# Chapter 6: Advanced Cryptanalysis

For there is nothing covered, that shall not be revealed;
neither hid, that shall not be known.
— Luke 12:2

没有甚麼掩盖的事不被揭露，也没有甚麼隐藏的事不被人知道。

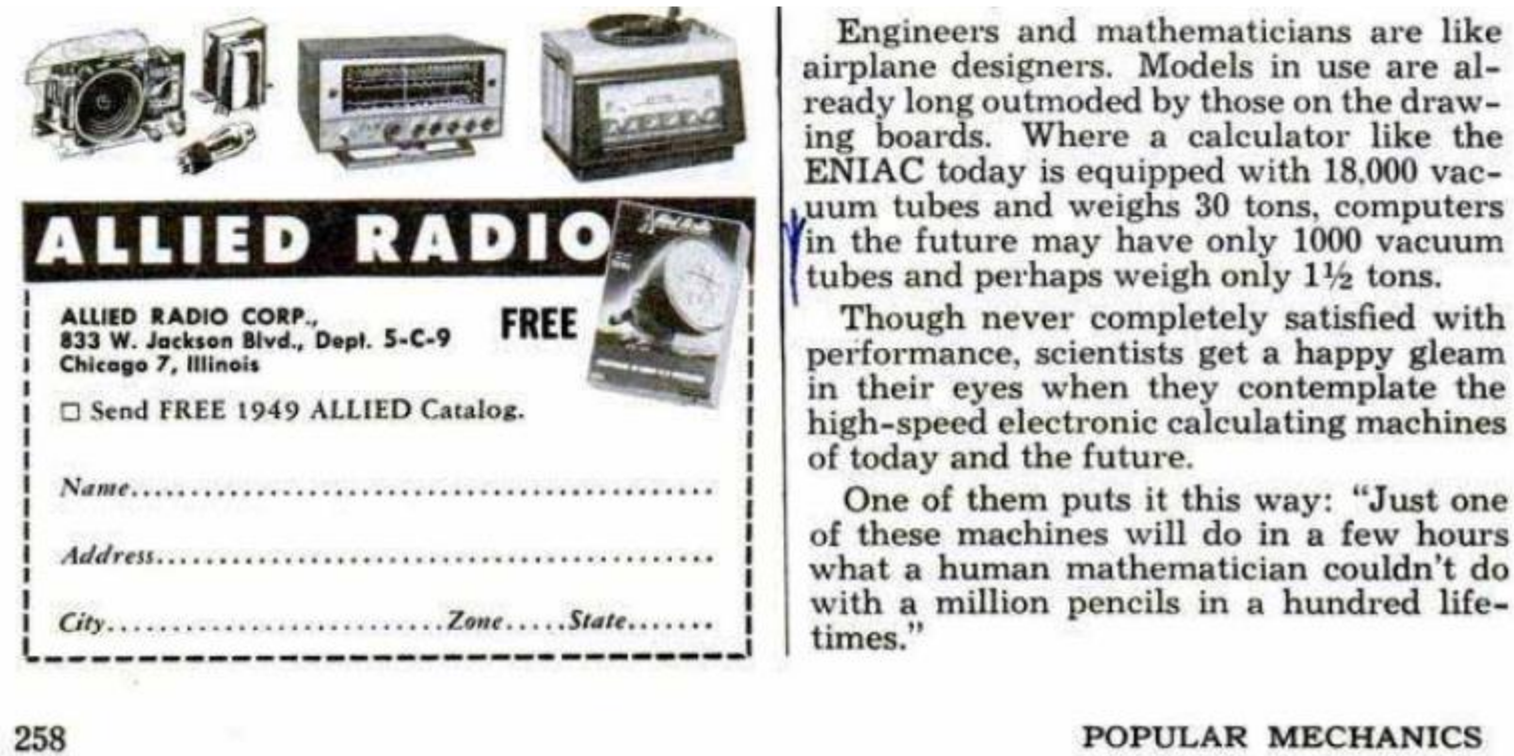
The magic words are squeamish ossifrage ('bone-breaker', from Latin)

— Solution to RSA challenge problem
posed in 1977 by Ron Rivest, who
estimated that breaking the message
would require 40 quadrillion (10^15) years.
It was broken in 1994.

More than 600 volunteers contributed CPU time from about 1,600 machines (two of which were fax machines) over six months.
In 2015, the same RSA-129 number was factored in about one day, with the CADO-NFS open source implementation of **number field sieve**, using a commercial cloud computing service for about **$30**.

# Advanced Cryptanalysis

Where a calculator like the ENIAC today is equipped with 18,000 vacuum tubes and weighs 30 tons, Computers in the future may have only 1000 vacuum tubes and perhaps weigh only 1.5 tons.


ALLIED RADIO

ALLIED RADIO CORP.,
833 W. Jackson Blvd., Dept. 5-C-9
Chicago 7, Illinois

FREE

☐ Send FREE 1949 ALLIED Catalog.

Name.........................................

Address.........................................

City..........................Zone.....State.......

Engineers and mathematicians are like airplane designers. Models in use are already long outmoded by those on the drawing boards. Where a calculator like the ENIAC today is equipped with 18,000 vacuum tubes and weighs 30 tons, computers in the future may have only 1000 vacuum tubes and perhaps weigh only 1½ tons.

Though never completely satisfied with performance, scientists get a happy gleam in their eyes when they contemplate the high-speed electronic calculating machines of today and the future.

One of them puts it this way: "Just one of these machines will do in a few hours what a human mathematician couldn't do with a million pencils in a hundred lifetimes."

258 POPULAR MECHANICS

https://ia600503.us.archive.org/21/items/PopularMechanics1949/Popular_Mechanics_03_1949.pdf

By the early 1960s vacuum tube computers were obsolete, superseded by second-generation transistorized computers.

