TRƯỜNG ĐẠI HỌC MỎ - ĐỊA CHẤT

ThS Đặng Văn Nam

HÀ NỘI – 2019

BÀI GIẢNG

AN TOÀN VÀ BẢO MẬT THÔNG TIN

(đang cập nhật…..)

2 | Tổng hợp và biên soạn: Đặng Văn Nam

MỤC LỤC

CHƯƠNG 1:TỔNG QUAN VỀ AN TOÀN VÀ BẢO MẬT THÔNG TIN

1.1 Giới thiệu chung về an toàn và bảo mật thông tin

1.1.1 Thông tin là gì?

Thông tin (Information) được sử dụng thường ngày. Con người có nhu cầu đọc báo, nghe đài, xem phim, học tập, nghiên cứu, đi tham quan, du lịch…. để nhận được thêm và bổ sung các thông tin mới. Thông tin mang lại cho chúng ta sự hiểu biết, nhận thức tốt hơn về những đối
tượng trong đời sống xã hội, trong thiên nhiên, … giúp cho họ thực hiện hợp lý công việc và
giao tiếp hàng ngày để đạt được hiệu quả và mục đích đặt ra một cách tốt nhất. Khi tiếp nhận được thông tin, con người thường phải được xử lý để tạo ra những

thông
tin mới, có ích hơn, từ đó có những phản ứng nhất định. Chẳng hạn như người tài xế chăm chú
quan sát người, xe cộ đi lạ trên đường, độ tốt xấu mặt đường, tính năng kỹ thuật cũng như vị trí
của xe để quyết định, cần tăng tốt độ hay hãm phanh, cần bẻ lái sang trái hay sang phải…nhằm
đảm bảo an toàn tối đa cho chuyến xe.
Thông tin có thể được phát sinh, được lưu trữ, được truyền, được tìm kiếm, được sao
chép, được xử lý, nhân bản. Thông tin cũng có thể biến dạng, sai lệch hoặc bị phá hủy.
Thông tin được thể hiện dưới nhiều dạng thức khác nhau như sóng ánh sáng, sóng âm,
điện từ, các ký hiệu viết trên giấy hoặc khắc trên gỗ, trên đá, trên các tấm kim loại… Về nguyên
tắc, bất kỳ cấu trúc vật chất nào hoặc bất kỳ dòng năng lượng nào cũng có thể mang thông tin.
Chúng được gọi là những vật mang tin. Dữ liệu (data) là biểu diễn của thông tin và được thể hiện
bằng các tín hiệu vật lý.
Thông tin chứa đựng ý nghĩa, còn dữ liệu là các dữ kiện không có cấu trúc và không có ý
nghĩa rõ ràng nếu nó không được tổ chức và xử lý. Cùng một thông tin, có thể được biểu diễn
bằng những ký hiệu vật lý khác nhau.
Thông tin là một khái niệm trừu tượng, tồn tại khách quan, có thể nhớ trong đối

tượng,
biến đổi trong đối tượng và áp dụng để điều khiển đối tượng. Thông tin làm tăng thêm hiểu biết
của con người, là nguồn gốc của nhận thức. Thông tin về một đối tượng chính là một dữ kiện về
đối tượng đó, chúng giúp ta nhận biết và hiểu được đối tượng.
Trong máy tính, các thông tin được biểu diễn bằng hệ đếm nhị phân. Tuy chỉ dùng 2 ký
số là 0 và 1 mà ta gọi là bit nhưng hệ nhị phân đã giúp máy tính biểu diễn – xử lý được thông tin
trên hầu hết các loại thông tin mà con người hiện đang sử dụng như văn bản, hình ảnh âm thanh,
video…

1.1.2 Tại sao phải đảm bảo an toàn và bảo mật thông tin

Ngày xưa, thông tin thường được lưu trữ và vận chuyển trên giấy tờ. Ngày nay, chúng
được lưu trữ dưới dạng số hóa và được vận chuyển bằng các hệ thống viễn thông hoặc các hệ
thống không dây. Cùng với sự xuất hiện của Internet và các mạng cục bộ đã giúp cho việc trao
đổi thông tin trở nên nhanh chóng, dễ dàng…Tuy nhiên, điều này cũng làm phát sinh ra nhiều
vấn đề mới. Những tin tức quan trọng nằm ở kho dữ liệu hay trên đường truyền có thể bị trộm

cắp, có thể bị làm sai lệch hoặc có thể bị giả mạo. Những bí mật kinh doanh, tài chính … là mục

tiêu của các đối thủ cạnh tranh. Những tin tức về an ninh quốc gia là mục tiêu của các tổ chức
tình báo trong và ngoài nước…

Câu hỏi đặt ra: ***Thông tin có quan trọng không? Và quan trọng đến mức nào?***

Phải khẳng định rằng, thông tin là một tài sản quan trọng, thông tin được xem là sự sống
còn đối với các doanh nghiệp, đối với an ninh quốc gia và các hoạt động quản lý khác…

- Với cá nhân: Các thông tin nhạy cảm, bí mật về cuộc sống đời tư, tài khoản cá nhân…nếu
các thông tin này bị lộ sẽ ảnh hưởng nghiêm trọng đến cuộc sống hàng ngày, nhân phẩm,
tài sản của mỗi cá nhân. Hậu quả để lại sẽ rất nghiêm trọng.
- Với các công ty, doanh nghiệp: Các thông tin về chiến lược kinh doanh, sản phẩm mới,
thông tin về khách hàng, về tài chính, nguồn vốn đầu tư…là một trong những thông tin
rất quan trọng ảnh hưởng trực tiếp tới sự tồn tại và phát triển của doanh nghiệp. Đây là
các thông tin mà các đối thủ cạnh tranh hoặc những kẻ phá hoại luôn tìm cách có được.

Nếu những thông tin này bị trộm cắp, giả mạo…thì hậu quả mà các công ty, doanh nghiệp gặp phải là không thể lường trước.

- Với quốc gia: Các tin tức quân sự, sơ đồ bố trí an ninh – quốc phòng, bí mật phòng thủ và
các kênh liên lạc,….Nếu các thông tin này bị các quốc gia thù địch, các tổ chức

phản
động…có được thì hậu quả có thể cự kỳ quan trọng, có thể ảnh hưởng tới sự tồn tại của
cả một chế độ, một quốc gia.
Với các thông tin được lưu trữ dưới dạng số hóa người ta có thể sao chép và thay đổi
chúng một cách nhanh chóng và dễ dàng. Người ta có thể tạo ra hàng ngàn mẩu tin giống nhau
và không thể phân biệt được nó với bản gốc.
Cùng với sự phát triển chung của thế giới thì các phương thức và thủ đoạn tấn công vào
hệ thống thông tin ngày càng tinh vi và phức tạp. Thông tin đang trở thành mục tiêu tấn công
hàng đầu. Theo số liệu của CERT (Computer Emegency Response Team) số lượng các vụ tấn
công trên Internet mỗi ngày một nhiều, qui mô của chúng ngày một lớn và phương pháp tấn công
ngày càng hoàn thiện.
Chính vì vậy, đảm bảo an toàn thông tin đã, đang và tiếp tục là một vấn đề cấp thiết trên
toàn thế giới và đặc biệt là tại Việt Nam.

Câu hỏi đặt ra: ***Đảm bảo an toàn thông tin như thế nào?***

Thông tin cần được bảo vệ ở mọi nơi, mọi lúc. Thông tin cần phải được đảm bảo an toàn khi lưu
trữ bên trong máy tính và khi truyền đi trên môi trường mạng.

- An toàn máy tính (Computer Security): là sự bảo vệ các thông tin cố định bên trong máy

tính, là khoa học về bảo đảm an toàn thông tin trong máy tính.

• An toàn truyền tin (Communication Security): Là sự bảo vệ thông tin trên đường truyền
tin (Thông tin đang được truyền từ hệ thống này sang hệ thống khác). Là khoa học về bảo
đảm thông tin trên đường truyền tin.

Ngày nay cả thế giới được kết nối với nhau thông qua hệ thống mạng máy tính (Internet)
mọi hoạt động, giao dịch - trao đổi…đều được thực hiện ở đây. Tuy nhiên, môi trường khó bảo
vệ an toàn thông tin nhất và cũng chính là môi trường dễ xâm nhập nhất đó là **môi trường mạng
và truyền tin**. Do đó, việc bảo vệ thông tin trong quá trình truyền thông tin trên mạng luôn được
đặt lên hàng đầu:

Câu hỏi đặt ra: ***Khi thông tin truyền trên mạng có thể bị tấn công bằng những hình thức
nào?***

ϖ Xem trộm thông tin: Khi Alice gửi thông tin cho Bob qua mạng, Trudy chặn các thông
điệp này, và xem được nội dung của thông tin mà Alice gửi cho Bob.

Hình 1.1: Hình thức tấn công xem trộng thông tin

ϖ Thay đổi thông tin: Trong quá trình Alice gửi thông tin tới cho Bob, Trudy chặn các
thông điệp này lại và ngăn không cho các thông điệp này đến đích. Sau đó Trudy

thay đổi
các nội dung của thông điệp và gửi tiếp cho Bob. Bob nghĩ rằng nhận được thông điệp
nguyên bản ban đầu của Alice mà không biết rằng chúng đã bị sửa đổi.

Hình 1.2: Hình thức tấn công thay đổi thông tin

ϖ Mạo danh: Trong trường hợp này Trudy giả là Alice gửi thông tin cho Bob. Bob không
biết điều này và nghĩ rằng thông điệp là của Alice

Hình 1.3: Hình thức tấn công mạo danh

ϖ Phát lại thông tin: Trudy sao chép lại thông điệp mà Alice gửi cho Bob. Sau đó một thời
gian Trudy gửi bản sao chép này cho Bob. Bob tin rằng thông điệp thứ hai vẫn là từ
Alice, nội dung hai thông điệp này là giống nhau. Thoạt đầu có thể nghĩ rằng việc phát lại
này là vô hại, tuy nhiên trong nhiều trường hợp cũng gây ra tác hại không kém so với
việc giả mạo thông điệp. Xét tình huống: Bob là bên ngân hàng, còn Alice là khách hàng.
Alice gửi thông điệp đề nghị Bob chuyển cho Trudy 1000$. Alice có áp dụng các biện
pháp như chữ ký điện tử với mục đích không cho Trudy mạo danh cũng như sửa thông
điệp. Tuy nhiên nếu Trudy sao chép và phát lại thông điệp thì các biện pháp bảo vệ này
không có ý nghĩa. Bob tin rằng Alice gửi tiếp một thông điệp mới để chuyển thêm

cho

Trudy 1000$ nữa.

Hình 1.4: Hình thức tấn công phát lại thông tin

1.1.3 Khái niệm về an toàn thông tin và đảm bảo an toàn thông tin

a. Khái niệm về an toàn thông tin

Thông tin được lưu trữ bởi các sản phẩm và hệ thống CNTT là một tài nguyên quan trọng
cho sự thành công của tổ chức, doanh nghiệp và cá nhân. Các thông tin lưu trữ trong hệ thống
thông tin cần được giữ bí mật, bảo vệ và không bị thay đổi khi không được phép. Trong khi các
sản phẩm và hệ thống thông CNTT thực hiện các chức năng của chúng, các thông tin cần được
kiểm soát để đảm bảo chúng được bảo vệ chống lại các nguy cơ, ví dụ như việc phổ biến và thay
đổi thông tin không mong muốn và trái phép, nguy cơ mất mát thông tin.

*An toàn thông tin bao gồm các hoạt động quản lý, nghiệp vụ và kỹ thuật đối với hệ thống
thông tin nhàm bảo vệ, khôi phục các hệ thống, các dịch vụ và nội dung thông tin đối với nguy
cơ tự nhiên hoặc do con người gây ra. Việc bảo vệ thông tin, tài sản và con người trong hệ thống
thông tin nhằm bảo đảm cho cá hệ thống thực hiện đúng chức năng, phục vụ đúng đối tượng một
cách sẵn sàng, chính xác và tin cậy. An toàn thông tin bao hàm cá nội dung bảo vệ và bảo mật*

thông tin, an toàn dữ liệu, an toàn máy tính và an toàn mạng. (theo Nghị định 64/2007/NĐ-CP)

b. Khái niệm về đảm bảo an toàn thông tin

Hệ thống thông tin thường tồn tại những điểm yếu dẫn đến những rủi ro có thể xảy ra,
làm tổn hại đến giá trị tài sản thông tin. Các đối tượng tấn công có chủ tâm đánh cắp, lợi dụng
hoặc phá hoại tài sản cảu cá chủ sở hữu, tìm cách khai thác các điểm yêu để tấn công, tạo ra các
nguy cơ và các rủi ro cho các hệ thống.

Với các biện pháp an toàn thông tin người dùng có được công cụ trong tay để nhận thức
được các điểm yếu, giảm thiểu các điểm yếu, ngăn chặn các nguy cơ tấn công, làm giảm các yêu
tố rủi ro. Như vậy, các biện pháp và kỹ thuật đảm bảo an toàn thông tin chính là mang lại sự tin
cậy cho các sản phẩm và hệ thống.

Đảm bảo an toàn thông tin là đảm bảo an toàn kỹ thuật cho hoạt động của các cơ sở hạ
tầng thông tin, trong đó bao gồm đảm bảo an toàn phần cứng và phần mềm hoạt động theo các
tiêu chuẩn kỹ thuật do nhà nước ban hành; ngăn ngừa khả năng lợi dụng mạng và các cơ sở hạ
tầng thông tin để thực hiện các hành vi trái phép gây hại cho cộng đồng, phạm pháp hay khủng
bố; đảm bảo các tính chất bí mật, toàn vẹn, chính xác, sẵn sàng phục vụ của thông tin trong lưu

trữ, xử lý và truyền tải trên mạng.

Như vậy, khái niệm đảm bảo an toàn thông tin bao hàm đảm bảo an toàn cho cả phần
cứng và phần mềm. An toàn phần cứng là đảm bảo hoạt động cho cơ sở hạ tầng thông tin. An
toàn phần mềm gồm các hoạt động quản lý, kỹ thuật nhằm bảo vệ hệ thống thông tin, đảm bảo

cho các hệ thống thực hiện đúng chức năng, phục vụ đúng đối tượng một cách sẵn sàn, chính
xác, tin cậy.

1.1.4 Mục tiêu của an toàn và bảo mật thông tin

An toàn và bảo mật thông tin nhằm đạt được 4 mục tiêu quan trọng:

- **Bảm đảm tính bí mật** (bảo mật): Đảm bảo thông tin không bị lộ đối với những người
dùng không được phép. Tính bí mật là tâm điểm điểm chính của mọi giải pháp an toàn
cho một sản phẩm/hệ thống CNTT. Một giải pháp an toàn là tập hợp các quy tắc xác định
quyền được truy cập đến với thông tin đang tìm kiếm, đối với một số người sử dụng
thông tin nhất định và một số lượng thông tin là tài sản nhất định. Trong trường hợp kiểm
soát truy cập cục bộ, nhóm người truy cập sẽ được kiểm soát xem họ đã truy cập những
số liệu nào. Tính bí mật là sự đảm bảo rằng các chức năng kiểm soát truy cập có hiệu lực.

Đảm bảo tính bí mật là nhằm loại bỏ những truy cập không được phép vào các khu vực
độc quyền của các các nhân, tổ chức.

• **Bảo đảm tính toàn vẹn** (bảo toàn): Đặc tính toàn vẹn được hiểu là chất lượng của thông
tin được xác định căn cứ vào độ xác thực khi phản ánh thực tế. Số liệu càng gần với thực
tế bao nhiêu thì chất lượng thông tin càng chuẩn bấy nhiêu. Để đảm bảo tính toàn vẹn của
thông tin phải sử dụng một loạt các biện pháp đồng bộ nhằm hỗ trợ và đảm bảo tính thời
sự kịp thời và sự đầy đủ trọn vẹn, cũng như sự bảo mật hợp lý cho thông tin. Bảo đảm
tính toàn vẹn chính là đảm bảo sự chính xác, hoàn chỉnh của thông tin; thông tin chỉ được
bổ sung, loại bỏ và sửa chữa bởi những người được cấp quyền.

ϖ **Bảo đảm tính xác thực** (chứng thực): Đảm bảo tính chính xác của người nhận và người
gửi thông tin.

• **Bảo đảm tính sẵn sàng:** Tính sẵn sàng của thông tin cũng rất quan trọng. Đó là khía
cạnh sống còn của an toàn thông tin, đảm bảo cho thông tin đến đúng địa chỉ khi có nhu
cầu, hoặc được yêu cầu. Tính sẵn sàng đảm bảo độ ổn định đáng tin cậy của thông tin,
cũng như đảm nhiệm chức năng là thước đo, xác định phạm vi tới hạn của an toàn

một hệ
thống thông tin. Đảm bảo tính sẵn sàng chính là đảm bảo thông tin có thể được truy xuất
bởi những người được cấp quyền bất kỳ khi nào họ cần.

1.1.5 Một số giải pháp đảm bảo an toàn và bảo mật thông tin

Có rất nhiều phương pháp được thực hiện để đảm bảo an toàn và bảo mật thông tin. Các
phương pháp này có thể được quy tụ vào ba nhóm sau:

- Bảo đảm an toàn thông tin bằng các biện pháp hành chính.
- Bảo đảm an toàn thông tin bằng các biện pháp kỹ thuật (phần cứng).
- Bảo đảm an toàn thông tin bằng các biện pháp thuật toán (phần mềm).

Ba nhóm trên có thể được ứng dụng riêng rẽ hoặc phối kết hợp.

Biện pháp hiệu quả nhất và kinh tế nhất hiện nay để đảm bảo an toàn thông tin trên mạng truyền
tin và mạng máy tính là các biện pháp thuật toán.

a. Phương pháp che – giấu, bảo đảm toàn vẹn và xác thực thông tin

- "Che" dữ liệu: Thay đổi hình dạng dữ liệu gốc để những người không được phép khó nhận ra.
- "Giấu" dữ liệu: Cất giấu dữ liệu này trong môi trường dữ liệu khác để những người không được
phép khó nhận ra.

+ Kỹ thuật sử dụng để thực hiện phương pháp này bao gồm: Mã hóa, hàm băm, ẩn - giấu tin, chữ
ký số, các giao thức bảo mật – xác thực thông tin.

b. Phương pháp kiểm soát lối vào ra của thông tin

- Kiểm soát, ngăn chặn thông tin vào ra hệ thống máy tính.

- kiểm soát, cấp quyền sử dụng các thông tin trong hệ thống máy tính.
- Kiểm soát, tìm diệt các loại virus, sâu, bọ máy tính …
+ Kỹ thuật sử dụng để thực hiện phương pháp này bao gồm: Thiết lập mật khẩu, tường lửa, mạng
riêng ảo, nhận dạng – xác thực thực thể, cấp quyền hạn truy cập.
c. Phát hiện và xử lý các lỗ hổng trong an toàn thông tin.
- Phát hiện và xử lý các lỗ hổng trong các thuật toán hay giao thức mật mã, giấu tin.
- Phát hiện và xử lý các lỗ hổng trong các giao dịch mạng.
- Phát hiện và xử lý các lỗ hổng trong các hệ điều hành.
- Phát hiện và xử lý các lỗ hổng trong các ứng dụng
Để đảm bảo an toàn thông tin hiệu quả điều trước tiên là phải lường trước hoặc dự đoán
trước các khả năng không an toàn, khả năng xâm phạm, các sự cố rủi ro có thể xảy ra đối với
thông tin dữ liệu được lưu trữ và trên môi trường truyền tin cũng như trên mạng. Tuy nhiên, ***Một thực tế là không có một biện pháp bảo đảm an toàn thông tin nào là an toàn tuyệt đối. Một hệ thống dù được bảo vệ chắc chắn đến đâu cũng không thể đảm bảo là tuyệt đối an toàn.*** Chúng
ta thường phải kết hợp nhiều biện pháp từ phần cứng đến phần mềm để làm sao thông tin có độ
an toàn cao nhất. An toàn thông tin là một lĩnh vực rộng lớn với nhiều nội dung như an toàn dữ
liệu, an toàn cơ sở dữ liệu, an toàn hệ điều hành, an toàn mạng máy tính…

1.2 Mật mã học

1.2.1 Một số khái niệm liên quan

Mật mã học (Cryptology) bao gồm hai lĩnh vực: Lập mã và thám mã

♣ Lập mã: Nghiên cứu các thuật toán và phương thức để đảm bảo tính bí mật và xác thực
của thông tin. Các sản phẩm của lĩnh vực này là các hệ mật mã, các hàm băm, các hệ chữ
ký điện tử, các cơ chế phân phối, quản lý khóa và các giao thức mật mã. Lập mã bao gồm
mã hóa và giải mã.

♣ Thám mã: Nghiên cứu các phương pháp phá mã hoặc tạo mã giả. Sản phẩm của lĩnh vực
này là các phương pháp thám mã, các phương pháp giả mạo chữ ký, các phương pháp tấn
công các hàm băm và các giao thức mật mã.

Mục tiêu của thám mã (phá mã) là tìm những điểm yếu hoặc không an toàn trong phương
thức mật mã hóa. Thám mã có thể được thực hiện bởi những kẻ tấn công ác ý, nhằm làm
hỏng hệ thống; hoặc bởi những người thiết kế ra hệ thống (hoặc những người khác) với ý
định đánh giá độ an toàn của hệ thống.

♣ Bản rõ (P-Plaintext) là những từ, những ký tự của văn bản gốc có thể hiểu được ý nghĩa
và nội dung của văn bản.

♣ Bản mã (C-Ciphertext) là những từ, những ký tự ở dạng bí mật khó có thể hiểu

được nội
dung của văn bản.
Thông thường không gian các bản rõ và không gian các bản mã là các văn bản được tạo
thành từ một bộ chữ cái A nào đó. Đó có thể là bộ chữ cái tiếng Anh, bộ mã ASCII, bộ
mã UNICODE hoặc đơn giản nhất là các bit 0 và 1.

♣ Mã hóa (Cryptography) là quá trình chuyển từ thông tin có thể đọc được (bản rõ) thành
thông tin khó có thể đọc được theo cách thông thường (bản mã).

♣ Giải mã là quá trình chuyển thông tin ngược lại từ bản mã thành bản rõ.

♣ Thuật toán mã hóa hay giải mã là thủ tục tính toán để thực hiện mã hóa hay giải mã.

♣ Khóa mã hóa là một giá trị làm cho thuật toán mã hóa thực hiện theo cách riêng biệt sinh
ra bản mã riêng. Thông thường khóa càng lớn thì bản mã càng an toàn. Phạm vi giá trị có
thể có của khóa được gọi là không gian khóa.

♣ Hệ mã hóa là một tập các thuật toán, các khóa nhằm che giấu thông tin cũng như làm rõ
nó.

Mô hình cơ bản của truyền tin bảo mật:

1.2.2 Sơ lược về lịch sử mật mã học

Mật mã học là một ngành có lịch sử từ hàng nghìn năm nay. Trong phần lớn thời gian
phát triển của mình (ngoại trừ vài thập kỷ trở lại đây), lịch sử mật mã học chính là

lịch sử của
những phương pháp mật mã học cổ điển - các phương pháp mật mã hóa với bút và giấy, đôi khi
có hỗ trợ từ những dụng cụ cơ khí đơn giản. Vào đầu thế kỷ 20, sự xuất hiện của các cơ cấu cơ
khí và điện cơ, chẳng hạn như máy Enigma, đã cung cấp những cơ chế phức tạp và hiệu quả hơn
cho việc mật mã hóa. Sự ra đời và phát triển mạnh mẽ của ngành điện tử và máy tính trong
những thập kỷ gần đây đã tạo điều kiện để mật mã học phát triển nhảy vọt lên một tầm cao mới.
Sự phát triển của mật mã học luôn luôn đi kèm với sự phát triển của các kỹ thuật phá
mã (hay thám mã). Các phát hiện và ứng dụng của các kỹ thuật phá mã trong một số trường hợp
đã có ảnh hưởng đáng kể đến các sự kiện lịch sử. Một vài sự kiện đáng ghi nhớ bao gồm việc
phát hiện ra bức điện Zimmermann khiến Hoa Kỳ tham gia Thế chiến II và việc phá mã thành
công hệ thống mật mã của Đức Quốc xã góp phần làm đẩy nhanh thời điểm kết thúc thế chiến II.

