**ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**

**-----o0o----**

****

**BÁO CÁO ĐỒ ÁN**

***MÔN : TOÁN RỜI RẠC NÂNG CAO***

**ĐỀ TÀI:**

**MÃ HÓA VĂN BẢN**

**BẰNG PHƯƠNG PHÁP MÃ HÓA DES**

***Giảng viên hướng dẫn:*** ThS.nGUYỄN ĐÌNH HIỂN

***Sinh viên tham gia :***

TRẦN THANH DƯƠNG 13520144

PHẠM NHẬT NAM 13520520

VÕ ĐÌNH NHÃ 13520568

NGUYỄN VĂN PHƯỚC 13520652

**Lớp: CS521.H11**

**TP. Hồ Chí Minh, ngày 23 tháng 12 năm 2016**

MỤC LỤC

[MỤC LỤC 2](#_Toc470207958)

[CHƯƠNG I. GIỚI THIỆU TỔNG QUAN 3](#_Toc470207959)

[1.1. Giới thiệu về mật mã học 3](#_Toc470207960)

[1.2. Các thành phần của một hệ mật mã 3](#_Toc470207961)

[1.3. Phân loại các hệ mật mã 3](#_Toc470207962)

[CHƯƠNG II. CHUẨN MẬT MÃ DES 5](#_Toc470207963)

[2.1. Lịch sử của DES 5](#_Toc470207964)

[2.2. Thuật toán và lưu đồ hoạt động của DES 5](#_Toc470207965)

[2.3. Thuật toán sinh khóa con 7](#_Toc470207966)

[2.4. Cấu trúc vòng lặp DES 8](#_Toc470207967)

[2.5. Cấu trúc cụ thể hàm f 8](#_Toc470207968)

[2.6. Cấu trúc của các S-Box 9](#_Toc470207969)

[2.7. Các thuộc tính của S-Box 9](#_Toc470207970)

[2.8. Các điểm yếu của DES 10](#_Toc470207971)

[2.8.1. Tính bù 10](#_Toc470207972)

[2.8.2. Khóa yếu 10](#_Toc470207973)

[2.8.3. Tấn công bằng phương pháp vét cạn(bruce-force attack) 11](#_Toc470207974)

[CHƯƠNG III. CÀI ĐẶT DES 12](#_Toc470207975)

[3.1. Phân tích bài toán 12](#_Toc470207976)

[3.2. Kết quả demo 12](#_Toc470207977)

[3.2.1. Thiết kế kiến trúc 12](#_Toc470207978)

[3.2.2. Thiết kế giao diện 12](#_Toc470207979)

[TÀI LIỆU THAM KHẢO 14](#_Toc470207980)

# GIỚI THIỆU TỔNG QUAN

## Giới thiệu về mật mã học

Kỹ thuật mật mã thông qua việc biến đổi hoặc mã hoá thông tin, biến đổi những thông tin nhạy cảm, vấn đề cơ mật thành những văn tự mã hoá có dạng hỗn loạn, làm cho tin tặc khó lòng mà đọc hiểu được, từ đó sẽ đạt được hai mục đích: một là, làm cho tin tặc không biết làm thế nào để giải mã nên cũng không thể thu được những thông tin có bất kỳ ý nghĩa nào trong chuỗi mật mã hỗn loạn đó; hai là làm cho tin tặc không có khả năng làm giả thông tin với chuỗi mật mã hỗn loạn như thế. Khoa học nghiên cứu kỹ thuật mật mã gọi là mật mã học.

Mật mã học bao gồm hai nhánh, là mật mã học lập mã và mật mã học phân tích. Mật mã học lập mã với ý là tiến hành mã hoá thông tin để thực hiện việc che giấu thông tin, còn mật mã học phân tích là ngành học nghiên cứu phân tích giải dịch mật mã. Hai cái đối lập với nhau, nhưng lại thúc đẩy lẫn nhau.

Dùng phương pháp mật mã có thể che dấu và bảo hộ những thông tin cơ mật, làm cho người chưa được uỷ quyền không thể lấy được thông tin, những thông tin được giấu kín kia được gọi là văn bản rõ, mật mã có thể đem văn bản rõ biến đổi thành một loại hình khác, gọi là văn bản mật. Sự biến đổi văn bản rõ thành văn bản mật gọi là mã hoá bảo mật, quá trình người thu nhận hợp pháp khôi phục từ văn bản mật trở thành văn bản rõ được gọi là quá trình giải mã (hoặc giải mật). Người thu nhận phi pháp có ý đồ phân tích từ văn bản mật ra thành văn bản rõ, gọi là giải dịch.

