nhiều kiểu dữ liệu

## Phân loại các hàm băm

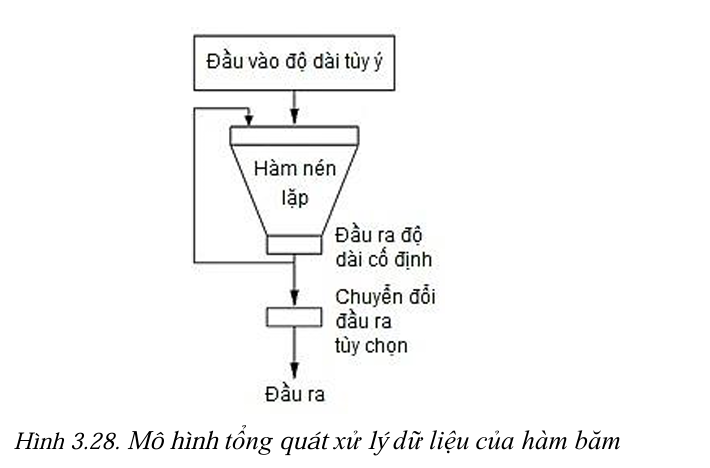
* + - Dựa trên khóa sử dụng
      * Hàm băm không khóa : chỉ gồm thông điệp. Các mã phát hiện sửa đổi (MDC – Modification Detection Code) được sử dụng rộng rãi nhất.
      * Hàm băm có khóa: thông điệp và khóa bí mật. Các mã xác thực thông điệp (MAC - Message Authentication Code) được sử dụng rộng rãi nhất



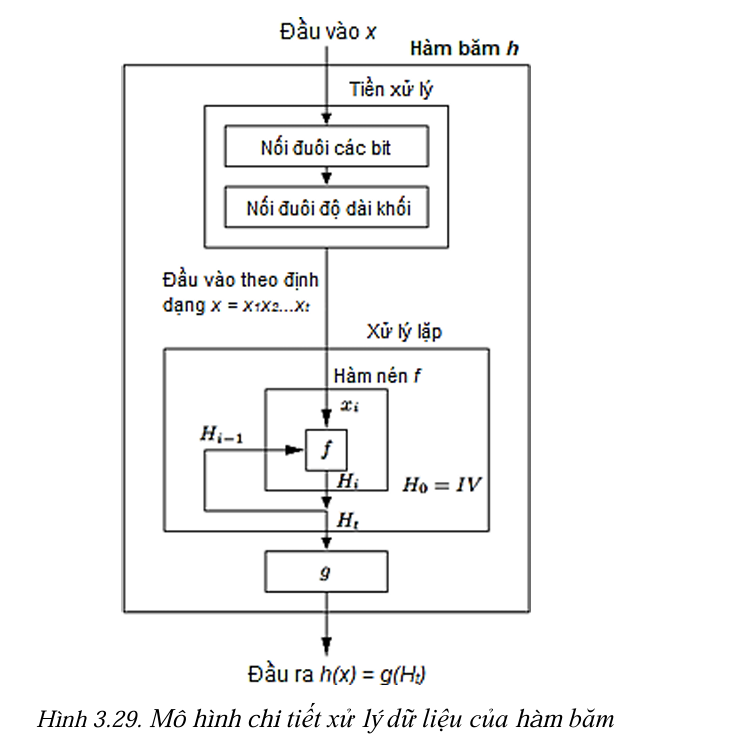
* + - Dựa trên chức năng
      * Mã phát hiện sửa đổi (MDC - Modification Detection Code) : MDC thường được sử dụng để tạo chuỗi đại diện cho thông điệp và dùng kết hợp với các kỹ thuật khác (như chữ ký số) để đảm bảo tính toàn vẹn của thông điệp. MDC thuộc loại hàm băm không khóa. MDC gồm 2 loại nhỏ:
        + Hàm băm một chiều (OWHF - One-way hash functions): Với hàm băm một chiều, việc tính giá trị băm là dễ dàng, nhưng việc khôi phục thông điệp từ giá trị băm là rất khó khăn;
        + Hàm băm chống đụng độ (CRHF - Collision resistant hash functions): Với hàm băm chống đụng độ, sẽ là rất khó để tìm được 2 thông điệp khác nhau nhưng có cùng giá trị băm.
      * Mã xác thực thông điệp (MAC - Message Authentication Code) : MAC cũng được dùng để đảm bảo tính toàn vẹn của thông điệp mà không cần một kỹ thuật bổ sung nào khác. MAC là loại hàm băm có khóa như đã đề cập ở trên, với đầu vào là thông điệp và một khóa bí mật.

