ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA CÔNG NGHỆ THÔNG TIN

Lưu Minh Phát - 20127061
Nguyễn Quốc Huy – 20127188
Trần Hoàng Minh Quang – 20127299
Đoàn Duy Phong - 20127405
Lê Cung Tiến - 20127682

BÁO CÁO

[GIÁO VIÊN HƯỚNG DẪN]
PGS.TS. Nguyễn Đình Thúc
Thầy Ngô Đình Hy

NHẬP MÔN MÃ HÓA MẬT MÃ

Thành phố Hồ Chí Minh – 2022

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA CÔNG NGHỆ THÔNG TIN

BÁO CÁO

Tìm hiểu về Differential cryptanalysis.

NHẬP MÔN MÃ HÓA MẬT MÃ

Thành phố Hồ Chí Minh – 2022

LỜI CẢM ƠN

Chúng em xin gửi lời cảm ơn chân thành đến thầy PGS.TS Nguyễn Đình Thúc, thầy Ngô Đình Hy và trường Đại Học Khoa Học Tự Nhiên đã luôn tạo điều kiện tốt nhất để chúng em có thể hoàn thành bài báo cáo một cách tốt đẹp.Trong suốt quá trình thực hiện, thầy đã luôn giúp đỡ và hướng dẫn để chúng em có cơ hội làm việc tốt nhất. Qua đồ án lần này, chúng em đã có thể hiểu thêm về mã hóa mật mã, cũng như có thể thực hiện những yêu cầu đơn giản. Hơn nữa, chúng em còn rèn luyện cho bản thân kĩ năng làm việc nhóm, kĩ năng phân chia công việc và quản lí thời gian tốt hơn.

Bài báo cáo đồ án thực hiện trong khoảng thời gian gần 2 tuần. Trong suốt quá trình thực hiện đồ án, do vẫn còn nhiều hạn chế về kinh nghiệm nên chúng em không thể nào tránh khỏi những sai sót. Chúng em rất mong nhận được những góp ý kiến, chỉ bảo để có thêm kinh nghiệm và cũng mong thầy cô bỏ qua.

Chúng em xin chân thành cảm ơn!

MỤC LỤC

PHÂN CÔNG CÔNG VIỆC

Giáo viên hướng dẫn môn Nhập Môn Mã Hóa Mật Mã

Giáo viên lý thuyết: PGS.TS Nguyễn Đình Thúc, thầy Ngô Đình Hy

Phân chia công việc:

Thành viên	Phân công
Đoàn Duy Phong	1.1 Phương pháp thực hiện tấn công.
Trần Hoàng Minh Quang	1.2 Điều kiện.
Lê Cung Tiến	1.3 Độ phức tạp trên lý thuyết.
Nguyễn Quốc Huy	1.4 Thuật toán dựa trên những định lý nào.
Lưu Minh Phát	2. Đưa ra ví dụ.

GIỚI THIỆU

I. Tìm hiểu.

a. Phương pháp thực hiện tấn công.

- Thám mã vi sai là một mô hình chung của phân tích mật mã có thể áp dụng chung cho các mật mã khối, nhưng nó cũng có thể dùng ở mật mã dòng và hàm băm mật mã. Thám mã vi sai là một dạng tấng công dựa vào các cặp bản rõ có lựa chọn, người tấn công phải thu được các bản mã đã được mã hóa từ tập các bản rõ đã được lựa chọn đó và phân tích giá trị khác nhau giữa 2 bản rõ ảnh hưởng tới giá trị khác nhau giữa 2 bản mã tương ứng.

- **Cách thức tấn công của thám mã vi sai:** Để thực hiện tấn công người tấn công phải có được bản rõ và bản mã của nó và từ đó phân tích giá trị khác nhau của các cặp bản mã tương ứng. Việc thống kê các giá trị khác nhau giữa các cặp bản mã tương ứng phụ thuộc vào S-Box, các giá trị đầu vào và đầu ra được gọi là vi sai. Các mẫu thống kê có được từ việc phân tích các đặc tính tự nhiên của S-Box. Người tấn công phân tích các vi sai của mỗi hộp S-Box.