## Các thành phần của một hệ mật mã

Một hệ mật mã là một bộ 5 (P, C, K, E, D) thoả mãn các điều kiện sau:

+ P là một tập hữu hạn các bản rõ có thể

+ C là tập hữu hạn các bản mã có thể

+ K (không gian khoá) là tập hữu hạn các khoá có thể

+ Đối với mỗi kK có một quy tắc mã ek: P → C và một quy tắc giải mã tương ứng dkD. Mỗi ek: P → C và dk: C → P là những hàm mà: dk(ek(x)) = x với mọi bản rõ xP.

Tính chất 4 là tính chất chủ yếu. Nội dung của nó là nếu một bản rõ x được mã hoá bằng ek và bản mã nhận được sau đó được giải mã bằng dk thì ta phải thu được bản rõ ban đầu x. Trong trường hợp này hàm mã hoá ek phải là hàm đơn ánh, nếu không việc giải mã sẽ không thể thực hiện được một cách tường minh.

## Phân loại các hệ mật mã

Hiện nay người ta đã thiết kế ra nhiều loại hệ thống mật mã, nếu như lấy khoá mật mã làm tiêu chuẩn có thể phân các hệ mật mã thành hai loại:

- **Hệ mật mã đối xứng** (còn gọi là mật mã khoá đơn hoặc là mật mã khoá riêng): Trong các hệ mật mã này, khoá mật mã mã hoá bảo mật giống với khoá giải mã hoặc trên thực tế là cùng đẳng cấp. Lúc này khoá mật mã cần phải có một đường truyền an toàn để truyền đưa khoá mật mã từ phía người truyền cho phía người nhận. Đặc điểm của mật mã đối xứng là bất luận khi gia công bảo mật hay là khi giải mã đều sử dụng cùng một khoá mật mã. Do đó tính an toàn của mật mã này là sự an toàn của khoá mật mã.Nếu như khoá mật mã bị tiết lộ, thì hệ thống mật mã này sẽ bị phá vỡ. Mật mã đối xứng có ảnh hưởng nhất là phép tính DES do cục tiêu chuẩn quốc gia Mỹ công bố vào năm 1977.

***Ưu điểm:*** Tính an toàn cao, tốc độ giải mã nhanh.

***Nhược điểm:***

+ Theo sự mở rộng của quy mô mạng lưới, việc quản lý khoá mật mã trở thành một việc khó khăn.

+ Không có cách nào giải quyết vấn đề xác nhận thông tin.

+ Thiếu năng lực kiểm tra tự động sự tiết lộ khoá mật mã.

**-Hệ mật mã bất đối xứng** (còn gọi là mật mã khoá công khai hoặc mật mã khoá đôi): Trong các hệ mật mã này quá trình mã hoá và giải mã có chìa khoá khác nhau, lúc này không cần có đường truyền an toàn để truyền đưa khoá mật mã mà chỉ cần bộ phát sinh khoá mã tại chỗ để tạo ra khoá giải mã đồng thời lấy đó để khống chế các thao tác giải mã. Mật mã bất đối xứng là một thể chế mật mã loại mới do W.Diffie và M.E Hellman đề xuất năm 1976. Do quá trình mã hoá và giải mã của thể chế mật mã bất đối xứng không như nhau và khoá mã bảo mật là công khai, hơn nữa, chỉ yêu cầu bảo mật khoá giải mã, cho nên mật mã bất đối xứng không tồn tại vấn đề quản lý khoá mật mã. Mật mã bất đối xứng còn một ưu điểm nữa là có thể có khả năng ký tên chữ số và một số chức năng mới. Mật mã bất đối xứng nổi tiếng nhất là thể chế mật mã RSA do ba người là Rivest, Shamir và Adleman đề xuất năm 1977. Khuyết điểm của mật mã bất đối xứng là: phép tính mật mã là tương đối phức tạp, tốc độ giải mã chậm.