Cho tới đầu thập kỷ 1970, các kỹ thuật liên quan tới mật mã học hầu như chỉ nằm trong
tay các chính phủ. Hai sự kiện đã khiến cho mật mã học trở nên thích hợp cho mọi người, đó là:
sự xuất hiện của tiêu chuẩn mật mã hóa DES và sự ra đời của các kỹ thuật mật mã

hóa khóa công
khai.

1.2.3 Phân loại các thuật toán và một số ứng dụng của mật mã học

Có nhiều cách khác nhau để chúng ta có thể phân loại các thuật toán mật mã học.

ϖ Phân loại các thuật toán mật mã học dựa vào:

- Dựa vào các dịch vụ an toàn bảo mật mà các thuật toán cung cấp.
- Dựa vào số lượng khóa sử dụng

a) Các thuật toán mã hóa khóa bí mật tương ứng với các hệ mã hóa khóa bí mật (hay khóa
đối xứng), do vai trò của người nhận và người gửi là như nhau, cả hai đề có thể mã hóa
hay giải mã thông điệp. Khóa sử dụng cho các thuật toán này là 1 khóa cho cả việc mã
hóa và giải mã.

b) Các thuật toán mã hóa khóa công khai tương ứng với các hệ mã hóa khóa công khai (hay
hệ mã hóa khóa bất đối xứng). Khóa sử dụng cho các thuật toán này là 2 khóa, một cho
việc mã hóa và một cho việc giải mã, khóa mã hóa được công khai hóa.

c) Các thuật toán tạo chữ ký điện tử. Các thuật toán tạo chữ ký điện tử tạo thành các hệ chữ
ký điện tử. Thông thường mỗi hệ chữ ký điện tử có cùng cơ sở lý thuyết với một hệ mã
hóa khóa công khai nhưng cách áp dụng khác nhau.

d) Các hàm băm. Các hàm băm là các thuật toán mã hóa không khóa hoặc có khóa thường

được sử dụng trong các hệ chữ ký.

ϖ Phân loại các thuật toán mã hóa dựa trên cách thức xử lý đầu vào của thuật toán.

a) Các thuật toán mã hóa khối (DES, AES…) xử lý bản rõ dưới đơn vị cơ bản là các khối có
kích thước giống nhau.

b) Các thuật toán mã hóa dòng (RC4…) coi bản rõ là một luồng bít, byte liên tục.

1.3 Cơ sở toán học

Lý thuyết mật mã là một ngành khoa học được xây dựng dựa trên cơ sở toán học, đặc biệt là lý
thuyết số.

1.3.1 Số học trên Modulo

ϖ Cho số nguyên a, n (n>0): ta định nghĩa a mod n là phần dư dương khi chia a cho n.

ϖ Quan hệ đồng dư: a đồng dư với b theo modulo n nếu chia a và b cho n, ta nhận được
cùng một số dư.

Ký hiệu: a≡b mod n (a≡b(modn))

Vd: 22≡14 mod 4

Một số tính chất với số học trên modulo:

(a+b) mod n = ((a mod n) + (b mod n)) mod n

(a-b) mod n = ((a mod n) - (b mod n)) mod n

(a.b) mod n = ((a mod n) * (b mod n)) mod n

(a*(b+c)) mod n =(((a*b) mod n) + ((a*c) mod n)) mod n

¬ Tất cả các phép tính trong các hệ mật mã đều thực hiện với một số modulo N nào

đó.

1.3.2 Ước số chung, bội số chung

Cho 2 số nguyên a và b (b ≠ 0) nếu có một số nguyên q sao cho a = b*q thì ta nói rằng a
chia hết cho b, ký hiệu b\a. Khi đó b là ước của a và a là bội của b.

Vd: a=18, b =6, ta có 18 = 6 * 3, khi đó 6\18. Ở đây 6 là ước của 18 và 18 là bội của 6.

Số nguyên **d** được gọi là ***ước chung*** của các số nguyên a1, a2, …, an , nếu nó là ***ước*** của
tất cả các số đó.

Ví dụ : a1 = 6 ; a2 = 12 ; a3 = 30 ◊ ước chung của (a1, a2, a3) là d1 = 1 ; d2 = 3 ; d3 = 6 ;

Số nguyên **m** được gọi là ***bội chung*** của các số nguyên a1, a2, …, an, nếu nó là ***bội*** của tất
cả các số đó.

Ví dụ : a1 = 2 ; a2 = 5 ; a3 = 6 ◊ Bội chung của (a1, a2, a3) là m1 = 30 ; d2 = 60 ; …

Một ước chung **d** >0 của các số nguyên a1, a2, …, an, trong đó mọi ước chung của a1, a2,
…, an đều là ước của **d**, thì **d** được gọi là ***ước chung lớn nhất*** (UCLN) của a1, a2, …, an. Ký
hiệu **d** = **gcd** (a1, a2, …, an) hay **d** = **UCLN** (a1, a2, …, an).

Một bội chung **m** >0 của các số nguyên a1, a2, …, an, trong đó mọi bội chung của a1, a2,
…, an đều là bội của **m**, thì **m** được goi là ***bội chung nhỏ nhất*** (BCNN) của a1, a2, …, an. Ký
hiệu **m** = **lcm** (a1, a2, …, an) hay **m** = **BCNN** (a1, a2, …, an).

Số nguyên tố cùng nhau :

Hai số a và b được gọi là 2 số nguyên tố cùng nhau nếu ước số chung lớn nhất của a và b

bằng 1, gcd (a, b) =1

Vd : a = 15, b = 8. Hai số a và b là nguyên tố cùng nhau, vì gcd (15, 8) = 1

Ký hiệu :

Zn = {0, 1, 2, ..., n-1} là tập các số nguyên không âm < n.

* *n*

Z **=** {e # 0, e ∈ **Zn**, e là nguyên tố cùng nhau với n}. Gọi là tập thặng dư thu gọn theo mod n,

có số phần tử là Φ(n)

Ví dụ:

Z 6 = {0, 1, 2, 3, 4, 5}. Khi đó số phần tử của **Z 6** là |**Z 6** |= 6.

Z 6 * = {1, 5}. Khi đó số phần tử của **Z 6** * là | **Z 6** * |= 2.

Z 7 = {0, 1, 2, 3, 4, 5, 6}. Khi đó số phần tử của **Z 7** là |**Z 7** |= 7.

Z 7 * = {1, 2, 3, 4, 5, 6}. Khi đó số phần tử của **Z 7** * là | **Z 7** * |= 6.

Z 8 = {0, 1, 2, 3, 4, 5, 6,7}. Khi đó số phần tử của **Z 8** là |**Z 8** |= 8.

Z 8 * = {1, 3, 5, 7}. Khi đó số phần tử của **Z 8** * là |**Z 8** * |= 4.

Z 26 = {0, 1, 2, 3,...,24, 25}. Khi đó số phần tử của **Z 26** là |**Z 26** |= 26.

Z 26 * = {1, 3, 5, 7, 9 ,11, 15, 17, 19, 21, 23, 25}. Khi đó số phần tử của **Z 26** * là |**Z 26** * |= 12.

1.3.3 Phần tử nghịch đảo

Cho a ∈ Zn, nếu tồn tại b ∈ Zn sao cho a*b ≡ 1 mod n ta nói b là phần tử nghịch đảo của

a trong Zn, ký hiệu là a-1, Một phần tử có phần tử nghịch đảo gọi là khả nghịch.

(Nói cách khác, phần tử b là phần tử nghịch đảo của a khi a*b chia cho n dư 1)

VD. Trong Z26, a = 7, khi đó b = 15 là phần tử nghịch đảo của a trong Z26. Vì

7*15 (mod 26) = 1

(a=19 \ b=11)

Trong Z36, a =17, khi đó b = 17

ϖ **Định lý về sự tồn tại của phần tử nghịch đảo:** Nếu UCLN (a, N) = 1 thì tồn tại duy nhất

một số b ∈ Zn là phần tử nhịch đảo của a, nghĩa là thỏa mãn a*b (mod N) = 1

ϖ Thuật toán Ơclit mở rộng xác định phần tử nghịch đảo.

Input: a, N với UCLN (a, N) = 1

Output: a-1

BEGIN

x0 = N, x1=a, b0=0, b1=1, i=1;

while (xi > 1)

{

y=xi-1 div xi;

xi+1=xi-1 – y*xi;

bi+1 = bi-1 – y*bi;

i=i+1;

}

x = bi-1;

if (x>0) then a-1 = x

else a-1 = (N+x);

END

Tính: **a-1 mod N**

Điều kiện: a và N phải là hai số nguyên tố cùng nhau (UCLN(a, N) = 1)

Nếu a=1 thì kết quả = 1 (không cần tính)

Tổng quát:

Giá trị điền sẵn:

x0 = N; b0 = 0

x1 = a; b1 = 1

x	**b**	**y**	
0	Điền giá trị N	0	Rỗng
1	Điền giá trị a	1	y1
i	xi	bi	yi
….	…	….	….

Giá trị x, a, b, y tính toán như sau:

xi = xi-2 mod xi-1

yi = xi-1 div xi

bi = bi-2 – (bi-1 * yi-1)

Dừng khi xi = 1, kết quả: bi

Nếu xi = 0 → Không tồn tại phần tử nghịch đảo:

Ví dụ: 75-1 mod 97

x	**b**	**y**	
0	97	0	
1	75	1	1
2	22	-1	3
3	9	4	2
4	4	-9	2

5	1	**22**

Kết quả: 75-1 mod 97 = 22

Ví dụ: 13-1 mod 77 = 6;

Ví dụ: 35-1 mod 79

x	**b**	**y**	
0	79	0	
1	35	1	2
2	9	-2	3
3	8	7	1
4	1	**-9**	

Kết quả: -9 mod 79 = 70 (cách tính mod với số âm: cộng liên tục với số chia cho đến khi được

kết quả dương)

Ví dụ: 39-1 mod 121 = 90

1.3.4 Thuật toán tính modulo của lũy thừa số mũ lớn

Phương pháp 1: Giới thiệu phương pháp bình phương liên tiếp:

Ví dụ: Tính 12588 mod 82

Khai triển số mũ 88 dưới dạng cơ số 2:

88 = 64 + 16 + 8 = 26 + 24 + 23

Tính liên tiếp các đồng dư bình phương như sau:

125 mod 82 = 43 → 20

1252 mod 82 = 45 → 21

1254 mod 82 = 452 mod 82 = 57 → 22

1258 mod 82 = 572 mod 82 = 51 → 23

12516 mod 82 = 512 mod 82 = 59 → 24

12532 mod 82 = 592 mod 82 = 37 → 25

12564 mod 82 = 372 mod 82 = 57 → 26

Lấy tích của các lũy thừa 26 + 24 + 23 rút gọn theo modulo 82

12588 mod 82 = 57*59*51 mod 82 = 51

Ví dụ: Tính 515307 mod 517

Khai triển số mũ 307 dưới dạng cơ số 2:

307 = 256 + 32 + 16 + 2 + 1 = 28 + 25 + 24 + 21 + 20

Tính liên tiếp các đồng dư bình phương như sau:

515 mod 517 = 515 → 20

5152 mod 517 = 4 → 21

5154 mod 517 = 42 mod 517 = 16 → 22

5158 mod 517 = 162 mod 517 = 256 → 23

17 | Tổng hợp và biên soạn: Đặng Văn Nam

51516 mod 517 = 2562 mod 517 = 394 → 24

51532 mod 517 = 3942 mod 517 = 136 → 25

51564 mod 517 = 1362 mod 517 = 401 → 26

515128 mod 517 = 4012 mod 517 = 14 → 27

515256 mod 517 = 142 mod 517 = 196 → 28

Lấy tích của các lũy thừa 28 + 25 + 24 + 21 + 20 rút gọn theo modulo 517

515307 mod 517 = 515 * 4 * 394 * 136 * 196 mod 517 = 26

Phương pháp 2: Tạo bảng

ax mod n

x	**a**	**d = 1 (giá trị khởi tạo)**
Giá trị x	Giá trị a	1
….	….	….

Nếu x lẻ tính lại các giá trị d, x, a như sau:

d = d*a mod n;

x = x div 2

a = (a*a) mod n

Nếu x chẵn: giữ nguyên d tính x, a như công thức trên

Lặp lại quy trình trên tới khi x = 0 thì dừng lại, và d là kết quả cần tính

Ví dụ: **Tính 12588 mod 82**

x	**a**	**d**	**Tính giá trị**
88	125	1	x = 88 chẵn: d = 1 (giữ nguy x = x div 2 = 4 a = a2 mod n = mod 82 = 45
44	45	1	x = 44 chẵn: d = 1 (giữ nguy x = x dive 2 = 2 a = a2 mod n = 82 = 57
22	57	1	x = 22 chẵn: d = 1 giữ nguyê x = x div 2 = 22 11 a = a2 mod n = 82 = 51

11	51	1	x = 11 lẻ: d = d*a mod 82 mod 82 = 51

			x = x div 2 = 1 5 a = a2 mod n = 82 = 59
5	59	51	x = 5 lẻ: d = d*a mod n mod 82 = 57 x = x div 2 = 5 a = a2 mod n = 82 = 37
2	37	57	x = 2 chẵn: d = 57 (giữ ngu x = x div 2 = 2 a = a2 mod n = 82 = 57
1	57	57	x = 1 lẻ: d = d*a mod n mod 82 = 51 x = x div 2 = 1 a = a2 mod n = 82 = 51
0	51	**51**	**x = 0 dừng!**

Kết quả 12588 mod 82 = 51

Ví dụ: **Tính 515307 mod 517**

x	**a**	**d**	**Tính giá trị**

307	515	1	x = 307 lẻ: d = d*a mod n mod 517 = 515 x = x div 2 = 3 = 153 a = a2 mod n = mod 517 = 4
153	4	515	x = 153 lẻ: d = d*a mod n mod 517 = 509 x = x div 2 = 1 = 76 a = a2 mod n = 517 = 16
76	16	509	x = 76 chẵn: d giữ nguyên x = x div 2 = 7 38 a = a2 mod n = 517 = 256
38	256	509	x = 38 chẵn: d giữ nguyên

x = x div 2 = 38 div 2 = 19

a = a2 mod n = 2562 mod 517 = 394			
19	394	509	x = 19 lẻ: d = d*a mod n 394 mod 517 = x = x div 2 = 19 9 a = a2 mod n = mod 517 = 136
9	136	467	x = 9 lẻ: d = d*a mod n 136 mod 517 = x = x div 2 = 9 a = a2 mod n = mod 517 = 401
4	401	438	x = 38 chẵn: d giữ nguyên x = x div 2 = 4 a = a2 mod n = mod 517 = 14
2	14	438	x = 38 chẵn: d giữ nguyên x = x div 2 = 2 a = a2 mod n = 517 = 196

1	196	438	x = 1 lẻ: d = d*a mod n 196 mod 517 = x = x div 2 = 1 a = a2 mod n = mod 517 = 158
0	158	**26**	**x = 0 dừng!**

Kết quả 515307 mod 517 = 26

1.3.5 Số nguyên tố lớn và bài toán kiểm tra số nguyên tố

Số nguyên tố: Số nguyên tố là số tự nhiên lớn hơn 1 và chỉ có hai ước 1 và chính nó.

Ví dụ: 2, 3, 5, 7, 53, 2521….

2 là số nguyên tố đầu tiên và là số nguyên tố chẵn duy nhất. Người ta đã chỉ ra rằng số

lượng các số nguyên tố là vô hạn.

Số nguyên tố có vai trò và ý nghĩa to lớn trong số học và lý thuyết mật mã. Bài toán kiểm

tra tính nguyên tố của một số nguyên dương N và phân tích một số N ra thành thừa số nguyên tố

là các bài toán rất được quan tâm.

Hệ mật mã thường sử dụng số nguyên tố lớn cỡ 512 bits và thậm chí lớn hơn như vậy.

ϖ 5 số nguyên tố lớn đã được tìm thấy (01/2017)

TT	**Giá trị**	**Kích thước**	**Thời gian**
1	277 232 917 - 1	23 249 425	26/12/2017

2	274 207 281 - 1	22238 618	1/2016 (9/2015)

3	257 885 161 - 1	17 425 170	1/2013
4	243 112 609 - 1	12 978 189	8/2008
5	242 643 801 - 1	12 837 064	4/2009
6	237 156 667 - 1	11 185 272	9/2008

(Great Internet Mersenne Prime Search(GIMPS)-Đội tìm kiếm số nguyên tố lớn Mersenne)

Để hình dung độ lớn của số nguyên tố lớn lớn thứ 1 (Mersenne 50) được tìm thấy, cần có

~ 4 740 trang giấy A4 để biểu diễn số đó với các chữ số trong hệ cơ số 10 (100 chữ số trên 1

dòng; 50 dòng trên 1 trang), nếu viết 2 chữ số trên 1 cm trải dài 118Km. Cần 4 chương trình

phần mềm, chạy trên 4 cấu hình phần cứng khác nhau, liên tục trong 6 ngày.

Để biểu diễn số nguyên tố Mersenne 48 cần có 4 647 trang giấy A4 để biểu diễn số đó

với các chữ số trong hệ cơ số 10, 75 chữ số một dòng và 50 dòng một trang. Nếu dùng giấy định

lượng 70g/m2, sẽ cần hơn 10 kg giấy để in thành tập dày khoảng 20 cm. Để tìm được số này,

GIMPS sử dụng 360 000 CPU, tốc độ tính toán đạt mức cao nhất là 150 nghìn tỷ phép tính trên

giây và chạy liên tục suốt 17 năm (từ năm 1996)

Số nguyên tố Mersenne là số nguyên tố có dạng p =2n – 1. Hiện tại chỉ có 50 số Mersenne được biết đến. 3, 7, 31, 127, 8191….

Giải thưởng GIMPS tiếp theo sẽ được dành cho người tìm được số nguyên tố lớn với 100

triệu chữ số, trị giá $150 000 USD.

Bài toàn kiểm tra số nguyên tố:

Các thuật toán để kiểm tra số nguyên tố được chia làm hai loại: Thuật toán tất định và

thuật toán xác suất. Các thuật toán tất định cho chúng ta biết chính xác câu trả lời một số nguyên

có phải là một số nguyên tố hay không còn một thuật toán xác xuất sẽ cho biết xác suất của một

số nguyên là một số nguyên tố là bao nhiêu.

Phương pháp cổ điển:

Ý tưởng: Kiểm tra tính nguyên tố của một số dương n, tiến hành kiểm tra lần lượt các số từ 2 tới

n/2 xem có giá trị nào là ước của n hay không, nếu không tìm thấy ước nào thì kết luận n là số

nguyên tố.

Thuật toán:

Input: N

Output: True or Fales

BEGIN

KT=True;

For i = 2 to sqrt(n) do

If (N mod i) = 0 then

{

KT=False;

Break;

21 | Tổng hợp và biên soạn: Đặng Văn Nam

}

END

Trong thuật toán trên, số phép tính để kiểm tra xem một số tự nhiên n có phải là số nguyên tố hay không cần khoảng n phép tính.

Như vậy, nếu kiểm tra một số n ~ 10 425 thì thuật toán trên phải tính khoảng 10425 phép tính.

- Nếu dùng 1 máy tính có tốc độ tính toán khoảng 100 triệu tỷ phép tính trong 1 giây (=

1017) thì thời gian thực hiện là khoảng 10425/1017 = 10408 giây.

- Mỗi ngày có khoảng 24 giờ * 60 phút * 60 giây = 86000(s) ≈ 100 000 s (lấy tròn) = 10 5

giây.

- Mỗi năm có khoảng 365 ngày ≈ 1000 ngày (lấy tròn) = 103 (ngày)*10 5 (giây)≈10 8 giây.

Như vậy nếu kiểm tra một số n ~ 10425 thì thuật toán trên phải tính khoảng 10408 /10 8 = **10400** năm.

Do đó với các số nguyên tố lớn, việc sử dụng phương pháp này là không thể. Hiện nay,

người ta có các phương pháp xác suất để kiểm tra tính nguyên tố của một số nguyên dương n

như phương pháp: Solovay-Strassen, Lehmann-Peralta, Miller-Rabin.

1.3.6 Độ phức tạp thuật toán.

Một chương trình máy tính thường được cài đặt dựa trên một thuật toán đúng để giải
quyết bài toán hay vấn đề. Tuy nhiên, ngay cả khi thuật toán đúng, chương trình vẫn có thể
không sử dụng được đối với một dữ liệu đầu vào nào đó vì thời gian để cho ra kết quả là quá lâu
hoặc sử dụng quá nhiều bộ nhớ (vượt quá khả năng đáp ứng của máy tính)
Khi tiến hành phân tích thuật toán nghĩa là chúng ta tìm ra một đánh giá về thời gian và
không gian cần thiết để thực hiện thuật toán. Không gian ở đây được hiểu là các yêu cầu về bộ
nhớ, thiết bị lưu trữ…. của máy tính để thuật toán có thể làm việc. Việc xem xét không gian của
thuật toán phụ thuộc phần lớn vào cách tổ chức dữ liệu của thuật toán. Trong phần này, khi nói
đến độ phức tạp của thuật toán, chúng ta chỉ đề cập đến những đánh giá về mặt thời gian.
1. Thời gian chạy trong trường hợp xấu nhất (worse-case running time) Thời gian chạy lớn nhất
của thuật toán đó trên tất cả các dữ liệu cùng cỡ
2. Thời gian chạy trung bình Là trung bình cộng thời gian chạy trên tất cả các bộ dữ liệu cùng cỡ.
3. Thời gian chạy trong trường hợp tốt nhất (best-case running time) Thời gian chạy ít nhất của
thuật toán đó trên tất cả các dữ liệu cùng cỡ
Đánh giá thời gian chạy thuật toán: – T(n) = số lượng phép toán sơ cấp cần phải thực hiện

(phép toán số học, phép toán logic, phép toán so sánh). Mỗi phép toán sơ cấp được thực hiện
trong một khoảng thời gian cố định.
Phân tích thuật toán là một công việc rất khó khăn, đòi hỏi phải có những hiểu biết sâu
sắc về thuật toán và nhiều kiến thức toán học khác. Đây là công việc mà không phải bất cứ ai
cũng làm được.

CHƯƠNG 2: MÃ HÓA DỮ LIỆU

2.1 Tổng quan về mã hóa dữ liệu

2.1.1 Khái niệm về mã hóa dữ liệu

Việc mã hoá phải theo quy tắc nhất định, quy tắc đó gọi là ***Hệ mã hóa***.

Hệ mã hóa được định nghĩa là bộ năm (**P**, **C**, **K**, **E**, **D**), trong đó:

P là tập hữu hạn các ***bản rõ*** có thể. **K** là tập hữu hạn các ***khoá*** có thể.	**C** là tập hữu hạn các ***bản mã*** có thể.	
E là tập các hàm lập mã.	**D** là tập các hàm giải mã.	
Với khóa lập mã **ke** $\in$ **K**, có hàm lập mã	**eke** $\in$ **E**, **eke**: **P**$\rightarrow$ **C**,	
Với khóa giải mã **kd** $\in$ **K**, có hàm giải mã **dkd** $\in$ **D**, **dkd**: **C**$\rightarrow$ **P**,		
sao cho	**dkd (eke (x)) = x**,	$\forall$ **x** $\in$ **P**.
Ơ đây **x** được gọi là ***bản rõ***,	**eke (x)** được gọi là ***bản mã***.	

2.1.2 Phân loại hệ mã hóa

Có hai hệ mã hóa chính:

- Hệ mã hóa khóa bí mật (Mã hóa khóa đối xứng).
- Hệ mã hóa khóa công khai (Mã hóa khóa bất đối xứng).

a. Hệ mã hóa khóa bí mật

Hệ mã hóa khoá bí mật là Hệ mã hóa mà biết được khóa lập mã thì có thể "***dễ***" tính được khóa giải
mã và ngược lại. Đặc biệt một số Hệ mã hóa có khoá lập mã và khoá giải mã trùng nhau (**ke** = **kd**),
như Hệ mã hóa "dịch chuyển" hay DES.

Hệ mã hóa khóa bí mật còn gọi là ***Mã hóa khóa đối xứng***, vì phải giữ bí mật cả 2 khóa.

Trước khi dùng Hệ mã hóa khóa đối xứng, người gửi và người nhận phải thoả thuận thuật toán mã
hóa và ***khoá chung*** (lập mã hay giải mã), khoá phải được giữ bí mật. Độ an toàn của Hệ mã hóa
loại này ***phụ thuộc vào khoá***.