## Mô hình xử lý dữ liệu

* + - Mô hình tổng quát xử lý dữ liệu của các hàm băm
      * Theo đó, thông điệp đầu vào với độ dài tùy ý đi qua hàm nén lặp nhiều vòng để tạo chuỗi đầu ra có kích thước cố định. Chuỗi này đi qua một khâu chuyển đổi định dạng tùy chọn để tạo ra chuỗi băm kết quả.



* + - Quá trình xử lý dữ liệu của các hàm băm.
      * B1 (Tiền xử lý) : thông điệp đầu vào x trước hết được nối đuôi thêm một số bit và kích thước khối, sau đó chia thành các khối có kích thước xác định. Kết quả của bước này là t khối dữ liệu có cùng kích thước có dạng x = x1x2…xt làm đầu vào cho bước 2.
      * B2 (Xử lý lặp) : từng khối dữ liệu xi được xử lý thông qua hàm nén f để tạo đầu ra là Hi. Kết quả là chuỗi đầu ra Ht
      * B3 (Chuyển đổi định dạng) : Ht được chuyển đổi định dạng bởi hàm g để tạo chuỗi giá trị băm hết quả h(x).



## Một số hàm băm thông dụng

* + - Hàm băm MD5(Message Digest) : kích thước 128bit. MD5 được sử dụng khá rộng rãi trong nhiều ứng dụng, như tạo chuỗi đảm bảo tính toàn vẹn thông điệp, tạo chuỗi kiểm tra lỗi, hoặc kiểm tra tính toàn vẹn dữ liệu (Checksum) và mã hóa mật khẩu trong các hệ điều hành và các ứng dụng. MD5 hiện nay được khuyến nghị không nên sử dụng do nó không còn đủ an toàn
    - Hàm băm SHA-1 (Secure Hash Algorithm) : kích thước 160bit. SHA1 được sử dụng rộng rãi để đảm bảo tính xác thực và toàn vẹn thông điệp.

## Ứng dụng

1. Lưu chữ mật khẩu : Phương pháp bảo vệ mật khẩu hiệu quả nhất là dùng hàm băm. Khi người sử dụng đăng ký mật khẩu, giá trị băm của mật khẩu được tính bằng một hàm băm nào đó (MD5 hay SHA-1,…) Giá trị băm được lưu trữ vào file hay cơ sở dữ liệu. Vì hàm băm là một chiều, nên dù biết được giá trị băm và loại hàm băm, hacker cũng không thể suy ra được mật khẩu. Khi người sử dụng đăng nhập, mật khẩu đăng nhập được tính giá trị băm và so sánh với giá trị băm đang được lưu trữ. Do tính chống trùng, chỉ có một mật khẩu duy nhất có giá trị băm tương ứng, nên không ai khác ngoài người sử dụng có mật khẩu đó mới có thể đăng nhập ứng dụng.
2. Đấu giá trực tuyến : Phương pháp lưu trữ mật khẩu bằng giá trị Hash cũng được áp dụng tương tự cho việc đấu giá trực tuyến bằng hình thức đấu giá bít mật. Giả sử Alice, Bob và Trudy cùng tham gia đấu giá, họ sẽ cung cấp mức giá của mình cho trọng tài. Các mức giá này được giữ bí mật cho đến khi cả ba đều nộp xong. Nếu ai là người đưa ra mức giá cao nhất thì thắng thầu
3. Download file : Khi chúng ta download file từ mạng internet, nếu chất lượng mạng không tốt thì có thể xảy ra lỗi trong quá trình download làm cho file tại máy client khác với file trên server. Hàm băm có thể giúp chúng ta phát hiện ra những trường hợp bị lỗi như vậy.
4. Chữ ký số
5. Bảo mật cho công nghệ Blockchain
6. Mã hóa dữ liệu

