**Tìm hiểu về Cryptography**

1. **Ba mục tiêu cơ bản của mật mã học**

**Mật mã học** (Cryptography) là ngành khoa học nghiên cứu về việc đảm bảo an toàn thông tin. Mật mã học gắn liền với quá trình mã hóa nghĩa là chuyển đổi thông tin từ dạng “có thể hiểu được” thành dạng “không thể hiểu được” và ngược lại là quá trình giải mã. Cryptography giúp đảm bảo những tính chất sau cho thông tin:

* **Tính bí mật** **(confidentiality):** thông tin chỉ được tiết lộ cho những ai được phép. Tính bí mật của thông tin có thể đạt được bằng cách giới hạn truy cập cả về mặt vật lý, ví dụ như tiếp cận trực tiếp tới thiết bị lưu trữ thông tin đó hoặc logic, ví dụ như truy cập thông tin đó từ xa qua môi trường mạng.
* **Tính toàn vẹn (integrity)**: đảm bảo tính toàn vẹn của thông tin, tức là thông tin chỉ được phép xóa hoặc sửa đổi bởi những đối tượng được phép và phải đảm bảo rằng thông tin vẫn còn chính xác khi được lưu trữ hay truyền đi. Về điểm này, nhiều người thường hay nghĩ tính “integrity” đơn giản chỉ là đảm bảo thông tin không bị thay đổi (modify) là chưa đầy đủ.
* **Tình khả dụng (availability):**  đảm bảo thông tin là sẵn sàng để đáp ứng sử dụng cho người có thẩm quyền.

1. **Nêu nguyên tắc Kirchoff, giải thích**

Luật Kirchoff, một giả thiết cơ bản của mật mã: *Toàn bộ cơ chế sinh mã và giải mã ngoại trừ thông tin về khóa là không bí mật với kẻ thù.* Rõ ràng khi đối phương không biết được mật mã đang sử dụng thuật toán mã hóa gì thì việc thám mã sẽ rất khó khăn. Nhưng chúng ta không thể tin vào độ an toàn của hệ mật mã chỉ dựa vào một giả thiết không chắc chắn là đối phương không biết thuận toán đang sử dụng. Vì vậy, khi trình bày một hệ mật bất kỳ, chúng ta đều giả thiết hệ mật đó được trình bày dưới luật Kirchoff.

**Ý nghĩa của luật Kirchoff:** sự an toàn của các hệ mật mã không phải dựa vào sự phức tạp của thuật toán mã hóa đang sử dụng.

1. **Secret key Cryptography**

Mã hóa đối xứng hay còn gọi là Secret Key Cryptography sử dụng cùng một khóa cho cả hai quá trình mã hóa và giải mã. Quá trình thực thiện như sau:

Diagram

Description automatically generated

Trong hệ thống mã hóa đối xứng, trước khi truyền dữ liệu, hai bên gửi và nhận phải thỏa thuận về khóa dùng chung cho quá trình mã hóa và giải mã, khóa được sinh ra bởi bộ sinh khóa. Sau đó, bên gửi sẽ mã hóa Plaintext bằng cách sử dụng khóa bí mật này và gửi thông điệp đã mã hóa cho bên nhận. Bên nhận sau khi nhận được thông điệp mã hóa sẽ sử dụng chính khóa bí mật mà hai bên thỏa thuận để giải mã và lấy ra Plaintext.

Những gì trong hình trên chính là quá trình tiền hành trao đổi thông tin giữa bên gửi và bên nhận thông qua việc sử dụng phương pháp mã hóa đối xứng. Trong quá trình này thì thành phần quan trọng nhất phải được giữ kín chính là khóa. Việc trao đổi, thỏa thuận về thuật toán được sử dụng trong việc mã hóa có thể tiến hành một cách công khai, nhưng bước thỏa thuận về khóa trong việc mã hóa và giải mã phải tiến hành bí mật. Phải có bước thỏa thuận về khóa, vì hai lý do cơ bản sau đây:

* Nêu như hai bên đã thống nhất thuật toán mã hóa, nhưng sau đó nơi gửi gửi thông điệp đã mã hóa tới cho nơi nhận mà không cho biết khóa đã sử dụng trong quá tình mã hóa. Nơi nhận không có khóa để giải mã cho nên sẽ không hiểu được nội dung trong tài liệu muốn nói gì. Vì thế bắt buộc ngoài việc trao đổi về thuật toán thì cần trao đổi về khóa.
* Khóa phải được trao đổi theo một kênh bí mật nào đó. Ví dụ trao đổi trực tiếp (mặt đối mặt) hay trao đổi gián tiếp (điện thoại, email, tin nhắn). Phải thực hiện qua kênh truyền bí mật vì rất có thể sẽ có bên thứ ba nghe lén cuộc trao đổi giữa hai bên và có được khóa, như vậy thông tin trao đổi sẽ bị kẻ khác biết được. Thêm vào đó hai bên buộc phải tin cậy lẫn nhau, không thể nhờ một người khác gửi hộ khóa được vì rất có thể người này sẽ dùng khóa đó để ăn cắp thông tin. Thuật toán mã hóa loại này có ưu điểm là nhanh, độ an toàn gần như tuyệt đối nếu không để lộ khóa bí mật. Chính vì vậy mấu chốt của việc sử dụng loại thuật toán mã hóa này là việc truyền khóa bí mật cho bên nhận một cách an toàn.
* Một số thuật toán mã hóa công khai phổ biến:
  + DES (Data Encryption Standard): Plaintext được mã hóa theo tứng khối 64bits và sử dụng một khóa 64 bits, nhưng thực tế thì chỉ có 56 bits là thực sự sử dụng để tạo khóa. Hiện tại DES không còn được đánh giá cao vì kích thước của khóa là quá nhỏ và nó dễ dàng bị phá vỡ chỉ trong 24 tiếng đồng hồ.
  + Triple DES (Advanced Encryption Standard)

1. **Block cipher là gì? Ứng dụng trong trường hợp nào, nêu ví dụ thuật toán**

* **Block Cipher** là khóa mật mã hay mật mã khối là một thuật toán mã hóa đối xứng dựa trên một khái niệm được gọi là mật mã sản phẩm lặp lại, trong đó nhiều thủ tục mã hóa được thực hiện trên các nhóm bit có độ dài cố định được gọi là khối. Mỗi vòng mã hóa làm cho việc giải mã trái phép trở nên khó khăn hơn, loại bỏ khả năng bị phá vỡ bởi bạo lực.

Phương pháp mật mã khối lần đầu được đề xuất bởi Claude Shannon vào năm 1949 trong một bài báo có ảnh hưởng có tên là Lý thuyết truyền thông của các hệ thống bí mật. Chúng là thành phần trung tâm của nhiều mật mã được sử dụng rộng rãi như DES và AES.

* Điều kiện để mật mã khối an toàn:
  + Độ dài khối (n-bit) phải đủ lớn để chống lại phương pháp tấn công thống kê, tuy nhiên điều này dẫn đến thời gian mã hóa sẽ tăng lên
  + Không gian khóa, tức là độ dài khóa phải đủ lớn để chống lại phương pháp tấn công vét cạn. Tuy nhiên khóa pải đủ ngắn để việc tạo khóa, phân phối khóa dễ dàng.
* Quá trình mã hóa gồm 2 thuật toán: mã hóa – ký hiệu E và giải mã – ký hiệu E-1. Cả hai thuật toán đều tác động lên một khối đầu vào n bit sử dụng một khóa k bit để cho ra một khối đầu ra n bit. Đối với bất kỳ khóa nào, giải mã là hàm ngược của mã hóa.
* **Ứng dụng của Block Cipher:** 
  + Block cipher thường được sử dụng trong việc mã hóa dữ liệu để lưu trữ chống lại cố gắng đọc dữ liệu của kẻ thù.
* Ví dụ thuật toán DES:
  + Thuật toán DES được thiết kế để mã hóa và giải mã các khối dữ liệu 64bit với một khóa cố độ dài 64 bit. Việc giải mã phải được thực hiện bởi một khóa tương ứng trong quá trình mã hóa, mỗi khóa gồm 64 bit, trong đó 56 bit được sử dụng dụng để mã hóa, giải mã còn 8 bit được dùng để phát hiện lỗi.
  + Quá trình mã hóa được mô tả như sau:

Diagram

Description automatically generated

Theo ta thấy ở trên, mã hóa DES phải thực hiện qua 16 vòng. Thông tin đầu vào là 64 bit, sẽ được chia thành 2 khối (block) trái (L) và phải (R). Sau đó, từ khóa (56 bit) người ta tạo ra các khóa con (subkey) 48 bit gọi là Ki. Hàm f ở trên thực chất là 1 hàm hoán vị.

Trong quá trình mã hóa, dữ liệu đầu vào phải thực hiện quá trình hoán vị đầu (initial permutation) và hoán vị cuối (final permutation) sau vòng thứ 16. Việc thực hiện hoán vị phục vụ cho quá trình đưa thông tin vào và lấy thống tin ra từ khối phần cứng, tạo điều kiện cho việc cài đặt phần cứng. Hàm cơ sở f cho phép đảm bảo tính bảo mật trong thuật toán DES này.

Việc thực hiện nhiều lần các bước lặp với tác dụng của hàm f là nhằm tăng cường tính phi tuyến và tính khuếch tán đã có trong hàm f.

Cấu trúc của vòng lặp DES thực hiện theo công thức sau:

(Li, Ri) = (Ri-1, Li-1 XOR f (Ri-1, Ki)

Trong đó (Li, Ri) là nửa trái và nửa phải lấy được của phép biến đổi vòng lặp thứ i.

