CS557: Cryptography

Modern Ciphers (Cryptanalysis-2)

S. Tripathy
IIT Patna

Present Class

- Cryptography
  - Modern Ciphers
    - Cryptanalysis
      - Linear Cryptanalysis
      - Differential Cryptanalysis

Piling-Up Lemma

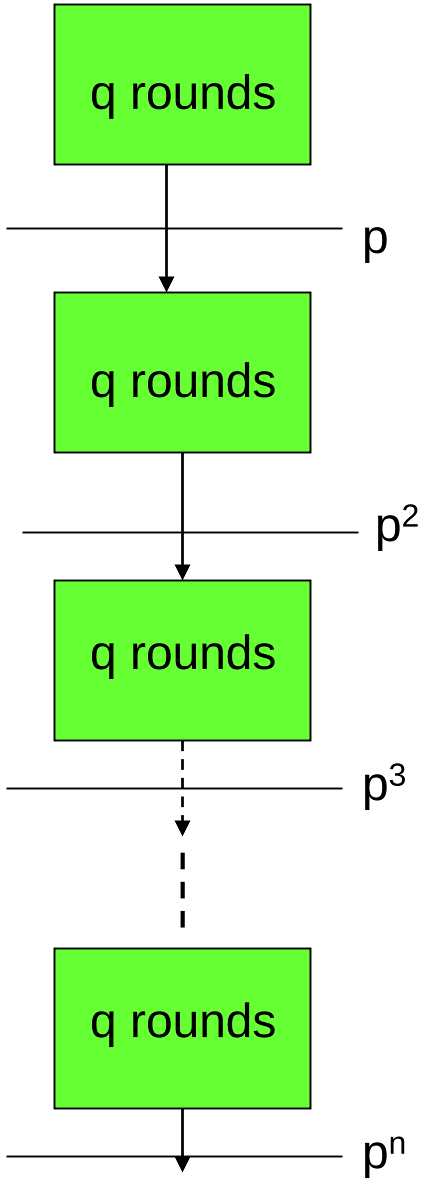
Matsui

- If $Pr(V_i = 0) = \frac{1}{2} + e_i$
- $Pr(V_1 \oplus V_2 \oplus \ldots \oplus V_n = 0) = \frac{1}{2} + 2^{n-1} \prod_{i=1}^{n} e_i$
- V_i's are independent random variables
- e_i is the bias $-\frac{1}{2} \leq e_i \leq \frac{1}{2}$