VD: Hệ mã hóa cổ điển, DES, 3DES, AES, IDEA…

Ưu điểm:

- Hệ mã hóa khóa đối xứng mã hóa và giải mã ***nhanh hơn*** Hệ mã hóa khóa công khai.

Hạn chế:

- Mã hóa khóa đối xứng chưa thật an toàn với lý do 2 khóa phải được giữ bí mật tuyệt đối.
- Vấn đề thỏa thuận và quản lý khóa chung là khó khăn và phức tạp. Người gửi và người

nhận phải luôn thống nhất với nhau về khoá. Việc thay đổi khoá là rất khó và dễ bị lộ.

Khóa chung phải được gửi cho nhau trên kênh an toàn.

Nơi sử dụng Hệ mã hóa khóa đối xứng:

• Hệ mã hóa khóa đối xứng thường được sử dụng trong môi trường mà khoá chung có thể

dễ dàng trao chuyển bí mật, chẳng hạn trong cùng một mạng nội bộ.

• Hệ mã hóa khóa đối xứng thường dùng để mã hóa những bản tin lớn, vì tốc độ mã hóa và

giải mã nhanh hơn Hệ mã hóa khóa công khai.

b. Hệ mã hóa khóa công khai

Hệ mã hoá khóa công khai là Hệ mã hóa có khóa lập mã và khóa giải mã khác nhau (**ke ≠ kd**),

biết được khóa này cũng "**khó**" tính được khóa kia.

Hệ mã hóa này còn được gọi là **Hệ mã hoá khóa công khai**, vì:

Khoá lập mã cho ***công khai***, gọi là ***khoá công khai (Public key)***.

Khóa giải mã giữ bí mật, còn gọi là ***khóa riêng (Private key)*** hay ***khóa bí mật***.

Một người bất kỳ có thể dùng khoá công khai để mã hoá bản tin, nhưng chỉ người nào có

đúng khoá giải mã thì mới có khả năng đọc được bản rõ.

Hệ mã hóa khoá công khai hay ***Hệ mã hóa phi đối xứng*** do Diffie và Hellman phát minh vào

những năm 1970.

VD: RSA, Elgammal…

Ưu điểm:

• Thuật toán được viết một lần, công khai cho nhiều lần dùng, cho nhiều người dùng, họ chỉ
cần giữ bí mật khóa riêng của mình.
• Người mã hoá dùng khóa công khai, người giải mã giữ khóa bí mật. Khả năng lộ khóa bí
mật khó hơn vì chỉ có một người giữ gìn.

Hạn chế:

• Hệ mã hóa khóa công khai: mã hóa và giải mã ***chậm hơn*** hệ mã hóa khóa đối xứng.

Nơi sử dụng Hệ mã hóa khoá công khai:

• Hệ mã hóa khóa công khai thường được sử dụng chủ yếu trên các mạng công khai như
Internet, khi mà việc trao chuyển khoá bí mật tương đối khó khăn.
• Đặc trưng nổi bật của hệ mã hoá công khai là khoá công khai (public key) và bản mã
(ciphertext) đều có thể gửi đi trên một kênh truyền tin ***không an toàn***.Nhưng vì có tốc
độ mã hóa và giải mã ***chậm***, nên hệ mã hóa khóa công khai chỉ dùng để mã hóa những
bản tin ngắn, ví dụ như mã hóa khóa bí mật gửi đi.
• Hệ mã hóa khóa công khai thường được sử dụng cho cặp người dùng thỏa thuận khóa bí
mật của Hệ mã hóa khóa riêng.

2.1.3 Tiêu chuẩn đánh giá hệ mã hóa

Để đánh giá một hệ mã hóa người ta thường đánh giá thông qua các tính chất sau:

a. Độ an toàn: Một hệ mã hóa được đưa vào sử dụng phải có độ an toàn cao. Ưu

điểm của mật
mã là có thể đánh giá được độ an toàn thông qua độ an toàn tính toán mà không phải cài đặt. Một
hệ mã hóa được coi là an toàn nếu để phá hệ mã hóa này phải dùng n phép toán, mà để giải quyết
n phép toán cần thời gian vô cùng lớn.
Một hệ mã hóa được gọi là tốt phải đảm bảo các tiêu chuẩn sau:
¬ Chúng phải có phương pháp bảo vệ mà chỉ dựa trên sự bí mật của các khoá, công
khai thuật toán.
¬ Khi cho khoá công khai Ke (khóa lập mã) và bản rõ P thì chúng ta dễ dàng tính được
Ke(P) = C. Ngược lại khi cho Kd (khóa giả mã) và bản mã C thì dễ dàng tính được Kd(M)=P. Khi không biết Kd thì không có khả năng để tìm được P từ C, nghĩa là khi
cho hàm f: X → Y thì việc tính y=f(x) với mọi x∈ X là dễ còn việc tìm x khi biết y lại
là vấn đề khó và nó được gọi là hàm một chiều.
¬ Bản mã C không được có các đặc điểm gây chú ý, nghi ngờ.
b. Tốc độ mã và giải mã: Khi đánh giá hệ mã hóa chúng ta phải chú ý đến tốc độ mã và giải mã.
Hệ mã hóa tốt thì thời gian mã và giải mã nhanh.

c. Phân phối khóa: Một hệ mã hóa phụ thuộc vào khóa, khóa này được truyền công khai hay
truyền khóa bí mật. Phân phối khóa bí mật thì chi phí sẽ cao hơn so với các hệ mật

có khóa công

khai. Vì vậy đây cũng là một tiêu chí khi lựa chọn hệ mã hóa.

2.2 Hệ mã hóa khóa bí mật

2.2.1 Hệ mã hóa khóa cổ điển

2.2.1.1 Hệ mã hóa dịch chuyển (Caesar)

Đây là hệ mã cổ điển và đơn giản nhất đã từng được sử dụng trong thực tế bởi hoàng đế

La mã Caesar. Hệ mã hóa dịch chuyển là một hệ mã hóa thay thế đơn âm làm việc trên bảng chữ

cái tiếng anh, bao gồm 26 ký tự.

A	**B**	**C**	**D**
0	**1**	**2**	**3**

Không gian bản rõ P ≡ bản mã C là các thông điệp được tạo thành từ bảng chữ cái A, số phần tử

của bảng chữ cái |A|=N.

Sơ đồ mã hóa:

Đánh số các chữ cái từ 0 tới N-1, Không gian khóa K=ZN. Với mỗi khóa k ϵ K, hàm mã hóa và

giải mã một ký tự có số thứ tự x sẽ được thực hiện như sau:

P≡C≡K≡ZN. Bản rõ x và bản mã là y. x,yϵZN

ϖ **Hàm mã hóa:** y = Ek(x) = (x+k) mod N. Ký tự bản rõ có số thứ tự là x sẽ được chuyển

thành ký tự có số thứ tự là (x+k) mod N trong bảng chữ cái A.

ϖ **Hàm giải mã:** x = Dk(y) = (y-k) mod N

Với hệ mã hóa Caesar có khóa lập mã và khóa giải mã là trùng nhau.

Ví dụ: P≡C≡K≡Z26

Bản rõ chữ:	TOI	NAY	THA	VIRUS
Bản rõ số: Chọn khóa kϵ K, k=3	19 14 8	13 0 24	19 7 0	21 8 17 20 18
Hàm mã hóa: y= Ek(x) = (x + k) mod N = (x+3) mod 26				
Bản mã số: Bản rõ chữ:	22 17 11 16 3 1 22 10 3	24 11 20 23 21		
WRL	QDB	WKD	YLUXV	

Hàm giải mã: x= Ek(y) = (y - k) mod N = (y - 3) mod 26

Độ an toàn của hệ mã hóa dịch chuyển: Rất thấp. Tập khóa K chỉ có 26 khóa (thực tế chỉ có 25

khóa có ích).

2.2.1.2 Hệ mã hóa Affine

Không gian các bản rõ và bản mã của hệ mã là các xâu được hình thành từ một bảng chữ cái A,

giả sử |A|=N. Khi đó không gian khóa của hệ mã hóa được xác định như sau:

K={(a,b): a,b ϵ ZN, UCLN(a,N)=1}

Để mã hóa tiến hành đánh số các chữ cái từ 0 tới N-1 và tiến hành mã hóa và giải mã từng ký tự

theo hàm sau:

ϖ **Hàm mã hóa:** y = Ek(x) = (a*x + b) mod N. Ký tự bản rõ có số thứ tự x sẽ được chuyển

thành ký tự có số thứ tự là (a*x+b) mod N trong bảng chữ cái.

ϖ **Hàm giải mã:** x = Dk(y) = a-1(y-b) mod N

Để giải mã cần tìm a-1 (do UCLN(a,N) = 1 nên luôn tồn tại phần tử nghịch đảo của a). Với hệ mã

hóa Affine có khóa lập mã (a,b); khóa giải mã (a-1, b)

Ví dụ: P≡C≡K≡Z26

Bản rõ chữ: TOI NAY THA VIRUS

Bản rõ số: 19 14 8 13 0 24 19 7 0 21 8 17 20 18

Chọn khóa lập mã (a,b) sao cho a,b ϵ Z26, UCLN(a,26)=1

Để a ϵ Z26 và nguyên tố cùng nhau với N=26 \

a={1,3,5,7,9,11,15,17,19,21,23,25}; số phần tử

Φ(a) = 12. Số lượng phần tử b ϵ Z26 là Φ(a) = 26. Không gian khóa (a,b) = 12*26 = 312.

Chọn khóa lập mã (a,b) = (3, 6). Hàm mã hóa y = Ek(x) = (a*x + b) mod N = (3*x+6) mod 26

Bản mã số: Bản mã chữ:	11 22 4 19 6 0 11 1 6 17 4 5 14 8		
LWE	TGA	LBG	REFOI

khóa giải mã (a-1, b) = (9,6)

Hàm giải mã: x = Dk(y) = a-1(y-b) mod N = 9*(y-6) mod 26

Độ an toàn của hệ mã hóa dịch chuyển: Rất thấp!

2.2.1.3 Hệ mã hóa Hill

Hệ mã hóa này dựa trên lý thuyết về đại số tuyến tính do Lester S.Hill đưa ra năm 1929.

Không gian bản rõ và bản mã đều là các xâu được thành lập từ một bảng chữ cái A, có số

phần từ là N. Các ký tự của bản chữ cái được đánh số từ 0 → N-1.

Với mỗi số nguyên M, khóa của hệ mã là một ma trận k vuông kích thước MxM

gồm các
phần tử gồm các số nguyên thuộc ZN.
Điều kiện để ma trận K có thể sử dụng làm khóa của hệ mã là K phải là một ma trận
không suy biến trên ZN. (Nói cách khác là tồn tại ma trận nghịch đảo của ma trận K trên ZN).
Để mã hóa một bản rõ, tiến hành chia bản rõ thành các xâu có độ dài M → chuyển các
xâu thành số thứ tự của các chữ cái trong bảng chữ cái dưới dạng một vector hàng M.

ϖ Hàm mã hóa: C = K*P mod N

ϖ Hàm giải mã: P = K-1 * C mod N

Ví dụ: P≡C≡K≡Z26

Bản rõ chữ:	TOI NAY THA VIRUS
Bản rõ số:	19 14 8 13 0 24 19 7 0 21 8 17 20 18

Chọn M=2 (N=26);

kh óa	K = 3 3 2 5

T O (19, 14) → C = (19,14)* 3 3 2 5	mod 26 = (7, 23) = H X

I N (8, 13) → ……= (24,11) = Y L

AY (0, 24) → ……= (20,16) = UQ

TH (19, 7) → ……= (19,14) = TO

AV (0, 21) → ……= (16,1) = QB

IR (8, 17) → ……= (6,5) = GF

US (20,18) → ……= (18,20) = SU

Bản mã số: Bản mã chữ:	7 23 24 11 20 16 19 14 16 1 6 5 18 20		
HXY	LUQ	TOQ	BGFSU

Quá trình giải mã:

Khóa giải mã: xác định ma trận nghịch đảo của ma trận khóa lập mã K là K-1 (K-1 * K mod 26 =

I ma trận đơn vị).

2.2.1.4 Hệ mã hóa Vigenere (Sinh viên tham khảo TL)

2.2.1.5 Hệ mã hóa Hoán vị

Các phương pháp mã hóa đã trình bày cho đến thời điểm này sử dụng phương thức thay một chữ

cái trong bản rõ bằng một chữ cái khác trong bản mã (phương pháp thay thế). Một cách thực hiện

khác là xáo trộn thứ tự của các chữ cái trong bản rõ. Do thứ tự của các chữ cái bị mất đi nên

người đọc không thể hiểu được ý nghĩa của bản tin dù các chữ đó không thay đổi.

Phương pháp này có những kỹ thuật sau:

ϖ Đảo ngược toàn bộ bản rõ: Bản rõ được viết theo thứ tự ngược lại tạo ra bản mã.

VD: Bản rõ: TOINAYTHAVIRUS

Bản mã: SURIVAHTYANIOT

ϖ Mã hóa theo sơ đồ: Chuyển bản rõ thành ma trận MxN ghi kết quả theo cột để thu được

bản mã.

VD: Bản rõ: TOINAYTHAVIRUS

T O I N

A Y T H

A V I R

U R

Bản mã: TAAUOYVRITINHR

Một cơ chế phức tạp hơn là chúng ta có thể hoán vị các cột trước khi xuất bản mã.

ϖ Khóa là tập hoán vị: Xác định khóa k = Л ϵ K, là một hoán vị của tập các ký tự của bản

rõ P.

2.2.2 Hệ mã khối và chuẩn mã hóa dữ liệu

Tại sao gọi là hệ mã khối?

- Các hệ mã hóa cổ điển đều có đặc điểm chung là từng ký tự bản rõ được mã hóa tách biệt

→ việc phá mã trở nên dễ dàng hơn.

- Trên thực tế, thường dùng kiểu mật mã trong đó từng khối ký tự của bản rõ được mã hóa

cùng một lúc như một đơn vị mã hóa đồng nhất. Quá trình mã hóa làm việc với các khối

dữ liệu có kích thước khác nhau (tối thiểu là 64 bit), khóa của hệ mã hóa cũng là một xâu

bít có độ dài cố định.

Điều kiện để hệ mã khối an toàn?

- Kích thước khối phải đủ lớn để chống lại phương án tấn công bằng phương pháp thống

kê. Tuy nhiên, khi kích thước khối tăng thì thời gian mã hóa sẽ tăng lên.

- Không gian khóa và chiều dài khóa phải đủ lớn để chống lại phương án tấn công bằng
phương pháp vét cạn. Khóa phải đủ ngắn để việc tạo khóa, phân phối và lưu trữ được dễ
dàng.

Khi thiết kế một hệ mã khối, phải đảm bảo hai yêu cầu:

- Sự hỗn loạn: Sự phụ thuộc giữa bản rõ và bản mã phải thực sự phức tạp để gây khó khăn
đối với việc tìm quy luật thám mã.
- Sự khuếch tán: Mỗi bit của bản rõ và khóa phải ảnh hưởng lên càng nhiều bit của bản mã
càng tốt.

Có nhiều hệ mã khối được phát triển qua thời gian như: DES (1975); AES (1998); CAST,
Blowfish, RC5, RC6, IDEA (1990), GOST, SAFER…Đặc điểm chung của các hệ mã khố là quá
trình mã hóa làm việc với các khối dữ liệu (thường ở dạng xâu bit) có ḱích thước khác nhau (tối
thiểu là 64 bit), khóa của hệ mã cũng là một xâu bít có độ dài cố định, Hoạt động của các hệ mã
khối thường được thực hiện qua một số lần lặp, mỗi lần lặp sẽ sử dụng một khóa con được sinh
ra từ khóa chính.

2.2.2.1 Chuẩn mã hóa dữ liệu DES

- Chuẩn mã hóa dữ liệu DES (Data Encryption Standard) là một chuẩn mã hóa được công
bố bởi ủy ban tiêu chuẩn quốc gia Hoa kỳ vào 15/2/1977. Hệ mã này được phát

triển bởi
IBM.
- DES có nhiều ưu điểm (nhanh, thuật toán công khai, dễ cài đặt) được sử dụng trong một
thời gian dài đến đầu những năm 90.
- DES là một thuật toán đầu vào là các khối 64 bit, đầu ra cũng là khối 64 bit. Khóa mã hóa
có độ dài 64 bit → không gian khóa 264

Thuật toán:

DES thực hiện trên từng khối 64 bit bản rõ. Sau khi thực hiện hoán vị khởi đầu, khối dữ liệu
được chia làm hai nửa trái – phải có chiều dài 32 bit. Tiếp đó có 16 vòng lặp giống hệt nhau
được thực hiện gọi là các hàm f, trong đó dữ liệu được kết hợp với khóa. Sau 16 vòng lặp, 2 nửa
trái – phải được kết hợp lại và hoán vị cuối cùng (Hoán vị ngược) được sử dụng để kết thúc thuật
toán.

Chi tiết thuật toán:

Input: Bản rõ P = p1 p2 ….p64. Khóa 64 bit K = k1 k2 ….k64 (bao gồm 8 bit kiểm tra tính chẵn
lẻ).

Output: Bản mã 64 bit C = c1 c2 ….c64.

Bước 1: Sinh khóa con: Sử dụng thuật toán sinh khóa con để sinh ra 16 khóa
K → K1, K2, ….. K16

Bước 2: Sử dụng bảng hoán vị IP để hoán vị các bit. Kết quả nhận được chia thành

2 nửa L0 (32

bit) R0 (32 bit)

IP (p1 p2 ….p64) → (L0, R0)

Bước 3: For I = 1 to 16 (16 vòng lặp)

Tính Li và Ri:

Li = Ri-1;

Ri = Li-1 Θ f(Ri-1,Ki) Với f(Ri-1,Ki) = P(S(E(Ri-1) Θ Ki))

Trong đó:

- Θ là kí hiệu của phép tuyển loại trừ (XOR) của hai xâu bit

A	**B**	**A AND B**	**A OR B**
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	1

- Hàm f là một hàm phi tuyến.
- E là hoán vị mở rộng ánh xạ Ri-1 từ 32 bít thành 48 bit.
- P là hoán vị cố định.

Tính f(Ri-1,Ki) = P(S(E(Ri-1) Θ Ki)) như sau:

a. Mở rộng Ri-1 = r1r2……r32 từ 32 bit thành 48 bit bằng cách sử dụng hoán vị

mở

rộng E.

E(Ri-1) → T (48 bit)

b. T Θ Ki → T’. Biểu diễn T’ thành 8 xâu, mỗi xâu 6 bit.

T’ = (B1, B2, B3, B4, B5, B6, B7, B8)

c. Đưa các xâu (B1, B2, B3, B4, B5, B6, B7, B8) vào các hộp S-Box tương ứng.

Xâu B1

đưa vào S1 chuyển từ 6 bit xuống 4 bit. Kết hợp lại để thu được T''.

(S1(B1), S2(B2), S3(B3), S4(B4), S5(B5), S6(B6), S7(B7), S8(B8)) → T''

d. Sử dụng hoán vị cố định P để chuyển đổi xâu T''

P(T'') → T'''

Bước 4: Đổi vị trí các khối cuối cùng (R16, L16) → (b1b2….b64). Sử dụng hóa vị ngược IP-1 để

thu được bản mã

IP-1 (b1b2….b64) → C

Hình 2.2: Sơ đồ mã hóa DES

A-Thuật toán sinh khóa con:

16 vòng lặp của DES chạy cùng thuật toán như nhau nhưng với 16 khóa con khác nhau.

Các khóa này đều được sinh ra từ khóa chính của DES bằng một thuật toán sinh khóa con.

Khóa chính (64 bit) đi qua 16 bước biến đổi, tại mỗi bước biến đổi một khóa con được

sinh ra với độ dài là 48 bit.

Mô tả thuật toán:

Input: Khóa K = k1k2….k64 (64 bit – bao gồm cả 8 bit kiểm tra tính chẵn lẻ)

Output: 16 khóa con 48 bit K → (K1, K2, ….. K16)

BEGIN

1. Định nghĩa vi, ($1 \leq i \leq 16$)

với: vi = 1 đối với i = (1,2,9,16);

vi = 2 đối với các trường hợp còn lại i = (3,4,5,6,7,8,10,11,12,13,14,15);.

2. Cho K qua bảng trật trự khóa PC-1.

PC-1 (K) → T.

Biểu diễn T thành 2 nửa 28 bit (C0, D0)

3. For i = 1 to 16 (16 vòng lặp)

Dịch trái (C0, D0) đi vi bit:

Ci-1 (dịch trái vi bit) → Ci

Di-1 (dịch trái vi bit) → Di

Cho xâu (Ci, Di) qua bảng trật tự nén PC-2 để thu được khóa Ki tương ứng.

END

Bảng trật tự khóa PC-1:

PC-1:						
Trái (C0)						
57	49	41	33	25	17	9
1	58	50	42	34	26	18
10	2	59	51	43	35	27
19	11	3	60	52	44	36
Phải (D0)						
63	55	47	39	31	23	15
7	62	54	46	38	30	22
14	6	61	53	45	37	29
21	13	5	28	20	12	4

Bảng trật tự khóa PC-2:

PC-2:			
14	17	11	24
15	6	21	10
26	8	16	7

41	52	31	37
51	45	33	48
34	53	46	42

Sơ đồ sinh khóa:

Hình 2.3: Sơ đồ sinh khóa con của DES

Ví dụ: Khóa mã hóa DES

K = 1A2B3C4D5E6F0011 (HEXA)

= 0001 1010 0010 1011 0011 1100 0100 1101 0101 1110 0110 1111 0000 0000 0001 0001

(64 bit)

Cho khóa K qua bảng trật tự khóa PC-1: PC-1(K) → T

T = 0000 0000 0011 1000 0010 0110 1001 0011 0011 0011 1100 0011 1111 0101

C0	D0	
0000 0000 0011 1000 0010 0110 1001	0011 0011 0011 1100 0011 1111 0101	
Vòng lặp 1 (i=1)		
LS1(v1 = 1)	C1	D1
0000 0000 0111 0000 0100 1101 0010	0110 0110 0111 1000 0111 1110 1010	
C1D1		
0000 0000 0111 0000 0100 1101 0010 0110 0110 0111 1000 0111 1110 1010		

PC-2 (C1D1)	0011 0000 0011 0010 0000 1000 1010 1111 1001 0110 1110 1000	
K1	**0011 0000 0011 0010** **0000 1000 1010 1111** **1001 0110 1110 1000**	
Vòng lặp 2 (i=2)		
LS1(v2=1)	C2	D2

33 | Tổng hợp và biên soạn: Đặng Văn Nam

0000 0000 1110 0000 1001 1010 0100	1100 1100 1111 0000 1111 1101 0100	
C2D2		
0000 0000 1110 0000 1001 1010 0100 1100 1100 1111 0000 1111 1101 0100		
PC-2 (C2D2)	0110 0000 0011 1000 1000 0100 0101 1011 0111 0110 1010 1010	
K2	**0110 0000 0011 1000** **1000 0100 0101 1011** **0111 0110 1010 1010**	
Vòng lặp 3 (i=3)		
LS1(v3=2)	C3	D3

0000 0011 1000 0010 0110 1001 0000	0011 0011 1100 0011 1111 0101 0011	
C3D3		
0000 0011 1000 0010 0110 1001 0000 0011 0011 1100 0011 1111 0101 0011		
PC-2 (C3D3)	0001 0000 1010 0100 0101 0000 0111 1100 0101 1001 0010 1101	
K3	**0001 0000 1010 0100 0101 0000 0111 1100 0101 1001 0010 1101**	
Vòng lặp 4 (i=4)		
LS1(v4=2)	C4	D4
0000 1110 0000 1001 0100 0000	1100 1111 0000 1111 1101 0100 1100	
C4D4		
0000 1110 0000 1001 0100 0000 1100 1111 0000 1111 1101 0100 1100		
PC-2 (C4D4)	0100 0100 0100 0100 0011 0010 1000 0010 0111 1000 1111 1110	