Do đó, việc bảo mật dữ liệu trên mạng nên dùng cơ chế bảo mật hỗn hợp kết hợp giữa mật mã đối xứng và mật mã bất đối xứng, tức là khi giải mã thì dùng mật mã đối xứng, khi truyền đưa khoá mật mã thì dùng mật mã bất đỗi xứng. Như thế tức là đã giải quyết được khó khăn trong việc quản lý khoá mật mã, vừa lại giải quyết được vấn đề tốc độ giải mã. Đây là một phương pháp tương đối tốt để giải quyết vấn đề an toàn thông tin khi truyền đưa trên mạng hiện nay.

# CHUẨN MẬT MÃ DES

### Lịch sử của DES

Vào những năm đầu thập kỷ 70, nhu cầu có một chuẩn chung về thuật toán mật mã đã trở nên rõ ràng. Các lý do chính là:

* Sự phát triển của công nghệ thông tin và của nhu cầu an toàn & bảo mật thông tin: sự ra đời của các mạng máy tính tiền thân của Internet đã cho phép khả năng hợp tác và liên lạc số hóa giữa nhiều công ty, tổ chức trong các dự án lớn của chính phủ Mỹ.
* Các thuật toán ‘cây nhà lá vườn’ (ad hoc) không thể đảm bảo được tính tin cậy đòi hỏi cao.
* Các thiết bị khác nhau đòi hỏi sự trao đổi thông tin mật mã thống nhất, chuẩn.

Một chuẩn chung cần thiết phải có với các thuộc tính như:

1. Bảo mật ở mức cao

2. Thuật toán được đặc tả và công khai hoàn toàn, tức là tính bảo mật không được phép dựa trên những phần che giấu đặc biệt của thuật toán.

3. Việc cài đặt phải dễ dàng để đem lại tính kinh tế

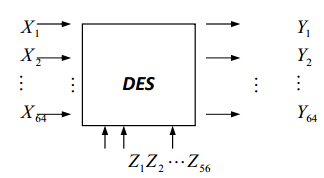
4. Phải mềm dẻo để áp dụng được cho muôn vàn nhu cầu ứng dụng

Năm 1973, Cục quản lý các chuẩn quốc gia của Mỹ đã có văn bản cổ động cho việc tạo lập các hệ mật mã chuẩn ở cơ quan đăng ký liên bang của Mỹ. Điều này đã dẫn đến sự công bố vào năm 1977 của cục An ninh Quốc gia Mỹ (NSA) về Data Encryption Standard, viết tắt là DES.

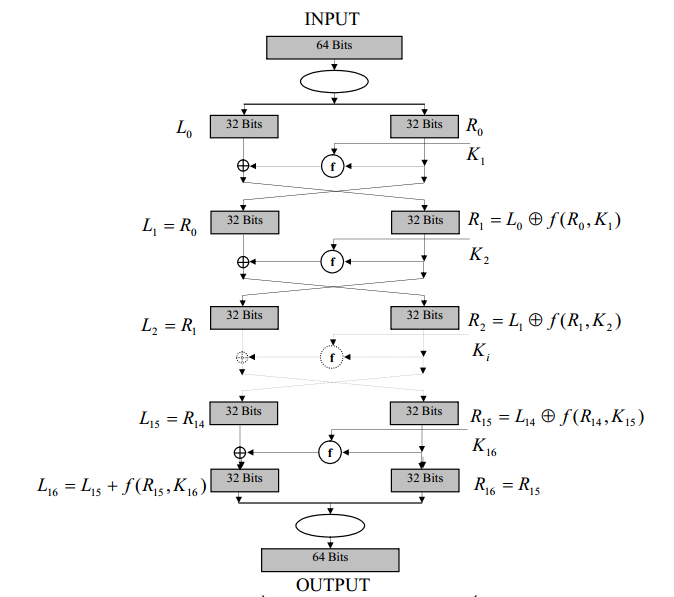
Thực chất, DES được phát triển bởi IBM như là sự sửa đổi của một hệ mã trước kia được biết với cái tên Lucipher. Trong khoảng 2 thập kỷ tiếp theo, DES là hệ mã được dùng rộng rãi nhất và cũng là gây ra nhiều nghi ngờ, tranh cãi trong lĩnh vực này: xung quanh các nguyên tắc thiết kế đảm bảo tính mật, chiều dài khóa tương đối ngắn và khả năng NSA còn che giấu cửa sau (backdoor) để có thể bẻ khóa, phá mã ít tốn kém hơn thông thường.