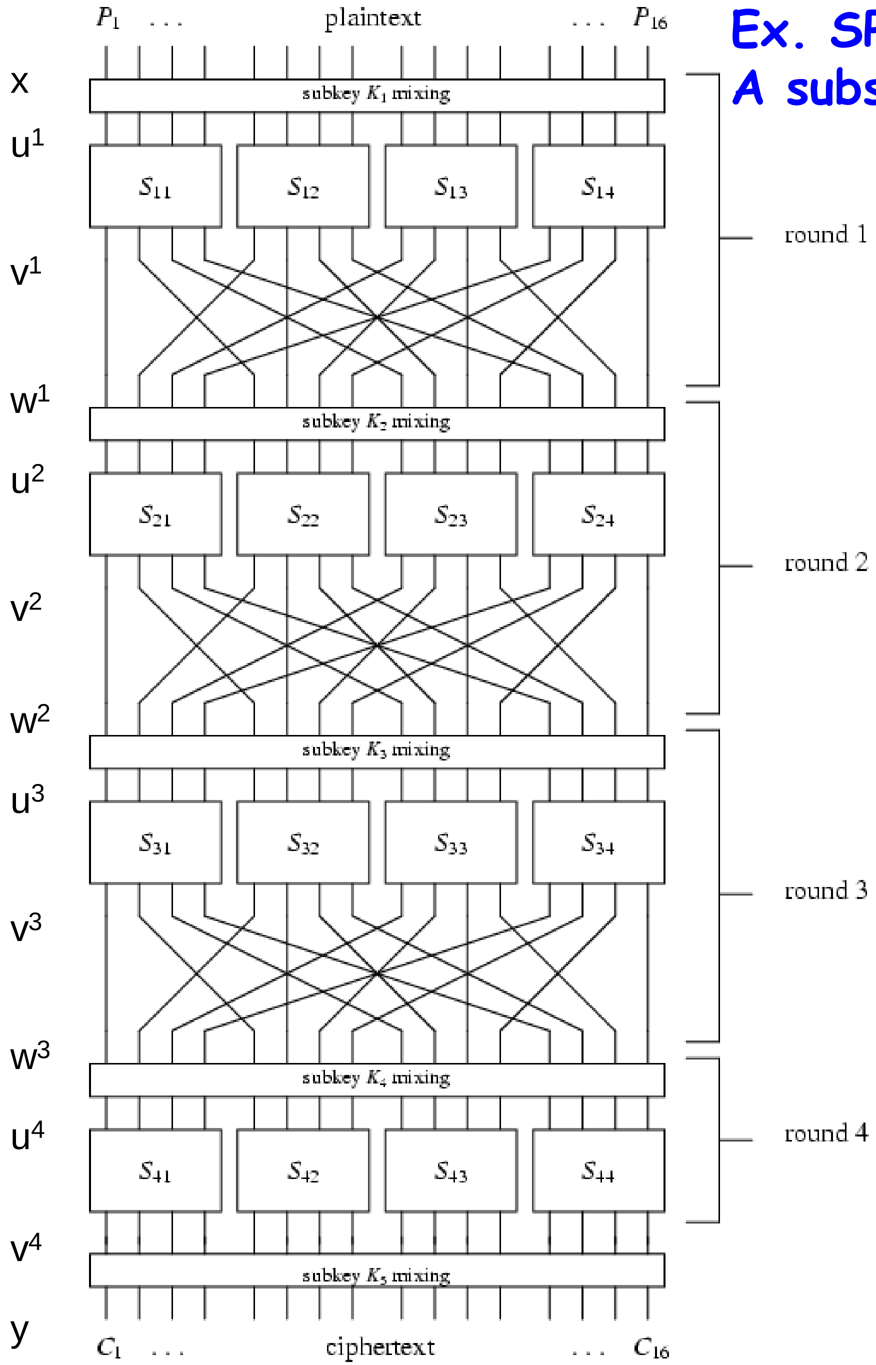
Use to combine linear equations if view each as independent random variable

Linear Bounds

- Bound a linear equation holds across q rounds: $0 < p \leq 1$
- Cipher has nq rounds
- Estimate upper bound $\leq p^n$
- 2^b possible plaintexts
- Round key bits, output of a round/input to next round not independent
- If $p^n \leq 2^{-b}$, no attack



Ex. SPN based Cipher:
A substitution-permutation network



Ex.: Substitution-Permutation Networks

- **Example:**
  - Suppose $l = m = Nr = 4$. Let π_S be defined as follows, where the input and the output are written in hexadecimal:

input	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	5	9	0	7

Let π_P be defined as follows:

input	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
output	1	5	9	13	2	6	10	14	3	7	11	15	4	8	12	16

