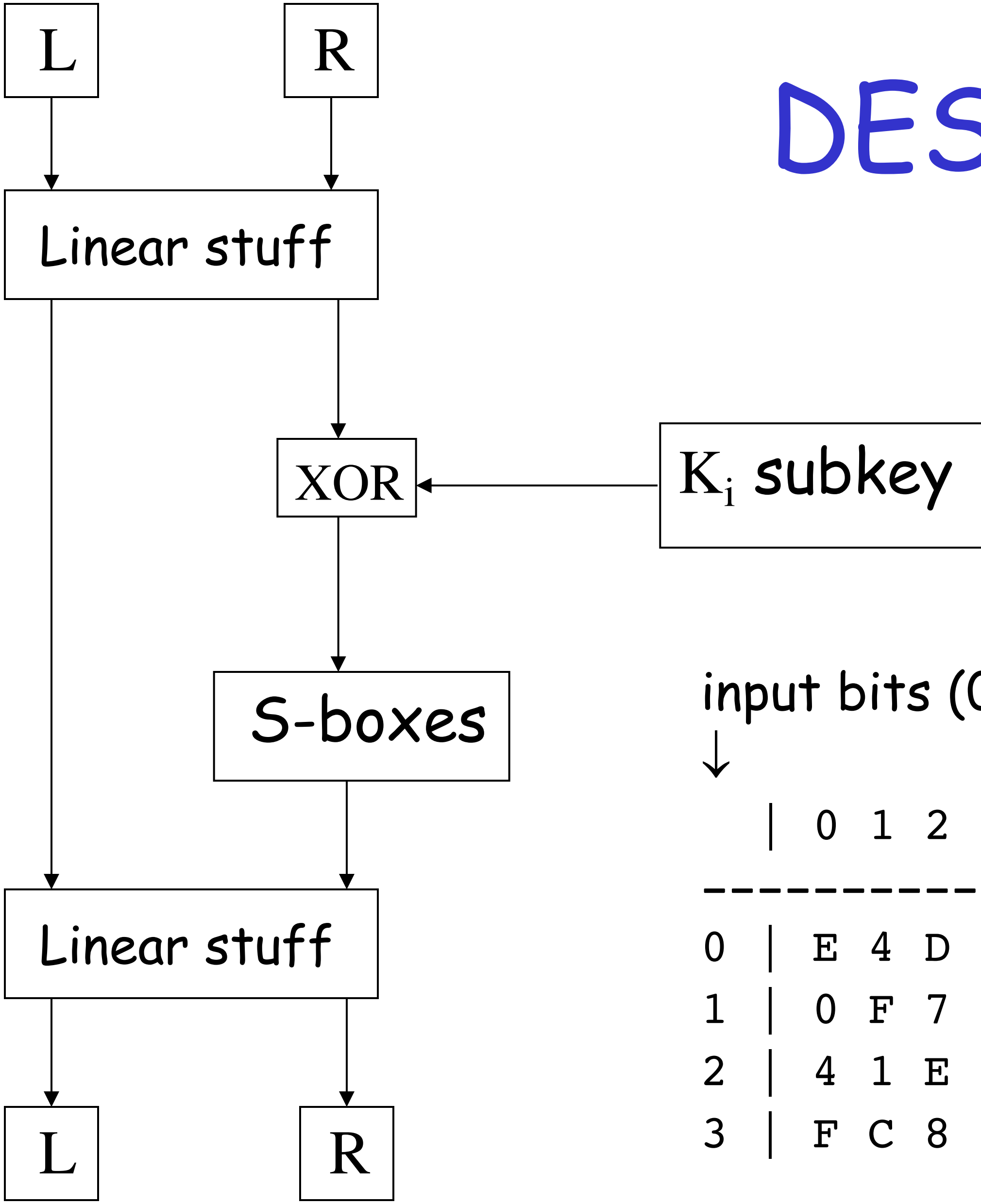
Chapter 6: Advanced Cryptanalysis

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
  - A chosen plaintext attack
- Linear cryptanalysis — Matsui, 1993
  - Perhaps not know to NSA in 1970s
  - A known plaintext attack



DES Overview

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

input bits (0,5)
↓
input bits (1,2,3,4)