### Thuật toán và lưu đồ hoạt động của DES

Các hình vẽ sau cung cấp sơ đồ khái quát và chi tiết của thuật toán sinh mã trong DES.



*Hình 2.1 Sơ đồ cơ bản của DES: đầu vào của DES là khối độ dài 64 bits, đầu ra 64 bits và khóa là 56 bits.*

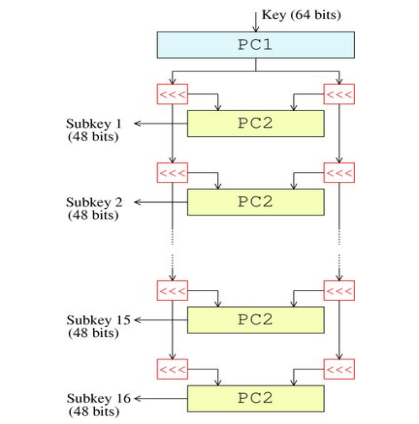


*Hình 2.2 Sơ đồ giải thuật sinh mã DES với cấu trúc 16 vòng lặp*

Sơ đồ hình vẽ 2.2 cho thấy DES được cấu tạo bởi 16 bước lặp với bước lặp cơ sở gọi hàm chuyển đổi phi tuyến f;16 bước lặp này được kẹp vào giữa hai tác tử giao hoán IP và IP-1. Hai tác từ này không có ý nghĩa gì về mặt bảo mật mà hoàn toàn nhằm tạo điều kiện cho việc cài đặt phần cứng, ‘chip hóa’ thuật toán DES. Hàm cơ sở f là nguồn gốc của sức mạnh bảo mật trong thuật toán DES này.

### Thuật toán sinh khóa con

16 vòng lặp của DES cùng gọi thực hiện f nhưng với các tham số khóa khác nhau. Tất cả 16 khóa khác nhau này, được gọi là khóa con, cùng sinh ra từ khóa chính của DES bằng một thuật toán sinh khóa con. Trong thuật toán sinh khóa con này (lập lịch khóa), khóa chính K, 64 bit, đi qua 16 bước biến đổi, tại mỗi bước này một khóa con được sinh ra với độ dài 48 bit.



*Hình 2.3 Sơ đồ thuật toán sinh khóa con (Key Scheduler) – Nguồn: Wikipedia*

Qua sơ đồ thuật toán sinh khóa con có thể thấy rằng thực sự chỉ có 56 bit của khóa chính được sử dụng, 8 bit còn lại là mã kiểm tra chẵn lẻ (parity bits) và bị lọc ra ở biến đổi PC1. Các bộ biến đổi PC1 và PC2 chỉ đơn giản là các bộ vừa chọn lọc vừa hoán vị

(PC = permuted choice =lựa chọn có hoán vị). Các biến đổi R1 và R2 (left rotate 1 bit và 2 bit) tương ứng là các phép đẩy bit trái 1 và 2 vị trí.

### Cấu trúc vòng lặp DES

Mỗi vòng lặp của DES thực hiện trên cơ sở công thức sau:

Trong đó là 2 nửa trái và phải thu được từ biến đối của vòng lặp thứ i.

Ta cũng có thể viết lại

Trong đó F là phép thay thế bằng , còn T là phép đổi chỗ hai thành phần L và R. Tức là mỗi biến đổi vòng lặp của DES có thể coi là một tích hàm số của F và T (trừ vòng cuối cùng không có T).

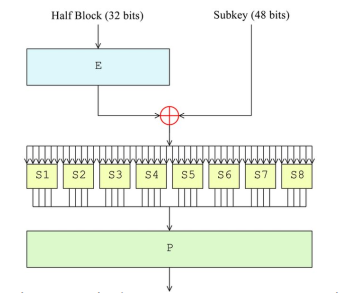
Ta có thể viết lại toàn bộ **thuật toán sinh mã DES** dưới dạng công thức tích hàm số như sau:

Thuật toán giải mã DES được xây dựng giống hệt như thuật toán sinh mã nhưng có các khóa con được sử dụng theo thứ tự ngược lại, tức là dùng khóa K16 cho vòng lặp 1, khóa K15 cho vòng lặp 2 ... Vì vậy, thuật toán giải mã có thể được viết lại dưới dạng công thức sau:

Mỗi hàm T hoặc F đều là các hàm có tính chất đối hợp (, hay).Do đó nếu ta thực hiện phép tích hàm hay thì sẽ thu được phép đồng nhất. Điều đó giải thích tại sao thuật toán giải mã lại giống hệt như sinh mã chỉ có khác về thứ tự trong chuỗi khóa con.