Finding Linear Relationships

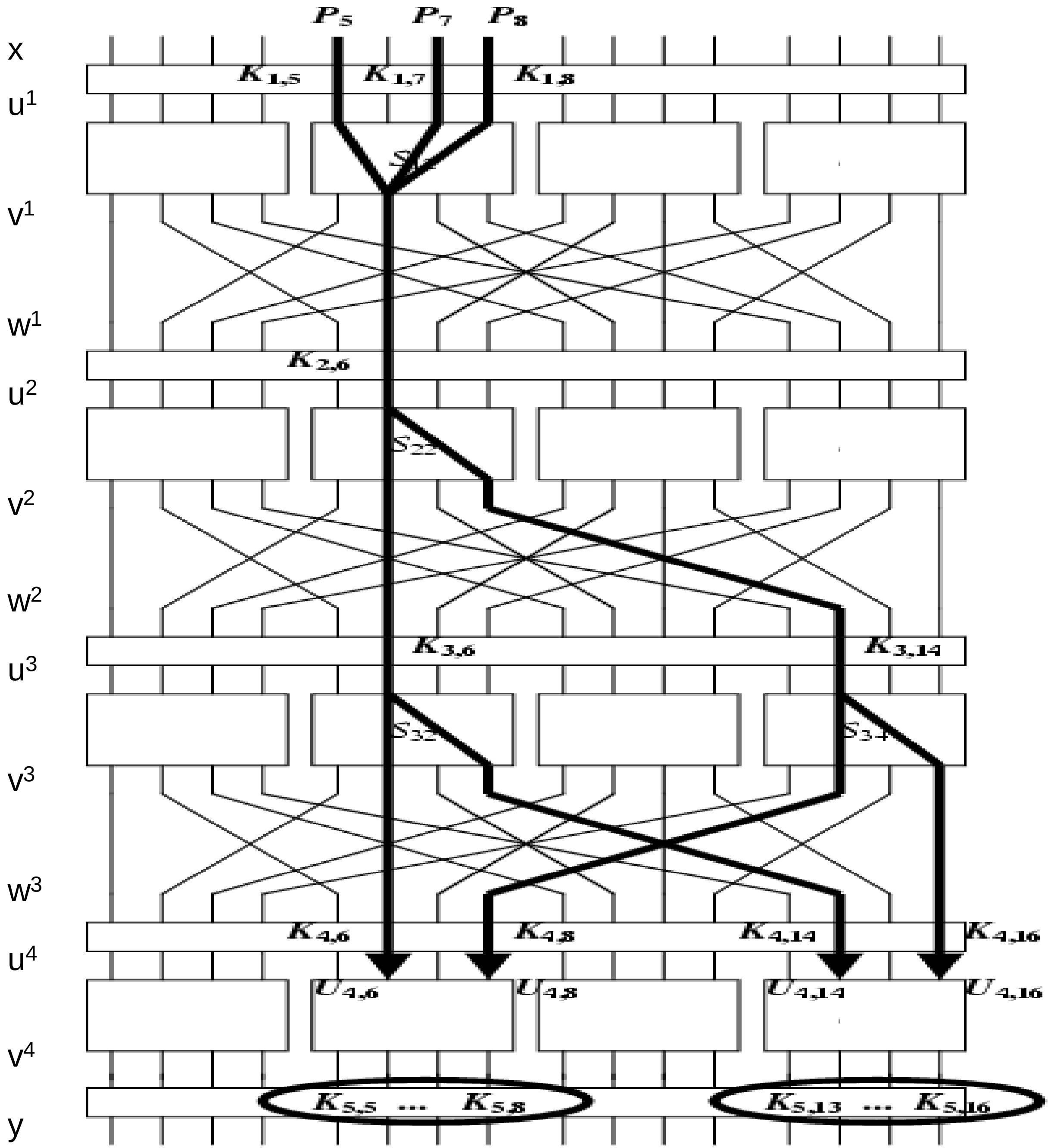
b1b2b3b4

a1a2a3a4

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

# of times equation holds: $a1Y1 \oplus a2Y2 \oplus a3Y3 \oplus a4Y4 = b1Z1 \oplus b2Z2 \oplus b3Z3 \oplus b4\,Z4$

A linear approximation of an SPN: Finding the Active S-boxes



Linear Cryptanalysis

- The approximation incorporates four active S-boxes:
  - In S_{12}, $T_1 = U_5^1 \oplus U_7^1 \oplus U_8^1 \oplus V_6^1$ has bias ¼
  - In S_{22}, $T_2 = U_6^2 \oplus V_6^2 \oplus V_8^2$ has bias -¼
  - In S_{32}, $T_3 = U_6^3 \oplus V_6^3 \oplus V_8^3$ has bias -¼
  - In S_{34}, $T_4 = U_{14}^3 \oplus V_{14}^3 \oplus V_{16}^3$ has bias -¼
- T_1, T_2, T_3, T_4 have biases that are high in absolute value. Further, we will see their XOR will lead to **cancellations of "intermediate"** random variables.