	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

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

Differential Cryptanalysis

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

Differential Cryptanalysis

row	column 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 $X_2 \oplus K = 001$
- $\text{Sbox}(X_1 \oplus K) = 10$ and $\text{Sbox}(X_2 \oplus K) = 01$

Differential Cryptanalysis

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:** $\text{Sbox}(X \oplus K) = 10$, $\text{Sbox}(X \oplus K) = 01$
- Know $X \oplus K \in \{000,101\}$, $X \oplus K \in \{001,110\}$
- Then $K \in \{110,011\} \cap \{011,100\} \Rightarrow K = 011$

Differential Cryptanalysis

- Attacking one S-box not very useful!
- 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 next round
   - Choose input so output is "good" for 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 thru 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

S-box Differential Analysis

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

- Input diff 000 not interesting
- Input diff 010 always gives output diff 01
- More biased, the better (for Trudy)

$X \oplus X$	Sbox(X) ⊕ Sbox(X) 00	01	10	11
000	8	0	0	0
001	0	0	4	4
010	0	8	0	0
011	0	0	4	4
100	0	0	4	4
101	4	4	0	0
110	0	0	4	4
111	4	4	0	0

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?

S-box Linear Analysis

- 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

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

	input	output y_0	y_1	$y_0 \oplus y_1$
	0	4	4	4
i	x_0	4	4	4
n	x_1	4	6	2
p	x_2	4	4	4
u	$x_0 \oplus x_1$	4	2	2
t	$x_0 \oplus x_2$	0	4	4
	$x_1 \oplus x_2$	4	6	6
	$x_0 \oplus x_1 \oplus x_2$	4	6	2

Linear Analysis

- For example,
  $y_1 = x_1$
  with prob. 3/4
- And
  $y_0 = x_0 \oplus x_2 \oplus 1$
  with prob. 1
- And
  $y_0 \oplus y_1 = x_1 \oplus x_2$
  with prob. 3/4

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

input	output 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

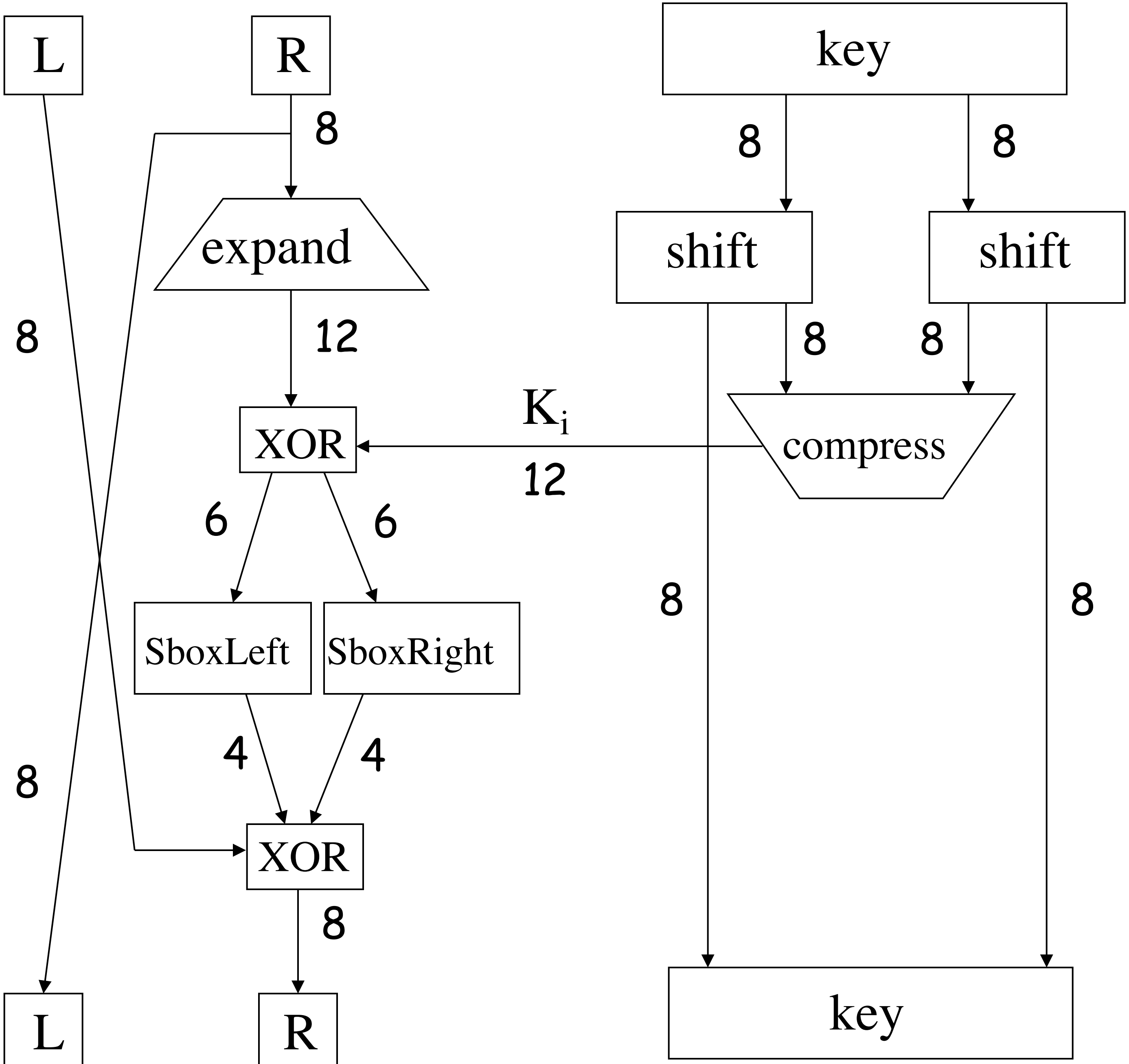
- Consider a single DES S-box
- Let $Y = \text{Sbox}(X)$
- Suppose $y_3 = x_2 \oplus x_5$ with high probability
  - I.e., a good linear approximation to output y_3
- Can we extend this so that we can solve linear equations for the key?
- As in differential cryptanalysis, we need to "chain" thru multiple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