K4	**0100 0100 0100 0100 0011 0010 1000 0010 0111 1000 1111 1110**	
Vòng lặp 5 (i=5)		
LS1(v5=2)	C5	D5
0011 1000 0010 0110 1001 0000 0000	0011 1100 0011 1111 0101 0011 0011	
C5D5		
0011 1000 0010 0110 1001 0000 0000 0011 1100 0011 1111 0101 0011 0011		
PC-2 (C5D5)	1110 0110 1000 0001 0000 0100 1110 0101 1011 1011 1011 0001	
K5	**1110 0110 1000 0001 0000 0100 1110 0101 1011 1011 1011 0001**	
Vòng lặp 6 (i=6)		
LS1(v6=2)	C6	D6
1110 0000 1001 1010 0100 0000 0000	1111 0000 1111 1101 0100 1100 1100	
C6D6		
1110 0000 1001 1010 0100 0000 0000 1111		

0000 1111 1101 0100 1100 1100		
PC-2 (C6D6)	0000 1010 1000 0010 0000 0011 1011 0011 0000 1110 0111 1011	
K6	**0000 1010 1000 0010 0000 0011 1011 0011 0000 1110 0111 1011**	
Vòng lặp 7 (i=7)		
LS1(v7=2)	C7	D7
1000 0010 0110 1001 0000 0000 0011	1100 0011 1111 0101 0011 0011 0011	
C7D7		
1000 0010 0110 1001 0000 0000 0011 1100 0011 1111 0101 0011 0011 0011		
PC-2 (C7D7)	0010 1001 0001 0000 0011 1010 0101 1111 1001 1011 0001 0110	
K7	**0010 1001 0001 0000 0011 1010 0101 1111 1001 1011 0001 0110**	
Vòng lặp 8 (i=8)		
LS1(v8=2)	C8	D8

0000 1001 1010 0100 0000 0000 1110	0000 1111 1101 0100 1100 1100 1111	
C8D8		
0000 1001 1010 0100 0000 0000 1110 0000 1111 1101 0100 1100 1100 1111		
PC-2 (C8D8)	1010 0100 0000 0000 1100 1000 0001 0101 0110 0101 1111 1100	
K8	**1010 0100 0000 0000** **1100 1000 0001 0101** **0110 0101 1111 1100**	
Vòng lặp 9 (i=9)		
LS1(v9=1)	C9	D9
0001 0011 0100 1000 0000 0001 1100	0001 1111 1010 1001 1001 1001 1110	

C9D9
0001 0011 0100 1000 0000 0001 1100 0001 1111 1010 1001 1001 1001 1110

PC-2 (C9D9)	0001 0000 0001 0001 1101 0010 1101 0100 0111 1110 1100 0101	
K9	**0001 0000 0001 0001** **1101 0010 1101 0100** **0111 1110 1100 0101**	
Vòng lặp 10 (i=10)		
LS1(v10=2)	C10	D10
0100 1101 0010 0000 0000 0111 0000	0111 1110 1010 0110 0110 0111 1000	
C10D10		
0100 1101 0010 0000 0000 0111 0000 0111 1110 1010 0110 0110 0111 1000		
PC-2 (C10D10)	0011 0100 0100 1000 0100 0001 0111 1010 1010 0010 1111 1001	
K10	**0011 0100 0100 1000** **0100 0001 0111 1010** **1010 0010 1111 1001**	
Vòng lặp 11 (i=11)		
LS1(v11=2)	C11	D11
0011 0100 1000 0000 0001 1100 0001	1111 1010 1001 1001 1001 1110 0001	
C11D11		

0011 0100 1000 0000 0001 1100 0001 1111 1010 1001 1001 1001 1110 0001		
PC-2 (C11D11)	0000 0011 0110 0001 0000 0100 1011 0011 1111 1101 0000 1011	
K11	**0000 0011 0110 0001** **0000 0100 1011 0011** **1111 1101 0000 1011**	
Vòng lặp 12 (i=12)		
LS1(v12=2)	C12	D12
1101 0010 0000 0000 0111 0000 0100	1110 1010 0110 0110 0111 1000 0111	
C12D12		
1101 0010 0000 0000 0111 0000 0100 1110 1010 0110 0110 0111 1000 0111		
PC-2 (C12D12)	0000 1000 0000 0101 1001 0101 0010 1110 0011 0111 0011 0010	
K12	**0000 1000 0000 0101** **1001 0101 0010 1110** **0011 0111 0011 0010**	
Vòng lặp 13 (i=13)		

LS1(v13=2)	C13	D13
0100 1000 0000 0001 1100 0001 0011	1010 1001 1001 1001 1110 0001 1111	
C13D13		
0100 1000 0000 0001 1100 0001 0011 1010 1001 1001 1001 1110 0001 1111		
PC-2 (C13D13)	0101 0101 0000 0000 0010 1001 1111 1101 0110 1001 0110 0110	
K13	**0101 0101 0000 0000 0010 1001 1111 1101 0110 1001 0110 0110**	
Vòng lặp 14 (i=14)		
LS1(v14=2)	C14	D14
0010 0000 0000 0111 0000 0100 1101	1010 0110 0110 0111 1000 0111 1110	
C14D14		
0010 0000 0000 0111 0000 0100 1101 1010 0110 0110 0111 1000 0111 1110		
PC-2 (C14D14)	1000 0011 1000 0000 1010 0000 0110 0100 1100 1010 1101 1010	

K14	**1000 0011 1000 0000 1010 0000 0110 0100 1100 1010 1101 1010**	
Vòng lặp 15 (i=15)		
LS1(v15=2)	C15	D15
1000 0000 0001 1100 0001 0011 0100	1001 1001 1001 1110 0001 1111 1010	
C15D15		
1000 0000 0001 1100 0001 0011 0100 1001 1001 1001 1110 0001 1111 1010		
PC-2 (C15D15)	1001 1000 0000 1010 1000 0110 1101 0101 1011 0100 0101 1111	
K15	**1001 1000 0000 1010 1000 0110 1101 0101 1011 0100 0101 1111**	
Vòng lặp 16 (i=16)		
LS1(v16=1)	C16	D16
0000 0000 0011 1000 0010 0110 1001	0011 0011 0011 1100 0011 1111 0101	
C16D16		

0000 0000 0011 1000 0010 0110 1001 0011 0011 0011 1100 0011 1111 0101	
PC-2 (C16D16)	0010 0001 0000 1110 0000 0010 1110 1001 1001 0111 0001 1101
K16	**0010 0001 0000 1110** **0000 0010 1110 1001** **1001 0111 0001 1101**

VD: Mã hóa DES với

P = ABCDEF0123456789 (HEXA)

= 1010 1011 1100 1101 1110 1111 0000 0001 0010 0011 0100 0101 0110 0111 1000 1001 (64 bit)

B-Hoán vị IP, IP-1

Hai hoán vị IP và IP-1 không có ý nghĩa gì về mặt mật mã mà hoàn toàn nhằm tạo điều kiện cho việc "Chíp hóa" thuật toán DES

Bảng hoán vị IP:

58	50	42	34	26	18	10	2
60	52	44	36	28	20	12	4
62	54	46	38	30	22	14	6
64	56	48	40	32	24	16	8
57	49	41	33	25	17	9	1
59	51	43	35	27	19	11	3

61	53	45	37	29	21	13	5
63	55	47	39	31	23	15	7

Bảng hoán vị ngược IP-1:

40	8	48	16
39	7	47	15
38	6	46	14
37	5	45	13
36	4	44	12
35	3	43	11
34	2	42	10
33	1	41	9

C – Hàm mở rộng (E):

Hàm mở rộng (E) tăng độ dài của Ri từ 32 bit lên 48 bit bằng cách thay đổi thứ tự của các bit

cũng như lặp lại các bit.

Nhằm 2 mục đích:

- Làm độ dài của Ri cùng cỡ với khóa Ki.
- Cho kết quả dài hơn để có thể nén trong suốt quá trình thay thế.

Bảng mô tả hàm mở rộng E:

32	1	2	3

4	5	6	7
8	9	10	11
12	13	14	15
16	17	18	19

20	21	22	23
24	25	26	27
28	29	30	31

D-Hộp S-Box

Với sơ đồ mã hóa DES mọi tính toán đều tuyến tính, chỉ duy nhất có các tính toán với hợp S là
phi tuyến.

Hộp S-Box là quan trọng nhất đối với độ mật của hệ mã, chính các hộp S tạo nên sự hỗ loạn và
sự khuếch tán của DES.

Tiêu chuẩn thiết kế hộp S như sau:

- Mỗi hàng trong mỗi hộp S là một hoán vị của các số nguyên từ 0 đến 15
- Không có hộp S nào là hàm Affine hay tuyến tính đối với các đầu vào của nó.
- Sự thay đổi của một bit đầu vào sẽ dẫn đến sự thay đổi ít nhất của hai bit đầu ra.
- Đối với hộp S bất kỳ và với đầu vào x (xâu bít có độ dài bằng 6 bit) bất kỳ, S(x) và S (x
O 001100) phải khác nhau ít nhất là 2 bit

3 thuộc tính của hộp S:

¬ Các bít vào luôn phụ thuộc không tuyến tính với các bit ra.

¬ Sửa đổi ở một bit vào làm thay đổi ít nhất là hai bit ra.

¬ Khi một bit vào được giữ cố định và 5 bit còn lại cho thay đổi thì hộp S thể hiện một tính
chất được gọi là “Phân phối đồng nhất”: So sánh số lượng bit 0 và 1 ở các đầu ra luôn ở
mức cân bằng. Tính chất này khiến cho việc phân tích theo lý thuyết thống kê để tìm cách

phá hộp S là vô ích.

Sau khi cộng mudulo với khóa K, kết quả thu được chuỗi 48 bit chia làm 8 khối đưa vào 8 hộp

S-Box.

Mỗi hộp S-Box có 6 bit đầu vào và 4 bit đầu ra:

Các bảng S-Box:

Bảng S1:

37 | Tổng hợp và biên soạn: Đặng Văn Nam

Bảng S2:

Bảng S3:

Bảng S4:

Bảng S5:

Bảng S6:

38 | Tổng hợp và biên soạn: Đặng Văn Nam

Bảng S7:

Bảng S8:

Cách tính:

B = B1B2 B3B4 B5B6 B7B8

= 111100 001101 100010 100011 011011 110111 110001 000011

B1 = b1b2b3b4b5b6 (6 bit) → S1 (B1)

¬ b1b6: xác định biểu diễn nhị phân của hàng r trong hộp Si ($0 \leq r \leq 3$)

¬ b2b3b4b5: xác định biểu diễn nhị phân của cột c trong hộp Si ($0 \leq c \leq 15$)

Xâu C(4bit) được định nghĩa là biểu diễn nhị phân của phần tử Si (r,c)

B1 = 111100

r = (10) = 2 → hàng 2

c = (1110) = 14 → cột 14

S1 (2,14) = 5 = (0101)

E-Hộp P-Box

Hoán vị qua hộp P-Box mang tính chất đơn ánh, nghĩa là một bit đầu vào sẽ cho một bit đầu ra,

không bit nào được sử dụng hai lần hay bị bỏ qua

Hộp P-Box thực chất chỉ làm chức năng sắp xếp đơn thuần.

16	7	20	21
1	15	23	26

2	8	24	14
19	13	30	6

Kết quả chạy:

P = 1010 1011 1100 1101 1110 1111 0000 0001 0010 0011 0100 0101 0110 0111 1000

IP(P) = 0110 0110 0000 0000 0110 0110 1111 1111 1000 0111 0101 0101 1000 0111 0101 0101

L0 = 0110 0110 0000 0000 0110 0110 1111 1111

L1 = R0 = 1000 0111 0101 0101 1000 0111 0101 0101

Vòng lặp 1:	
E(R0)	1100 0000 1110 1010 1010 1011 1100 0000 1110 1010 1010 1011
K1	0011 0000 0011 0010 0000 1000 1010 1111 1001 0110 1110 1000

E(R0) O K1	1111 0000 1101 1000 1010 0011 0110 1111 0111 1100 0100 0011
S-Box	0101 1000 0110 1111 1001 0111 1001 1111
P-Box	1010 1011 0110 1011 1011 1000 0110 1111
L2=R1=POL0	1100 1101 0110 1011 1101 1110 1001 0000
Vòng lặp 2:	
E(R1)	0110 0101 1010 1011 0101 0111 1110 1111 1101 0100 1010 0001
K2	0110 0000 0011 1000 1000 0100 0101 1011 0111 0110 1010 1010
E(R1) O K2	0000 0101 1001 0011 1101 0011 1011 0100 1010 0010 0000 1011
S-Box	0000 0110 1010 0111 0010 0010 1111 0011
P-Box	1100 0010 0111 0010 0001 1101 1001 0101
L3=R2=POL1	0100 0101 0010 0111 1001 1010 1100 0000
Vòng lặp 3:	

E(R2)	0010 0000 1010 1001 0000 1111 1100 1111 0101 0110 0000 0000
K3	0001 0000 1010 0100 0101 0000 0111 1100 0101 1001 0010 1101
E(R2) O K3	0011 0000 0000 1101 0101 1111 1011 0011 0000 1111 0010 1101
S-Box	1011 1111 1110 1001 0111 0111 1001 1000
P-Box	1110 1010 1010 1101 0110 0011 1101 1111
L4=R3=POL2	0010 0111 1100 0110 1011 1101 0100 1111
Vòng lặp 4:	
E(R3)	1001 0000 1111 1110 0000 1101 0101 1111 1010 1010 0101 1110
K4	0100 0100 0100 0100 0011 0010 1000 0010 0111 1000 1111 1110
E(R3) O K4	1101 0100 1011 1010 0011 1111 1101 1101 1101 0010 1010 0000

S-Box	0011 0010 1000 1110 1001 0011 0000 0111
P-Box	0110 0001 0110 0010 0011 1011 0110 0010
L5=R4=POL3	0010 0100 0100 0101 1010 0001 1010 0010
Vòng lặp 5:	
E(R4)	0001 0000 1000 0010 0000 1011 1101 0000 0011 1101 0000 0100
K5	1110 0110 1000 0001 0000 0100 1110 0101 1011 1011 1011 0001
E(R4) O K5	1111 0110 1000 0011 0000 1111 0011 0101 1000 0110 1011 0101
S-Box	0110 0000 1111 0011 1101 1110 1010 1001
P-Box	1011 1101 0110 0101 1000 1111 0000 1101
L6=R5=POL4	1001 1010 1010 0011 0011 0010 0100 0010
Vòng lặp 6:	
E(R5)	0100 1111 0101 0101 0000 0110 1001 1010 0100 0010 0000 0101

K6	0000 1010 1000 0010 0000 0011 1011 0011 0000 1110 0111 1011
E(R5) O K6	0100 0101 1101 0111 0000 0101 0010 1001 0100 1100 0111 1110

S-Box	1010 1011 0010 1011 1010 0011 1001 1000
P-Box	1100 1011 1110 1000 0110 0010 1100 0101
L7=R6=POL5	1110 1111 1010 1101 1100 0011 0110 0111
Vòng lặp 7:	
E(R6)	1111 0101 1111 1101 0101 1011 1110 0000 0110 1011 0000 1111
K7	0010 1001 0001 0000 0011 1010 0101 1111 1001 1011 0001 0110
E(R6) O K7	1101 1100 1110 1101 0110 0001 1011 1111 1111 0000 0001 1001
S-Box	1110 0100 1110 0011 1101 1101 0100 0000

P-Box	1011 0001 1101 0101 1010 0011 0001 1100
L8=R7=POL6	0010 1011 0111 0110 1001 0001 0101 1110
Vòng lặp 8:	
E(R7)	0001 0101 0110 1011 1010 1101 0100 1010 0010 1010 1111 1100
K8	1010 0100 0000 0000 1100 1000 0001 0101 0110 0101 1111 1100
E(R7) O K8	1011 0001 0110 1011 0110 0101 0101 1111 0100 1111 0000 0000
S-Box	0010 1101 1000 0000 1010 0100 1001 1101
P-Box	0000 1011 0000 1000 0100 1011 1011 1001
L9=R8=POL7	1110 0100 1010 0101 1000 1000 1101 1110
Vòng lặp 9:	
E(R8)	0111 0000 1001 0101 0000 1011 1100 0101 0001 0110 1111 1101

K9	0001 0000 0001 0001 1101 0010 1101 0100 0111 1110 1100 0101
E(R8) O K9	0110 0000 1000 0100 1101 1001 0001 0001 0110 1000 0011 1000
S-Box	0101 0110 1000 0001 0100 0100 0001 1111
P-Box	1100 1010 0000 0110 1000 1001 0011 1010
L10=R9=POL8	1110 0001 0111 0000 0001 1000 0110 0100
Vòng lặp 10:	
E(R9)	0111 0000 0010 1011 1010 0000 0000 1111 0000 0011 0000 1001
K10	0011 0100 0100 1000 0100 0001 0111 1010 1010 0010 1111 1001
E(R9) O K10	0100 0100 0110 0011 1110 0001 0111 0101 1010 0001 1111 0000
S-Box	1010 1110 1010 0011 1000 0111 0111 0000
P-Box	1100 0011 1111 1000 0010 0111 0001 1100

L11=R10=POL9	0010 0111 0101 1101 1010 1111 1100 0010
Vòng lặp 11:	
E(R10)	0001 0000 1110 1010 1111 1011 1101 0101 1111 1110 0000 0100
K11	0000 0011 0110 0001 0000 0100 1011 0011 1111 1101 0000 1011
E(R10) O K11	0001 0011 1000 1011 1111 1111 0110 0110 0000 0011 0000 1111
S-Box	1101 1001 0111 1110 0011 1001 1000 0100
P-Box	0011 0100 1100 1001 1111 0000 1110 0111
L12=R11=POL10	1101 0101 1011 1001 1110 1000 1000 0011
Vòng lặp 12:	
E(R11)	1110 1010 1011 1101 1111 0011 1111 0101 0001 0100 0000 0111
K12	0000 1000 0000 0101 1001 0101 0010 1110 0011 0111 0011 0010

E(R11) O K12	1110 0010 1011 1000 0110 0110 1101 1011 0010 0011 0011 0101
S-Box	0011 1111 0001 0000 0101 0000 1000 1001
P-Box	0110 1100 0000 1100 0100 1010 0001 0011
L13=R12=POL11	0100 1011 0101 0001 1110 0101 1101 0001
Vòng lặp 13:	
E(R12)	1010 0101 0110 1010 1010 0011 1111 0000 1011 1110 1010 0010
K13	0101 0101 0000 0000 0010 1001 1111 1101 0110 1001 0110 0110
E(R12) O K13	1111 0000 0110 1010 1000 1010 0000 1101 1101 0111 1100 0100
S-Box	0101 1110 1111 0110 1011 0011 0110 1000

P-Box	0110 1101 0111 1001 1011 0101 1001 0110

L14=R13=POL12	1011 1000 1100 0000 0101 1101 0001 0101
Vòng lặp 14:	
E(R13)	1101 1111 0001 0110 0000 0000 0010 1111 1010 1000 1010 1011
K14	1000 0011 1000 0000 1010 0000 0110 0100 1100 1010 1101 1010
E(R13) O K14	0101 1100 1001 0110 1010 0000 0100 1011 0110 0010 0111 0001
S-Box	1011 1111 0100 1010 0101 1010 0100 1111
P-Box	0111 1000 1111 1111 0100 1010 0111 0010
L15=R14=POL13	0011 0011 1010 1110 1010 1111 1010 0011
Vòng lặp 15:	
E(R14)	1001 1010 0111 1101 0101 1101 0101 0101 1111 1101 0000 0110
K15	1001 1000 0000 1010 1000 0110 1101 0101 1011 0100 0101 1111

E(R14) O K15	0000 0010 0111 0111 1101 1011 1000 0000 0100 1001 0101 1001
S-Box	1110 0001 0001 1010 0100 1010 1101 0000
P-Box	0001 0110 1111 0100 1100 0010 0100 0001
L16=R15=POL14	1010 1110 0011 0100 1001 1111 0101 0100
Vòng lặp 16:	
E(R15)	0101 0101 1100 0001 1010 1001 0100 1111 1110 1010 1010 1001
K16	0010 0001 0000 1110 0000 0010 1110 1001 1001 0111 0001 1101
E(R15) O K16	0111 0100 1100 1111 1010 1011 1010 0110 0111 1101 1011 0100
S-Box	0011 0011 0111 0001 0001 1100 1000 1010
P-Box	1111 1100 0000 0011 0100 0010 0000 1111
R16=POL15	1100 1111 1010 1101 1110 1101 1010 1100

R16 = 1100 1111 1010 1101 1110 1101 1010 1100

L16 = 1010 1110 0011 0100 1001 1111 0101 0100

C = IP-1(R16L16) = 0101 1100 1100 1000 1111 1111 1101 1101 0010 1010 1011 0101 0100 0110

1101 1101 = 5CC8FFDD2AB546DD (HEXA)

b) Qui trình giải mã DES

Qui trình giải mã của DES tương tự như qui trình lập mã, nhưng theo

dùng các khóa thứ tự ngược lại: **k16 , k15, … , k1** .

Xuất phát (đầu vào) từ bản mã **y**, kết quả (đầu ra) là bản rõ **x**.

Quy trình giải mã DES tương tự như quy trình lập mã, nhưng các khóa được dùng theo thứ tự

ngược lại: K16 … K1

2.2.2.2 Một số điểm yếu và phương pháp phá mã hệ mã hóa DES

ϖ Một số điểm yếu của DES:

¬ Không gian khóa: DES có 264 khóa.

¬ Khóa yếu: Khóa yếu là các khóa mà theo thuật toán sinh khóa con thì tất cả 16 khóa con

đều như nhau: K1 = K2 = … = K16 (mã hóa và giải mã với khóa yếu là giống hệt nhau)

Có 4 khóa yếu (Hexa):

Có 6 cặp khóa nửa yếu: y = DES(x,K1) và y = DES(x,K2)

Nghĩa là với hai khóa khác nhau nhưng mã hóa ra cùng 1 bản mã từ cùng một bản rõ.

¬ Tính bù:

Nếu ký hiệu U’ là phần bù của U theo từng bit (1 thay bằng 0 và ngược lại). DES

có tính
chất sau:
y = DES(x,k) → y’ = DES(x’,k’)
nếu biết y được mã hóa từ x với khóa k → y’ mã hóa hóa từ x’ với khóa k’: điểm yếu của
DES qua đó đối phương có thể loại bỏ đi 1 số khóa phải thử khi tiến hành giải mã theo
kiểu vét cạn.

ϖ Các phương pháp phá mã DES

¬ Phá mã bằng phương pháp vét cạn khóa: Trong giới nghiên cứu, nhiều đề xuất về các hệ
thống phá mã DES được đề ra. Năm1977, Diffie và Hellman dự thảo một hệ thống có giá
khoảng 20 triệu đô la Mỹ và có khả năng phá khóa DES trong 1 ngày. Năm1993, Wiener
dự thảo một hệ thống khác có khả năng phá mã trong vòng 7 giờ với giá 1 triệu đô la Mỹ.
Những điểm yếu của DES được thực sự chứng minh vào cuối những năm 1990. Vào năm
1997, công ty bảo mật RSA đã tài trợ một chuỗi cuộc thi với giải thưởng 10.000 đô la Mỹ
cho đội đầu tiên phá mã được một bản tin mã hóa bằng DES. Đội chiến thắng trong cuộc
thi này là dự án DESCHALL với những người dẫn đầu bao gồm Rocke Verser, Matt
Curtin và Justin Dolske. Họ đã sử dụng hàng nghìn máy tính nối mạng để phá mã. Khả

năng phá mã DES được chứng minh thêm lần nữa vào năm 1998 khi tổ chức Electronic
Frontier Foundation (EFF), một tổ chức hoạt động cho quyền công dân trên Internet, xây
dựng một hệ thống chuyên biệt để phá mã với giá thành 250 000 USD. Động cơ thúc đẩy
EFF trong hành động này là nhằm chứng minh DES có thể bị phá vỡ trên lý thuyết cũng
như trên thực tế. Hệ thống này đã tìm được khóa DES bằng phương pháp bạo lực trong
thời gian hơn 2 ngày;

¬ Phá mã vi sai: hai nhà toán học người Do thái – Biham và Shamir phát minh ra vào năm
1990. Đây là một kỹ thuật sử dụng những phỏng đoán khác nhau trong bản rõ để đưa ra
những thông tin trong bản mã.

¬ Phá mã tuyến tính:

¬ Phá mã Davies:

2.2.2.3 Giới thiệu Triple DES (3DES) và chuẩn mã hóa cao cấp AES

a) Giới thiệu Triple DES

Thực tế hiện nay, Hệ mã hóa DES đã không còn an toàn. Vì vậy việc tìm kiếm các hệ mã
khác thay thế cho DES là một điều cần thiết. Một trong những cách thức được xem xét đầu tiên
là tận dụng DES nhưng sử dụng mã hóa nhiều lần.