Linear Cryptanalysis

- Using **Piling-up lemma**, $T_1 \oplus T_2 \oplus T_3 \oplus T_4$ has bias equal to $2^3(1/4)(-1/4)^3=-1/32$.
  - Note: we assume the four round v are independent.
- Then T_1, T_2, T_3, T_4 can be expressed in terms of plaintext bits, bits of u^4 (input to the last round) and key bits as follows:

$$
\begin{aligned}
T_1 &= U_5^1 \oplus U_7^1 \oplus U_8^1 \oplus V_6^1 = X_5 \oplus K_5^1 \oplus X_7 \oplus K_7^1 \oplus X_8 \oplus K_8^1 \oplus V_6^1 \\
T_2 &= U_6^2 \oplus V_6^2 \oplus V_8^2 \qquad = V_6^1 \oplus K_6^2 \oplus V_6^2 \oplus V_8^2 \\
T_3 &= U_6^3 \oplus V_6^3 \oplus V_8^3 \qquad = V_6^2 \oplus K_6^3 \oplus V_6^3 \oplus V_8^3 \\
T_4 &= U_{14}^3 \oplus V_{14}^3 \oplus V_{16}^3 \qquad = V_8^2 \oplus K_{14}^3 \oplus V_{14}^3 \oplus V_{16}^3
\end{aligned}
$$

Linear Cryptanalysis

- XOR the right side and we get

$$X_5 \oplus X_7 \oplus X_8 \oplus V_6^3 \oplus V_8^3 \oplus V_{14}^3 \oplus V_{16}^3$$
$$\oplus K_5^1 \oplus K_7^1 \oplus K_8^1 \oplus K_6^2 \oplus K_6^3 \oplus K_{14}^3 \qquad (3.1)$$

- Then replace V_i^3 by U_i^4 and key bits:

$$V_6^3 = U_6^4 \oplus K_6^4 \qquad V_8^3 = U_{14}^4 \oplus K_{14}^4$$
$$V_{14}^3 = U_8^4 \oplus K_8^4 \qquad V_{16}^3 = U_{16}^4 \oplus K_{16}^4$$

- Now substitute them into 3.1:

$$X_5 \oplus X_7 \oplus X_8 \oplus U_6^4 \oplus U_8^4 \oplus U_{14}^4 \oplus U_{16}^4$$
$$\oplus K_5^1 \oplus K_7^1 \oplus K_8^1 \oplus K_6^2 \oplus K_6^3 \oplus K_{14}^3 \oplus K_6^4 \oplus K_8^4 \oplus K_{14}^4 \oplus K_{16}^4 \quad (3.2)$$

Linear Cryptanalysis

- The expression obtained only involves plaintext bits, bits of u^4 and key bits.
- Suppose the key bits are fixed. Then

  $$K_5^1 \oplus K_7^1 \oplus K_8^1 \oplus K_6^2 \oplus K_6^3 \oplus K_{14}^3 \oplus K_6^4 \oplus K_8^4 \oplus K_{14}^4 \oplus K_{16}^4$$

  has the (fixed) value 0 or 1.
- It follows that $$X_5 \oplus X_7 \oplus X_8 \oplus U_6^4 \oplus U_8^4 \oplus U_{14}^4 \oplus U_{16}^4 \qquad (3.3)$$
  has bias -1/32 or 1/32 where the sign depends on the key bits (=0 or =1).
- Suppose that we have T plaintext-ciphertext pairs (denoted by $\mathcal{T}$), all use the same unknown key, K. The attack will allow us to obtain the eight key bits,

  $$K_5^5, K_6^5, K_7^5, K_8^5, K_{13}^5, K_{14}^5, K_{15}^5, K_{16}^5$$

- There are 2^8=256 possibilities for the eight key bits. We refer to a binary 8-tuple as a **candidate subkey**.