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)

One Round of TDES



TDES Fun Facts

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

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

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

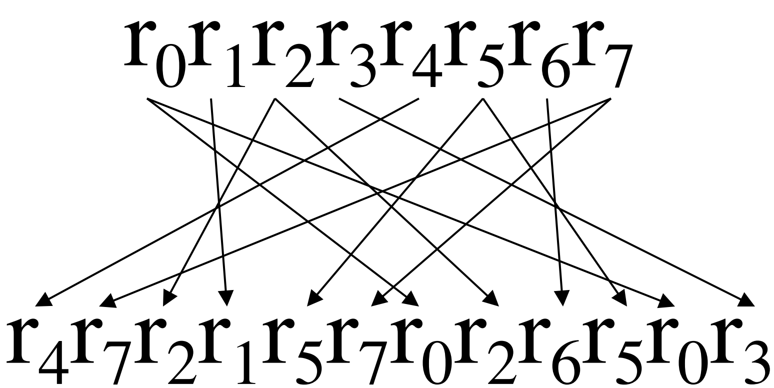
- Ciphertext = (L_4, R_4)
- $F(R_{i-1}, K_i) = \text{Sboxes}(\text{expand}(R_{i-1}) \oplus K_i)$ where $\text{Sboxes}(x_0x_1x_2\ldots x_{11}) = (\text{SboxLeft}(x_0x_1\ldots x_5), \text{SboxRight}(x_6x_7\ldots x_{11}))$

TDES Key Schedule

- 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 perm

- 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

- **Right S-box**
- SboxRight

	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

- **Left S-box**
- SboxLeft

	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

Differential Cryptanalysis of TDES

TDES

- TDES SboxRight

	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 $X \oplus X = 001000$
- Then $\text{SboxRight}(X) \oplus \text{SboxRight}(X) = 0010$ with probability 3/4

Differential Crypt. of TDES

- 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. $(000000\ 001000)$
- 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 $3/4$