### Cấu trúc cụ thể hàm f

Sơ đồ biến đổi cụ thể của hàm f được minh họa trong hình 2.4. Trước hết, 32 bit của thành phần được mở rộng thành 48 bit thông qua biến đổi E (expansion: mở rộng với sự lặp lại một số bit) rồi đem XOR với 48 bit của khóa . Tiếp theo, 48 bit kết quả sẽ được phân thành 8 nhóm 6 bit. Mỗi nhóm này sẽ đi vào một biến đổi đặc biệt gọi là biến đổi S-box (có 8 S-box khác nhau ứng với mỗi nhóm 6 bit) và cho ra kết quả là 8 nhóm 4 bit. Từ đó, 32 bit hợp thành (sau khi qua 8 S-box khác nhau) sẽ được hoán vị lại theo hàm hoán vị P để đưa ra kết quả cuối cùng của hàm f (tức nhân của).

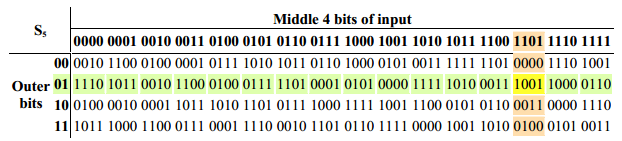


*Hình 2.4 Cấu trúc của biến đổi hàm f, bước lặp cơ sở của DES. Nguồn: Wikipedia*

### Cấu trúc của các S-Box

Như ta biết mỗi một trong 8 nhóm 6 bit sẽ đi vào mỗi trong 8 bộ biến đổi S1,S2 ... S8.

Mỗi S-box bao gồm 4 bảng biến đổi dòng, thực chất là một biến đổi hoán vị cho 16 tổ hợp của 4 bits. Trong 6 bits đầu vào thì hai bit ngoài cùng (bit 1 và 6) được dùng để chỉ định 1 trong 4 bảng biến đổi dòng này; vì thế chúng được gọi là các bit điều khiển trái và phải (CL và CR).Còn lại 4 bit chính (các bit 2-5) của nhóm 6 bit đầu vào sẽ là tổ hợp 4 bits bị biến đổi.



*Hình 2.5 Bảng biến đổi S5: đầu vào 6 bits 011011 sẽ được biến đổi thành 1001 (ô vàng)*

### Các thuộc tính của S-Box

Các nguyên tắc thiết kế của 8 S-box được đưa vào lớp thông tin mật ‘Classified information’ ở Mỹ. Mặc dù vây, NSA đã tiết lộ 3 thuộc tính của S-boxes.

* Các bít vào (output bit) luôn phụ thuộc không tuyến tính vào các bít ra (input bit).
* 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 con lại cho thay đổi thì S-boxes thể hiện một tính chất được gọi là ‘phân bố đồng nhất ‘ (uniform distribution): so sánh số lượng bit số 0 và 1 ở các đầu ra luôn ở mức cân bằng. Tính chất này khiến cho việc áp dụng phân tích theo lý thuyết thống kê để tìm cách phá S-boxes là vô ích.

Thực tế, sau 8 vòng lặp tất cả các bit ra của DES sẽ chịu ảnh hưởng của tất cả các bit vào và tất cả các bit của khóa. Hơn nữa sự phụ thuộc này là rất phức tạp. Tuy nhiên sau này một số tấn công mới đã được đề xuất và cho thấy 8 vòng lặp này là chưa đủ để bảo mật (điều này cho thấy NSA đã biết trước các dạng tấn công này nên mới qui định số vòng lặp là 16 ngay từ đầu).

Chính cấu tạo của S-box đã gây tranh luận mạnh mẽ trong các thập kỷ 70-90 về khả năng cơ quan NSA (National Security Agency), Mỹ, vẫn còn che dấu các một số đặc tính của S-box hay cài bên trong những cửa bẫy (trapdoor) mà qua đó họ có thể dễ dàng phá giải mã hơn người bình thường (biết các bí mật này có thể giản lược không gian khóa để tìm kiếm vét cạn nhanh hơn). Sự phát hiện sau đó của các tấn công mới, rất mạnh như tấn công vi phân, đã củng cố sự nghi ngờ của giới khoa học.