- **Các bước thực hiện:**

  Bước 1: chọn 1 bản rõ P và tạo ra một bản rõ khác P' = P ^ a (a là chênh lệch đầu vào)

  Bước 2: mã hóa hai bản rõ (P,P') để lấy hai bản mã (C,C')

  Bước 3: kiểm tra C^C'= b (b là sự khác biệt giữa các cắp bản mã và bản rõ hay còn gọi là vi sai).

  Bây giờ các cặp này có thể dùng để thử các tính toán, phân tích khác nhau cho khóa, có thể tìm ra vài khóa hợp lệ và trong các khóa hợp lệ đó chắc chắn có khóa chính xác.

b. Điều kiện.

- Trong mật mã hiện đại, phân tích mật mã vi sai là một ví dụ điển hình của một cuộc tấn công bằng văn bản được chọn. Đây cũng là một kỹ thuật hiếm hoi có thể chuyển đổi từ văn bản gốc đã chọn sang văn bản gốc đã biết (do nó hoạt động với các cặp văn bản).
- Khi ứng dụng thám mã vi sai (Diferential Cryptanalysis) phải dựa trên bản rõ được lựa chọn. Để thực hiện được như vậy, người tấn công phải dựa trên tập các bản rõ được lựa chọn để thu được các bản mã đã được mã hóa. Từ đó, phân tích sự khác nhau giữa hai bản rõ ảnh hưởng như thế nào đến sự khác nhau giữa hai bản mã tương ứng. Hay hiểu là, phương pháp cơ bản của thám mã vi sai là sử dụng các cặp bản mã, bản rõ có liên quan dựa vào giá trị khác nhau là hằng số.

- Phân tích thám mã vi sai thường yêu cầu một bội số nhỏ 1/p của các cặp bản rõ được chọn, khi sử dụng một đặc trưng với xác suất p, để đảm bảo rằng có đủ nhiều cặp đúng xuất hiện trong dữ liệu.
- Cần có được sự khác nhau của các cặp bản rõ cần thiết, người tấn công thực hiện tính toán giá trị khác nhau giữa các cặp bản mã tương ứng nhằm khám phá ra được các mẫu thống kê quy luật trong sự phân bổ cảu chúng. Để đưa ra được cặp kết quả giá trị khác nhau, hay được gọi là vi sai.
- Để việc thám mã được diễn ra suôn sẻ, phải đảm bảo các điều kiện sau:
  + Giá trị khác nhau đầu vào cần được lựa chọn cẩn thận,
  + Việc phân tích bên trong của thuật toán cần phải thực hiện đồng thời.
  + Phụ thuộc vào xác suất mà một đặc tính vi phân của mật mã nắm giữ.
  + Xác suất trong việc truyền chênh lệch bên trong các hộp S đơn lẻ hoặc các phần tử phi tuyến có thể có khác trong mật mã càng cao nhất có thể.
  + Số lượng hộp S đang hoạt động hoặc các phần tử phi tuyến tính khác trong bất kỳ đặc tính vi phân nào của mật mã phải càng ít càng tốt.
  ⇨ Theo dõi đặc tính vi sai là một đường dẫn vi sai với xác suất hiện cao thồn qua các giai đoạn của việc mã hóa.
- Đối với các hoạt động phi tuyến tính (chẳng hạn như hộp S), chúng ta cũng có thể nghiên cứu sự cải tiến của những khác biệt. Mà khi chênh lệch của đầu vào bằng 0, hai đầu vào bằng nhau và dẫn đến hai đầu ra cũng bằng nhau, có chênh lệch bằng 0.
- Khi chênh lệch đầu vào khác 0, chúng ta không thể dự đoán chênh lệch đầu ra, vì nó có thể có nhiều chênh lệch đầu ra khác nhau đối với bất kỳ chênh lệch đầu vào nào. Tuy nhiên, có thể dự đoán thông tin thống kê về chênh lệch đầu ra với chênh lệch đầu vào.

c. Độ phức tạp trên lý thuyết.