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

TDES Differential Attack

- Select P and P with $P \oplus P = \text{0x0002}$

$(L_0,R_0) = P$	$(L_0,R_0) = P$	$P \oplus P = \text{0x0002}$
$L_1 = R_0$ $R_1 = L_0 \oplus F(R_0,K_1)$	$L_1 = R_0$ $R_1 = L_0 \oplus F(R_0,K_1)$	With probability 3/4 $(L_1,R_1) \oplus (L_1,R_1) = \text{0x0202}$
$L_2 = R_1$ $R_2 = L_1 \oplus F(R_1,K_2)$	$L_2 = R_1$ $R_2 = L_1 \oplus F(R_1,K_2)$	With probability $(3/4)^2$ $(L_2,R_2) \oplus (L_2,R_2) = \text{0x0200}$
$L_3 = R_2$ $R_3 = L_2 \oplus F(R_2,K_3)$	$L_3 = R_2$ $R_3 = L_2 \oplus F(R_2,K_3)$	With probability $(3/4)^2$ $(L_3,R_3) \oplus (L_3,R_3) = \text{0x0002}$
$L_4 = R_3$ $R_4 = L_3 \oplus F(R_3,K_4)$	$L_4 = R_3$ $R_4 = L_3 \oplus F(R_3,K_4)$	With probability $(3/4)^3$ $(L_4,R_4) \oplus (L_4,R_4) = \text{0x0202}$
$C = (L_4,R_4)$	$C = (L_4,R_4)$	$C \oplus C = \text{0x0202}$

TDES Differential Attack

- Choose P and $\tilde{P}$ with $P \oplus \tilde{P} = 0x0002$
- If $C \oplus \tilde{C} = 0x0202$ then

$$R_4 = L_3 \oplus F(R_3, K_4) \qquad \tilde{R}_4 = \tilde{L}_3 \oplus F(\tilde{R}_3, K_4)$$

$$R_4 = L_3 \oplus F(L_4, K_4) \qquad \tilde{R}_4 = \tilde{L}_3 \oplus F(\tilde{L}_4, K_4)$$

  and $(L_3, R_3) \oplus (\tilde{L}_3, \tilde{R}_3) = 0x0002$

- Then $L_3 = \tilde{L}_3$ and $C=(L_4, R_4)$ and $\tilde{C}=(\tilde{L}_4, \tilde{R}_4)$ are both known
- Since $L_3 = R_4 \oplus F(L_4, K_4)$ and $\tilde{L}_3 = \tilde{R}_4 \oplus F(\tilde{L}_4, K_4)$, for correct choice of subkey K_4 we have

$$R_4 \oplus F(L_4, K_4) = \tilde{R}_4 \oplus F(\tilde{L}_4, K_4)$$

TDES Differential Attack

- Choose P and P with $P \oplus P = \text{0x0002}$
- If $C \oplus C = (L_4, R_4) \oplus (L_4, R_4) = \text{0x0202}$
- Then for the correct subkey K_4

  $R_4 \oplus F(L_4, K_4) = R_4 \oplus F(L_4, K_4)$

  which we rewrite as

  $R_4 \oplus R_4 = F(L_4, K_4) \oplus F(L_4, K_4)$

  where the only unknown is K_4
- Let $L_4 = l_0 l_1 l_2 l_3 l_4 l_5 l_6 l_7$. Then we have

$$\begin{aligned} 0010 = {} & \text{SBoxRight}(l_0 l_2 l_6 l_5 l_0 l_3 \oplus k_{13} k_{14} k_{15} k_9 k_{10} k_{11}) \\ & \oplus \text{SBoxRight}(l_0 l_2 l_6 l_5 l_0 l_3 \oplus k_{13} k_{14} k_{15} k_9 k_{10} k_{11}) \end{aligned}$$

TDES Differential Attack

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

TDES Differential Attack

- Experimental results
- 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,
  $k_{13}k_{14}k_{15}k_9k_{10}k_{11} \in \{000001, 001001, 110000, 111000\}$
- Actual key is K=1010 1001 1000 0111

Linear Cryptanalysis of TDES

Linear Approx. of Left S-Box

- TDES left S-box or SboxLeft

	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 3/4
- 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 $3/4$

TDES Linear Cryptanalysis

- Known $P=p_0p_1p_2\ldots p_{15}$ and $C=c_0c_1c_2\ldots c_{15}$

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

TDES Linear Cryptanalysis

- Experimental results
- 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 $K = 1010\ 0011\ 0101\ 0110$

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