### Các điểm yếu của DES

##### Tính bù

Ký hiệu là phần bù của (ví dụ 0100101 và 1011010 là bù của nhau) thì DES có tính chất sau:

Do đó nếu biết MÃ y được mã hóa từ TIN x với khóa z thì ta suy ra được mã hóa từ TIN với khóa . Tính chất này chính là một điểm yếu của DES bởi vì nhờ đó kẻ địch có thể loại trừ một nửa số khóa cần phải thử khi tiến hành phép thử-giải mã theo kiểu tìm kiếm vét cạn không gian khóa.

##### Khóa yếu

Các khóa yếu là các khóa mà theo thuật toán sinh khóa con thì tất cả 16 khóa con đều như nhau

điều đó khiến cho phép sinh mã và giải mã đối với các khóa yếu này là giống hệt nhau

Có tất cả 4 khóa yếu như sau:

1. [00000001 00000001… … 00000001]
2. [11111110 11111110 ... ... 11111110 ]
3. [11100000 11100000 11100000 11100000 11110001 11110001 11110001 11110001]
4. [00011111 00011111 00011111 00011111 00001110 00001110 00001110 00001110]

Đồng thời có 10 khóa yếu với thuộc tính tồn tại Z, Z’ sao cho

hay là

##### Tấn công bằng phương pháp vét cạn(bruce-force attack)

DES có khóa. Nếu như biết một cặp plaintext-ciphertext thì chúng ta có thể thử tất cả khả năng này để tìm ra khóa cho kết quả khớp. Giả sử như một phép thử mất quãng s(trên một máy PC thông thường), thì chúng ta sẽ thử mất s tức là 7300 năm!

Nhưng nhớ rằng đấy mới chỉ là sử dụng các máy tính thông thường, còn có các máy tính được chế tạo theo nguyên lý xử lý song song. Chẳng hạn nếu như làm được một thiết bị với con chip mật mã DES chạy song song thì bây giờ mỗi con chip chỉ phải chịu trách nhiệm tính toán với phép thử. Chip mã DES ngày nay có thể xử lý tới tốc độ là 4.5 x bits/s tức là có thể làm được hơn phép mã DES trong một giây.

Diffie và Hellman (1977) đã ước lượng rằng có thể chế được một máy tính chuyên dụng để vét cạn không gian khóa DES trong1/2 ngày với giá cho chiếc máy này là 20 triệu đô la.Giá này được tính toán lại và giảm xuống 200,000$ vào năm 1987. Vì vậy DES đã bị phê bình ngay từ khi ra đời vì có kích thước khóa quá ngắn!

Hiện nay đã có những thiết kế cụ thể cho loại máy tính chuyên dụng phá khóa này dựa trên kỹ thuật xử lý song song tiên tiến và cho biết một thiết bị kiểu này có giá khoảng 10,000$ có thể cho kết quả trong 1 ngày.

***Các dạng tấn công khác:***

*Differential Cryptanalysis*(Phương pháp vi sai): Được công bố lần đầu bởi E. Biham và A. Shamir vào cuối những năm 90 (thế kỷ trước), tuy nhiên thực tế đã được biết đến từ lâu nhưng không công bố bởi IBM và NSA (Cục An ninh Quốc gia Mỹ). Để phá được DES với đầy đủ 16 vòng lặp, tấn công này cần tới bản rõ chọn trước (chosen plaintext). Để có được khối lượng bản rõ này là không thể xảy ra trên thực tế, điều đó cũng cho thấy là DES đã được thiết kế ban đầu để tránh được tấn công này.

*Linear Cryptanalysis*(Phương pháp thử tuyến tính) Phương pháp này được phát hiện bởi Matsui vào năm 1997, và cần bản rõ chọn trước.Tuy nhiên cũng là con số khá lớn nên phá mã tuyến tính cũng không phải là một phương pháp khả thi.

# CÀI ĐẶT DES

#### Phân tích bài toán

Chương trình được cài đặt để thực hiện ứng dụng mã hóa thông tin nhằm mã hóa và giải mã văn bản sử dụng hệ mã DES.