- Differential cryptanalysis là một loại tấn công mật mã được sử dụng để phân tích sự khác biệt trong kết quả của một hàm mật mã cho các cặp tin nhắn đầu vào khác nhau theo cách đặc biệt. Mục tiêu của tấn công này là xác định trạng thái bên trong của hàm mật mã, điều này có thể được sử dụng để khôi phục khóa bí mật đang được sử dụng bởi hàm.
- Độ phức tạp của phân tích đại lượng phụ thuộc vào hàm mật mã cụ thể đang bị tấn công, cũng như đặc điểm của sự khác biệt được sử dụng trong tấn công. Chung chung, phân tích đại lượng được coi là một phương pháp tấn công thực tế, có nghĩa là nó có thể được sử dụng để phá vỡ hệ thống mật mã thực tế trong khoảng thời gian hợp lý. Tuy nhiên, độ phức tạp chính xác của tấn công sẽ phụ thuộc vào hệ thống đang bị tấn công và tài nguyên có sẵn của người tấn công.

- Một yếu tố khác ảnh hưởng đến độ phức tạp của Differential cryptanalysis là số vòng chức năng mật mã bị tấn công. Nói chung, hàm càng có nhiều vòng thì càng khó thực hiện phân tích mật mã vi phân. Điều này là do mỗi vòng của hàm sẽ tăng thêm độ phức tạp cho trạng thái bên trong của hàm, khiến việc xác định khóa bí mật trở nên khó khăn hơn.

- Nhìn chung, độ phức tạp của Differential cryptanalysis có thể rất khác nhau tùy thuộc vào chức năng mật mã cụ thể bị tấn công và các tài nguyên sẵn có cho kẻ tấn công. Nó thường được coi là một phương pháp tấn công thực tế đối với nhiều hệ thống mật mã trong thế giới thực, nhưng độ phức tạp chính xác của cuộc tấn công sẽ phụ thuộc vào các đặc điểm cụ thể của hệ thống bị tấn công.

- Differential cryptanalysis được đánh giá là phương pháp đầu tiên cho phép bẻ khóa DES với mức độ phức tạp của bài toán nhỏ hơn 2^55 . Theo Biham, với sự trợ giúp của phương pháp đã cho có thể dẫn đến thám mã thành công DES với độ phức tạp 2^47 , nhưng cần phải có 2^47 bản rõ chọn lựa.

- Mặc dù 2^47 rõ ràng rất nhỏ hơn rất nhiều so với 2^55 , nhưng sự cần thiết phải có 2^47 bản rõ chọn lựa, làm cho phương pháp thám mã đã cho trở nên thuần túy lý thuyết

d. Thuật toán dựa trên những định lý nào.

- Cụ thể, thuật toán thám mã vi sai dựa trên nguyên tắc có thể tính toán sự khác biệt giữa các giá trị đầu ra của một hàm mật mã khi đầu vào có một số lượng nhỏ khác biệt.
- Khi điều này khả thi, người tấn công có thể sử dụng những điểm khác biệt này để phỏng đoán thông tin về cách hoạt động và có thể sử dụng thông tin này để tấn công hệ thống.
    + Để áp dụng thuật toán thám mã vi sai, người tấn công cần có một số dữ liệu đầu vào và đầu ra cho chức năng và cần tìm ra một số lượng nhỏ sự khác biệt giữa các đầu vào này.
    + Sau đó, kẻ tấn công tính toán sự khác biệt giữa các đầu ra tương ứng và sử dụng những khác biệt này để suy ra thông tin về cách hoạt động.
- Thám mã vi sai có thể được áp dụng cho nhiều loại hệ thống mật mã, bao gồm mã hóa khối (Block Cipher) và mã hóa luồng (Stream Cipher). Nó đã được sử dụng để tấn công thành công một số hệ thống mật mã và kết quả là các nhà thiết kế mật mã đã phải thực hiện các biện pháp mới để bảo vệ hệ thống của mình.

II. Ví dụ.

0b-0h; 1b-1h; 2b-2h; 3b-3h; 4b-4h; 5b-5h; 6b-6h; 7b-7h; 8b-8h; 9b-9h; 10b-Ah; 11b-Bh; 12b-Ch; 13b-Dh; 14b-Eh; 15b-Fh

Ví dụ mô tả cách tấn công Differential Cryptanalysis:

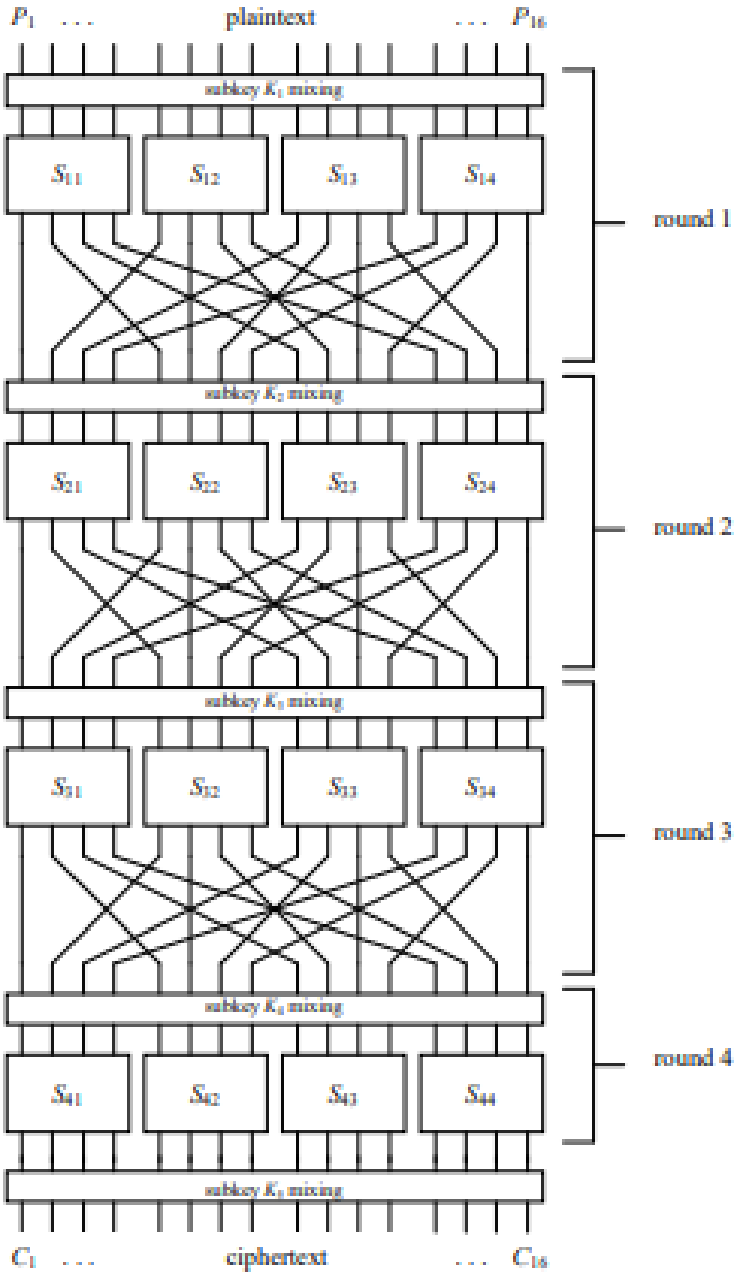

Figure 1. Basic Substitution-Permutation Network (SPN) Cipher

Kích thước S-Box là 4x4 (hex)

Input (X)	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
Output	E	4	D	1	2	F	B	8	3	A	6	C	D	9	0	7

(S(X))																

$$S(X \oplus \Delta X) \oplus S(X)$$

❖ VD1:

X=0000(b);

ΔX=1011

$X \oplus \Delta X$ = 0000(b) ⊕ 1011(b) = 1011(b)

S(1011(b) = B(h)) = C(h) = 1100(b)

S(X = 0(h)) = E(h) = 1110(b)

1011(b) ⊕ 1110(b) = 0010(b) = 2(h)

❖ VD2:

X=0001;

ΔX=1011;

$X \oplus \Delta X$ = 1010(b) = A(h)

S(A(h)) = 6(h) = 0110(b)

S(X = 1(h)) = 4(h) = 0100(b)

0110(b) ⊕ 0100(b) = 0010(b) = 2(h)

X Input	Y Output	ΔY		
		ΔX=1011	ΔX=1000	ΔX=0100
0000	1110	0010	1101	1100
0001	0010	0010	1110	1011
0010	1101	0111	1011	0110
0011	0001	0010	1101	1001
0100	0010	0101	1111	1100
0101	1111	1111	0110	1011
0110	1011	0010	1011	0110
0111	1000	1101	1111	1001