# RSA公钥密码体制

● 密钥生成

1. 选择两个大素数$p, q$
2. 计算$n=pq$, $\varphi(n)=(p-1)(q-1)$
3. 随机选取$e$ $(e<n)$, 且$\gcd(e, \varphi(n))=1$
4. 采用欧几里得算法，求解$d$, 使得$ed\equiv 1 (\bmod \varphi(n))$
5. 公钥是$(n,e)$, 私钥是$(\varphi(n),d)$

● 加密算法

$$c = m^\wedge e \bmod n$$

● 解密算法

$$m = c^\wedge d \bmod n$$

RONALD (RON) LINN RIVEST
United States – 2002
together with Leonard M. Adleman and Adi Shamir, for their ingenious contribution to making public-key cryptography useful in practice.
https://amturing.acm.org/award_winners/rivest_1403005.cfm

# Advanced Cryptanalysis

*n* (RSA-129) = 1 1438 1625 7578 8886 7669 2357 7997 6146 6120 1021 8296 7212 4236 2562 5618 4293 5706 9352 4573 3897 8305 9712 3563 9587 0505 8989 0751 4759 9290 0268 7954 3541 (129 digits)

*e* = 9007

*C* = 9686 9613 7546 2206 1477 1409 2225 4355 8829 0575 9991 1245 7431 9874 6951 2093 0816 2982 2514 5708 3569 3147 6622 8839 8962 8013 3919 9055 1829 9451 5781 5154

*Scientific American*, August 1977
www.jstor.org/stable/24954008 https://dx.doi.org/10.1038/scientificamerican0877-120

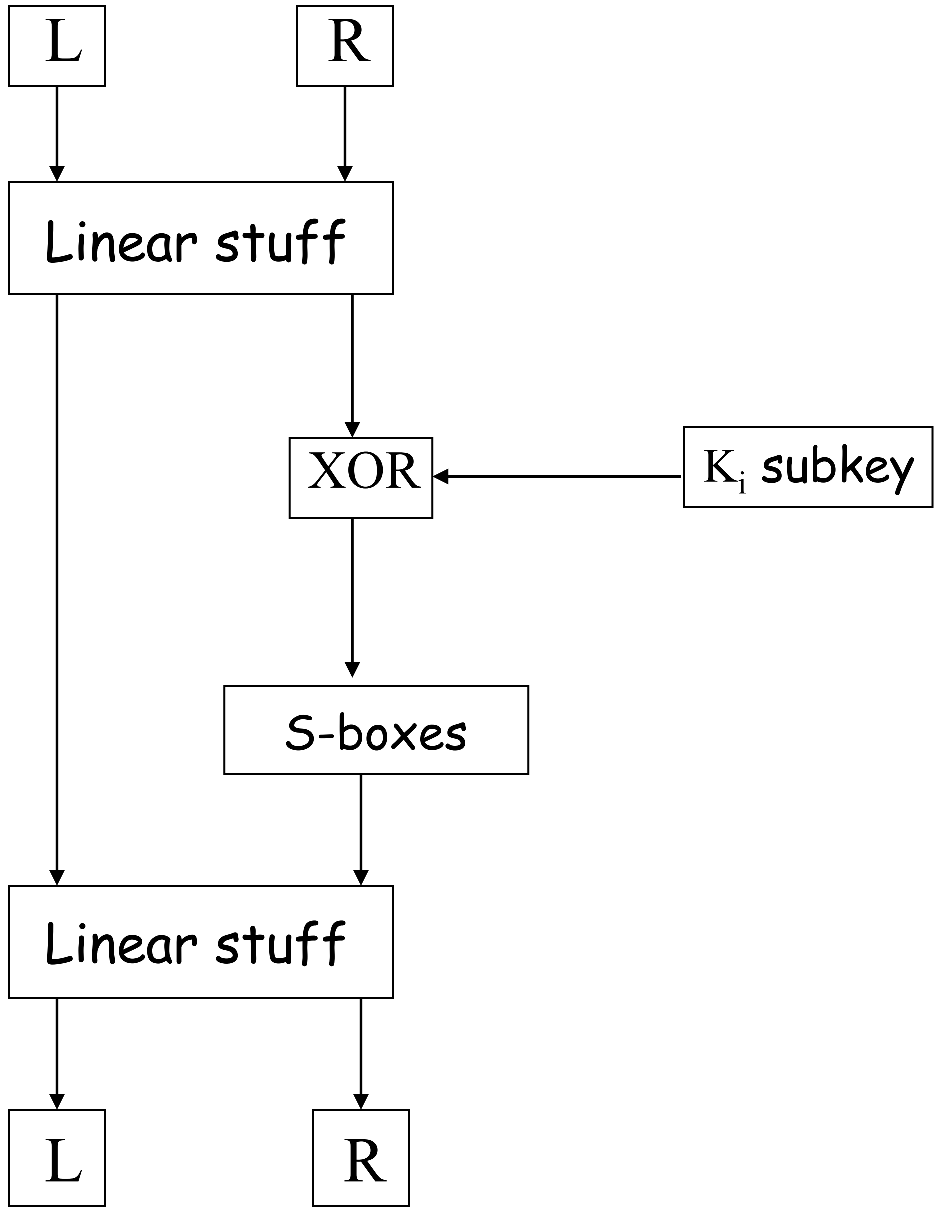
*n* (RSA-129) = 3490 5295 1084 7650 9491 4784 9619 9038 9813 3417 7646 3849 3387 8439 9082 0577 *
3 2769 1329 9326 6709 5499 6198 8190 8344 6141 3177 6429 6799 2942 5397 9828 8533

# Linear and Differential Cryptanalysis

# Introduction

- Both linear and differential cryptanalysis developed to attack DES
- Applicable to other block ciphers
- Differential — Biham and Shamir, 1990
  - Apparently known to NSA in 1970s
  - For analyzing ciphers, not a practical attack
  - A chosen plaintext attack
- Linear cryptanalysis — Matsui, 1993
  - Perhaps not know to NSA in 1970s
  - Slightly more feasible than differential
  - A known plaintext attack

# DES Overview




- 8 S-boxes
- Each S-box maps 6 bits to 4 bits
- Example: S-box 1

# S-box

- S-box是DES加密算法中唯一的非线性结构
- S-box映射：6bit → 4bit
- S-box结构是人为设计的
- 举例：DES中S-box1

(1,2,3,4)-th input bits

(0,5)-th input bits

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | E | 4 | D | 1 | 2 | F | B | 8 | 3 | A | 6 | C | 5 | 9 | 0 | 7 |
| 1 | 0 | F | 7 | 4 | E | 2 | D | 1 | A | 6 | C | B | 9 | 5 | 3 | 4 |
| 2 | 4 | 1 | E | 8 | D | 6 | 2 | B | F | C | 9 | 7 | 3 | A | 5 | 0 |
| 3 | F | C | 8 | 2 | 4 | 9 | 1 | 7 | 5 | B | 3 | E | A | 0 | 6 | D |

# Linear function

$$f(a \oplus b) = f(a) \oplus f(b)$$

Since $\mathbf{expand}$ is linear, if $X_1 \oplus X_2 = 0000\ 0010$ then

$$\begin{aligned} \mathbf{expand}(X_1) \oplus \mathbf{expand}(X_2) &= \mathbf{expand}(X_1 \oplus X_2) \\ &= \mathbf{expand}(0000\ 0010) \\ &= 000000\ 001000. \end{aligned}$$