Linear Cryptanalysis

- For each $(x, y) \in \tau$ and for each candidate subkey, compute a partial decryption of y and obtain the resulting value for $u^4_{(2)}, u^4_{(4)}$
- Then we compute the value

$$x_5 \oplus x_7 \oplus x_8 \oplus u^4_6 \oplus u^4_8 \oplus u^4_{14} \oplus u^4_{16} \qquad (3.4)$$

- We maintain an array of counters indexed by the 256 possible candidate subkeys, and increment the counter corresponding to a particular subkey when (equation 3.4) has the value 0.
- At the end, we expect most counters will have a value close to T/2, but the correct candidate subkey will close to $T/2 \pm T/32$.

» DONE

Linear Cryptanalysis Origins

- Linear cryptanalysis first defined on Feal by Matsui and Yamagishi, 1992.
-
  - Matsui later published a linear attack on DES.
- 16-round DES can be attacked using 2^{43} known plaintexts - get 26 bits, brute force the remaining 30 bits
    - $2^{43} = 9 \times 10^{12}$ = 9 trillion known plaintext blocks
-

Differential Cryptanalysis

Notation

- P = plaintext
- C = ciphertext
- (P1,P2) = plaintext pair
- (C1,C2) = ciphertext pair
- $\Delta P = P1 \oplus P2$
- $\Delta C = C1 \oplus C2$
- Characteristic: $\Omega = (\lambda_{i1},\lambda_{o1},\lambda_{i2},\lambda_{o2},\ldots.\lambda_{ir},\lambda_{or})$
    - $\lambda_{ij} = \oplus$ of inputs to round j
    - $\lambda_{oj} = \oplus$ of outputs from round j
    - If pr_j = probability λ_{oj} occurs given λ_{ij}
    - then probability of $\Omega = \Pi\ pr_j$'s (upper bound)

Differential Cryptanalysis

- Differential cryptanalysis originally defined on DES
  - Eli Biham and Adi Shamir, Differential Cryptanalysis of the Data Encryption Standard, Springer Verlag, 1993.

- The main difference from linear attack is that differential attack involves comparing the XOR of two inputs to the XOR of the corresponding outputs.
- Differential attack is a **chosen-plaintext attack**.

- We consider inputs x and x* having a specified XOR value denoted by $x' = x \oplus x^*$

- We decrypt y and y* using all possible key and determine if their XOR has a certain value. Whenever it does, increment the corresponding counter. At the end, we expect the largest one is the most likely subkey.

Differential Cryptanalysis

- It is easy to see that any set $\Delta(x')$ contains 2^m pairs, and that

$$\Delta(x') = \{(x, x \oplus x') : x \in \{0,1\}^m\}$$

- For each pair in $\Delta(x')$, we can compute the output XOR of the S-box. Then we can tabulate the distribution of output XORs. There are 2^m output XORs which are distributed among 2^n possible values.
  - A non-uniform output distribution will be the basis for a successful attack.

Differential Cryptanalysis

x	x*	y	y*	y’
0000	1011	1110	1100	0010
0001	1010	0100	0110	0010
0010	1001	1101	1010	0111
0011	1000	0001	0011	0010
0100	1111	0010	0111	0101
0101	1110	1111	0000	1111
0110	1101	1011	1001	0010
0111	1100	1000	0101	1101
1000	0011	0011	0001	0010
1001	0010	1010	1101	0111
1010	0001	0110	0100	0010
1011	0000	1100	1110	0010
1100	0111	0101	1000	1101
1101	0110	1001	1011	0010
1110	0101	0000	1111	1111
1111	0100	0111	0010	0101

Distribution table for x’=1011

0000	0	1000	0
0001	0	1001	0
0010	8	1010	0
0011	0	1011	0
0100	0	1100	0
0101	2	1101	2
0110	0	1110	0
0111	2	1111	2