# Chữ ký số

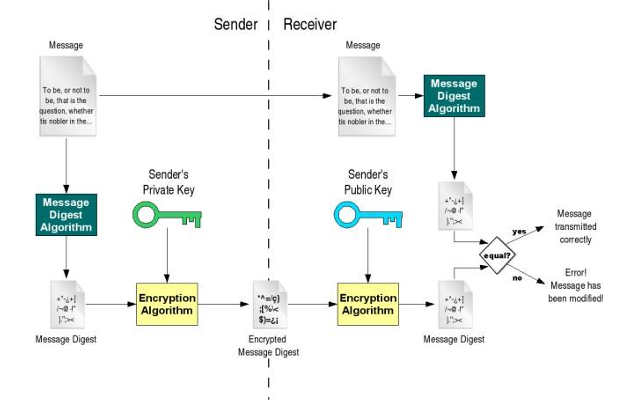
## Một số khái niệm

* + - Chữ ký số: Chuỗi dữ liệu liên kết với thông điệp và thực thể tạo ra thông điệp
    - Giải thuật tạo chữ ký số: Phương pháp sinh chữ ký số
    - Giải thuật kiểm tra chữ ký số: Phương pháp xác minh tính xác thực
    - Hệ chữ ký số: Bao gồm cả giải thuật tạo và kiểm tra chữ ký số
    - Quá trình tạo chữ ký số gồm:
      * Giải thuật tạo chữ ký số
      * Phương pháp chuyển dữ liệu thông điệp thành dạng có thể ký được
    - Quá trình kiểm tra chữ ký số gồm:
      * Giải thuật kiểm tra chữ ký số
      * Phương pháp khôi phục dữ liệu từ thông điệp

## Tính chất chữ ký số

* + - Tính xác thực (Authentication): Xác định được người ký
    - Tính toàn vẹn (Integrity): Phát hiện được sự thay đổi của dữ liệu
    - Tính không thể chối bỏ (Non-repudiation): Người ký không thể phủ nhận
    - Tính duy nhất: Mỗi chữ ký chỉ dùng được một lần

## Quá trình ký và kiểm tra



*Quá trình tạo chữ ký và kiểm tra chữ ký số*

* + - Quy trình chi tiết tạo chữ ký số (Bên người gửi - Sender):

Các bước thực hiện:

* + - * Tính toán chuỗi đại diện (message digest/hash value) của thông điệp bằng giải thuật băm
      * Sử dụng khóa riêng để ký chuỗi đại diện tạo chữ ký số
      * Ghép thông điệp ban đầu với chữ ký số thành thông điệp đã ký
      * Gửi thông điệp đã ký cho người nhận
    - Quy trình chi tiết kiểm tra chữ ký số (Bên người nhận - Receiver):

Các bước thực hiện:

* + - * Tách riêng chữ ký số và thông điệp gốc
      * Tính toán chuỗi đại diện MD1 từ thông điệp gốc
      * Giải mã chữ ký số bằng khóa công khai để có MD2
      * So sánh MD1 và MD2:
        + Nếu MD1 = MD2: Chữ ký hợp lệ
        + Nếu MD1 ≠ MD2: Chữ ký không hợp lệ

## Các giải thuật chữ ký số

* + - Giải thuật chữ ký số RSA
      * Đặc điểm
        + Có thể dùng cho cả mã hóa và chữ ký số
        + Quá trình kiểm tra nhanh
        + Được sử dụng phổ biến nhất hiện nay
      * Khác biệt khi dùng cho mã hóa:
        + Người nhận sở hữu cặp khóa (công khai, riêng)
        + Mã hóa dùng khóa công khai người nhận
        + Giải mã dùng khóa riêng người nhận
      * Khác biệt khi dùng cho chữ ký số:
        + Người gửi sở hữu cặp khóa (công khai, riêng)
        + Ký bằng khóa riêng người gửi
        + Kiểm tra bằng khóa công khai người gửi
      * Ví dụ triển khai

***# Sinh khóa***

***p = 61 # số nguyên tố thứ nhất***

***q = 53 # số nguyên tố thứ hai***

***n = p \* q # 3233***

***e = 17 # số mũ công khai***

***d = 2753 # số mũ bí mật***

***# Ký số***

***message = 123***

***signature = pow(message, d, n) # (123^2753) mod 3233***

***# Kiểm tra***

***verify = pow(signature, e, n) # Nếu verify == message thì chữ ký hợp lệ***

* + - Giải thuật chữ ký số DSA

DSA gồm 3 khâu chính:

* + - * Sinh khóa:
        + Chọn số ngẫu nhiên x (0 < x < q)
        + Tính y = g^x mod p
        + Khóa công khai: (q, p, g, y)
        + Khóa riêng: x
      * Quá trình ký:
        + H là hàm băm, m là thông điệp gốc
        + Tính H(m) từ thông điệp gốc
        + Tạo số ngẫu nhiên k (0 < k < q)
        + Tính r = (g^k mod p) mod q
        + Nếu r = 0, chọn một k mới và tính lại r;
        + Tính s = k^-1(H(m) + xr) mod q
        + Nếu s = 0, chọn một k mới và tính lại r và s
        + Chữ ký là cặp (r, s).
      * Quá trình kiểm tra:
        + Kiểm tra điều kiện 0 < r, s < q
        + Loại bỏ chữ ký nếu r và s không thỏa mãn 0 < r , s < q
        + Tính H(m) từ thông điệp nhận được
        + Tính các giá trị w, u1, u2, v
        + Tính w = s^-1 mod q
        + Tính u1 = H(m) \* w mod q
        + Tính u2 = r \* w mod q
        + Tính v = ((g^u1 \* y^u2) mod p) mod q
        + Xác thực khi v = r

## So sánh RSA và DSA:

* + - Độ an toàn tương đương
    - Ưu điểm DSA:
      * Sinh khóa nhanh hơn
      * Ký nhanh hơn
    - Nhược điểm DSA:
      * Kiểm tra chậm hơn RSA
      * Chỉ dùng được cho chữ ký số
    - RSA phổ biến hơn do:
      * Đa năng (mã hóa và chữ ký số)
      * Quá trình kiểm tra nhanh hơn

## Công nghệ hỗ trợ

* + - Hàm băm phổ biến
      * SHA-256
      * SHA-3
      * RIPEMD-160
      * Whirlpool
    - Yêu cầu độ dài khóa
      * RSA: Tối thiểu 2048 bit
      * DSA: Tối thiểu 2048 bit cho p và 224 bit cho q

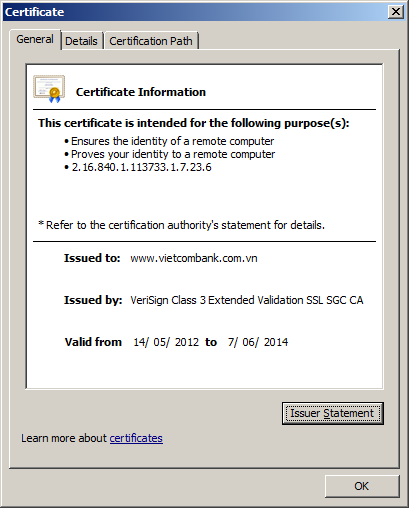
## Ứng dụng thực tế

* + - Giao dịch ngân hàng điện tử
    - Ký kết hợp đồng điện tử
    - Khai báo thuế điện tử
    - Chứng từ điện tử trong thương mại
    - Xác thực văn bản, tài liệu số

# Chứng chỉ số

## Giới thiệu

* + - Chứng chỉ số (Digital certificate), còn gọi là chứng chỉ khóa công khai (Public key certificate), hay chứng chỉ nhận dạng (Identity certificate) là một tài liệu điện tử sử dụng một chữ ký số để liên kết một khóa công khai và thông tin nhận dạng của một thực thể. Ba thành phần cơ bản nhất của một chứng chỉ số gồm:
      * Chữ ký số: là chữ ký của một bên thứ 3 tin cậy cung cấp chứng chỉ số, thường gọi là CA – Certificate Authority
      * Khóa công khai: là khóa công khai trong cặp khóa công khai và khóa riêng của thực thể
      * Thông tin nhận dạng: là tên, địa chỉ, tên miền hoặc các thông tin định danh của thực thể.
    - Chứng chỉ số có thể được sử dụng để xác minh chủ thể thực sự của một khóa công khai.

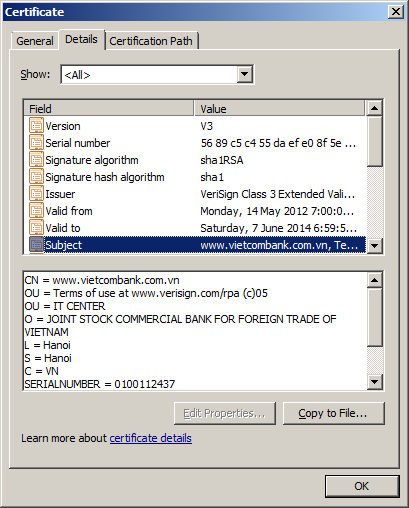


*Giao diện kiểm tra thông tin một chứng chỉ số do bên thứ 3 là một đơn vị của*

*công ty Verisign cấp cho tên miền www.vietcombank.com.vn của ngân hàng TMCP Ngoại thương Việt Nam*