# Overview of Differential Cryptanalysis

# Differential Cryptanalysis

- Recall that all of DES is **linear** except for the S-boxes
- Differential attack focuses on overcoming this nonlinearity
- Idea is to compare input and output **differences**
- For simplicity, first consider only one round and only one S-box

# Operations of a sample s-box

- Suppose a cipher has 3-bit to 2-bit S-box

| | column | | | |
|---|---|---|---|---|
| row | 00 | 01 | 10 | 11 |
| 0 | 10 | 01 | 11 | 00 |
| 1 | 00 | 10 | 01 | 11 |

- Sbox (abc) is the element in row a column bc
- Example: Sbox(010) = 11

# Operations of a sample S-box

| row | 00 | 01 | 10 | 11 |
|---|---|---|---|---|
| 0 | 10 | 01 | 11 | 00 |
| 1 | 00 | 10 | 01 | 11 |

- Suppose $X_1$=110, $X_2$ = 010, K = 011
- Then $X_1 \oplus$ K=101 and $X_2 \oplus K$ = 001
- $\text{Sbox}(X_1 \oplus K) = \text{Sbox}(1\ 01) = 10$
  $\text{Sbox}(X_2 \oplus K) = \text{Sbox}(0\ 01) = 01$

# Differential Cryptanalysis on an S-box

| row | column 00 | 01 | 10 | 11 |
|---|---|---|---|---|
| 0 | 10 | 01 | 11 | 00 |
| 1 | 00 | 10 | 01 | 11 |

- Suppose
  - **Unknown key:** K
  - **Known inputs:** X = 110, X = 010
  - **Known outputs:** Sbox(X ⊕ K) = 10, Sbox(X ⊕ K) = 01
- Know $X \oplus K \in \{000,101\}$, $X \oplus K \in \{001,110\}$
- Then $K \in \{110,011\} \cap \{011,100\} \Rightarrow K = 011$
- Like a known plaintext attack on S-box

# Differential Cryptanalysis

- Attacking one S-box not very useful!
  - And Trudy can't always see input and output
- To make this work we must do 2 things:

1. Extend the attack to **one round**
   - Have to deal with all S-boxes
   - Choose input so only one S-box "active"
2. Then extend attack to (almost) **all rounds**
   - Output of one round is input to the next round
   - Choose input so output is "good" for the next round

# Differential Cryptanalysis

- We deal with input and output differences
- Suppose we know inputs $X$ and $X$
  - For $X$, the input to S-box is $X \oplus K$
  - For $X$, the input to S-box is $X \oplus K$
  - Key $K$ is unknown
  - **Input difference:** $(X \oplus K) \oplus (X \oplus K) = X \oplus X$
- Input difference is independent of key $K$
- **Output difference:** $Y \oplus Y$ is (almost) input difference to next round
- Goal is to "chain" differences through multiple rounds

# Differential Cryptanalysis

- If we obtain known output difference from known input difference…
  - May be able to chain differences thru rounds
  - It's OK if this only occurs with some probability
- If input difference is 0…
  - …output difference is 0
  - Allows us to make some S-boxes "inactive" with respect to differences

# Overview of Linear Cryptanalysis

# Linear Cryptanalysis

- Like differential cryptanalysis, we target the nonlinear part of the cipher
- But instead of differences, we **approximate** the nonlinearity with **linear equations**
- For DES-like cipher we need to approximate S-boxes by linear functions
- How well can we do this?

# Linear Analysis on an S-box

| row | column 00 | 01 | 10 | 11 |
|---|---|---|---|---|
| 0 | 10 | 01 | 11 | 00 |
| 1 | 00 | 10 | 01 | 11 |

- Input $x_0x_1x_2$ where $x_0$ is row and $x_1x_2$ is column
- Output $y_0y_1$

- Count of 4 is unbiased
- Count of 0 or 8 is best for Trudy

| 相等的次数 | $y_0$ | $y_1$ | $y_0 \oplus y_1$ |
|---|---|---|---|
| 0 | 4 | 4 | 4 |
| $x_0$ | 4 | 4 | 4 |
| $x_1$ | 4 | 6 | 2 |
| $x_2$ | 4 | 4 | 4 |
| $x_0 \oplus x_1$ | 4 | 2 | 2 |
| $x_0 \oplus x_2$ | 0 | 4 | 4 |
| $x_1 \oplus x_2$ | 4 | 6 | 6 |
| $x_0 \oplus x_1 \oplus x_2$ | 4 | 6 | 2 |

# Linear Analysis

| row | column 00 | 01 | 10 | 11 |
|---|---|---|---|---|
| 0 | 10 | 01 | 11 | 00 |
| 1 | 00 | 10 | 01 | 11 |

- For example,

  $y_1 = x_1$

  with probability 3/4
- And

  $y_0 = x_0 \oplus x_2 \oplus 1$

  with probability 1
- And

  $y_0 \oplus y_1 = x_1 \oplus x_2$

  with probability 3/4

| 相等的次数 | $y_0$ | $y_1$ | $y_0 \oplus y_1$ |
|---|---|---|---|
| 0 | 4 | 4 | 4 |
| $x_0$ | 4 | 4 | 4 |
| $x_1$ | 4 | 6 | 2 |
| $x_2$ | 4 | 4 | 4 |
| $x_0 \oplus x_1$ | 4 | 2 | 2 |
| $x_0 \oplus x_2$ | 0 | 4 | 4 |
| $x_1 \oplus x_2$ | 4 | 6 | 6 |
| $x_0 \oplus x_1 \oplus x_2$ | 4 | 6 | 2 |

# Linear Cryptanalysis

- Consider a single DES S-box
- Let $Y = \text{Sbox}(X)$
- Suppose $y_3 = x_2 \oplus x_5$ with high probability
  - i.e., a good linear approximation to output $y_3$
- Can we extend this so that we can solve linear equations for the key?
- As in differential cryptanalysis, we need to "chain" through multiple rounds