Một trong những giải pháp hiện nay đang sử dụng rộng rãi là mã hóa DES 3 lần, còn gọi
là Triple DES (3DES).
Mô hình sử dụng đơn giản nhất của 3DES là mã hóa 3 lần sử dụng 3 khóa K1, K2, K3.
Bản mã C = DESK3 (DESK2(DESK1(P)); Khóa của 3DES là 168 bit.
Các biến thể khác:
C = DESK3(DES-1K2 (DESk1(P))
C = DESK1 (DESK2(DESK1(P));
Các chứng minh về mặt lý thuyết và các tấn công đối với 3DES cho thấy hệ mã này vẫn
còn an toàn trong một tương lai dài nữa, tuy nhiên thực tế cho thấy nó chậm hơn so với AES 6
lần.

b) Chuẩn mã hóa cao cấp AES ***(Sinh viên tham khảo về quá trình mã hóa – giải mã AES***
trong tài liệu)

AES là một chuẩn mã hóa do Ủy ban tiêu chuẩn của Hoa Kỳ đưa ra vào năm 2001, do 2
nhà mật mã học người Bỉ phát triển là Joan Daemen và Vincent Rijmen.
AES xử lý các khối dữ liệu đầu vào có chiều dài 128 bit, khóa sử dụng trong hệ mã có độ
dài 128, 192 hoặc 256 bit. (AES-128; AES-192; AES-256)

2.3 Hệ mã hóa khóa công khai

2.3.1 Nguyên tắc cấu tạo của hệ mã hóa khóa công khai

Các hệ mã hóa khóa bí mật (khóa đối xứng) có các điểm yếu sau:
♣ Nếu số lượng người sử dụng lớn thì số khóa sẽ tăng lên rất nhanh, chẳng hạn với

n người
sử dụng thì số khóa là n*(n-1)/2 do đó rất khó quản lý, phức tạp và không an toàn.
♣ Dựa trên các hệ mã này không thể xây dựng các khái niệm và dịch vụ như chữ ký điện tử,
dịch vụ xác thực người dùng cho các ứng dụng thương mại điện tử.
Vào năm 1975 Diffie và Hellman trong một công trình của mình đã đề xuất ý tưởng cho phép
xây dựng nên các hệ mã hoạt động theo nguyên tắc mới gắn liền với các bên truyền tin chứ
không gắn với các cặp truyền tin.
Trong hệ mã hóa khóa công khai, mỗi bên truyền tin sẽ có hai khóa, một khóa gọi là khóa bí
mật và một khóa được gọi là khóa công khai. Khóa bí mật là khóa được dùng để giải mã và phải
được giữ bí mật (Kd), khóa công khai là khóa dùng để mã hóa và được công khai hóa để tất cả
mọi người muốn tham gia vào truyền tin có thể sử dụng (Ke).

ϖ Nguyên tắc chung:

Các hệ mã hóa khóa công khai được xây dựng dựa trên các hàm một chiều (hàm một phía).
Hàm một chiều f: X → Y là hàm mà nếu biết x ϵ X, ta có thể dễ dàng tính được y = f(x). Nhưng với bất kỳ y ϵ Y việc tìm x ϵ X sao cho y = f(x) là "khó". Có nghĩa là việc tìm hàm ngược f-1 là bài toán khó.

ϖ Các yêu cầu của hệ mã hóa khóa công khai:

- Việc sinh khóa Kd, Ke phải dễ dàng.
- Nếu có Ke, việc mã hóa C = E(Ke, P) là dễ dàng.

• Nếu có C = E(Ke,P) và khóa Kd thì việc giải mã tìm bản rõ P cũng dễ dàng.

• Nếu biết Ke, C thì việc dò tìm Kd là "khó".

• Việc khôi phục bản rõ P từ bản mã C là rất khó

ϖ Một số hệ mã hóa khóa công khai:

KnapSack, RSA, El Gamal, Các hệ mã dựa trên đường cong Elliptic…

2.3.2 Hệ mã RSA (1977)

a) Mã hóa RSA

Hệ mã RSA (Rivest – Shamir - Adleman) là một thuật toán mã hóa nổi tiếng và được ứng dụng

rộng rãi trong thực tế.

• Tạo cặp khóa (bí mật, công khai):

- Chọn bí mật hai số nguyên tố lớn p, q (p ≠ q) tính N = p*q, công khai N. Đặt P ≡ C ≡

ZN

- Tính Φ(N) = (p-1)*(q-1), giữ bí mật Φ(N)

- Chọn một số e < Φ(N), nguyên tố cùng nhau với Φ(N).

UCLN (e, Φ(N)) = 1; e – được gọi là số mũ lập mã.

- Xác định d: d là phần tử nghịch đảo của e theo modulo Φ(N):

e*d ≡ 1 mod Φ(N); d – được gọi là số mũ giải mã.

¬ Khóa công khai (khóa lập mã): Ke = (e, N)

¬ Khóa bí mật (Khóa giải mã): Kd = (d, p, q)

•	Mã hóa – giải mã: Với bản rõ x ϵ P và bản mã y ϵ C định nghĩa:

¬ Hàm mã hóa y = xe (mod N)

¬ Hàm giải mã x = yd (mod N)

• Ví dụ:

Chọn p = 11; q = 47

Bản rõ x = 26

N = p*q = 11*47 = 517 (công khai N)

Φ(N) = (p-1)*(q-1) = 10 * 46 = 460 (bí mật Φ(N))

Xác định e, sao cho UCLN (e,460) = 1: Chọn e = 3

Xác định d, sao cho e*d ≡ 1 mod Φ(N): d = 307 (307*3 mod 460 = 1)

¬ Khóa công khai: Ke = (e,N) = (3, 517)

¬ Khóa bí mật: Kd = (d,p,q) = (307, 11, 47)

Mã hóa:

y = xe mod N = 263 mod 517 = 515

Giải mã:

x = yd mod N = 515307 mod 517 = 26

Vd: p=31, q = 11, e = 19, d = 79

b) Độ an toàn của hệ mã hóa RSA

• Hệ mã hóa RSA là tất định, tức là với một bản rõ x và một khóa công khai Ke thì chỉ tạo

ra được một bản mã y duy nhất.

• Độ an toàn của hệ mã hóa RSA dựa vào hai vấn đề của toán học: Bài toán RSA và bài

toán phân tích một số nguyên ra thừa số nguyên tố.

Phá mã RSA:

Cách 1: y = xe mod N; Giả sử kẻ tấn công có được bản mã y, khóa công khai Ke (e, N):

→ xe = y mod N. Lúc này kẻ tấn công gặp phải bài toán logarit rời rạc. Logarit rời rạc là

một bài toán khó (chưa có thuật toán hiệu quả nào được phát triển)

Cách2: Phân tích N = p*q, → xác định được Φ(N); tính được d (theo thuật toán Euclide).

Người ta đã chứng minh được rằng, bài toán phân tích một số nguyên N thành thừa số

nguyên tố là một bài toán khó (Độ phức tạp đa thức)

Bảng dưới đây đưa ra các thông số về thời gian và số lượng phép toán trên bit để thực

hiện việc phân tích một số nguyên N ra thừa số nguyên tố áp dụng thuật toán tốt nhất trên

máy tính có tốc độ xử lý 106 phép tính/giây.

Số chữ số thập phân	**Số phép tính trên bit**	**Thời gian**
50	1,4 . 1010	3,9 giờ
75	9,0 . 1012	104 ngày
100	2,3 . 1015	74 năm
200	1,2 . 1023	3,8 . 109 năm
300	1,5 . 1029	4,9 . 1015 năm
500	1,3 . 1039	4,2 . 1025 năm

•	Khóa của hệ mã hóa RSA: Khóa dùng trong các hệ thống này cần phải dài hơn trong các hệ thống mật mã khóa đối

xứng. Tại thời điểm năm 2002, độ dài 1024 bít được xem là giá trị tối thiểu cho hệ thống
sử dụng thuật toán RSA.
Năm 2003, công ty RSA Security cho rằng khóa RSA 1024 bít có độ an toàn tương đương với khóa 80 bít, khóa RSA 2048 bít tương đương với khóa 112 bít và khóa RSA
3072 bít tương đương với khóa 128 bít của hệ thống mật mã khóa đối xứng. Họ cũng
đánh giá rằng, khóa 1024 bít có thể bị phá vỡ trong khoảng từ 2006 tới 2010 và khóa
2048 bít sẽ an toàn tới 2030. Các khóa 3072 bít cần được sử dụng trong trường hợp thông
tin cần giữ bí mật sau 2030.

- Hiện tượng lộ bản rõ:

Vd: N = p*q = 5*7, mã hóa x=6, chọn e=17 → y = xe mod N = 617 mod 35 = 6 = x
Tương tự: N = p*q = 109 *97, chọn e = 865
Với mọi x luôn có xe mod N = x

c) Tốc độ mã hóa và giải mã của RSA

Tốc độ thực hiện của hệ RSA là một trong những điểm yếu so với các hệ mật mã khóa
đối xứng. Vì các quá trình mã hóa và giải mã của hệ RSA đều thực hiện các phép tính có các
toán hạng là những số nguyên cực lớn. RSA có tốc độ mã hóa và giải mã chậm hơn đáng kể so
với hệ mã hóa DES và các thuật toán mã hóa khóa đối xứng khác. Thường thì Mã hóa RSA

chậm hơn DES ít nhất là 100 lần khi cài đặt bằng phần mềm, và có thể chậm hơn từ 1000 đến 10

000 lần khi cài đặt bằng phần cứng (còn tùy thuộc vào cách cài đặt). Do đó, người ta thường sử

dụng một thuật toán mã hóa hóa khóa đối xứng để mã hóa văn bản cần gửi và sử dụng RSA để

mã hóa khóa bí mật của nó (thông thường khóa ngắn hơn nhiều so với văn bản)

2.3.3 Hệ mã hóa El Gamal

a) Mã hóa ElGamal

Hệ mã này được El Gamal đưa ra vào năm 1985.

Sơ đồ tạo cặp khóa (bí mật, công khai):

Chọn p là số nguyên tố sao cho bài toán logarith rời rạc trong Zp là "khó" giải.

Cho $\alpha \epsilon Zp^*$ là

phần tử nguyên thủy. và số ngẫu nhiên a (bí mật) $P = Zp^*$, $C = Zp^* \times Zp^*$

Tính $\beta = \alpha a \bmod p$

Ta định nghĩa: $K = \{(p, \alpha, a, \beta): \beta = \alpha a \bmod p\}$

Các giá trị Ke = (p, α, β) được công khai, Kd = (a) được giữ bí mật

•	Lập mã: Với khóa lập mã Ke = (p, α, β), chọn số ngẫu nhiên bí mật $r \epsilon Zp-1$

y = EKe(x) = (y1, y2) trong đó:

$y1 = \alpha r \bmod p$

$y2 = x*\beta r \bmod p$

• Giải mã: $x = DKd(y) = DKd(y1, y2) = y2(y1a)-1 \bmod p$

VD1: cho p = 2579, α = 2, a = 765. Khi đó $\beta = 2765 \bmod 2579 = 949$

Mã hóa bản rõ: x = 1299

• Mã hóa: Chọn ngẫu nhiên r = 853, y = EKe(x) = (y1, y2)

y1 = αr mod p = 2853 mod 2579 = 435

y2 = x*βr mod p = 1299*949853 mod 2579 = 2396

bản mã y = (435, 2396)

• Giải mã: x = 2369*(435765)-1 mod 2579 = 1299

VD2: p = 97, α = 5, a = 58. Khi đó β = 558 mod 97 = 44

Mã hóa bản rõ: x = 3

Khóa công khai: Ke = (97,5,44)

Khóa bí mật: Kd = (58)

• Mã hóa: chọn ngẫu nhiên r = 36, y = EKe(x) = (y1, y2)

y1 = αr mod p = 536mod 96 = 50

y2 = x*βr mod p = 3*4436 mod 97 = 31

bản mã y = (50, 31)

• Giải mã: x = 31*(5058)-1 mod 97 = 3

b) Độ an toàn của hệ mã El Gamal

• Hệ mã El Gamal là không tất định, với một bản rõ x, một khóa công khai Ke thu được

các bản mã y khác nhau (phụ thuộc vào số ngẫu nhiên r).

• Hệ mã Elgamal cần gấp đôi bộ nhớ để chứa bản mã. Kích thước thông tin sau khi mã hóa

sẽ tăng gấp đôi so thông tin gốc.

• Độ an toàn của hệ mã Elgamal dựa trên tính khó giải của bài toán logarit rời rạc. Để đảm

bảo an toàn số nguyên tố p phải lớn (1024 bít trở lên). Ngoài ra số ngẫu nhiên r không

được sử dụng để mã hóa nhiều hơn một thông điệp.

c) Tốc độ thực hiện Elgamal

Phương pháp tính toán của hệ mã Elgamal rất phức tạp, vì đòi hỏi nhiều phép tính lũy
thừa modulo trong quá trình mã hóa và giải mã, nên hiệu suất thực hiện kém.

CHƯƠNG 3: CHỮ KÝ ĐIỆN TỬ VÀ HÀM BĂM

3.1 Tổng quan về chữ ký điện tử

3.1.1 Khái niệm về chữ ký điện tử

Để chứng thực nguồn gốc hay hiệu lực của một tài liệu, lâu nay người ta dùng chữ ký tay,
ghi vào phía dưới của mỗi tài liệu. Như vậy, người ký phải trực tiếp ký tay vào tài liệu. Ngày nay
các tài liệu được số hóa, người ta cũng có nhu cầu chứng thực nguồn gốc hay hiệu lực của các tài
liệu đó. Rõ ràng không thể ký tay vào tài liệu vì chúng không được in trên giấy → Chữ ký điện
tử.

Chữ ký điện tử hay chữ ký số có thể so sánh tương tự hoàn toàn với chữ ký tay hay không? Thực ra không phải hoàn toàn tương tự. Chữ ký tay là dấu vết của con người tác động
lên cùng bản giấy đã mang chứa văn bản (in/viết sẵn). Phần chữ ký tay và phần văn bản có sẵn là
độc lập, không có quan hệ ràng buộc nào. Do các qui luật của thế giới vật lý, người ta không thể
đánh tráo chữ ký theo kiểu đơn giản là xé bỏ phần tờ giấy chứa chữ ký và ghép nối vào một phần
giấy mang chữ ký tạo mới khác. Tuy nhiên trong thế giới số hóa, các qui luật vật lý

này không có
mặt, và bất cứ lập trình viên nào cũng có thể tha hồ cắt ghép văn bản số hóa mà không bị phát
hiện.
Một số định nghĩa về chữ ký điện tử như sau:
¬ Là một định danh điện tử được tạo ra bởi máy tính, được các tổ chức sử dụng nhằm đạt
được tính hiệu quả và có hiệu lực như là chữ ký tay.
¬ Là một cơ chế xác thực hóa cho phép người tạo ra thông điệp đính kèm một mã số vào
thông điệp giống như là việc ký lên một văn bản bình thường.
Theo khoản 1 Điều 21 Luật Giao dịch điện tử năm 2005: chữ ký điện tử là chữ ký được tạo
lập dưới dạng từ, chữ, số, ký hiệu, âm thanh hoặc các hình thức khác bằng phương tiện điện tử,
gắn liền hoặc kết hợp một cách lôgíc với thông điệp dữ liệu, có khả năng xác nhận người ký
thông điệp dữ liệu và xác nhận sự chấp thuận của người đó đối với nội dung thông điệp dữ liệu
được ký.
Có thể hiểu đơn giản chữ ký điện tử là một đoạn thông tin đi kèm theo dữ liệu điện tử nhằm
mục đích xác nhận người ký thông điệp dữ liệu và xác nhận sự chấp thuận của người đó đối với
nội dung thông điệp dữ liệu được ký. Chữ ký điện tử có thể được chứng thực bởi một tổ chức
cung cấp dịch vụ chứng thực chữ ký điện tử.

Nguyên lý tạo chữ ký điện tử là khác hẳn và phức tạp hơn. Đó là, khi có một văn bản ở dạng
nhị phân X, người ta phải tạo ra một chữ ký ở dạng nhị phân S sao cho S phụ thuộc hàm vào X,
tức là S=f (X); hơn nữa quan hệ hàm này là bí mật (có tham số khóa bí mật) đối với người ngoài.
Do đó nếu có kẻ nào thử đánh tráo (tức giả mạo) chữ ký, quan hệ hàm S=f (X) sẽ không còn
đúng và bị phát hiện. Tuy nhiên việc phát hiện xem một văn bản có chữ ký có là chuẩn hay bị giả

mạo lại phải là một thao tác mà ai cũng làm được dễ dàng, không cần đến khóa bí mật kia (do
người chủ chữ ký nắm giữ). Vì vậy hệ thống chữ ký điện tử được xây dựng trên nguyên tắc sử
dụng hai thuật toán riêng rẽ cho việc tạo chữ ký và kiểm định chữ ký, thông qua việc sử dụng
cặp 2 hàm toán học đối lập nhau, một cần khóa bí mật còn một thì không. Chính do điều này,
mật mã khóa công khai đã được khai thác để giúp hiện thực điểm chốt của cơ chế đặc biệt này.
Người ta tạo ra chữ ký điện tử trên các tài liệu số giống như tạo ra bản mã trên tài liệu với
khóa lập mã. Để kiểm tra chữ ký điện tử thuộc về một tài liệu số, người ta giải mã chữ ký điện tử
bằng khóa giải mã và so sánh với tài liệu gốc.
Một sơ đồ chữ ký điện tử thường chứa hai thành phần: Thuật toán ký Sig() và thuật

toán

xác minh Ver().

Các chữ ký điện tử được sinh ra và sử dụng bởi các hệ chữ ký (sơ đồ) điện tử.

Dưới đây định

nghĩa một sơ đồ chữ ký điện tử:

Định nghĩa:

Một sơ đồ chữ ký điện tử là bộ 5 (P, A, K, S, V), trong đó:

- *P là tập hợp hữu hạn các văn bản có thể.*
- *A là tập hợp hữu hạn các chữ ký có thể.*
- *K là tập hợp hữu hạn các khóa có thể.*
- *S là tập các thuật toán ký.*
- *V là tập các thuật toán kiểm tra chữ ký.*

Với mỗi khóa k ϵ K, có thuật toán ký Sigk ϵ S, Sigk : P → A,

Có thuật toán kiểm tra chữ ký Verk ϵ V, Verk : P x A → {đúng, sai}, thỏa mãn điều kiện

sau với mọi x ϵ P, y ϵ A:

Người ta thường dùng hệ mã hóa khóa công khai để lập sơ đồ chữ ký điện tử.

Trong đó,

khóa bí mật được dùng làm khóa ký, khóa công khai dùng làm khóa kiểm tra chữ ký.

3.1.2 Phân loại "Chữ ký điện tử"

- Phân loại chữ ký theo đặc trưng kiểm tra chữ ký:

1. Chữ ký khôi phục thông điệp
2. Chữ ký kèm theo thông điệp

- Phân loại chữ ký theo mức độ an toàn

1. Chữ ký không thể phủ nhận

2. Chữ ký một lần

• Phân loại chữ ký theo ứng dụng đặc trưng

1. Chữ ký "mù"

2. Chữ ký "nhóm"

3. Chữ ký "bội"

3.1.3 Ưu nhược điểm của chữ ký điện tử

ϖ Những ưu điểm khi sử dụng chữ ký điện tử:

♣ Trước hết, sử dụng chữ ký điện tử là điều kiện bảo đảm tính pháp lý của các giao dịch
điện tử, cho phép các giao dịch có thể thực hiện trong môi trường điện tử. Khác với văn
bản giấy với chữ ký bằng tay, những văn bản điện tử có thể chuyển theo đường truyền
internet trong một thời gian rất ngắn. Như vậy, việc sử dụng chữ ký điện tử và thực hiện
những giao dịch điện tử cho phép tiết kiệm thời gian, sức lực và tăng hiệu quả lao động.

♣ Ngăn chặn khả năng giả mạo chữ ký (theo nghĩa tạo ra một chữ ký điện tử y hệt như chữ
ký đang được sử dụng và có thể kiểm tra bằng cách thông thường bởi mã khóa công
khai). Theo nghiên cứu của các chuyên gia, khả năng giả mạo chữ ký là 1/10, trong khi
đối với chữ ký tay, khả năng này có thể tăng đến 60-70%.

♣ Ngăn chặn khả năng làm giả tài liệu. Sau khi tài liệu điện tử đã được ký bằng

chữ ký điện
tử thì không thể thay đổi. Bất cứ sự thay đổi nào, dù nhỏ nhất, cũng có thể bị phát hiện do
chữ ký điện tử được tạo ra bởi cặp khóa bí mật và khóa công khai. Khi nội dung tài liệu
thay đổi, khóa công khai sẽ không còn tương thích với khóa bí mật, hay nói cách khác,
người nhận sẽ không thể dùng khóa công khai để giải mã bí mật. Như vậy, khi tài liệu đã
được ký bằng chữ ký điện tử, người ta sẽ không thể thay đổi một phần (giả mạo từng
phần) hay toàn bộ (giả mạo toàn bộ) tài liệu mà vẫn dưới chữ ký đó.

♣ Cho phép xác định tác giả văn bản và tính nguyên gốc của văn bản. Về lý thuyết, khi văn
bản điện tử đã được ký bởi chữ ký điện tử thì không thể thay đổi. Nếu thay đổi dù chỉ
một ký tự trong văn bản, việc kiểm tra chữ ký sẽ không mang lại kết quả trùng khớp, và
văn bản đó, đương nhiên, sẽ không có hiệu lực. Như vậy. chữ ký số có thể là công cụ xác
định tác giả tài liệu điện tử cũng như sự vẹn toàn của chúng và một văn bản điện tử được
ký bởi chữ ký số có thể là căn cứ pháp lý để bảo vệ quyền lợi hợp pháp cho người tham
gia giao dịch điện tử.

ϖ Những hạn chế khi sử dụng chữ ký điện tử trong giao dịch điện tử:

♣ Sự lệ thuộc vào máy móc và chương trình phần mềm: Như đã nói ở trên, chữ ký điện tử
là một chương trình phần mềm máy tính. Để kiểm tra tính xác thực của chữ ký cần có hệ
thống máy tính và phần mềm tương thích. Đây là hạn chế chung khi sử dụng văn bản
điện tử và chữ ký điện tử.
♣ Tính bảo mật không cao: Nếu chữ ký bằng tay được thực hiện trên giấy, được ký trực tiếp
và luôn đi kèm với vật mang tin, chữ ký tay không thể chuyển giao cho người khác, thì
chữ ký số không như vậy. Chữ ký số là một bộ mật mã được cấp cho người sử dụng, đây

là phần mềm máy tính không phụ thuộc vào vật mang tin. Chính vì vậy, trở ngại lớn nhất
khi sử dụng chữ ký số là khả năng tách biệt khỏi chủ nhân của chữ ký. Nói cách khác,
chủ nhân của chữ ký số không phải là người duy nhất có được mật mã của chữ ký. Tồn
tại một số nhóm đối tượng có thể có được mật mã, đó là: bộ phận cung cấp phần mềm; bộ
phận cài đặt phần mềm, những người có thể sử dụng máy tính có cài đặt phần mềm.
Ngoài ra, mật mã có thể bị đánh cắp. Cũng có thể, chủ nhân chữ ký số chuyển giao cho
người khác mật mã của mình. Như vậy, tính bảo mật của chữ ký điện tử không cao.