***Đầu vào của chương trình***

* Văn bản:Nhập trực tiếp văn bản cần mã hóa từ giao diện của chương trình.
* Khóa:Nhập theo đúng quy cách mã hóa tức gồm 8 ký tự bất kỳ có phân biệt chữ hoa hay chữ thường.

#### Kết quả demo

###### Thiết kế kiến trúc

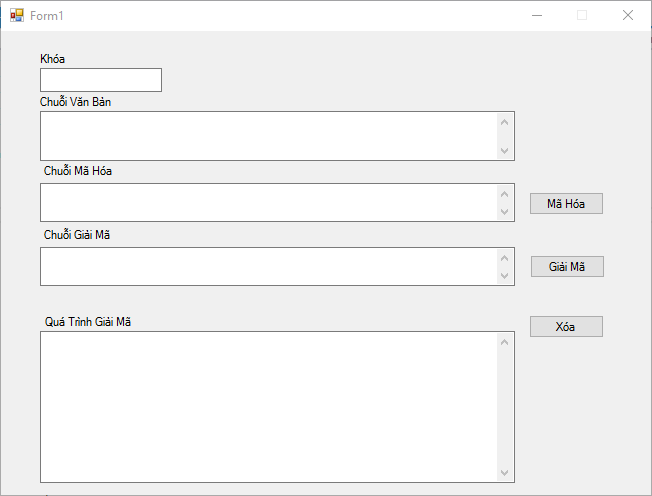
Ngôn ngữ cài đặt: C#

Môi trường cài đặt: Visual Studio 2015

Giao diện: Windows Form

###### Thiết kế giao diện

Khi khởi động chương trình, người dùng sẽ nhìn thấy giao diện như hình dưới:



*Hình 3.1 Giao diện chương trình cài đặt thuật toán DES*

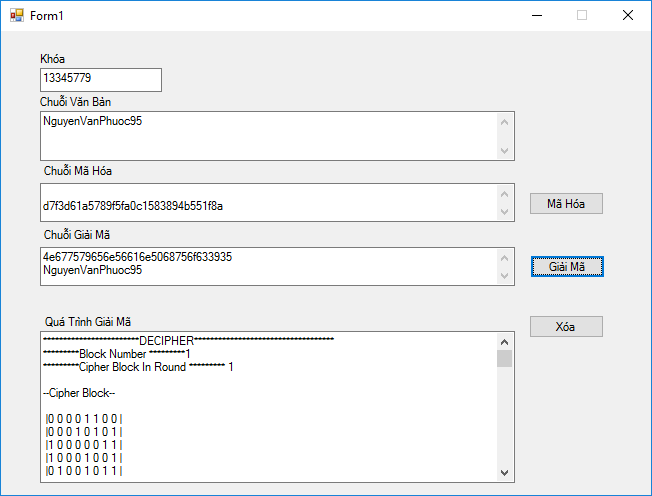
Người dùng nhập Khóa và Chuỗi văn bản muốn giải mã vào ô tương ứng.Lưu ý Khóa và Chuỗi văn bản có phân biệt chữ hoa và chữ thường, chiều dài Khóa tối đa 8 ký tự.

Sau khi nhập xong, người dùng click nút “Mã hóa” để chương trình tiến hành mã hóa chuỗi văn bản nhập vào.

Click nút “Giải mã” để chương trình tiến hành giải mã ngược từ chuỗi đã mã hóa thành chuỗi người dùng đã nhập vào.

Click nút “Xóa” để tiến hành nhập vào và mã hóa chuỗi văn bản khác.

Ô “Quá trình giải mã” thể hiện cụ thể từng bước mã hóa của chương trình.



*Hình 3.2 Giao diện khi mã hóa và giải mã thành công*

TÀI LIỆU THAM KHẢO

[1]Trần Minh Văn, Bài giảng An toàn và bảo mật thông tin(2008), Khoa Công nghệ thông tin, Đại học Nha Trang.

[2]TS.Nguyễn Khanh Văn, Giáo trình An toàn và Bảo mật thông tin(2012), Viện Công nghệ thông tin-Truyền thông, Đại học Bách khoa Hà Nội.

[3] <https://vi.wikipedia.org/wiki/DES_(mã> hóa)

[4] http://des.online-domain-tools.com