## Nội dung

* + - Các trường thông tin cụ thể theo chuẩn chứng chỉ số X.509 gồm:
      * Serial Number: Số nhận dạng của chứng chỉ số
      * Subject: Thông tin nhận dạng một cá nhân hoặc một tổ chức
      * Signature Algorithm: Giải thuật tạo chữ ký
      * Signature Hash Algorithm: Giải thuật tạo chuỗi băm cho tạo chữ ký
      * Signature: Chữ ký của người/tổ chức cấp chứng chỉ
      * Issuer: Người/tổ chức có thẩm quyền/tin cậy cấp chứng chỉ
      * Valid-From: Ngày bắt đầu có hiệu lực của chứng chỉ
      * Valid-To: Ngày hết hạn sử dụng chứng chỉ
      * Key-Usage: Mục địch sử dụng khóa (chữ ký số, mã hóa,…)
      * Public Key: Khóa công khai của chủ thể
      * Thumbprint Algorithm: Giải thuật băm sử dụng để tạo chuỗi băm cho khóa công khai
      * Thumbprint: Chuỗi băm tạo từ khóa công khai
    - Các mục thông tin của trường Subject gồm:
      * CN (Common Name): Tên chung, nhưng một tên miền được gán chứng chỉ
      * OU (Organisation Unit): Tên bộ phận/phòng ban
      * O (Organisation): Tổ chức/Cơ quan/công ty
      * L (Location): Địa điểm/Quận huyện
      * S (State/Province): Bang/Tỉnh/Thành phố
      * C (Country): Đất nước.



*Nội dung chi tiết của một chứng chỉ số*

## Ứng dụng

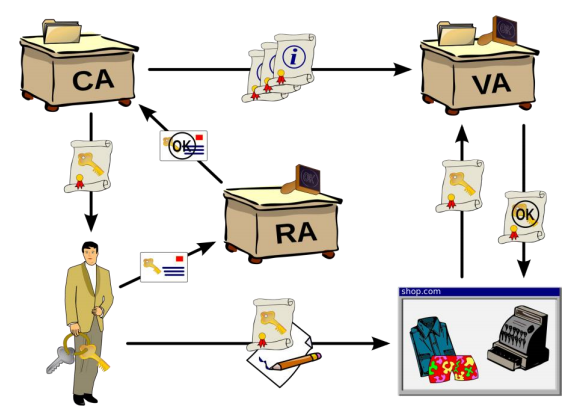
* + - Sử dụng chứng chỉ số trong đảm bảo an toàn giao dịch trên nền web: với chứng chỉ số, một website có thể được cấu hình để hoạt động theo chế độ “an toàn” (HTTPS), trong đó toàn bộ thông tin trao đổi giữa máy chủ và máy khách được đảm bảo tính bí mật (sử dụng mã hóa khóa đối xứng), tính toàn vẹn và xác thực (sử dụng hàm bămcó khóa MAC). Ngoài ra, các máy chủ và máy khách có thể xác thực thông tin nhận dạng của nhau sử dụng chứng chỉ số.
    - Chứng chỉ số cũng có thể được sử dụng để bảo mật thông tin truyền trong nhiều ứng dụng khác, như email, truyền file,…
    - Sử dụng chứng chỉ số có thể ngăn chặn hiệu quả dạng tấn công người đứng giữa do các bên tham gia giao dịch có thể xác thực thông tin nhận dạng của nhau. Nếu các bên sử dụng thêm chữ ký số thì có thể ngăn chặn việc sửa đổi các thông điệp trao đổi trên đường truyền.
    - Chứng chỉ số có thể được sử dụng trong trao đổi khóa.

# PKI

**a, Giới thiệu**

Hạ tầng khóa công khai (Public-key infrastructure - PKI) là một tập các phần cứng, phần mềm, nhân lực, chính sách và các thủ tục để tạo, quản lý, phân phối, sử dụng, lưu trữ và thu hồi các chứng chỉ số. Một PKI gồm các thành phần sau:

* Certificate Authority (CA): Cơ quan cấp và kiểm tra chứng chỉ số;
* Registration Authority (RA): Bộ phận tiếp nhận, kiểm tra thông tin nhận dạng của
* người dùng theo yêu cầu của CA;
* Validation Authority (VA): Cơ quan xác nhận thông tin nhận dạng của người dùng
* thay mặt CA;
* Central Directory (CD): Là nơi lưu danh mục và lập chỉ số các khóa;
* Certificate Management System: Hệ thống quản lý chứng chỉ;
* Certificate Policy: Chính sách về chứng chỉ.