♣ Vấn đề bản gốc, bản chính: Nếu đối với tài liệu giấy, chữ ký được ký một lần và chỉ có
một bản duy nhất (được coi là bản gốc). Bản gốc được ký bằng chữ ký sẽ không thể cùng
lúc ở hai chỗ khác nhau. Có thể tin tưởng rằng, nếu bản gốc duy nhất mất đi thì sẽ không
thể có bản thứ hai giống hệt như vậy. Nhưng với văn bản điện tử đã được ký bằng chữ ký
số, người ra có thể copy lại và bản copy từ bản chính và bản copy từ bản copy không có
gì khác biệt so với bản chính duy nhất được ký. Đây là một thách thức đối với công tác
văn bản và cả nền hành chính. Khái niệm bản gốc, bản chính trong văn bản hành chính sẽ
phải xem xét lại đối với văn bản điện tử.
♣ Sự có thời hạn của chữ ký điện tử. Chữ ký điện tử là chương trình phần mềm được cấp có
thời hạn cho người sử dụng. Về lý thuyết, văn bản sẽ có hiệu lực pháp lý khi được ký
trong thời hạn sử dụng của chữ ký. Tuy nhiên, thực tế hiệu lực pháp lý của văn bản hoàn
toàn có thể bị nghi ngờ khi chữ ký số hết thời hạn sử dụng. Đây cũng là một hạn chế và
thách thức rất lớn đối với việc sử dụng chữ ký điện tử.
Như vậy, sự ra đời của chữ ký điện tử là cơ sở khẳng định giá trị pháp lý của văn bản
điện tử và cho phép thực hiện những giao dịch điện tử. Tuy nhiên, để chữ ký điện

tử trở nên phổ
biến, thông dụng tạo tiền đề cho việc văn bản điện tử phát huy những tính năng vượt trội của
mình và có thể thay thế tài liệu giấy, cần nghiên cứu và khắc phục những hạn chế của chữ ký
điện tử, đồng thời cần có thêm những công cụ khác để khẳng định giá trị pháp lý của tài liệu điện
tử. Điều này cần sự đầu tư nghiên cứu của các nhà khoa học, các chuyên gia và sự hỗ trợ của các
nhà quản lý.

Hình 3-1: Một số nhà cung cấp dịch vụ chứng thư số tại Việt Nam

3.1.4 Ví dụ ứng dụng chữ ký điện tử trong giao dịch điện tử

Chữ ký điện tử được sử dụng trong các giao dịch điện tử như kê khai thuế qua mạng, Hải
quan điện tử, giao dịch ngân hàng, chứng khoán…

Hình 3-2: ***Mô hình ứng dụng chữ ký điện tử trong kê khai thuế***

3.2 Chữ ký RSA

a) Sơ đồ chữ ký

Tạo cặp khóa (khóa ký-khóa bí mật Ks; khóa kiểm tra chữ ký – khóa công khai Kd)

- Chọn bí mật số nguyên tố lớn p,q. Tính N = p*q (công khai N), Đặt P ≡ A ϵ ZN
- Tính bí mật Φ(N) = (p-1)*(q-1).
- Chọn khóa công khai e < Φ(N); nguyên tố cùng nhau với Φ(N); (UCLN(e, Φ(N)) = 1)
- Xác định khóa bí mật d là phần tử nghịch đảo của e theo modulo Φ(N):

d*e ≡ 1 mod Φ(N);

Cặp khóa như sau:

→ Khóa ký (khóa bí mật): Ks = (d, p, q)

→ Khóa kiểm tra chữ ký (khóa công khai): Kd = (e, N)

ϖ Ký số:

Chữ ký x ϵ P là y = Sigks(x) = xd mod N; y ϵ A

ϖ Kiểm tra chữ ký: Verkd(x, y) = đúng ⎧ x ≡ ye mod N

Ví dụ:

Chọn p = 11; q = 47

Tài liệu cần ký x = 26

N = p*q = 11*47 = 517 (công khai N)

Φ(N) = (p-1)*(q-1) = 10 * 46 = 460 (bí mật Φ(N))

Xác định e, sao cho UCLN (e,460) = 1: Chọn e = 3

Xác định d, sao cho e*d ≡ 1 mod Φ(N): d = 307 (307*3 mod 460 = 1)

¬ Khóa ký (Khóa bí mật): Ks = (d, p, q) = (307, 11, 47)

¬ Khóa kiểm tra chữ ký (Khóa công khai): Kd = (e, N) = (3, 517)

Thực hiện ký lên tài liệu:

y = xd mod N = 26307 mod 517 = 445

Thực hiện kiểm tra chữ ký (x, y) = (26, 445):

x = yd mod N ◊ 26 = 4453 mod 517 ◊ chữ ký đúng.

Sơ đồ mã hóa RSA và sơ đồ chữ ký RSA có sự tương ứng:

Chỉ ra những điểm giống và khác nhau CỦA MÃ HÓA RSA VỚI CHỮ KÝ RSA?

Để đảm bảo an toàn người gửi thường mã hóa thông điệp ký x trong quá trình gửi đi. Để

thực hiện điều này có hai cách:

1. Thực hiện việc ký lên thông điệp x trước sau đó mã hóa:

Bên gửi: Người gửi (G) ký vào tài liệu x bằng chữ ký y = SigG(x). sau đó mã hóa thông điệp x

và chữ ký y trước khi gửi đi tới người nhận (N). z = EG (x, y). G gửi z cho N.

Bên nhận: N nhận được z, giải mã z để thu được x, y: (x, y) = DN(z). Tiếp theo thực hiện việc

kiểm tra chữ ký: VerN(x,y) ≡ Đúng

Mã hóa Eke(x)	Bản mã y

Bản rõ x

Khóa công khai Ke

Giải mã Dkd(y)	Bản rõ

Khóa bí mật Kd

x

Hình 3-2(a): ***Sơ đồ mã hóa RSA***

Ký số y =Sigks(x)	(x,y)

Thông điệp x

Khóa bí mật

(Khóa ký Ks)

Kiểm tra Verkd(x, y)

Khoa công khai

(Khóa kiểm tra Kd)

Đúng: x ≡ Verkd(y)

Hình 3-2(b): ***Sơ đồ chữ ký RSA***

Sai: x <> Verkd(y)

Hình 3-3(a): ***Sơ đồ minh họa sơ đồ ký – mã hóa***

2. Thực hiện mã hóa thông điệp trước sau đó ký:

Hình 3-3(b): ***Sơ đồ minh họa sơ đồ mã hóa – ký***

Bên gửi: G tiến hành mã hóa thông điệp x: u = EG(x). Sau đó ký lên bản mã u bằng chữ ký V =

SigG(u). Sau đó G gửi (u,v) cho N.

Bên nhận: N nhận được (u,v). tiến hành giải mã u để thông điệp x: x = DN(u). và kiểm tra chữ

ký u bằng VerN(x,y) ≡ Đúng?

Lựa chọn phương án nào là tốt hơn?

- Giả sử tin tặc (H) lấy trộm được thông tin trên đường truyền khi G gửi đến N.

Trong trường hợp 1: H lấy được Z

Mã hóa

u =EG(x)

Thông điệp x Ký số v =SigG(u)
Bên gửi G
Bên nhận N Đúng: u ≡ Verkd(v) Sai: u <> Verkd(v)
(u,v)

(u)

Giải mã

x =DN(u)

Kiểm tra

Verkd(u,v)

(u,v)

Ký số

y =SigG(x)

Thông điệp x Mã hóa z =EG(x,y)
Bên gửi G
Bên nhận N Đúng: x ≡ Verkd(y) Sai: x <> Verkd(y)
(z)

(x,y)

Giải mã

(x,y) =DN(z)

Kiểm tra

Verkd(x,y)

(x,y)

Trong trường hợp 2: H lấy được (u,v)

Có hai dạng tấn công:

- Tấn công vào thông điệp x (xem nội dung thông điệp, thay đổi nội dung thông

điệp….):
Với dạng tấn công này H đều phải giải mã thông tin lấy được. Trường hợp 1; giải mã Z để thu
được x; trường hợp 2; giải mã u để thu được x; dạng tấn công vào thông điệp trong cả hai trường
hợp đều là "khó".
- Tấn công vào chữ ký (giả mạo chữ ký): Trường hợp 1: để tấn công vào chữ ký y, H phải
giải mã Z, mới nhận được chữ y để tấn công → "khó". Trường hợp 2: để tấn công vào chữ ký v,
H đã có sẵn v. H thay thế chữ ký v trên u, bằng chữ ký của H là v' = SigH(u), rồi gửi (u, v') đến
N. N kiểm thử thấy sai, gửi lại phản hồi đến cho G, G có thể chứng minh chữ ký đó là giả mạo.
→ Trong trường hợp 2, H có thể giả mạo chữ ký mà không cần giải mã. Vì vậy hãy ký trước, sau
đó mã hóa cả chữ ký và thông điệp.

3.3 Chuẩn chữ ký điện tử - DSS

Chuẩn chữ ký điện tử - DSS (Digital Signature Standard) được sửa đổi từ hệ chữ ký
Elgamal. Được công tại hội nghị tiêu chuẩn xử lý thông tin liên bang (FIPS) 19/05/1994 và trở
thành chuẩn vào 01/12/1994.
Thông thường tài liệu số được mã hóa và giải mã một lần, nhưng ký lại liên quan đến đến
pháp luật, chữ ký có thể phải kiểm thử sau nhiều năm → chữ ký phải được bảo vệ cẩn thận.

Sơ đồ:

ϖ Tạo cặp khóa

- Chọn số nguyên tố lớn **p** sao cho bài toán logarit rời rạc trong Zp là "khó giải".
- Chọn **q** là ước nguyên tố của p-1. Tức là:

p-1 = t*q hay p = t*q +1

(để đảm bảo an toàn: p cỡ 512 bit, q cỡ 160 bit trở lên)

- Chọn g ϵ Z*p là căn bậc q của 1 mod p (g là phần tử sinh của Z*p). Tính α = gt mod p
- Chọn khóa bí mật a ϵ Z*p. Tính khóa công khai h ≡ αa mod p.
- Đặt P = Z*p, A = Z*p x Z*p

K = {(p, q, α, a, h)/ a ϵ Z*p , h ≡ αa mod p }

Với mỗi khóa (p, q, α, a, h) → k' = a (bí mật)

→ k'' = (p, q, α, h) (công khai)

ϖ Ký số:

Dùng 2 khóa ký: Khóa a và khóa ngẫu nhiên bí mật r ϵ Z*p

Chữ ký trên x ϵ Z*p là Sigk'(x, r) = (y1, y2) trong đó:

y1 = (αr mod p) mod q

y2 = (x + a*y1)*r -1 mod q

ϖ Kiểm tra chữ ký: (x, y1, y2)

Tính e1 = x * y2-1 mod q

e2 = y1 * y2-1 mod q

Verk''(x, y1, y2) = đúng ⎾ (αe1 * he2 mod p) mod q = y1

VD:

- Chọn p =7649.
- chọn q = 239 là ước nguyên tố của p -1, t = 32

p = q*t + 1 = 32 * 239 + 1 = 7649

(trong thực thế thường xác định q trước sau đó tính p = q * t + 1)

- chọn phần tử sinh g = 3 ϵ Z*7649. Tính α = gt mod p = 332 mod 7649 = 7098

- Chọn khóa bí mật a = 85, tính h = αa mod p = 709885 mod 7649 = 5387

→ khóa ký (bí mật): k’ = (a) = 85

→ khóa kiểm tra chữ ký: (công khai) k’’ = (p, q, α, h) = (7649, 239, 7098, 5387)

ϖ Ký số:

Dùng 2 khóa ký a và số ngẫu nhiên r = 58 ϵ Z*7649. r -1 mod q = 136.

Tài liệu cần ký: x = 1246.

Xác định chữ ký x như sau: Sigk’ (x, r) = (y1, y2)

y1 = (αr mod p) mod q = (709858 mod 7649) mod 239 = 115

y2 = (x + a*y1)*r -1 mod q = (1246 + 85*115) * 136 mod 239 = 87

→ Sigk’ (x, r) = (y1, y2) = (115, 87)

ϖ Kiểm tra chữ ký:

y2-1 mod q = 87-1 mod 239 = 11

e1 = x * y2-1 mod q = 1246 * 11 mod 239 = 83

e2 = y1 * y2-1 mod q = 115 * 11 mod 239 = 70

kiểm tra:

(αe1 * he2 mod p) mod q = (709883 * 538770 mod 7649) mod 239 = 115 = y1

3.4 Đại diện thông điệp và hàm băm

3.4.1 Đại diện thông điệp

Khi thực hiện chữ ký điện tử, tồn tại một số vấn đề:

→ “Chữ ký điện tử” thực hiện trên từng bit tài liệu, nên độ dài của chữ ký điện tử “ít nhất” cũng

bằng độ dài của tài liệu. Trong khi đó trên thực tế, ta cần phải ký vào các bản tin có kích thước

lớn → phải tốn nhiều bộ nhớ để lưu trữ "chữ ký", và tốn nhiều thời gian để lưu trữ trữ ký trên
mạng.
→ Với một số chữ ký "an toàn" (DSS), tốc độ ký và kiểm tra chữ ký lại chậm vì dùng nhiều
phép tính số học phức tạp.
→ Thực tế xảy ra trường hợp: Với nhiều bản tin đầu vào khác nhau sử dụng sơ đồ chữ ký giống
nhau (có thể khác nhau), nhưng lại cho ra chữ ký giống nhau, làm cho việc xác thực thông tin trở
nên phức tạp.
Giải quyết vấn đề trên như thế nào?
Cách 1: Với thông điệp có kích thước lớn ta "chặt" bản tin thành nhiều đoạn nhỏ. Sau đó ký lên
các đoạn đó độc lập nhau.
Cách 2: Thay vì phải ký lên tài liệu, dùng "hàm băm" để tạo đại diện cho tài liệu, sau đó ký lên
đại diện này.
- Đại diện của tài liệu chính là giá trị của hàm băm trên tài liệu đó, nó còn được gọi là "Tóm
lược" hay "bản thu gọn" của tài liệu.
- Với mỗi tài liệu, qua hàm băm chỉ có thể tính ra một đại diện – giá trị băm tương ứng – duy
nhất. Đại diện của tài liệu được xem như là đặc thù của tài liệu, giống như dấu vân tay của
con người.

Hình 3-4: ***Sơ đồ kết hợp hàm băm trong sử dụng chữ ký điện tử***

3.4.2 Hàm băm

Định nghĩa: Một hàm băm H sẽ lấy ở đầu vào một thông tin x có kích thước biến thiên và sinh

kết quả là một chuỗi có độ dài cố định, được gọi là cốt của bức điện.

Vd: MD5( ): 7215ee9c7d9dc229d2921a40e899ec5f (hexa)

MD5(c): 4a8a08f09d37b73795649038408b5f33(hexa)

MD5(cong): 3de6a8f9608ddd4ba89f97b36d7587d6 (hexa)

MD5(congnghethongtin): 75e77803c24aba6ce2d1a57e6170219a (hexa)

ϖ Đặc tính của hàm băm:

¬ Hàm băm là một hàm một chiều, theo nghĩa giá trị của hàm băm là duy nhất và từ giá trị

này khó có thể suy ngược lại nội dung hay độ dài ban đầu của tài liệu gốc.

¬ Với tài liệu đầu vào (bản tin gốc) x, chỉ thu được giá trị băm duy nhất z = h(x).

¬ Nếu dữ liệu trong bản tin x bị thay đổi dù chỉ một bit để trở thành bản tin x', thì giá trị

băm h(x') ≠ h(x) → hai thông điệp khác nhau thì giá trị băm cũng khác nhau.

Vd: MD5(congnghethongtin): 75e77803c24aba6ce2d1a57e6170219a (hexa)

MD5(congnghethongton): 463c3274f8daa23a4671bc56c34fd0d8 (hexa)

MD5(congnghethongtiN): ad6060fdc5c0cb2739c518ea9859f174 (hexa)

ϖ Ứng dụng của hàm băm:

¬ Ứng dụng trong chữ ký điện tử. Giảm chi phí về thời gian "ký", bộ nhớ lưu trữ chữ ký,

thời gian truyền chữ ký trên mạng (chữ ký trên đại diện tài liệu sẽ nhỏ hơn rất nhiều so

với trên tài liệu gốc).

¬ Hàm băm dùng để xác định tính toàn vẹn của dữ liệu.

¬ Hàm băm dùng để bảo vệ một số dữ liệu đặc biệt như mật khẩu.

ϖ ***Đụng độ:***

Rõ ràng là với không gian giá trị băm nhỏ hơn không gian tin về mặt kích thước thì chắc

chắn sẽ tồn tại đụng độ (collision), nghĩa là có hai bản rõ X≠X' mà giá trị băm của chúng giống

nhau nghĩa là h(X)=h(X'). *Điều này có thể thấy rõ ràng qua nguyên lý Diricle - Nếu có n+1 con*

thỏ được thả vào n cái chuồng thì phải tồn tại ít nhất một cái chuồng mà trong đó có ít ra là hai

con thỏ ở chung.

Việc chế tạo các hàm băm phi đụng độ là rất khó. Nhiều hàm băm được phát minh bởi

các nhóm có tên tuổi trên thế giới sau một thời gian xuất hiện đã bị những người khác chỉ ra

những đụng độ tồn tại và không được công nhận là an toàn nữa.

3.4.3 Một số hàm băm thông dụng

ϖ **Hàm băm MD:** do Ronald Rivest phát triển, bao gồm MD2, MD4(1990), MD5 (1991).

MD5 là một trong những hàm băm được sử dụng rộng rãi. MD5 có thể xử lý các khối

thông tin có độ dài không giới hạn để tạo ra mã băm có chiều dài 128 bit.

Với chiều dài 128 bit, việc tìm ra hai khối thông tin để có cùng một giá trị băm không

còn là điều bất khả thi. Do đó, độ an toàn của MD5 đang bị đe dọa nghiêm trọng, và thời
gian tới mức độ phổ biến của MD5 có thể sẽ giảm đi.

ϖ **Hàm băm SHA (Secure Hash Function):** Năm 1995, Viện công nghệ và tiêu chuẩn
quốc gia Hoa Kỳ thiết kế ra thuật toán băm băm an toàn (SHA) sử dụng cho chuẩn chữ
ký điện tử DSS.

SHA được thiết kế dựa trên những nguyên tắc của MD4/MD5, tạo ra 160 bit gí trị băm
(SHA-1). Về sau có nhiều nâng cấp đối với SHA, chủ yếu là tăng chiều dài mã băm, từ
đó xuất hiện các phiên bản khác nhau của SHA, bao gồm:

SHA-2: SHA-256 (Mã băm có chiều dài 256 bit)

SHA-384 (Mã băm có chiều dài 384 bit)

SHA-512 (Mã băm có chiều dài 512 bit)

ϖ **RIPEMD:** 128/160/256/320 bit

ϖ **Whirlpol:** 512 bit

ϖ **Tiger:** 128/160/192 bit

VD: congnghephanmem

MD5: 2544c39a539d8013b058e15011ed6331 (128bit - hexa)

SHA-1: b2448e58c8dfd1843d21e4151df9848d8db18619 (160bit-hexa)

RIPEMD(160): 8e982abee503288447063728b648dc3fb6c5224e (160bit-hexa)

Whirlpol:

a38ed5634769c9b89675d836a63a9dfe310ed6bd60ca137297d11fe45d384cdcc7e84
3898979789a

0f2769694bd65602c02c1d44f3e0a0832ce140757bd8bbd6 (512bit-hexa)

Tiger 128: e9fefadf50976b2aaac816c98e97813e0697df36d38f033e (128bit-hexa)

CHƯƠNG 4: ẨN – GIẤU TIN

4.1 Khái niệm và phân loại

Ẩn – giấu tin là một kỹ thuật nhúng các mẩu tin vào một đối tượng mang tin sao cho khó
có thể phát hiện ra mẩu tin được nhúng.

Một trong những yêu cầu cơ bản của ẩn - giấu tin là đảm bảo tính chất ẩn của thông tin
được giấu, đồng thời không làm ảnh hưởng đến chất lượng của dữ liệu gốc.

Sự khác biệt giữa mã hóa và giấu tin?

Mã hóa: làm thay đổi hình dạng của thông tin ban đầu, thông tin hiện rõ là nó có được mã hóa
hay không?

Giấu tin: Khó nhận biết được sự tồn tại của thông tin được giấu trong môi trường chứa tin.

Phân loại ẩn - giấu tin

So sánh giữa giấu tin và thủy vân số:

Giấu tin	**Thủy vân số**
- Mục đích là bảo vệ thông tin được giấu. - Giấu được càng nhiều thông tin càng tốt, ứng dụng trong	- Mục đích là bảo vệ môi trường giấu tin. - Chỉ cần thông tin đầy đủ để đặc trưng cho bản quyền sở

truyền dữ liệu thông tin mật. - Thông tin được giấu phải ẩn, không cho người khác thấy được bằng mắt thường. - Chỉ tiêu quan trọng nhất là dung lượng của tin được giấu	hữu. - Thông tin giấu có thể ẩn (thủy vân ẩn) hoặc hiện (thủy vân hiện). - Chỉ tiêu quan trọng nhất là tính bền vững của tin được giấu.

4.2 Môi trường giấu tin

a. Giấu tin trong Audio

Ẩn – Giấu tin

(Information hiding)

Giấu tin mật

(Steganography)

Thủy vân số

(Watermarking)

Kỹ thuật giấu tin trong audio phụ thuộc vào hệ thống thính giác của con người (HAS –

Human Auditory System). HAS cảm nhận được tín hiệu ở dải tần rộng và công suất thay đổi lớn,

nhưng lại kém trong việc phát hiện sự khác biệt nhỏ giữa dải tần và công suất. Điều này có nghĩa

là các âm thanh to, cao tần có thể che giấu được các âm thanh nhỏ thấp một cách dễ dàng. Kênh

truyền tin cũng là một vấn đề. Kênh truyền hay băng thông chậm sẽ ảnh hưởng đến chất lượng
thông tin sau khi giấu. Giấu thông tin trong audio yêu cầu rất cao về tính đồng bộ và tính an toàn
của thông tin.

b. Giấu tin trong video

Giấu tin trong video cũng được quan tâm và được phát triển mạnh mẽ cho nhiều ứng
dụng như điều khiển truy cập thông tin, xác thực thông tin và bảo vệ bản quyền tác giả. Ta có thể
lấy một ví dụ là hệ thống chương trình trả tiền xem theo video clip của các thuật toán trước đây
thường cho phép giấu ảnh vào trong video, nhưng gần đây kỹ thuật cho phép giấu các âm thanh
và ảnh vào trong video.

c. Giấu tin trong ảnh

Giấu thông tin trong ảnh hiện nay chiếm tỷ lệ lớn nhất trong các chương trình ứng dụng,
các phần mềm, và hệ thống giấu tin trong đa phương tiện bởi lượng thông tin trao đổi bằng ảnh
là rất lớn. Hơn nữa giấu thông tin trong ảnh cũng đóng vai trò hết sức quan trọng đối với hầu hết
các ứng dụng bảo vệ an toàn thông tin như: xác thực thông tin, xác định xuyên tạc thông tin, bảo
vệ bản quyền tác giả, điều khiển truy cập, giấu thông tin mật... Chính vì thế mà vấn đề này đã
nhận được sự quan tâm rất lớn của các cá nhân, tổ chức, trường đại học, và các

viện nghiên cứu
trên thế giới.
Thông tin sẽ được giấu cùng với dữ liệu ảnh nhưng chất lượng ảnh ít thay đổi và chẳng ai
biết được đằng sau ảnh đó mang những thông tin có ý nghĩa. Ngày nay, khi ảnh số đã được sử
dụng rất phổ biến, thì giấu thông tin trong ảnh đã đem lại rất nhiều những lợi ích quan trọng trên
nhiều lĩnh vực của đời sống xã hội.

4.3 Mô hình giấu tin cơ bản

Giấu tin vào các phương tiện chứa và tách lấy thông tin là hai quá trình trái ngược nhau:

a. Giấu tin vào phương tiện chứa

Quy trình giấu thông tin vào phương tiện chứa được mô tả qua sơ đồ khối dưới đây trong đó:

- Thông tin cần giấu tùy theo mục đích của người sử dụng, nó có thể là thông điệp (với các
tin bí mật) hay là các logo, hình ảnh bản quyền.
- Phương tiện chứa: các file ảnh, text, audio,... là môi trường để nhúng tin.
- Bộ nhúng thông tin: là những chương trình thực hiện việc giấu tin
- Đầu ra: là các phương tiện chứa đã có tin giấu trong đó.

b. Tách thông tin

Tách thông tin từ các phương tiện chứa diễn ra theo quy trình ngược ra là các thông tin đã được
giấu vào phương tiện chứa. Phương tiện chứa sau khi tách lấy thông tin có thể được sử dụng,

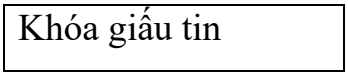
quản lý theo những yêu cầu khác nhau.