1000	0011	0010	1101	0110
1001	1010	0111	1110	0011
1010	0110	0010	1011	0110
1011	1100	0010	1101	1011
1100	0101	1101	0111	0110
1101	1001	0010	0110	0011
1110	0000	1111	1011	0110
1111	0111	0101	1111	1011

		Output Difference															
		0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
Input Difference	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	2	0	0	0	2	0	2	4	0	4	2	0	0
	2	0	0	0	2	0	6	2	2	0	2	0	0	0	0	2	0
	3	0	0	2	0	2	0	0	0	0	4	2	0	2	0	0	4
	4	0	0	0	2	0	0	6	0	0	2	0	4	2	0	0	0
	5	0	4	0	0	0	2	2	0	0	0	4	0	2	0	0	2
	6	0	0	0	4	0	4	0	0	0	0	0	0	2	2	2	2
	7	0	0	2	2	2	0	2	0	0	2	2	0	0	0	0	4
	8	0	0	0	0	0	0	2	2	0	0	0	4	0	4	2	2
	9	0	2	0	0	2	0	0	4	2	0	2	2	2	0	0	0
	A	0	2	2	0	0	0	0	0	6	0	0	2	0	0	4	0
	B	0	0	8	0	0	2	0	2	0	0	0	0	0	2	0	2
	C	0	2	0	0	2	2	2	0	0	0	0	2	0	6	0	0
	D	0	4	0	0	0	0	0	4	2	0	2	0	2	0	2	0
	E	0	0	2	4	2	0	0	0	6	0	0	0	0	0	2	0
	F	0	2	0	0	6	0	0	0	0	4	0	2	0	0	2	0

S_{12}: ΔX=B → ΔY=2 với xác suất là 8/16

S_{23}: ΔX=4 → ΔY=6 với xác suất là 6/16

S_{32}: ΔX=2 → ΔY=5 với xác suất là 6/16

S_{33}: ΔX=2 → ΔY=5 với xác suất là 6/16

Chênh lệch đầu vào của mật mã tương đương với chênh lệch đầu vào của vòng đầu tiên và được cho bởi:

$$\Delta P = \Delta U_1 = [0000\ 1011\ 0000\ 0000]$$

Một lần nữa, chúng ta đang sử dụng Ui để đại diện cho đầu vào của vòng S-box thứ i và Vi để đại diện cho đầu ra của vòng thứ i:

$$\Delta V_1 = [0000\ 0010\ 0000\ 0000]$$

Xem xét cặp khác biệt cho S12 được liệt kê ở trên và theo hoán vị vòng 1 với xác suất 8/16 = 1/2 với sự khác biệt của bản rõ ΔP.:

$$\Delta U_2 = [0000\ 0000\ 0100\ 0000]$$

Bây giờ vi sai vòng thứ hai sử dụng cặp chênh lệch cho S23 dẫn đến kết quả:

$$\Delta V_2 = [0000\ 0000\ 0110\ 0000]$$

Và

$$\Delta U_3 = [0000\ 0010\ 0010\ 0000]$$

Với xác suất 6/16 cho ΔU_2 và xác suất 8/16 × 6/16 = 3/16 cho ΔP. Khi xác định xác suất cho trước chênh lệch bản rõ ΔP, chúng ta đã giả định rằng vi phân của vòng đầu tiên không phụ thuộc vào vi phân của vòng thứ 2 và do đó, xác suất xảy ra của cả hai được xác định bằng tích của các xác suất.

Sau đó, chúng ta có thể sử dụng sự khác biệt cho các hộp S của vòng thứ ba, S32 và S33, và hoán vị của vòng thứ ba để đi đến

$$\Delta V_3 = [0000\ 0101\ 0101\ 0000]$$

Và

$$\Delta U_4 = [0000\ 0110\ 0110\ 0000]$$

Với xác suất của ΔU_3 là $(6/16)^2$ và sự khác biệt của bản rõ ΔP có xác suất là 8/16 × 6/16 × $(6/16)^2$ = 27/1024. Trong đó một lần nữa chúng ta giả định tính độc lập giữa các cặp hộp S khác nhau trong tất cả các vòng.