*Lưu đồ cấp và sử dụng chứng chỉ số trong PKI*

* Đăng ký, xét duyệt và cấp chứng chỉ số:
  + Người dùng có yêu cầu cấp chứng chỉ số tạo một cặp khóa, gồm 1 khóa công khai và 1 khóa riêng
  + Người dùng tạo yêu cầu cấp chứng chỉ số (Certificate request), trong đó tích hợp khóa công khai và thông tin định danh của mình. Yêu cầu cấp chứng chỉ số thường được lưu dưới dạng 1 file văn bản theo định dạng của chuẩn X.509
  + Người dùng gửi yêu cầu cấp chứng chỉ số đến Bộ phận tiếp nhận (RA). RA kiểm tra các thông tin trong yêu cầu cấp chứng chỉ số, nếu hợp lệ thì chuyển yêu cầu đến Cơ quan cấp chứng chỉ (CA)
  + CA sẽ thực hiện việc xác minh các thông tin nhận dạng của chủ thể và nếu xác minh thành công thì cấp chứng chỉ số cho người yêu cầu. Chứng chỉ số được CA ký bằng khóa riêng của mình để đảm bảo tính xác thực và toàn vẹn và thường được lưu dưới dạng 1 file văn bản theo định dạng của chuẩn X.509
  + Sau khi phát hành chứng chỉ số cho người dùng, CA chuyển thông tin về chứng chỉ số đã cấp cho thành phần VA để xác nhận thông tin nhận dạng theo yêu cầu
  + Người dùng cài đặt chứng chỉ số vào hệ thống và có thể bắt đầu sử dụng trong các ứng dụng của mình.
* Sử dụng và kiểm tra chứng chỉ số:
  + Người dùng tạo đơn hàng, ký vào đơn hàng bằng khóa riêng, gửi đơn hàng đã ký và chứng chỉ số cho nhà cung cấp
  + Nhà cung cấp chuyển chứng chỉ số của người dùng cho VA để kiểm tra, nếu chứng chỉ số hợp lệ thì tiến hành xác thực chữ ký số của người dùng sử dụng khóa công khai của người dùng lấy từ chứng chỉ số. Nếu chữ ký của người dùng xác thực thành công thì đơn hàng được duyệt.

**b, Ứng dụng**

1. **Xác thực danh tính (Authentication)**
   * PKI giúp xác thực người dùng, thiết bị và hệ thống thông qua chứng chỉ số.
   * Được sử dụng trong đăng nhập hệ thống, truy cập mạng VPN, và các giao dịch trực tuyến.
2. **Mã hóa dữ liệu (Data Encryption)**
   * Đảm bảo rằng dữ liệu chỉ có thể được đọc bởi những người được ủy quyền.
   * Áp dụng trong bảo mật email, lưu trữ dữ liệu, và truyền thông an toàn.
3. **Chữ ký số (Digital Signature)**
   * Đảm bảo tính toàn vẹn của tài liệu và xác thực nguồn gốc.
   * Được dùng trong hợp đồng điện tử, tài liệu pháp lý, và giao dịch thương mại điện tử.
4. **Bảo mật giao dịch trực tuyến (Secure Transactions)**
   * Hỗ trợ các giao thức bảo mật như SSL/TLS trong giao dịch ngân hàng và mua sắm trực tuyến.
   * Ngăn chặn các cuộc tấn công giả mạo và đánh cắp thông tin.
5. **Quản lý quyền truy cập (Access Control)**
   * Kiểm soát quyền truy cập vào hệ thống, dữ liệu và ứng dụng.
   * Được sử dụng trong doanh nghiệp, chính phủ, và các hệ thống IoT.
6. **Bảo mật email (Secure Email Communication)**
   * Cung cấp mã hóa và xác thực email nhằm ngăn chặn giả mạo và rò rỉ dữ liệu.
   * Được ứng dụng trong các tổ chức chính phủ và doanh nghiệp.

=> Do đó PKI là nền tảng quan trọng trong bảo mật hệ thống thông tin, giúp đảm bảo tính xác thực, tính toàn vẹn và tính bảo mật của dữ liệu trong nhiều lĩnh vực.