Sau khi nhận được đối tượng phương tiện chứa có giấu thông tin, quá trình giải mã được

thực hiện thông qua một bộ giải mã tương ứng với bộ nhúng thông tin cùng với khóa của quá

trình nhúng. Kết quả thu được gồm phương tiện chứa gốc và thông tin đã giấu. Bước tiếp theo

thông tin đã giấu sẽ được xử lý kiểm định so sánh với thông tin ban đầu.



Môi trường chứa

Bộmã giải (audio, ảnh, video)

Môi trường

chứa đã được

giấu tin

Thông tin giấu Kiểm

định

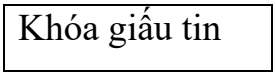
Thông tin giấu

Bộ nhúng

thông tin

Môi trường chứa

(audio, video, ảnh)



Môi trường

chứa đã được

giấu tin

Phân

phối

4.4 Tính chất của ẩn giấu tin

Để đánh giá chất lượng của một phương pháp ẩn giấu tin, người ta dựa vào một số tiêu chí sau:

a. Bảo đảm tính vô hình

Ẩn giấu tin sẽ làm biến đổi môi trường mang tin. Tính vô hình thể hiện mức độ biến đổi
của môi trường mang tin. Phương pháp nào ẩn giấu tin tốt, sẽ làm cho thông tin mật trở nên vô
hình trên môi trường mang tin, người dùng khó có thể nhận ra trong môi trường mang tin có ẩn
chứa thông tin mật được giấu.

Chú ý rằng với ẩn tin thì trong thực tế không phải khi nào cũng cố gắng để đạt được tính vô hình
cao nhất, ví dụ trong truyền hình, người ta gắn hình ảnh mờ gọi là thủy ấn để bảo vệ bản quyền
bản tin.

b. Khả năng chống giả mạo

Mục đích của giấu tin là truyền đi thông tin mật. Nếu không thể do thám tin mật, thì kẻ
địch cũng cố tìm cách làm sai lệch tin mật, làm giả mạo tin mật để gây bất lợi cho đối phương.

Phương pháp giấu tin tốt phải đảm bảo tin mật không bị tấn công một cách chủ động trên cơ sở
những hiểu biết thuật toán nhúng tin và có vật mang tin (nhưng không biết khóa giấu tin).

Đối với ẩn tin thì khả năng chống giả mạo là yêu cầu vô cùng quan trọng, vì có

như vậy mới bảo
vệ được bản quyền, minh chứng tính pháp lý của sản phẩm.

c. Dung lượng giấu

Dung lượng giấu được tính bằng tỷ lệ của lượng thông tin cần giấu so với kích thước của
vật mang tin. Các phương pháp đều cố gắng giấu được nhiều tin nhưng vẫn giữ được bí mật. Tuy
nhiên trong thực tế người ta luôn phải cân nhắc giữa dung lượng và các tiêu chí khác như tính vô
hình, tính bền vững.

d. Tính bền vững

Sau khi ẩn giấu tin vào môi trường mang tin, bản thân môi trường mang tin có thể phải
qua các biến đổi khác nhau, tính bền vững là thước đo sự nguyên vẹn của tin mật sau những biến
đổi như vậy.

e. Độ phức tạp tính toán

Độ phức tạp của thuật toán ẩn giấu tin và giải tin (tách tin) cũng là một chỉ tiêu quan
trọng để đánh giá một phương pháp ẩn giấu tin. Chỉ tiêu này cho chúng ta biết tài nguyên (thời
gian và bộ nhớ) tốn bao nhiêu dùng cho một phương pháp ẩn giấu tin. Với chủ nhân ẩn giấu tin
thì thời gian thực hiện phải nhanh, nhưng với kẻ thám tin phải là bài toán khó. Ví dụ bài toán
tách tin từ thủy vân để đánh dấu bản quyền cần phải là bài toán khó, thì mới chịu được sự tấn

công của tin tặc nhằm phá hủy thủy vân.

4.5 Ví dụ xây dựng ứng dụng thủy vân trên ảnh số

(Trích kết quả đề tài nghiên cứu khoa học cấp cơ sở mã số T14-15)

4.5.1 Các thuật toán thủy vân trong ảnh số

Các thuật toán thực hiện thủy vân trên ảnh đều tương đối dễ dàng, và đã được nghiên cứu
trong môn xử lý ảnh số. Các yêu cầu về tính thẩm mỹ được đề cao. Tất nhiên cũng cần có
yêu cầu là khó sử dụng các công cụ xử lý ảnh để loại bỏ thủy vân. Các thuật toán thủy
vân trên ảnh có thể chia ra làm hai loại chính, đó là thủy vân dựa vào biến đổi miền không gian ảnh, và biến đổi miền tần số. Một số thuật toán như:

- Thuật toán giấu thủy vân vào các bít có trọng số thấp.
- Thuật toán thủy vân ghép nổi.
- Thuật toán thủy vân ảnh trên miền DCT.
- Thuật toán thủy vân trên miền DWT.
- Thuật toán thủy vân sử dụng biến đổi Karhunen-Loeve.
- …

Chương trình thủy vân số được xây dựng trên cơ sở thuật toán giấu thủy vân vào các
bit có trọng số thấp, thuật toán được mô tả như sau:

Một ý tưởng tự nhiên của thủy vân với ảnh số, cũng giống như giấu tin, đó là sẽ sử dụng
các bit có trọng số thấp (Least Significan Bit – LSB) để giấu thủy vân.

Các bit có trọng số thấp được hiểu là các bit mà nếu thay đổi giá trị của chúng thì

sẽ ít làm thay đổi đến chất lượng ảnh. Ví dụ, với ảnh bitmap 256 màu, màu của mỗi điểm
ảnh được biểu diễn bởi 8 bit, nếu ta thay đổi giá trị bit thứ 8 của mã màu, thì mã màu
cũng chỉ thay đổi giá trị có 1 đơn vị, nên nhìn chung thì các bức ảnh không bị ảnh hưởng
nhiều.

Ta có thể minh họa thuật toán như sau:

Xét thủy vân chuỗi bit là 0111.

Xét bức ảnh là chuỗi bit: 11001101 11000001 11110000 11110010.

Để nhúng thủy vân vào bức ảnh, ta sẽ chia bức ảnh thành các khối 8 bit, và đặt giá trị bit
cuối cùng của khối bằng giá trị của bit thủy vân tương ứng.

Với minh họa trên, chúng ta có bức ảnh sau khi nhúng thủy vân là:

1100110**0** 1100000**1** 1111000**1** 1111001**1**

Để tách thủy vân, đơn giản ta chỉ làm ngược lại quy trình trên, tức là tách ra các bit cuối của từng khối 8 bit, ta sẽ thu được thủy vân ban đầu.

Muốn tăng tính an toàn của hệ thống, có thể nhúng liên tiếp thủy vân vào các khối 8 bit
liền nhau, bởi thường thì dung lượng bức ảnh sẽ lớn hơn nhiều lần so với độ dài của thủy
vân. Ưu điểm của thuật toán trên là đơn giản, và dung lượng giấu cao. Tuy nhiên, nhược
điểm là do quá đơn giản nên rất dễ bị tấn công. Kẻ tấn công chỉ cần thay đổi ngẫu nhiên
giá trị của các bit có trọng số thấp là thủy vân đã bị phá hủy.

4.5.2 Xây dựng chương trình ứng dụng

Chương trình xây dựng trên môi trường window, sử dụng ngôn ngữ VB.Net. Dưới đây là
kết quả chạy chương trình với các chức năng chính.

Hình 4.1: Giao diện chính của chương trình

a) Chức năng nhúng/tách thông tin bản quyền là một đoạn thông điệp:

Chức năng này thực hiện việc nhúng một đoạn thông điệp vào file ảnh, đoạn thông điệp sẽ có tác dụng tạo bản quyền của người dùng. Để thực hiện chức năng này yêu cầu
xác định file ảnh cần nhúng thông điệp bản quyền thông qua nút chọn file, ảnh ban đầu sẽ
được hiển thị trong phần ảnh gốc. Tiếp sau người dùng phải nhập đoạn thông điệp muốn
nhúng vào ảnh, xác định đường dẫn lưu ảnh chứa thông điệp và thiết lập mật khẩu bảo vệ
(Mật khẩu bảo vệ có chiều dài lớn hơn 5 ký tự). Quá trình thực hiện theo chức năng này
theo giao diện như hình dưới đây:

Hình 4.2: Chạy chương trình với chức năng tạo bản quyền bằng đoạn thông điệp .

Chương trình được chạy với ảnh gốc có tên: NCKH.bmp, độ phân giải 1500 x
1125 pixel. Bản quyền nhúng vào file ảnh là một đoạn thông điệp có nội dung: “Bản
quyền thuộc Khoa CNTT Tác giả: Đặng Văn Nam Ngày: 09/05/2014”. Sau khi chạy
chức năng này thu được một ảnh mới có tên: NKH_BQ_Mes.bmp có độ phân giải tương

tự như NCKH.bmp. Bức ảnh trước và sau khi nhúng thông điệp bản quyền hầu như không thay đổi. Nếu bằng mắt thường người dùng khó có thể phân biệt được ảnh nào là
ảnh có chứa thông điệp bản quyền.

Hình 4.3: Ảnh gốc NCKH.bmp (1500x1125 pixel)

Hình 4.4: Ảnh sau khi nhúng thông tin bản quyền NCKH_BQ_Mes.bmp

Tách thông tin để kiểm tra bản quyền của ảnh: Chương trình xây dựng chức năng cho
phép người dùng kiểm tra một ảnh có chứa thông điệp hay không? Nếu có thông điệp sẽ
tách và hiển thị thông điệp ra để kiểm tra.

Quy trình tách thông điệp được thực hiện như sau: Người dùng chọn tới file ảnh cần kiểm tra thông điệp bản quyền. File ảnh sau khi được chọn sẽ hiển thị trong phần
"Ảnh kiểm tra bản quyền". Tiếp đó người dùng phải nhập mật khẩu bảo vệ để xác thực
đúng đối tượng. Trong trường hợp file ảnh không chứa thông điệp bản quyền hoặc mật
khẩu bảo vệ nhập vào không đúng, chương trình sẽ kiểm tra và hiển thị thông báo như
sau:

Ngược lại, nếu File ảnh có chứa thông điệp và mật khẩu người dùng nhập vào là chính
xác, chương trình sẽ thực hiện việc tách thông điệp ra khỏi file ảnh và hiển thị nội dung

của thông điệp để người dùng xác thực bản quyền. Kết quả chạy được minh họa như hình
dưới đây:

Hình 4.5: Kết quả chạy chức năng tách bản quyền là một đoạn thông điệp

b) Chức năng nhúng/tách thông tin bản quyền là một file văn bản.

Chương trình hỗ trợ người dùng sử dụng một file văn bản (dạng *.txt) chứa các thông tin về bản quyền và nhúng nội dung của file văn bản này vào ảnh gốc. Để thực hiện
chức năng này yêu cầu người dùng xác định file ảnh cần đánh dấu bản quyền; Sau khi lựa
chọn, file ảnh sẽ được hiển thị trong phần "ảnh gốc". Tiếp theo, cần xác định file văn bản
(có dạng .txt) cần nhúng, sau khi xác định nội dung của file văn bản sẽ được đọc và hiển
thị trong phần "Nội dung file text". Người dùng xác định tên, vị trí lưu ảnh sau khi nhúng
và thiết lập mật khẩu bảo vệ.

Trong ví dụ dưới đây, sử dụng file văn bản có tên file_Bquyen.txt kích thước 157 byte,
nội dung của file văn bản này được trình bày như hình dưới đây.

Tương tự như thực hiện với chức năng trên ảnh gốc sử dụng là NCKH.bmp và kết quả sau khi nhúng file văn bản có tên NCKH_BQ_Text.bmp hầu như không có sự khác
biệt khi quan sát bằng mắt thường.

Hình 4.6: Chạy chương trình với chức năng tạo bản quyền bằng file văn bản

Hình 4.7: Nội dung file văn bản file_Bquyen.txt

Quá trình tách file văn bản này được thực hiện ngược lại:

Hình

4.8: Chạy chương trình với chức năng tách file văn bản

Người dùng xác định file ảnh có chứa thông tin bản quyền cần tách. Sau khi xác định, file ảnh chứa bản quyền sẽ được hiển thị trong phần "Ảnh chứa bản quyền". Người

dùng nhập vào mật khẩu bảo vệ và chương trình sẽ thực hiện việc kiểm tra và tách thông

điệp bản quyền hiển thị nội dung của file văn bản đã được lưu. Nếu người dùng muốn lưu

lại nội dung này, chọn vào phần lưu file text để xác định tên, đường dẫn lưu file. Quy

trình tách file văn bản bản quyền được minh họa như hình trên.

c) Chức năng nhúng/tách thông tin bản quyền là một file ảnh:

Chức năng này cho phép người dùng tạo bản quyền ảnh bằng file ảnh (nhúng file ảnh vào file ảnh gốc). Người dùng lựa chọn file ảnh gốc và file ảnh bản quyền, thiết lập

mật khẩu và thực hiện việc tạo bản quyền. file ảnh bản quyền phải đảm bảo kích thước

cho phép. Chương trình sẽ hiển thị ảnh gốc và ảnh bản quyền để người dùng quan sát,

chọn đường dẫn lưu file ảnh chứa bản quyền và thiết lập mật khẩu bảo vệ để thực hiện.

Hình 4.9: chạy chương trình với chức năng tạo bản quyền là file ảnh

Chương trình sẽ kiểm tra kích thước ảnh bản quyền và ảnh gốc nếu không phù hợp sẽ hiển thị thông báo như dưới đây:

Để tách file ảnh bản quyền, người dùng thực hiện tương tự như việc tách file văn bản.

Quy trình được minh họa như hình dưới đây:

72 | Tổng hợp và biên soạn: Đặng Văn Nam

Hình 4.10: Chạy chương trình với chức năng tách bản quyền là file ảnh

73 | Tổng hợp và biên soạn: Đặng Văn Nam

TÀI LIỆU THAM KHẢO

[1]-*An toàn và bảo mật thông tin*, Đại học Công nghệ-Đại học Quốc gia Hà Nội.

[2]-*Giáo trình An toàn và bảo mật thông tin*, Nguyễn Hữu Tuân, Đại học Hàng hải.

[3]-*Giáo trình An toàn và bảo mật thông tin,* Trần Văn Dũng, Đại học Giao thông vận

tải.

[4]-*Cryptography and Network Security: Principles and Practice*, William Stallings.

Prentice Hall, 2005.

[5]-*Cryptography Theory and pracetice*. Douglas R.Stinson. CRC Press, 1995.

74 | Tổng hợp và biên soạn: Đặng Văn Nam

PHỤ LỤC I:

DANH SÁCH CÁC NHÀ CUNG CẤP DỊCH VỤ CHỮ KÝ SỐ TẠI VIỆT NAM

1. Tập đoàn Bưu chính Viễn thông Việt Nam (VNPT-CA) Trụ sở chính: Nhà

Internet, lô II, Làng Quốc tế Thăng Long, quận Cầu Giấy, Hà Nội
Điện thoại: (84-4) 3793 2924/3793 0506; Fax: (84 4) 3793 0506
Website: www.vnpt-ca.vn

2. Công ty cổ phần Công nghệ Thẻ Nacencomm (CA2)
Trụ sở chính: Tầng 5,02 Chùa Bộc, quận Đống Đa, Hà Nội.
Điện thoại: (84-4) 3576 5146; Fax: (84-4) 3576 5147
Website: www.cavn.vn

3. Công ty TNHH Hệ thống Thông tin FPT (FPT.CA)
Trụ sở chính: Tầng 22 tòa nhà Keangnam Landmark Tower - E6 Phạm Hùng, Hà Nội
Điện thoại: +84 4

35626000

Fax: +84 4 35624850

Website: dichvudientu.fpt.com.vn

4. Tập đoàn Viễn thông Quân đội Viettel (VIETTEL-CA)

Trụ sở chính: Tòa nhà N2 Viettel, Km 2 Đại lộ Thăng Long, Mễ Trì, huyện Từ Liêm, Hà Nội

Điện thoại: 1900 9099; Fax: (84-4) 6288 0005

Website: www.viettel-ca.vn

5. Công ty Cổ phần Bkav (BKAV-CA)

Trụ sở chính: Tòa nhà Bkav, Yên Hòa, quận Cầu Giấy, Hà Nội

Điện thoại: (84-4)7305 0060/1900545414; Fax: (84-4) 3782 2135

Website: www.bkavca.vn

6. Công ty Công nghệ

và Truyền thông CK (CK-CA)

Trụ sở chính: 4/122 Định Công, quận Hoàng Mai, Hà Nội

Điện thoại: (84-4) 3664 7888; Fax: (84-4) 3664 7770

Website: www.ckca.vn

7. Công ty cổ phần Truyền thông Newtelecom (Newtel-CA)

Trụ sở chính: 08 Láng Hạ, quận Ba Đình, Hà Nội

Điện thoại: (84-4) 3772 7766; Fax: (84-4) 3772 7755

Website: newtel-ca.vn

8. Công ty cổ phần Chữ ký số Vi Na (Vina-CA)

Trụ sở chính: 41A Nguyễn Phi Khanh, phường Tân Định, quận

1, Tp. HCM.
Điện thoại: (84-8) 3820 2261; Fax: (84-8) 3820 2262
Website: www.smartsign.com.vn

PHỤ LỤC II:

(Trích báo cáo tại Hội nghị khoa học Trường lần thứ 19)

NGHIÊN CỨU CÁC PHƯƠNG PHÁP XÁC THỰC BẰNG CÔNG NGHỆ SINH TRẮC HỌC

Đặng Văn Nam

Khoa CNTT, Trường đại học Mỏ - Địa chất.

Tóm tắt:

Xác thực bằng công nghệ sinh trắc học đang chứng tỏ ưu thế so với các phương pháp xác thực truyền thống bởi mức độ an toàn và độ tin cậy cao, tạo điều kiện thuận lợi cho người sử dụng. Các phương pháp xác thực phổ biến hiện nay bằng công nghệ này có thể chỉ ra là: xác thực bằng vân tay(Fingerprint), giọng nói (Voice), khuôn mặt (Face), hình học bàn tay (Hand), mống mắt (Iris), ADN…Xác thực bằng phương pháp sinh trắc học đang được ứng dụng rộng rãi trong nhiều lĩnh vực từ các lĩnh vực yêu cầu độ chính xác và bảo mật cao như trong quân đội,

chính phủ điện tử, giao dịch ngân hàng tới các lĩnh vực thường ngày như an ninh doanh nghiệp,
phòng chống tội phạm-khủng bố, trong các dịch vụ xã hội…
Tại Việt Nam, Các ứng dụng xác thực bằng công nghệ sinh trắc học vẫn còn khá mới mẻ
và chưa phổ biến. Bài báo này sẽ giới thiệu lý do tại sao cần phải phát triển các hệ thống xác
thực sinh trắc học, cũng như các phương pháp xác thực bằng công nghệ sinh trắc học đang
được sử dụng rộng rãi hiện nay.

1. Lý do phát triển hệ thống xác thực sinh trắc học.

Trong thực tế ngày nay yêu cầu xác thực thông tin là rất phổ biến và ngày càng trở nên
quan trọng. Việc xác thực thông tin sẽ cho phép những người có quyền có thể truy cập vào
những dữ liệu, cũng như thông tin được bảo mật. Ví dụ đơn giản nhất có thể thấy đó là việc đặt
các hệ thống bảo vệ cho máy tính, yêu cầu người dùng phải xác thực thông tin hợp lệ mới được
quyền sử dụng, hay các tài khoản ngân hàng yêu cầu phải xác thực đúng thông tin của khách
hàng trước khi có thể thực hiện được các giao dịch…

Các phương pháp xác thực chính hiện nay mà chúng ta đang sử dụng đó là xác thực
thông qua mật mã (sử dụng User name và Password để đăng nhập vào hệ thống thư điện tử, blog,
tài khoản trên mạng, đăng nhập vào hệ thống máy tính, hệ thống mạng…), xác

thực thông qua
Tokens (các thẻ thông minh - Smart cards hoặc thẻ đeo của nhân viên - ID badges chứa các
thông tin xác thực), giấy tờ xác minh (giấy chứng minh thư, hộ chiếu…), thiết bị vật lý (chìa
khóa)…
Tuy nhiên các phương pháp xác thực này hiện đang bộc lỗ những khó khăn cho người
dùng và khả năng bảo đảm an toàn dữ liệu. Việc sử dụng User Name và Password đang là
phương pháp xác thực phổ biến. Tuy nhiên, với phương pháp này độ bảo mật không cao, do các
thông tin về User name và Password có thể bị đánh cắp trên đường truyền, hơn nữa việc phải
nhớ quá nhiều tài khoản đăng nhập cũng là một khó khăn lớn. Việc sử dụng Tokens thì cũng
mang lại những rắc rối nhất định đó là đi đâu chúng ta cũng phải mang nó theo và nếu có làm
mất thì cũng gây phiền hà không nhỏ. Thông tin tài khoản cá nhân dùng để xác thực và Tokens
có thể bị trộm cắp, lạm dụng hoặc làm giả mạo, các loại giấy tờ tuỳ thân, thẻ ngân hàng, hộ chiếu
cũng vậy....Tất cả những khó khăn của các phương pháp xác thực truyền thống đã thúc đẩy các
nhà khoa học nghiên cứu để đưa ra một phương pháp xác thực mới, đảm bảo đủ độ an toàn, đạt
được mức độ tin cậy cần thiết và tiện dụng cho người sử dụng và đó là các phương

pháp xác thực
bằng công nghệ sinh trắc học. Vậy công nghệ sinh trắc học là gì?

Công nghệ Sinh trắc học (Biometric) là một công nghệ sử dụng những thuộc tính sinh lý
hoặc các mẫu hành vi, các đặc điểm sinh học đặc trưng như dấu vân tay, mẫu mống mắt,
giọng nói, khuôn mặt, dáng đi,... để nhận diện con người.

Hay nói cách khác Sinh trắc học là phương pháp tự động nhận biết một người dựa trên một đặc tính sinh lý hoặc hành vi. Đó là các đặc tính có thể đo được như khuôn mặt, dấu vân tay, chữ viết tay, mống mắt, võng mạc, tĩnh mạc và giọng nói. Công nghệ sinh trắc học đang trở thành nền tảng trong việc xác thực đảm bảo độ an toàn cao.

Trong khi khả năng bảo mật bảo mật còn nhiều hạn chế,
cùng với mức độ gian lận trong việc xác thực ngày một
tăng cao thì xác thực bằng công nghệ sinh trắc học càng
trở nên cấp thiết. [1]

Các giải pháp dựa trên công nghệ sinh trắc học có thể cung cấp cho các giao dịch tài chính
và dữ liệu cá nhân. Các công nghệ sinh trắc học có thể được sử dụng trong quân đội, trong các
ứng dụng thương mại, an ninh doanh nghiệp, chính phủ điện tử, ngân hàng điện tử, trong đầu tư
và các giao dịch tài chính khác, trong thực thi pháp luật, y tế và trong các dịch vụ xã hội…

2. Giới thiệu một số phương pháp xác thực bằng công nghệ sinh trắc học

Xác thực bằng công nghệ sinh trắc học đã và đang được ứng dụng, cũng như nghiên cứu
hoàn thiện. Nhiều phương pháp đã được sử dụng và đem lại những kết quả rất khả quan. Các
phương pháp xác thực bằng công nghệ sinh trắc học ngày càng chứng tỏ vai trò và hiệu quả của
nó so với các phương pháp xác thực thông thường.

Hiện nay có nhiều phương pháp xác thực bằng công nghệ sinh trắc học. Hình 2.1

minh họa
các phương pháp sinh trắc học đang được sử dụng và nghiên cứu. Xác thực bằng công nghệ sinh
trắc học có thể chia làm hai loại:
+ Các phương pháp xác thực dựa vào đặc điểm sinh lý của con người như: Xác thực dựa vào
khuôn mặt, dấu vân tay, bàn tay, mống mắt, ADN
+ Các phương pháp xác thực dựa vào hành vi như: Xác thực dựa vào sự gõ bàn phím, chữ
ký, giọng nói.
Trong khuôn khổ báo cáo này tác giả sẽ tập trung trình bày các phương pháp xác thực dựa
vào đặc điểm sinh lý con người như: Xác thực dựa vào khuôn mặt, dấu vân tay, hình học bàn tay,
mống mắt, ADN.