# Linear Cryptanalysis of DES

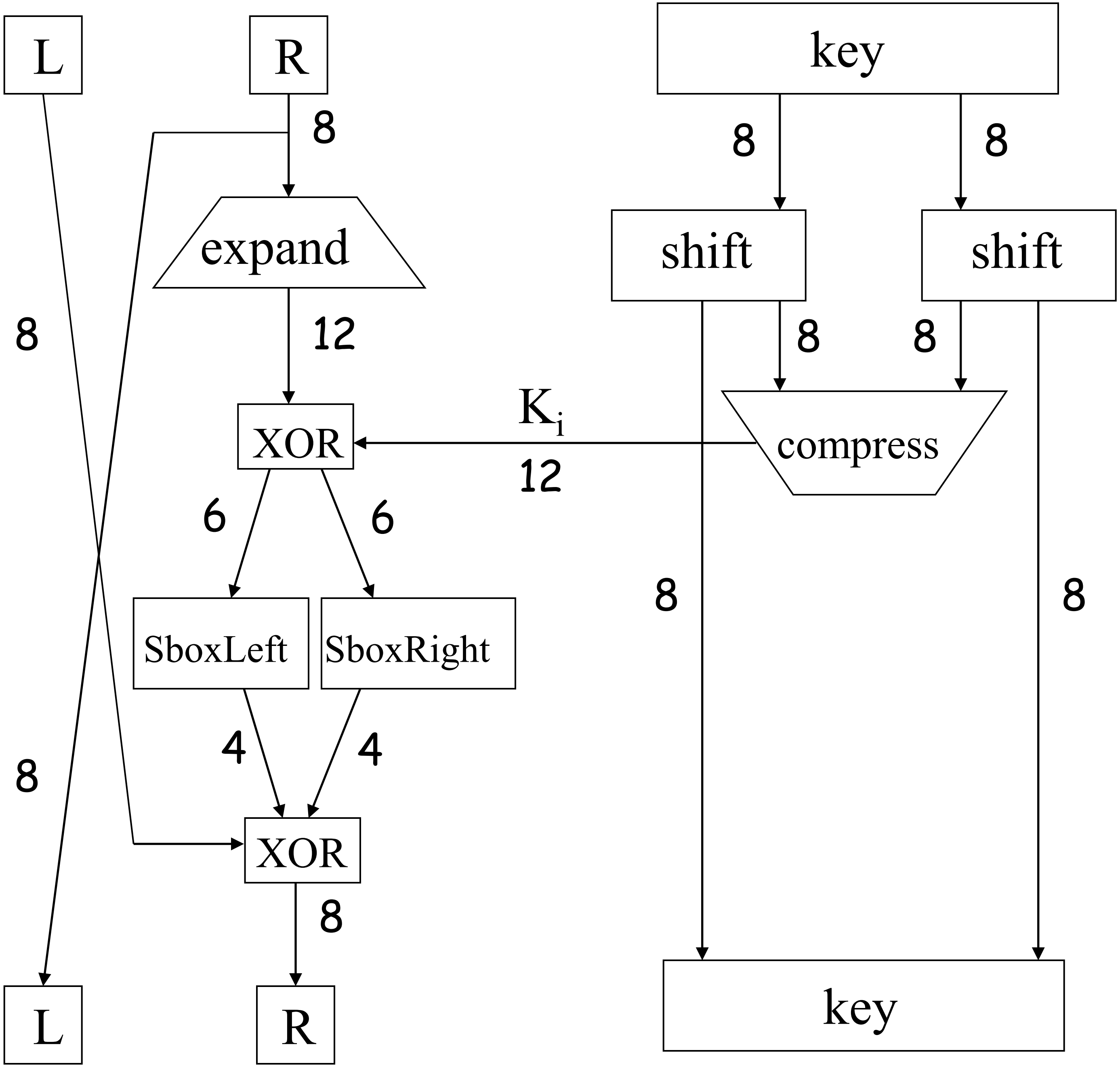
- DES is linear except for S-boxes
- How well can we approximate S-boxes with linear functions?
- DES S-boxes designed so there are no good linear approximations to any one output bit
- But there are linear combinations of output bits that can be approximated by linear combinations of input bits

# Tiny DES

# Tiny DES (TDES)

- A much simplified version of DES
  - 16 bit block
  - 16 bit key
  - 4 rounds
  - 2 S-boxes, each maps 6 bits to 4 bits
  - 12 bit subkey each round
- Plaintext = $(L_0, R_0)$
- Ciphertext = $(L_4, R_4)$
- No useless junk

# One Round of TDES



# Fun Facts on TDES

- TDES is a Feistel Cipher
- $(L_0, R_0)$ = plaintext
- For $i = 1$ to 4

$$L_i = R_{i-1}$$

$$R_i = L_{i-1} \oplus F(R_{i-1}, K_i)$$

- Ciphertext = $(L_4, R_4)$
- $F(R_{i-1}, K_i) = \text{Sboxes}(\text{expand}(R_{i-1}) \oplus K_i)$

  where $\text{Sboxes}(x_0x_1x_2\ldots x_{11}) = (\ \text{SboxLeft}(x_0x_1\ldots x_5),\ \text{SboxRight}(x_6x_7\ldots x_{11})\ )$

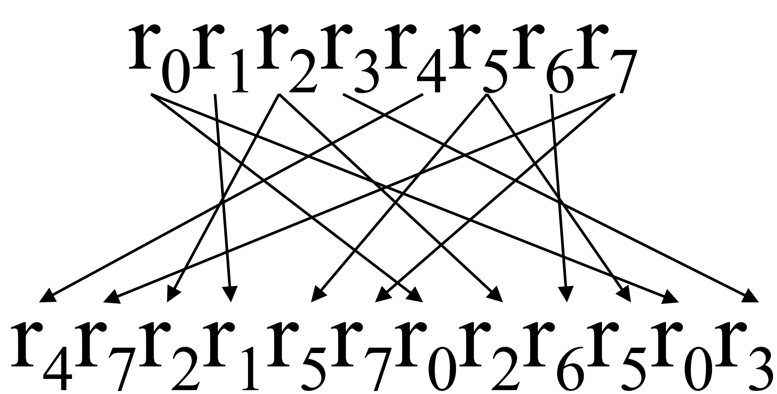
德籍密码学家Horst Feistel在IBM工作的时发明Feistel Cipher
www.cnblogs.com/flydean/p/14888005.html

# Key Schedule of TDES（密码分配/编排）

- Key: $K = k_0k_1k_2k_3k_4k_5k_6k_7k_8k_9k_{10}k_{11}k_{12}k_{13}k_{14}k_{15}$
- Subkey
  - Left: $k_0k_1\ldots k_7$ rotate left 2, select 0,2,3,4,5,7
  - Right: $k_8k_9\ldots k_{15}$ rotate left 1, select 9,10,11,13,14,15
- Subkey $K_1 = k_2k_4k_5k_6k_7k_1k_{10}k_{11}k_{12}k_{14}k_{15}k_8$
- Subkey $K_2 = k_4k_6k_7k_0k_1k_3k_{11}k_{12}k_{13}k_{15}k_8k_9$
- Subkey $K_3 = k_6k_0k_1k_2k_3k_5k_{12}k_{13}k_{14}k_8k_9k_{10}$
- Subkey $K_4 = k_0k_2k_3k_4k_5k_7k_{13}k_{14}k_{15}k_9k_{10}k_{11}$