Trong quá trình tấn công, nhiều cặp bản rõ có ΔP = [0000 1011 0000 0000] sẽ được mã hóa. Nhiều cặp bản rõ có ΔP = [0000 1011 0000 0000] sẽ được mã hóa. Với xác suất cao, 27/1024, đặc điểm khác biệt được minh họa sẽ xảy ra. Chúng ta gọi các cặp như vậy cho ΔP là các cặp đúng. Các cặp khác biệt trong bản rõ mà đặc tính không xuất hiện được gọi là các cặp sai.

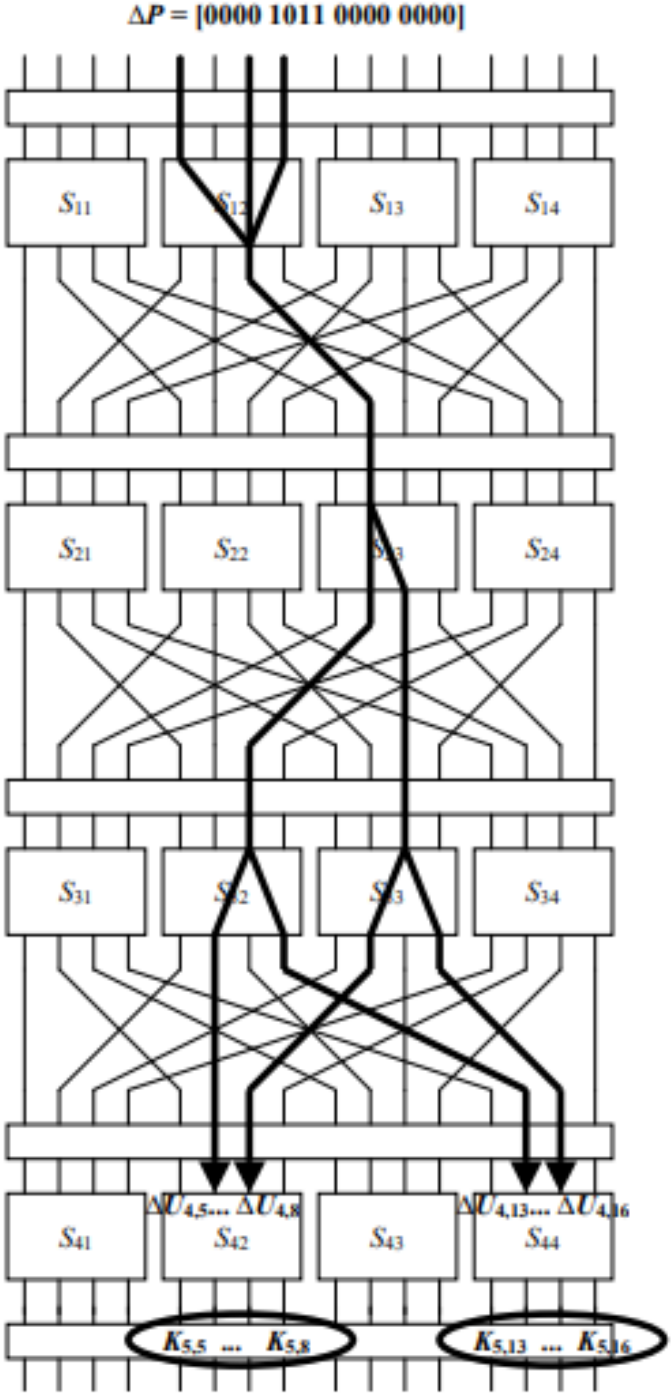

Figure 5. Sample Differential Characteristic

TÀI LIỆU THAM KHẢO

https://websitehcm.com/tham-ma-vi-sai-differential-cryptanalysis/#Cach_thuc_tan_cong_cua_tham_ma_vi_sai

https://www.tutorialspoint.com/what-is-differential-cryptanalysis-in-information-security

https://www.geeksforgeeks.org/differential-and-linear-cryptanalysis/

https://fit.mta.edu.vn/files/FileMonHoc/B%C3%A0i%203_M%E1%BA%ADt%20m%C3%A3%20kh%C3%B3a%20b%C3%AD%20m%E1%BA%ADt.pdf
https://www.cs.bgu.ac.il/~biham/Reports/Talks/Differential_Cryptanalysis_of_DES-like_Cryptosystems.pdf
https://www.crypto-textbook.com/differential-cryptanalysis/