Hình II-1: Các phương pháp sinh trắc học

2.1 Xác thực bằng khuôn mặt (Face recognition technology)

Xác thực bằng nhận dạng khuôn mặt [2] là công nghệ sinh trắc học ít xâm phạm vào cơ thể
con người nhất và tương đối nhanh. Nó làm việc dựa vào thông số của một bộ phận cơ thể người
đó là gương mặt.

Thay vì phải yêu cầu người dùng đặt ngón tay của họ lên một thiết bị chuyên dụng như đối
với công nghệ xác thực bằng vân tay hay phải định vị chính xác đôi mắt ở phía trước của một

máy quét như đối với công nghệ xác thực bằng mống mắt thì hệ thống nhận dạng bằng khuôn
mặt sẽ âm thầm chụp ảnh khuôn mặt của một người khi họ bước vào một khu vực xác định.
Phương pháp xác thực này không có sự xâm phạm đến cơ thể và không làm gián đoạn công việc
của họ, và trong nhiều trường hợp những người này hoàn toàn không nhận thức được quá trình
nhận dạng đang diễn ra. Chính điều này tạo cho chúng ta không có cảm giác như mình đang bị
"theo dõi" hay sự riêng tư của mình bị ảnh hưởng.
Nhận dạng khuôn mặt sẽ phân tích những đặc trưng của các ảnh về khuôn mặt của một
người được nhập vào thông qua một máy quay video kỹ thuật số. Nó sẽ thực hiện việc đo cấu
trúc tổng thể khuôn mặt, trong đó có khoảng cách giữa mắt, mũi, miệng và quai hàm. Những
phép đo này được lưu trữ trong cơ sở dữ liệu và được sử dụng để so sánh khi người dùng đứng
trước camera. Phương pháp sinh trắc học đã được sử dụng rộng rãi, tuy nhiên độ chính xác trong
việc xác thực bằng phương pháp này còn cần phải xem xét một cách cẩn thận hơn.
Mỗi một khuôn mặt có nhiều điểm đặc trưng, các đỉnh lồi lên và những vùng trũng xuống,
tất cả những điểm đó tạo nên đặc điểm khuôn mặt của một người duy nhất. Trên mỗi khuôn mặt
của con người có khoảng 80 điểm nút. Những điểm nút này được đo bởi công nghệ

nhận dạng
khuôn mặt bao gồm các đặc tính như khoảng cách giữa đôi mắt, chiều rộng của mũi, chiều sâu
của những lỗ mắt, hình dạng của xương gò má và chiều dài đường quai hàm.

Hình II-2: Minh họa các điểm nút trên khuôn mặt

Công nghệ nhận dạng khuôn mặt làm việc như thế nào?

4 giai đoạn của quá trình sau đây sẽ minh họa cách vận hành các hệ thống sinh trắc học nhận
dạng thông qua khuôn mặt:

- Lấy mẫu (Capture): Các mẫu vật lý hay hành vi được chụp bởi hệ thống trong quá trình
kết nạp vào tập dữ liệu mẫu của đối tượng.

- Chọn lọc (Extraction): Dữ liệu duy nhất được chiết xuất từ tập dữ liệu mẫu đưa vào và
một mẫu tương ứng với một người được tạo ra.

- So sánh (Comparison): So sánh mẫu mới đưa vào với tập dữ liệu mẫu duy nhất đã được
chọn lọc và lưu trữ trong cơ sở dữ liệu.

- Phù hợp (Matching): Trên cơ sở các kết quả so sánh với các mẫu, hệ thống sẽ quyết
định xem đối tượng có phù hợp hay không?

Với hệ thống này, cần có các máy quay để chụp khuôn mặt. Căn cứ vào đó hệ thống sẽ xác
định vị trí của khuôn mặt và thực hiện việc kiểm tra tính hợp lệ với tập dữ liệu mẫu lưu trữ trong
cơ sở dữ liệu. Hệ thống nhận dạng thông qua khuôn mặt có thể đưa ra một quyết định trong vòng

chưa đầy 5 giây. Các hệ thống nhận dạng sinh trắc học khác nói chung cũng làm việc theo 4 giai
đoạn trên. Do đó, các phần tiếp theo khi trình bày về các phương pháp xác thực sinh trắc học
khác sẽ không đề cập đến các giai đoạn này.

Ứng dụng:

Hiện nay các hệ thống nhận dạng thông qua khuôn mặt được sử dụng như một công cụ
tiềm năng cho việc ngăn chặn tội phạm khủng bố, hệ thống nhận dạng khuôn mặt cũng đang
được sử dụng trong nhiều khu vực yêu cầu mức độ an ninh cao. Các phần mềm đã được phát

triển cho mạng máy tính và trong các ngân hàng để có thể tự động nhận dạng khuôn mặt cho các
mục đích xác minh người sử dụng

Đánh giá:

Một trong những mặt tích cực nhất của hệ thống nhận dạng khuôn mặt mà ta có thể kể ra đó
là nó không xâm phạm đến đối tượng được nhận dạng. Việc xác thực có thể được thực hiện từ
một khoảng cách khá xa, hơn nữa nó không đòi hỏi người dùng phải chờ một khoảng thời gian
dài hoặc làm bất cứ điều gì nhiều hơn việc nhìn vào máy quay.

Tuy nhiên, các hệ thống nhận dạng khuôn mặt cũng còn tồn tại một số khó khăn trong quá
trình xác thực, điều này làm cho kết quả của nó có thể sai lệch. Chẳng hạn như việc

xác định các
thông số như khoảng cách giữa đôi mắt, chiều rộng của mũi, chiều sâu của những lỗ mắt, hình
dạng của xương gò má và chiều dài đường quai hàm…sẽ gặp bất lợi khi khuôn mặt của đối
tượng bị che đi bởi mũ, kính mắt, râu hoặc trang điểm…

Sử dụng:

Qua các phần trình bày ở trên, ta có thể thấy các hệ thống nhận dạng khuôn mặt lý tưởng để
áp dụng tại các khu vực có mật độ giao thông cao, các nơi công cộng…

2.2 Xác thực bằng dấu vân tay (Fingerprint recognition technology)

Dấu vân tay thực sự là rất quan trọng. Những năm 1998, các sản phẩm sử dụng công nghệ
nhận dạng sinh học dựa trên dấu vân tay đã chiếm tới 78% tổng doanh số của công nghệ sinh
trắc học nói chung.

Vậy tại sao vân tay lại được sử dụng để nhận dạng?

Thông qua các nghiên cứu chuyên sâu đã được thực hiện trên dấu vân tay của con người. Hai
kết quả quan trọng nhất trong số các kết quả thu được của quá trình nghiên cứu trên là :

- Dấu vân tay của một con người sẽ không bao giờ thay đổi cấu trúc tự nhiên của nó sau
khoảng thời gian 1 năm từ khi sinh ra.
- Dấu vân tay của một cá nhân là duy nhất. Ngay cả với anh em sinh đôi thì dấu vân tay
của họ cũng không giống nhau. Trong thực tế chưa tìm thấy trường hợp nào mà hai

người
lại có cùng dấu vân tay.
Nguyên lý hoạt động của công nghệ nhận dạng vân tay: Khi đặt ngón tay lên trên một thiết
bị nhận dạng dấu vân tay (hình 4), ngay lập tức thiết bị này sẽ quét hình ảnh ngón tay đó và đối
chiếu các đặc điểm của ngón tay đó với dữ liệu đã được lưu trữ trong hệ thống. Quá trình xử lý
dữ liệu sẽ được thiết bị chuyển sang các dữ liệu số và ra thông báo rằng dấu vân tay đó là hợp lệ
hay không hợp lệ để cho phép hệ thống thực hiện các chức năng tiếp theo.
Đặc điểm của dấu vân tay mặc dù chỉ gồm 7 loại (vòng móc đơn, vòng móc kép, vòng tập
trung ở giữa, vòng cung, vòng cung hình lều, vòng xoắn, vòng bất thường) nhưng sự thể hiện của
nó trên mỗi một bàn tay của con người lại muôn hình muôn vẻ và không bao giờ có sự giống
nhau tuyệt đối. Chính tính độc nhất về cấu tạo hình dạng vân tay của mỗi người, các nhà sinh
trắc học sẽ biến nó thành chiếc chìa khóa riêng mà chỉ có chính họ mới có thể sử dụng, giúp họ
tránh được nhiều phiền toái trong cuộc sống như bị chộm cắp, lạm dụng hoặc giả mạo như đối
với các loại giấy tờ tùy thân, thẻ ngân hàng, hộ chiếu…đảm bảo an ninh và bảo mật.

Hình II-3: Nhận dạng bằng dấu vân tay

Hệ thống sinh trắc học sẽ ghi nhận mẫu vân tay của người dùng và lưu trữ tất cả

những dữ
liệu đặc biệt này thành một mẫu nhận diện được số hóa toàn phần. Có hai phương pháp để lấy
dấu vân tay:
- Cách thứ nhất (cổ điển): Sao chép lại hình dạng vân tay (lăn tay, hay chạm tay vào một
vật gì đó) thông qua các máy quét ghi nhận và xử lý. Tuy nhiên, phương pháp này hiện

nay không đảm bảo được mức độ an ninh và bảo mật cao, bởi vì nó có thể bị vô hiệu hóa
trong trường hợp kẻ gian dùng một bao tay cao su giả mạo dấu vân tay hay cũng không
loại trừ trường hợp bọn chúng "mượn" luôn ngón tay của nạn nhân để phục vụ mục đích
riêng.
- Cách thứ 2 được xem là "đọc" dòng điện dưới ngón tay thông qua hệ thống khuyếch đại
xung điện, rồi chuyển thành vân tay. Cách này giúp giảm thiểu khả năng giả mạo vân tay,
nhưng lại đòi hỏi trình độ công nghệ cao và chi phí sử dụng lớn hơn nhiều so với phương
pháp trên. Đây cũng là cách những thiết bị kỹ thuật số như điện thoại, máy tính xách tay,
ổ cứng USB sử dụng để làm cổng mã hóa bảo vệ dữ liệu.
Ngày nay việc ứng dụng công nghệ sinh trắc học về nhận dạng vân tay được sử dụng khá phổ

biến vì nó đáp ứng yêu cầu bảo vệ dữ liệu, đảm bảo an ninh an toàn với độ chính xác tương đối
cao.

Hình II-4: Các thiết bi chuyên dụng dùng để nhận dạng vân tay

2.3 Xác thực bằng mống mắt (Iris recognition technology)

Vân mống mắt (iris) là phần nằm giữa tròng trắng (sclera) và con ngươi (pupil). Nó có cấu
trúc rất phức tạp và duy nhất. Nhiệm vụ chính của mống mắt là điều chỉnh sự co giản của tròng
trắng và con ngươi với ánh sáng.

So với những kỹ thuật nhận dạng sinh trắc học khác, nhận dạng vân mống mắt có nhiều đặc điểm
nổi trội hơn.

- Tính duy nhất: Ý tưởng dùng vân mống mắt cho nhận dạng được Frank Burch, một bác
sĩ khoa mắt đề xuất vào năm 1936 với lý do là vân mống mắt có tính duy nhất gần như
tuyệt đối. Khả năng hai người có vân mống mắt giống nhau là 10-35. Ngay cả hai người
sinh đôi cũng có vân mống mắt khác nhau.

- Tính tin cậy và ổn định: Mống mắt là một phần của mắt được bảo vệ bởi mi mắt và
nước mắt. Không như vân tay và lòng bàn tay, nó khó bị làm tổn thương và lỗi nhận dạng
do sẹo không bao giờ xảy ra. Bên cạnh đó nó không thay đổi trong suốt cuộc đời của
chúng ta.

- Không thể giả mạo: Đường kính của con ngươi và đường kính của mống mắt luôn thay
đổi để thích nghi với ánh sáng. Dụng cụ đo đồng tử dựa trên sự thay đổi này có thể cung
cấp cho việc kiểm tra sự giả mạo.

Hình II-5: Nhận dạng bằng mống mắt

Năm 1987, hai bác sĩ khoa mắt Aran Safir và Leonard Flom lấy bằng sáng chế ý tưởng này và
yêu cầu TS Joln Daugman, một giáo sư chuyên về lĩnh vực thị giác máy tính Đại học
Cambrigdge nghiên cứu. Năm 1993 TS Joln Daugman công bố giải thuật nhận dạng vân mống
mắt mà phần lớn các hệ thống ngày nay đều dùng phương pháp này. Sau đó Wildes và Boles đề
xuất hai giải thuật khác và cũng rất thành công. Những năm gần đây Học viện tự động thuộc

Viện khoa học của Trung Quốc (Chinese Academy of Sciences' Institute of Automation) cũng
đưa ra một vài phương pháp nhận dạng mống mắt.

Hình II-6: Các thiết bị quét võng mạc

Các thuật toán để áp dụng sinh trắc học vào nhận dạng vân mống mắt người

i. Định vị tâm, đường bao con ngươi, đường bao vân mống mắt
ii. Chuyển vân mống mắt về dạng hình chữ nhật
iii. Chuyển hình chữ nhật về dạng ảnh nhị phân
iv. So sánh và quyết định

So Trùng *(Matching)*

Vấn đề của việc so sánh giữa các mẫu Iriscode khi chúng ta muốn xác thực một người
dùng. Ảnh mắt của người cần xác thực sẽ được chụp và tạo ra Iriscode. Bộ Iriscode này sẽ được
so sánh với cơ sở dữ liệu Iriscodes đã tạo từ trước. Chúng ta áp dụng khoảng cách Hamming
iriscode của người cần xác thực lên từng mẫu iriscode trong cơ sở dữ liệu. Nếu kết quả bằng 0
coi như hai iriscode là của cùng một người và ngược lại. Nhưng trên thực tế điều này không luôn
xảy ra do nhiều nguyên nhân như ảnh bị xoay, ánh sáng do lúc chụp ảnh... nên kết quả thu được
gọi là trùng khớp (match) chỉ dưới một ngưỡng nào đó. Phương pháp xác định ngưỡng dựa trên
phương pháp thống kê độc lập.

2.4 Xác thực thông qua hình học bàn tay (Hand recognition technology)

Hình học bàn tay là một trong số các phương pháp sinh trắc học sử dụng hình dạng của bàn
tay con người làm công cụ cho việc xác thực. Các đặc điểm của bàn tay, chẳng hạn như chiều dài
ngón tay, chiều rộng và độ cong, cũng như các tính năng độc đáo có thể được sử dụng để nhận
dạng. Thiết bị quét hình học bàn tay yêu cầu người dùng đặt tay lên một bề mặt với 5 chốt. Mặt
phẳng này để cho các máy quét có thể nhận được các dữ liệu của bàn tay sau từng lần quét. Các
thông số thu được của máy quét sau đó được so sánh với cơ sở dữ liệu để xác

minh.

Hình II- 7: Nhận dạng thông qua hình học bàn tay

Hình II-8: Tĩnh mạch lưng bàn tay.	Một kiểu quét sinh trắc học khác có thể được sử dụng để xác định mạng lưới tĩnh mạch lưng bàn tay. Điều này về cơ bản cho thấy các mạch máu trên mặt sau của bàn tay và yếu tố này có thể là hữu ích phục vụ cho việc xác thực.

Các nghiên cứu về việc sử dụng hình học bàn tay để thực hiện nhận dạng được nghiên cứu bởi

Đại học Michigan. Trong hệ thống của họ người dùng tương tác với hệ thống kiểm định tay để

cấp quyền truy cập đến các dịch vụ. Một vấn đề có thể gặp phải trong khi phát triển hệ thống

nhận dạng bằng hình học bàn tay đó là có thể có nhiều hơn một người có cùng một kiểu hình học

bàn tay tương đối giống nhau. Điều này đã làm hạn chế các ứng dụng sử dụng phương pháp nhận

dạng thông qua hình học bàn tay. Tuy nhiên, việc kết hợp hình học bàn tay với một hình thức bảo

mật sinh trắc học khác, chẳng hạn như nhận dạng bằng vân tay, sẽ tạo ra được một

hệ thống nhận
dạng rất an toàn.

2.5 Xác thực thông qua ADN (ADN recognition technology)

Những năm gần đây, xác thực thông qua ADN đã được nhắc đến rất nhiều. Như chúng ta
đều biết thông qua ADN có thể xác định chính xác nguồn gốc của một người, xem bố mẹ của
người đó là ai thông qua các cặp nhiễm sắc thể di truyền từ cha mẹ cho con cái, ngoài ra trong
cấu trúc ADN còn có các cặp nhiễm sắc thể đặc trưng cho từng cá thể, và trên cơ sở những cặp
nhiễm sắc thể đặc trưng này người ta đã đưa ra phương pháp xác thực thông qua AND. Do đó,
trước khi tìm hiểu các vấn đề liên quan đến công nghệ xác thực sinh trắc học thông qua ADN ta
tìm hiểu một số vấn đề cơ bản về ADN và trả lời câu hỏi tại sao ADN lại có thể sử dụng để xác
thực cá nhân con người.

Hình II-9: Cấu trúc ADN	Acid Deoxyribo Nucleic (ADN) là vật liệu di truyền trong hầu hết các sinh vật bao gồm cả con người. ADN có thể được hiểu một cách đơn giản là nơi chứa mọi thông

	tin chỉ dẫn cần thiết để tạo nên các đặc tính sự sống của mỗi sinh vật. Mỗi một con người có thể được nhận dạng bởi các đặc điểm di truyền ADN của họ. ADN nằm trong nhân của tế bào cũng như ty thể. ADN phục vụ như một mã di truyền và là duy nhất cho một sinh vật. Mã ADN của một sinh vật bao gồm bốn cơ sở: Adenine (A), Guanine (G), Cytosine (C) và Thymine (T).

Các ứng dụng của ADN trong thực tiễn:

ADN trong vấn đề tội phạm

Khoa học hình sự có thể sử dụng ADN thu nhận từ máu, tinh dịch hay lông, tóc của hung

thủ để lại trên hiện trường mà điều tra, giám định vụ án. Lĩnh vực này gọi là kỹ thuật vân tay

ADN (*genetic fingerprinting*) hay *ADN profiling* (kỹ thuật nhận diện ADN).

ADN trong khoa học máy tính

Dù có nguồn gốc từ sinh học, ADN lại đóng một vai trò quan trọng trong khoa học máy
tính, vừa là một vấn đề đang được tập trung nghiên cứu vừa là một phương pháp tính toán, gọi là
tính toán ADN.

Hệ thống nhận dạng sinh trắc học thông qua ADN hoạt động như thế nào:

Các tế bào có chứa vật liệu di truyền ADN chia sẻ thông tin thông qua các nhiễm sắc thể.

Con người có 23 cặp nhiễm sắc thể đó là ADN của một người và gen của họ. Trong tổng số 46
cặp nhiễm sắc thể, có 23 cặp do cha mẹ di truyền cho con cái. 99,7% trong số ADN của con cái
chung với cha mẹ của chúng. 0,3% còn lại của một người là biến ADN lặp đi lặp lại một mã duy
nhất tương ứng với người đó. Mã hóa này lặp di lặp lại là cơ sở của sinh trắc học ADN. Nhận
dạng ADN sử dụng hồ sơ di truyền, còn gọi là kỹ thuật vân tay ADN (*genetic fingerprinting*), để
cô lập và xác định các khu vực ADN này. Nó là duy nhất cho từng cá nhân, có thể sử dụng để
xác định hoặc xác minh danh tính của một người. [6]

Các bước cơ bản của kỹ thuật nhận dạng vân tay ADN bao gồm:

1. Cô lập các ADN (mẫu có thể có nguồn gốc từ máu, nước bọt, tóc, tinh dịch, hoặc mô).

2. Phân các mẫu ADN thành các đoạn ngắn hơn có chứa số lặp đi lặp lại được biết đến như
biến số lặp đi lặp lại – giống hệt trình tự lặp lại của ADN.
3. Tổ chức các phân đoạn ADN theo kích cỡ.
4. So sách các phân đoạn ADN từ các mẫu khác nhau.
Nhận dạng bằng sinh trắc học ADN hiện là phương pháp có độ chính xác cao nhất, với
phương pháp sinh trắc học ADN chúng ta không lo bị làm giảm mạo. Tuy nhiên, phương pháp
này hiện nay được ứng dụng rất hạn chế, chủ yếu ở các nước tiên tiến bởi vì để thực hiện được
phương pháp này đòi hỏi một trình độ khoa học công nghệ nhất định, hệ thống máy móc tối tân,
hơn nữa chi phí thực hiện lại cao, thời gian cho việc nhận dạng là tương đối lâu và việc nhận
dạng chỉ được thực hiện trong phòng thí nghiệm... Chính vì những lý do đó mà phương pháp
nhận dạng này chưa được nghiên cứu phát triển mở rộng như với các phương pháp nhận dạng
sinh trắc học khác. Hy vọng rằng trong tương lai bằng sự phát triển của khoa học công nghệ
những khó khăn còn tồn tại sẽ được khắc phục để phương pháp này có thể sử dụng rộng rãi hơn.

Hình II-10: Xác thực sinh trắc học ADN.

3. Kết luận

Qua những kết quả trình bày ở trên chúng ta có thể thấy được tại sao lại cần phải phát triển

các hệ thống xác thực sinh trắc học và bài báo cũng đã giới thiệu một số phương pháp xác thực
bằng công nghệ sinh trắc học hiện đang được sử dụng phổ biến hiện nay. Xác thực bằng vân tay
là phương pháp xác thực ra đời sớm nhất và hiện là phương pháp xác thực sinh trắc học phổ biến
nhất, doanh thu từ việc phát triển các ứng dụng xác thực bằng công nghệ này đang chiếm một tỷ
lệ lớn (năm 2003 chiếm 34% tổng số doanh thu của công nghệ sinh trắc học). Ngày nay phương
pháp xác thực thông qua ADN đang có những bước phát triển mới bởi đây là phương pháp hiện
có độ tin cây cao nhất. Tuy nhiên, phương pháp xác thực này đòi hỏi các máy móc thiết bị hiện
đại, thời gian xác thực lâu, hơn nữa chi phí cao nên hiện tại vẫn chưa thực sự phổ biến.

Hình II-11: Doanh thu của công nghệ sinh trắc học, năm 2003

Tại Việt Nam, đã có một số lĩnh vực áp dụng công nghệ sinh trắc học trong việc xác thực,
tuy nhiên vẫn còn khá kiêm tốn. Với xu thế phát triển chung cùng với những lợi ích to lớn mà
công nghệ sinh trắc học mang lại, có thể khẳng định rằng trong những năm tới việc áp dụng các
phương pháp xác thực sinh trắc học ở Việt nam sẽ ngày càng trở nên quan trọng và phát triển
không ngừng. Như chúng ta có thể triển khai các hệ thống nhận dạng khuôn mặt ở

các sân bay,
các nơi tập trung đông người để đảm bảo an ninh, ngăn ngừa nguy cơ khủng bố..hay triển khai
hệ thống hộ chiếu sinh trắc học thay thế cho hệ thống hiện tại để đảm bảo an ninh, thuận lợi cho
người dùng…Hiện nay, hộ chiếu sinh trắc học đã được các nước liên minh châu Âu (EU) sử
dụng và cho những kết quả rất khả quan.

TÀI LIỆU THAM KHẢO

[1] - http://www.biometrics.org, http://www.biometrics.gov
[2] - http://www.inttelix.com/face-recognition-technology.php
[3] - http://misbiometrics.wikidot.com
[4] - Fingerprint Recognition, Andrew Ackeman, Professor Rafail Ostrovsky.
[5] - Biometrics Overview, National Science and Technology Council (NSTC).
[6] - ADN as a Biometric Identifier, Presented by: Shannon Soltysiak, Hamed Valizadegan.
[7] – The National Biometrics Challenge, National Science and Technology Council (NSTC).

SUMMARY

Research authentication methods using biometrics technologies

Dang Van Nam, *Ha Noi University of Mining and Geology*

Abstract

Authentication using biometric technology are demonstrated advantages compared with
traditional authentication methods by the level of safety and reliability, creating favorable
conditions for the user. The authentication methods currently popular in this

technology can
show that: authentication by fingerprint, voice, face, hand geometry, iris, ADN ... Validating by
biometric methods are widely applied in many fields from the fields of requirements for accuracy
and security as the military, E-Government, bank transaction come to quotidian fields such as
enterprise security, protect against delinquent- terror, in social welfares…

In Vietnam, the application authentication using biometric technology is still relatively
new and not popular. This paper will discuss why the need to develop biometric authentication
systems, as well as the authentication method using biometric technology is being widely used
today.