6 E F

B 0 7

001,10

le, wh

1011,

$x), y^* = \pi_S(x^*),$

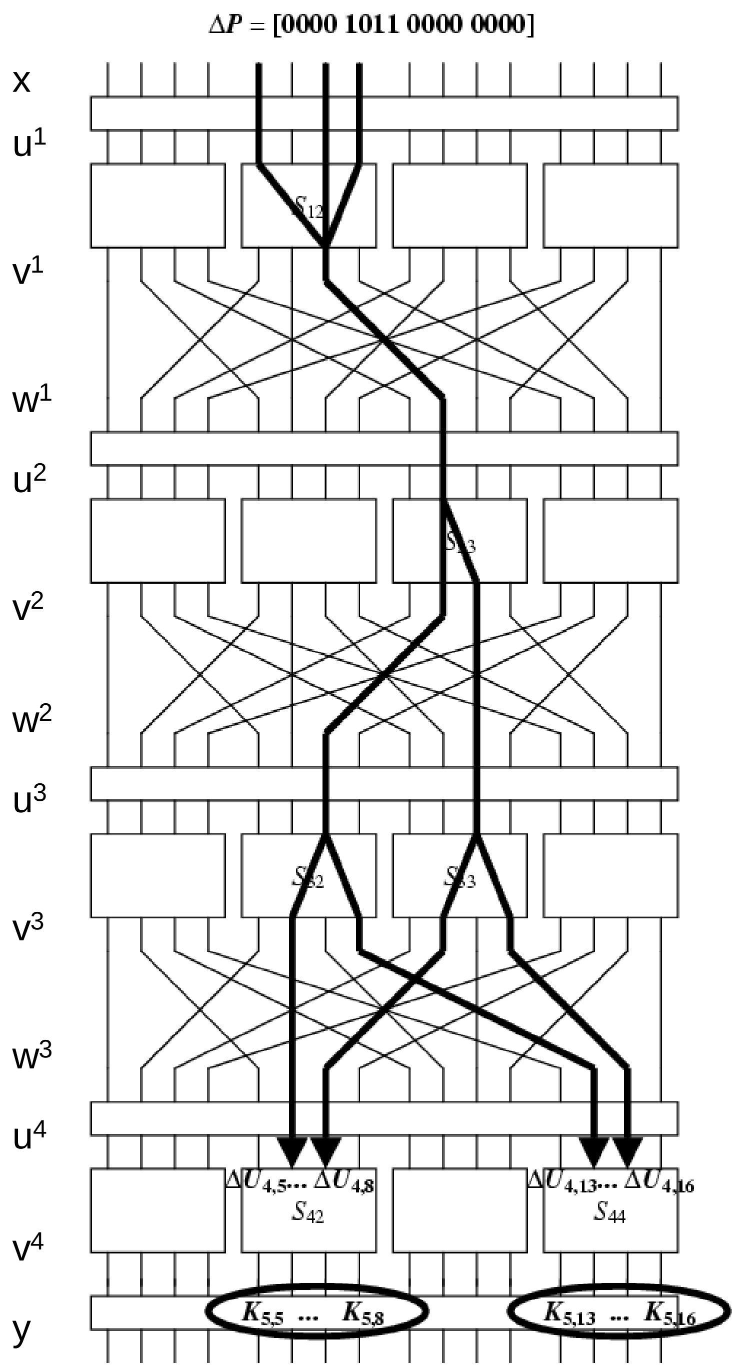
y^*

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

Difference distribution table: values of $N_D(x',y')$

$N_D(x',y')$ counts the number of pairs with input XOR equal to x' and output XOR equal to y'.

Ex.: A differential trail for a SPN



Differential Cryptanalysis
Attack Overview

- Determine key bits of last round:
  - Choose pairs (P1,P2) such that ΔP provides λ_{i1} .
  - Decrypt ciphertext with key guess for last round
  - Count # of (C1,C2) pairs such that match characterstic
  - Assume correct key bits is guess with highest count.

Reading Assignment:

Ref: A Tutorial on Linear and Differential Cryptanalysis By H.M. Hey

*https://www.engr.mun.ca/~howard/PAPERS/**ldc_tutorial**.pdf*