# TDES expansion permutation

- Expansion permutation: 8 bits to 12 bits




- We can write this as

$$\text{expand}(r_0r_1r_2r_3r_4r_5r_6r_7) = r_4r_7r_2r_1r_5r_7r_0r_2r_6r_5r_0r_3$$

# TDES S-boxes

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | C | 5 | 0 | A | E | 7 | 2 | 8 | D | 4 | 3 | 9 | 6 | F | 1 | B |
| 1 | 1 | C | 9 | 6 | 3 | E | B | 2 | F | 8 | 4 | 5 | D | A | 0 | 7 |
| 2 | F | A | E | 6 | D | 8 | 2 | 4 | 1 | 7 | 9 | 0 | 3 | 5 | B | C |
| 3 | 0 | A | 3 | C | 8 | 2 | 1 | E | 9 | 7 | F | 6 | B | 5 | D | 4 |

- Right S-box
- SboxRight

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 6 | 9 | A | 3 | 4 | D | 7 | 8 | E | 1 | 2 | B | 5 | C | F | 0 |
| 1 | 9 | E | B | A | 4 | 5 | 0 | 7 | 8 | 6 | 3 | 2 | C | D | 1 | F |
| 2 | 8 | 1 | C | 2 | D | 3 | E | F | 0 | 9 | 5 | A | 4 | B | 6 | 7 |
| 3 | 9 | 0 | 2 | 5 | A | D | 6 | E | 1 | 8 | B | C | 3 | 4 | 7 | F |

- Left S-box
- SboxLeft

# Differential Cryptanalysis of TDES

# SboxRight 性质1

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | C | 5 | 0 | A | E | 7 | 2 | 8 | D | 4 | 3 | 9 | 6 | F | 1 | B |
| 1 | 1 | C | 9 | 6 | 3 | E | B | 2 | F | 8 | 4 | 5 | D | A | 0 | 7 |
| 2 | F | A | E | 6 | D | 8 | 2 | 4 | 1 | 7 | 9 | 0 | 3 | 5 | B | C |
| 3 | 0 | A | 3 | C | 8 | 2 | 1 | E | 9 | 7 | F | 6 | B | 5 | D | 4 |

- For $X$ and $X$ suppose $X \oplus X = 00\ 1000$
- Then $\text{SboxRight}(X) \oplus \text{SboxRight}(X) = 0010$ with probability $3/4$

# SboxRight 性质2

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | C | 5 | 0 | A | E | 7 | 2 | 8 | D | 4 | 3 | 9 | 6 | F | 1 | B |
| 1 | 1 | C | 9 | 6 | 3 | E | B | 2 | F | 8 | 4 | 5 | D | A | 0 | 7 |
| 2 | F | A | E | 6 | D | 8 | 2 | 4 | 1 | 7 | 9 | 0 | 3 | 5 | B | C |
| 3 | 0 | A | 3 | C | 8 | 2 | 1 | E | 9 | 7 | F | 6 | B | 5 | D | 4 |

- For $X$ and $X$ suppose $X \oplus X = 00\ 0000$
- Then $\text{SboxRight}(X) \oplus \text{SboxRight}(X) = 0000$ with probability $1$

# Differential cryptanaysis of TDES

- The game plan…
- Select P and P so that

  $P \oplus P = 0000\ 0000\ 0000\ 0010 = 0x0002$
- Note that P and P differ in exactly 1 bit
- Let's carefully analyze what happens as these plaintexts are encrypted with TDES

# TDES

- If $Y \oplus Y = 001000$ then with probability $3/4$ $\text{SboxRight}(Y) \oplus \text{SboxRight}(Y) = 0010$
- $Y \oplus Y = 001000 \Rightarrow (Y \oplus K) \oplus (Y \oplus K) = 001000$
- If $Y \oplus Y = 000000$ then for any S-box, we have $\text{Sbox}(Y) \oplus \text{Sbox}(Y) = 0000$
- Difference of $(0000\ 0010)$ is expanded by TDES expand perm to diff. $(000000\ 001000)$
- **The bottom line:** If $X \oplus X = 00000010$ then $F(X, K) \oplus F(X, K) = 00000010$ with prob. $3/4$

# TDES

- From the previous slide
  - Suppose $R \oplus R = 0000\ 0010$
  - Suppose $K$ is unknown key
  - Then with probability $3/4$

    $F(R,K) \oplus F(R,K) = 0000\ 0010$
- The bottom line? With probability 3/4…
  - Input to next round same as current round
- So we can chain thru multiple rounds

# TDES第一轮差分推导

- 根据上述分析，可以通过以下方案攻击TDES
- Select P and P so that

  $P \oplus P = 0000\ 0000\ \mathbf{0000\ 0010} = 0x0002$
- Note that P and P differ in exactly 1 bit
- Let’s carefully analyze what happens as these plaintexts are encrypted with TDES

# TDES第一轮差分推导

- 对于第一轮加密
- 明文为P $= (L_0, R_0)$和P $= (L_0', R_0')$

$$L_0 \oplus L_0' = \text{0x00}$$

$$R_0 \oplus R_0' = \text{0x02}$$

- 第一轮加密结果为

$$(L_1, R_1) = (R_0, L_0 \oplus F(R_0, K_1))$$

$$(L_1', R_1') = (R_0', L_0' \oplus F(R_0', K_1))$$

- 第一轮加密结果差分得

$$L_1 \oplus L_1' = R_0 \oplus R_0' = \text{0x02}$$

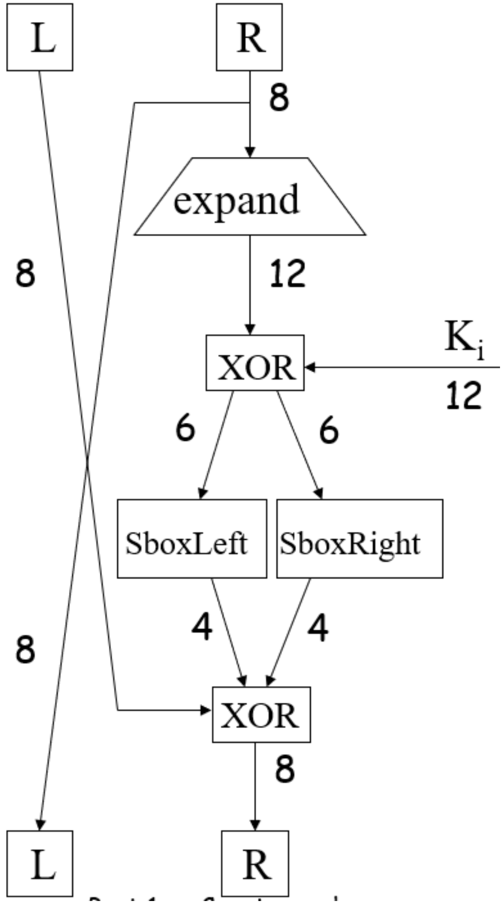
$$R_1 \oplus R_1' = F(R_0, K_1) \oplus F(R_0', K_1)$$