1. **Stream Cipher là gì? Ứng dụng trong trường hợp nào, nêu ví dụ thuật toán**

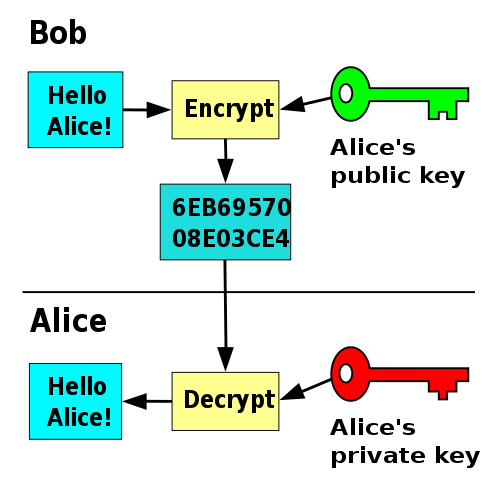
* **Stream Cipher** hay mật mã dòng là một loại mật mã đối xứng. Ý tưởng của thuật toán này là plaintext sẽ được phân rã thành các bit, từng bit này sẽ được mã hóa bằng mỗi bit khóa để cho kết quả là một bit mã.
* Stream Cipher đi kèm với nhiều lợi ích như tốc độ mã hóa nhanh, độ phức tạp thuật toán thấp và rất dễ sử dụng. Vì vậy nó phù hợp để mã hóa những thông tin có tuổi đời ngắn. Ví dụ như OTP.
* Ví dụ thuật toán RC4 (Rivest Cipher 4)
  + Trong mật mã học, RC4 (Rivest Cipher 4) là một mật mã dòng. Ưu điểm của RC4 chính là đơn giản, dễ thực hiện trong phần mềm, tốc độ mã hóa cao. Tuy nhiên RC4 ngày nay được coi là hệ mã có nhiều lỗ hổng và không an toàn. RC4 là stream cipher có độ dài key không ổn định (tùy người dùng) với đơn vị thao tác trên byte (8 bit). Dựa trên thuật toán hoán vị ngẫu nhiên với chu kỳ lặp lại của mã lên đến, mỗi byte của dữ liệu sẽ được thao tác từ 8 đến 16 toán tử để ra thành chuỗi mã hóa cho nên thời gian vận hành nhanh, dễ thực hiện trong phần mềm.
  + Để hiểu được thuật toán RC4 ta tìm hiểu các giai đoạn thực hiện của thuật toán RC4 gồm 3 giai đoạn chính là **giai đoạn khởi tạo, giai đoạn hoán vị Key-scheduling Algorithm (KSA)**, và **giai đoạn Pseudo-Random Generation Algorithm (PRGA).**
    - **Giai đoạn khởi tạo:** 
      * S(Inittial Vector – IV) mọi phần tử của S (256 bytes) được gán giá trị mặc định tăng dần từ 0 đến 255
      * VD: S[0] = 0; S[1] = 1; S[2] = 2; ….
      * Bên cạnh đó ở giai đoạn này T cũng được tạo ra từ khóa bí mật với cách thức tạo khóa như sau:
        + Nếu khóa bí mật có độ dài lớn là 256 bytes thì T là khóa bí mật
        + Nếu khóa bí mật có độ dài lớn nhỏ hơn 256 bytes thì khóa bí mật sẽ viết liền lặp lại cho đến khi đủ 256 bytes.
    - **Giai đoạn KSA** Vector S và vector T sau khi được khởi tạo được đưa vào hàm hoán vị KSA với hàm KSA được định nghĩa như sau:
      * Tiến hành lặp với mỗi giá trị I có giá trị từ 0 đến 255
      * Mỗi vòng lặp tiến hành 2 bước sau:
        + Bước 1: gán j = (j +S[i] + T[j]) mod 256
        + Bước 2: tiến hành hoán vị S[i] và S[j]
    - Sau khi vector S hoán vị xong sẽ tiếp tục đưa vào **giai đoạn hoán vị PRGA** để phát sinh thành dòng key với kích thước bất kì. Với hàm PRGA được định nghĩa như sau:
      * Tiến hành lặp cho đến khi dừng sinh key
      * Mỗi vòng lặp thực hiện các bước sau:
      * Gán i = (i+1) mod 256
      * Gán j = (j + S[i]) mod 256
      * Hoán vị giá trị S[i] và S[j]
      * Gán giá trị t = (S[i] + S[j]) mod 256
      * Xuất ra giá trị k = S[t]
      * Tiếp tục vòng lặp cho đến khi dừng sinh key
    - Mỗi khóa k được XOR với byte tiếp theo của plaintext để tạo thành ciphertext (để giải mật mã ta XOR byte tiếp theo của cipher text với khóa k)

1. **Message authentication code (MAC) là gì, tính chất của nó là gì, ứng dụng trong trường hợp nào, nêu ví dụ thuật toán**

* Mã xác thực thông báo (Message Authentication Code) là một đoạn mã được tạo ra nhằm mục đích xác thực tính toàn vẹn và nguồn gốc của thông báo. Đây là một cơ chế quan trọng để duy trì tính toàn vẹn và không thể chối bỏ của dữ liệu. với các giao thức trực tuyến, mã xác thực thông báo được sử dụng phổ biến và với việc kết hợp với các cơ chế khác (như chữ ký số) để đảm bảo tính xác thực giữa các bên tham gia giao dịch.
* Mục đích của hàm MAC là đảm bảo để hai (hay nhiều) bên tham gia giao dịch khi có chung khóa bí mật có thể giao dịch với nhau, kèm theo khả năng có thể phát hiện được thay đổi của thông báo trong quá trình vận chuyển, nhằm tránh các tần công làm thay đổi thông báo
* Các hàm mã xác thực thông báo được sử dụng rộng rãi trong các sản phẩm bảo mật và an toàn thông tin để đảm bảo tính xác thực và toàn vẹn của thông tin. Các hàm mã xác thực thông báo thông dụng hay được sử dụng là: HMAC-SHA-1, HMAC-SHA-224, HMAC-SHA-256,…

1. **Public key cryptography**

* Public key cryptography thường được gọi là mã hóa khóa bất đối xứng. Khác với mã đối xứng, mã hóa bất đối xứng sử dụng một cặp khóa: **khóa công khai (public key)** và **khóa bí mật (private key)**. Hai khóa này được xây dựng sao cho từ một khóa, rất khó có cách để sinh ra được khóa còn lại. Một khóa sẽ dành để mã hóa, khóa còn lại dùng để giải mã. Chỉ có người sở hữu nắm được khóa bí mật trong khi khóa công khai được phổ biến rộng rãi.
* Hình vẽ sau minh họa cho việc mã hóa và giải mã:



1. **Cryptographic hash function là gì, có tính chất gì, nêu ví dụ thuật toán**

* Cryptographic hash function hay hàm băm mật mã là một thuật toán có thể chạy trên dữ liệu như một file riêng lẽ hoặc mật khẩu để tạo ra một giá trị gọi là checksum.
* Công dụng chính của hàm băm mật mã là để xác minh tính xác thực của một phần dữ liệu. Hai file chỉ có thể được coi là giống hệt nhau nếu checksum được tạo ra từ mỗi file, sử dụng cùng hàm băm mật mã, giống hệt nhau
* Một số hàm băm mật mã thường được sử dụng bao gồm MD5 và SHA-1

1. **Digital Signature là gì? Có tính chất gì? ứng dụng trong trường hợp nào?**

* Digital Signature hay chữ ký số là một kỹ thuật xác thực cho phép người chủ nội dung của một thông điệp được quyền đính kèm một đoạn dữ liệu số như này chữ ký đánh dấu của người chủ với nội dung đã tạo ra. Về cơ bản, chữ kí số sẽ được tạo ra bằng cách hash nội dung thông điệp sau đấy mã hóa chuỗi hash bằng khóa bí mật của người chủ nội dung.
* Chữ ký số, digital signature, giúp ta đạt được những mục đích sau:
  + Non-repudiation: có thể hiểu là giúp cho người nhận khi kiểm tra nội dung đã được ký chữ ký số kèm theo sẽ biết chắc người ký không thể chối cãi về những gì đã tạo ra ở thời điểm bắt đầu tạo ra chữ ký số.
  + Integrity: chữ ký số giúp kiểm tra tính toàn vẹn dữ liệu của nội dung được gửi đi là không bị thay đổi hay chỉnh sửa kể từ lúc tạo ra chữ ký số và ký vào văn bản gửi đi
  + Authenticity: chữ ký số dùng để chứng thực nguồn gửi nội dung thông điệp. Thường thì thông tin về người chủ của chữ ký sẽ được thêm vào kèm với nội dung chữ ký số để giúp người nhận chứng thực được ai đã gửi thông điệp đi