# TDES第一轮差分推导

$$
\begin{aligned}
R_1 \oplus R_1' &= F(R_0, K_1) \oplus F(R_0', K_1) \\
&= \mathrm{Sboxes}(\mathrm{expand}(R_0) \oplus K_1) \oplus \\
&\quad\ \mathrm{Sboxes}(\mathrm{expand}(R_0') \oplus K_1)
\end{aligned}
$$

$$
\begin{aligned}
&= \big(\mathrm{SboxesL}(A \oplus K_1), \mathrm{SboxesR}(B \oplus K_1)\big) \oplus \\
&\quad\ \big(\mathrm{SboxesL}(A' \oplus K_1), \mathrm{SboxesR}(B' \oplus K_1)\big)
\end{aligned}
$$

$$
\begin{aligned}
&= \big(\mathrm{SboxesL}(A \oplus K_1) \oplus \mathrm{SboxesL}(A' \oplus K_1)\big), \\
&\quad\ \big(\mathrm{SboxesR}(B \oplus K_1) \oplus \mathrm{SboxesR}(B' \oplus K_1)\big)
\end{aligned}
$$




- A为expand($R_0$)左半部分，B为expand($R_0$)右半部分

# TDES第一轮差分推导

- 由于expand为线性函数

$$\text{expand}(R_0) \oplus \text{expand}(R_0') = \text{expand}(R_0 \oplus R_0')$$
$$= 000000\ \mathbf{001000}$$

- $A \oplus A' = 000000$ and $B \oplus B' = 001000$

- 根据性质1： $B \oplus K_1 \oplus B' \oplus K_1 = 001000$

$\text{SboxesR}(B \oplus K_1) \oplus \text{SboxesR}(B' \oplus K_1) = 0010$ with probability 3/4

- 根据性质2： $A \oplus K_1 \oplus A' \oplus K_1 = 000000$

$\text{SboxesL}(A \oplus K_1) \oplus \text{SboxesL}(A' \oplus K_1) = 0000$ with probability 1

# TDES第一轮差分推导

- 综上，第一轮加密差分结果为

$$L_1 \oplus L_1' = R_0 \oplus R_0' = 0x02$$

$$R_1 \oplus R_1' = F(R_0, K_1) \oplus F(R_0', K_1) = 0x02 \quad (3/4\text{概率})$$

- 通过同样的方法，可以计算出后续所有轮次的差分结果。

# Differential Attack of TDES

By the choice of $P$ and $\tilde{P}$, we have

$$R_0 \oplus \tilde{R}_0 = 0000\ 0010 \quad \text{and} \quad L_0 \oplus \tilde{L}_0 = 0000\ 0000.$$

$$R_1 \oplus \tilde{R}_1 = 0000\ 0010 \quad \text{with probability } 3/4.$$

$$\begin{aligned}
R_2 \oplus \tilde{R}_2 &= (L_1 \oplus F(R_1, K_2)) \oplus (\tilde{L}_1 \oplus F(\tilde{R}_1, K_2)) \\
&= (L_1 \oplus \tilde{L}_1) \oplus (F(R_1, K_2) \oplus F(\tilde{R}_1, K_2)) \\
&= (R_0 \oplus \tilde{R}_0) \oplus (F(R_1, K_2) \oplus F(\tilde{R}_1, K_2)) \\
&= 0000\ 0010 \oplus 0000\ 0010 \\
&= 0000\ 0000 \text{ with probability } (3/4)^2 = 9/16 = 0.5625.
\end{aligned}$$

Mark Stamp
Information Security: Principles and Practice, 2nd Edition
https://dx.doi.org/10.1002/9781118027974

# Differential Attack of TDES

- Select $P$ and $P$ with $P \oplus P = 0x0002$

| | | |
|---|---|---|
| $(L_0, R_0) = P$ | $(L_0, R_0) = P$ | $P \oplus P = 0x0002$ |
| $L_1 = R_0$<br>$R_1 = L_0 \oplus F(R_0,K_1)$ | $L_1 = R_0$<br>$R_1 = L_0 \oplus F(R_0,K_1)$ | with probability 3/4<br>$(L_1,R_1) \oplus (L_1,R_1) = 0x0202$ |
| $L_2 = R_1$<br>$R_2 = L_1 \oplus F(R_1,K_2)$ | $L_2 = R_1$<br>$R_2 = L_1 \oplus F(R_1,K_2)$ | with probability $(3/4)^2$<br>$(L_2,R_2) \oplus (L_2,R_2) = 0x0200$ |
| $L_3 = R_2$<br>$R_3 = L_2 \oplus F(R_2,K_3)$ | $L_3 = R_2$<br>$R_3 = L_2 \oplus F(R_2,K_3)$ | with probability $(3/4)^2$<br>$(L_3,R_3) \oplus (L_3,R_3) = 0x0002$ |
| $L_4 = R_3$<br>$R_4 = L_3 \oplus F(R_3,K_4)$ | $L_4 = R_3$<br>$R_4 = L_3 \oplus F(R_3,K_4)$ | with probability $(3/4)^3$<br>$(L_4,R_4) \oplus (L_4,R_4) = 0x0202$ |
| $C = (L_4,R_4)$ | $C = (L_4,R_4)$ | $C \oplus C = 0x0202$ |

# Differential Attack of TDES

根据以上计算结果，可以得到：

- 选择初始明文： $P \oplus P = 0x0002$
- 四轮加密结果： $(L_4,R_4) \oplus (L_4,R_4) = 0x0202$ $((3/4)^3$概率)

即

$L_4 \oplus L_4 = 0x02$

$R_4 \oplus R_4 = L_3 \oplus L_3 \oplus F(R_3,K_4) \oplus F(R_3,K_4)$

$= F(L_4,K_4) \oplus F(L_4,K_4)=0x02$

# Differential Attack of TDES

- $F(L_4, K_4) \oplus F(L_4, K_4)=0x02$

按照同样的方式展开

$SboxesL(A \oplus K_4) \oplus SboxesL\ (A \oplus K_4)=0000$

$SboxesR(B \oplus K_4) \oplus SboxesR\ (B \oplus K_4)=0010$

假设$L_4 = l_0l_1l_2l_3l_4l_5l_6l_7$, $L_4 = l_0l_1l_2l_3l_4l_5l_6l_7$

A为expand($L_4$)左半部分，B为expand($L_4$)右半部分

$A= l_4l_7l_2l_1l_5l_7$ $B= l_0l_2l_6l_5l_0l_3$

$A= l_4l_7l_2l_1l_5l_7$ $B= l_0l_2l_6l_5l_0l_3$

$K_4 = k_0k_2k_3k_4k_5k_7\ k_{13}k_{14}k_{15}k_9k_{10}k_{11}$

# Differential Attack of TDES

对于SboxesL($l_4l_7l_2l_1l_5l_7 \oplus k_0k_2k_3k_4k_5k_7$) $\oplus$

SboxesL ($l_4l_7l_2l_1l_5l_7 \oplus k_0k_2k_3k_4k_5k_7$)=0000

根据性质2

- For $X$ and $X$ suppose $X \oplus X = 000000$
- Then SboxRight($X$) $\oplus$ SboxRight($X$) = 0000 with probability 1

对于任意$k_0k_2k_3k_4k_5k_7$都满足以上条件

所以无法从SboxesL中提取到任何信息

# Differential Attack of TDES

对于SboxesR($l_0l_2l_6l_5l_0l_3 \oplus k_0k_2k_3k_4k_5k_7$) $\oplus$
SboxesR ($l_0l_2l_6l_5l_0l_3 \oplus k_0k_2k_3k_4k_5k_7$)=0010

根据性质1

- For $X$ and $X$ suppose $X \oplus X = 001000$
- Then $\mathrm{SboxRight}(X) \oplus \mathrm{SboxRight}(X) = 0010$ with probability 3/4

对于和原密钥一样的$k_0k_2k_3k_4k_5k_7$一定符合以上条件
对于和原密钥不一样的$k_0k_2k_3k_4k_5k_7$有一定概率符合以上条件
可以增大选择明文的数量降低非原密钥的概率，从而找到原密钥。

# Algorithm to recover subkey bits

## Algorithm to find right 6 bits of subkey $K_4$

```
count[i] = 0, for i = 0,1,. . .,63
for i = 1 to iterations
    Choose P and P with P ⊕ P = 0x0002
    Obtain corresponding C and C
     if C ⊕ C = 0x0202
        for K = 0 to 63
            if 0010 == (SBoxRight( l0l2l6l5l0l3 ⊕K) ⊕ SBoxRight( l0l2l6l5l0l3 ⊕K))
                ++count[K]
            end if
        next K
     end if
next i
```

All K with max count[K] are possible (partial) $K_4$

# Test sample 1

- Choose 100 pairs $P$ and $P$ with $P \oplus P = 0x0002$
- Found 47 of these give $C \oplus C = 0x0202$
- Tabulated counts for these 47
  - Max count of 47 for each
    $K \in \{000001, 001001, 110000, 111000\}$
  - No other count exceeded 39
- Implies that $K_4$ is one of 4 values, that is,

$$k_{13}k_{14}k_{15}k_9k_{10}k_{11} \in \{000001, 001001, 110000, 111000\}.$$

So, $k_{13}k_{14}k_9k_{10}k_{11} \in \{00001, 11000\}$

- Actual key is $K = 1010\ 1001\ 1000\ 0111$

$\quad 0\ 1\ \ 2\ 3 \quad 4\ 5\ 6\ 7 \quad 8\ 9\ 10\ 11 \quad 12\ 13\ 14\ 15$

# Test sample 2

- Choose 100 pairs P and P with P ⊕ P= 0x0002
- Found 51 of these give C ⊕ C = 0x0202
- Tabulated counts for these 51
  - Max count of 51 for each

    K ∈ {010100, 011100, 100101, 101101}
- Implies that $K_4$ is one of 4 values, that is,

  $k_{13}k_{14}k_{15}k_9k_{10}k_{11} \in \{010100, 011100, 100101, 101101\}$.
  So, $k_{13}k_{14}k_9k_{10}k_{11} \in \{01100, 10101\}$
- Actual key is K=0100 0001 0101 1101

  0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

# Differential Attack of TDES

- To complete the recovery of K, we could exhaustively search over the remaining $2^{16-5}=2^{11}$ unknown key bits, and for each of these try both of the possibilities. For each of these $2^{12}$ putative keys K, we would try to decrypt the ciphertext, and for the correct key, we will recover the plaintext.
- The total expected work to recover the entire key K by this method is about $2^{11}$ encryptions, plus the work required for the differential attack, which is insignificant in comparison. As a result, we can recover the entire 16-bit key with a work factor of about $2^{11}$ encryptions, which is much better than an exhaustive key search, since an exhaustive search has an expected work of $2^{15}$ encryptions. This shows that a shortcut attack exists, and as a result TDES is insecure.

# Linear Cryptanalysis of TDES

# Linear Approximation of Left S-Box

- SboxLeft 性质

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 6 | 9 | A | 3 | 4 | D | 7 | 8 | E | 1 | 2 | B | 5 | C | F | 0 |
| 1 | 9 | E | B | A | 4 | 5 | 0 | 7 | 8 | 6 | 3 | 2 | C | D | 1 | F |
| 2 | 8 | 1 | C | 2 | D | 3 | E | F | 0 | 9 | 5 | A | 4 | B | 6 | 7 |
| 3 | 9 | 0 | 2 | 5 | A | D | 6 | E | 1 | 8 | B | C | 3 | 4 | 7 | F |

- Notation: $y_0y_1y_2y_3 = \text{SboxLeft}(x_0x_1x_2x_3x_4x_5)$
- For this S-box, $y_1=x_2$ and $y_2=x_3$ both with probability ¾
- *Can we "chain" this thru multiple rounds?*

# TDES Linear Relations

- Recall that the expansion perm is

  $\text{expand}(r_0r_1r_2r_3r_4r_5r_6r_7) = r_4r_7\mathbf{r_2r_1}r_5r_7r_0r_2r_6r_5r_0r_3$
- And $y_0y_1y_2y_3 = \text{SboxLeft}(x_0x_1x_2x_3x_4x_5)$ with $y_1=x_2$ and $y_2=x_3$ each with probability $3/4$
- Also, $\text{expand}(R_{i-1}) \oplus K_i$ is input to Sboxes at round $i$
- Then $y_1=r_2\oplus k_m$ and $y_2=r_1\oplus k_n$ both with prob $3/4$
- New right half is $y_0y_1y_2y_3\ldots$ plus old left half
- **Bottom line:** New right half bits: $r_1 \leftarrow r_2 \oplus k_m \oplus l_1$ and $r_2 \leftarrow r_1 \oplus k_n \oplus l_2$ both with probability $3/4$