1. **PKI là gì, bao gồm những thành phần nào, vai trò của từng thành phần**

* PKI (Public Key Infrastructure) là một khuôn khổ cho việc quản lý, phân phối, sử dụng, lưu trữ và thu hồi chứng chỉ số đáng tin cậy. PKI liên kết các khóa công khai với danh tính tương ứng của chúng. Hầu hết mọi người đều biết và PKI từ các chứng chỉ SSL, nhưng PKI đã bảo mật các trang web, mã hóa các tệp và xác thực và mã hóa email trước đó trong thời gian dài.
* PKI cho phép nhưng người tham gia xác thực lẫn nhau và sử dụng thông tin từ các chứng thực khóa công khai để mật mã hóa và giải mã thông tin trong quá trình trao đổi. Thông thường, PKI bao gồm phần mềm máy khách (client), phần mềm máy chủ (server), phần cứng (như thẻ thông minh smart-card) và các quy trình hoạt động liên quan. Người sử dụng cũng có thể ký các văn bản điện tử với private-key của mình và mọi người đều có thể kiểm tra với public key của người đó. PKI cho phép các giao dịch điện tử được diễn ra đảm bảo tính bí mật, toàn vẹn và xác thực lẫn nhau mà không cần phải trao đổi thông tin từ trước
* Để hiểu rõ hơn, PKI gồm những thành phần chính sau đây:
  + Thực thể cuối (End Entity – EE): Đối tượng sử dụng chứng nhận (chứng thư số): có thể là một tổ chức, một con người cụ thể hay một dịch vụ trên máy chủ,…
  + Tổ chức chứng nhận (Certificate Authority – CA): Có nhiệm vụ phát hành, quản lý và hủy bỏ các chứng thư số. Là thực thể quan trọng trong một PKI mà được thực thể cuối tín nhiệm. Gồm tập hợp các con người và các hệ thống máy tính có độ an toàn cao.
  + Chứng nhận khóa công khai (Public Key Certificate): Một chứng nhận khóa công khai thể hiện hay chứng nhận sự ràng buộc của danh tính và khóa công khai của thực thể cuối. Chứng nhận khóa công khai chứa đủ thông tin cho những thực thể khác có thể xác nhận hoặc kiểm tra danh tính của chủ nhân chứng nhận đó. Định dạng được sử dụng rộng rãi nhất của chứng nhận số dựa trên tiêu chuẩn IETFX.509
  + Tổ chức đăng kí chứng nhận (Registration Authority – RA): Mục đích chính của RA là để giảm tải công việc của CA. Xác thực cá nhân, chủ thể đăng ký chứng thư số. Kiểm tra tính hợp lệ của thông tin do chủ thể cung cấp. Xác nhận quyền của chủ thể đối với những thuộc tính của chứng thư số được yêu cầu. Kiểm tra xem chủ thể có thực sự sở hữu khóa riêng đang được đăng ký hay không (chứng minh sở hữu). Tạo cặp khóa bí mật, công khai (nếu chủ thể yêu cầu). Phân phối bí mật được chia sẻ đến thực thể cuối (ví dụ khóa công khai của CA). Thay mặt chủ thể thực thể cuối khởi tạo quá trình đăng ký với CA. Lưu trữ khóa riêng. Khởi sinh quá trình khôi phục khóa. Phân phối thẻ bài vật lý (thẻ thông minh).
  + Kho lưu trữ chứng nhận (Certificate Repository – CR): Hệ thống (có thể tập trung hoặc phân tán) lưu trữ chứng thư và danh sách các chứng thư bị thu hồi. Cung cấp cơ chế phân phối chứng thư và danh sách thu hồi chứng thư (CRLs – Certificate Revocatio Lists)
* Tất cả quá trình được mô tả như sau:

**Diagram

Description automatically generated**

(1): Người dùng gửi yêu cầu phát hành thẻ chứng thư số và khóa công khai của nó đến RA.

(2): Sau khi xác nhận tính hợp lệ định danh của người dùng thì RA sẽ chuyển yêu cầu này đến CA.

(3): CA phát hành thẻ chứng thư số cho người dùng.

(4): Sau đó người dùng “ký” thông điệp trao đổi với thẻ chứng thư số mới vừa nhận được từ CA và sử dụng chúng (thẻ chứng thực số + chữ ký số) trong giao dịch.

(5): Định danh của người dùng được kiểm tra bởi đối tác thông qua sự hỗ trợ của VA (Validation Authority – cơ quan xác thực bên thứ ba có thể cung cấp thông tin thực thể này thay mặt cho CA)

(6): Nếu chứng thư số của người dùng được xác nhận tính hợp lệ thì đối tác mới tin cậy người dùng và có thể bắt đầu quá tình trao đổi thông tin với nó (VA nhận thông tin về thẻ chứng thư số đã được phát hành từ CA)

1. **Digital certificate (chứng thư số) là gì? Nêu các trường thông tin trong chứng thư số**

* Digital Certificate hay chứng thư số 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.
* Chứng chỉ số bao gồm những trường chính sau:
  + Serial Number: Số nhận dạng của chứng chỉ số.
  + Subject: Thông tin nhận một cá nhận 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-Form: Ngày bắt đầu có hiệu lực chứng chỉ.
  + Valid-To: Ngày hết hạn sử dụng chứng chỉ.
  + Key-Usage: Mục đích sử dụng khóa (chữ ký số, mã hóa, …).
  + Public Key: Khóa công khai của chủ thể
  + Thumbprint Algorithm: Giải thuật hash sử dụng để tạo chuỗi băm cho khóa công khai.
  + Thumbprint: Chuỗi băm tạo từ khóa công khai.

1. **Nêu quy trình chứng thực một chứng thư số**