# Recall TDES subkeys

- Key: $K = k_0k_1k_2k_3k_4k_5k_6k_7k_8k_9k_{10}k_{11}k_{12}k_{13}k_{14}k_{15}$
- Subkey $K_1 = k_2k_4\mathbf{k_5k_6}k_7k_1k_{10}k_{11}k_{12}k_{14}k_{15}k_8$
- Subkey $K_2 = k_4k_6\mathbf{k_7k_0}k_1k_3k_{11}k_{12}k_{13}k_{15}k_8k_9$
- Subkey $K_3 = k_6k_0\mathbf{k_1k_2}k_3k_5k_{12}k_{13}k_{14}k_8k_9k_{10}$
- Subkey $K_4 = k_0k_2k_3k_4k_5k_7k_{13}k_{14}k_{15}k_9k_{10}k_{11}$

# TDES线性分析

- 对于第一轮加密

明文： $(L_0,R_0) = (p_0p_1p_2p_3p_4p_5p_6p_7, p_8p_9p_{10}p_{11}p_{12}p_{13}p_{14}p_{15})$

密文：$(L_1,R_1) = (R_0 , L_0 \oplus F(R_0,K_1))$

$R_1 = L_0 \oplus F(R_0,K_1)$

$= L_0 \oplus (\mathrm{SboxesL}(A \oplus K_1), \mathrm{SboxesR}(B \oplus K_1))$

A为expand($R_0$)左半部分， B 为expand($R_0$)右半部分

- $\mathrm{expand}(p_8p_9p_{10}p_{11}p_{12}p_{13}p_{14}p_{15}) = \underbrace{p_{12}p_{15}p_{10}p_9p_{13}p_{15}}_{A}\ \underbrace{p_8p_{10}p_{14}p_{13}p_8p_{11}}_{B}$

# TDES线性分析

- 根据SboxLeft的性质
- $y_0y_1y_2y_3$ = $\mathrm{SboxLeft}(x_0x_1x_2x_3x_4x_5)$ with $y_1=x_2$ and $y_2=x_3$ each with probability ¾

$$K_1 = k_2k_4k_5k_6k_7k_1\ k_{10}k_{11}k_{12}k_{14}k_{15}k_8$$

$$A = p_{12}p_{15}p_{10}p_9p_{13}p_{15}$$

- Then $y_1=p_{10}\oplus k_5$ and $y_2=p_9\oplus k_6$ both with prob ¾

- 密文：$(L_1, R_1)= (.p_9p_{10}\ldots.. , .\ p_{10}\oplus k_5 \oplus p_1\quad p_9\oplus k_6 \oplus p_2 \ldots..)$（3/4概率）
- 通过该方法可以分析出所有轮加密的结果

# Linear Cryptanalysis of TDES

| $(L_0,R_0)=(p_0\ldots p_7,p_8\ldots p_{15})$ | Bit 1, Bit 2 (numbering from 0) | probability |
|---|---|---|
| $L_1 = R_0$<br>$R_1 = L_0 \oplus F(R_0,K_1)$ | $p_9$, $p_{10}$<br>$p_1 \oplus p_{10} \oplus k_5$, $p_2 \oplus p_9 \oplus k_6$ | 1<br>3/4 |
| $L_2 = R_1$<br>$R_2 = L_1 \oplus F(R_1,K_2)$ | $p_1 \oplus p_{10} \oplus k_5$, $p_2 \oplus p_9 \oplus k_6$<br>$p_2 \oplus k_6 \oplus k_7$, $p_1 \oplus k_5 \oplus k_0$ | 3/4<br>$(3/4)^2$ |
| $L_3 = R_2$<br>$R_3 = L_2 \oplus F(R_2,K_3)$ | $p_2 \oplus k_6 \oplus k_7$, $p_1 \oplus k_5 \oplus k_0$<br>$p_{10} \oplus k_0 \oplus k_1$, $p_9 \oplus k_7 \oplus k_2$ | $(3/4)^2$<br>$(3/4)^3$ |
| $L_4 = R_3$<br>$R_4 = L_3 \oplus F(R_3,K_4)$ | $p_{10} \oplus k_0 \oplus k_1$, $p_9 \oplus k_7 \oplus k_2$ | $(3/4)^3$ |
| $C = (L_4, R_4)$ | $k_0 \oplus k_1 = c_1 \oplus p_{10},$ $k_7 \oplus k_2 = c_2 \oplus p_9$ | $(3/4)^3$ |

# Test sample 3

- Use 100 known plaintexts, get ciphertexts.
  - Let $P=p_0p_1p_2\ldots p_{15}$ and let $C=c_0c_1c_2\ldots c_{15}$
- Resulting counts
  - $c_1 \oplus p_{10} = 0$ occurs 38 times
  - $c_1 \oplus p_{10} = 1$ occurs 62 times
  - $c_2 \oplus p_9 = 0$ occurs 62 times
  - $c_2 \oplus p_9 = 1$ occurs 38 times
- Conclusions
  - Since $k_0 \oplus k_1 = c_1 \oplus p_{10}$ we have $k_0 \oplus k_1 = 1$
  - Since $k_7 \oplus k_2 = c_2 \oplus p_9$ we have $k_7 \oplus k_2 = 0$
- Actual key is $K = 1010\ 0011\ 0101\ 0110$

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

# Linear Cryptanalysis of TDES

- Only recover the equivalent of two bits of information. To recover the entire key K, we could do an exhaustive key search for the remaining unknown bits. This would require an expected work of about $2^{13}$ encryptions and the work for the linear attack, which is negligible in comparison.
- In TDES, the number of rounds is small, and the one-round success probabilities are not sufficiently diminished during encryption. Also, the TDES S-boxes are poorly designed, resulting in limited confusion. Finally, the TDES expand permutation—the only source of diffusion in the cipher—does a poor job of mixing the bits of one round into the next round. All of these combine to yield a cipher that is highly susceptible to both linear and differential attacks.

# To Build a Better Block Cipher…

- How can cryptographers make linear and differential attacks more difficult?
  1. **More rounds** — success probabilities diminish with each round
  2. **Better confusion** (S-boxes) — reduce success probability on each round
  3. **Better diffusion** (permutations) — more difficult to chain thru multiple rounds
- Limited mixing and limited nonlinearity, means that more rounds required: TEA (Tiny Encryption Algorithm).
- Strong mixing and nonlinearity, then fewer (but more complex) rounds: AES (Advanced Encryption Standard).

# Homework

- 跟踪调试运行针对TDES的差分和线性攻击代码
  http://chengqingli.com/mate/codes/TDES_cryptanalysis_Error_Correct.zip
- 使用Python或Matlab等语言重写上述代码
- 自行选定密钥*K*，按Test sample 1的格式输出差分攻击结果
- 自行选定密钥*K*，按Test sample 3的格式输出线性攻击结果