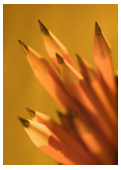
# Substitution-permutation ciphers

Linear cryptanalysis

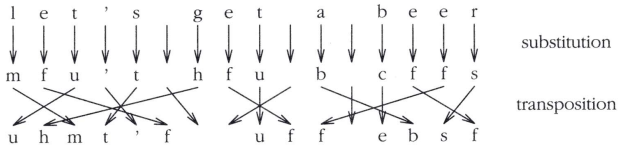


**CS310 Foundations of Cryptography**

Department of Computer Science
Wellesley College

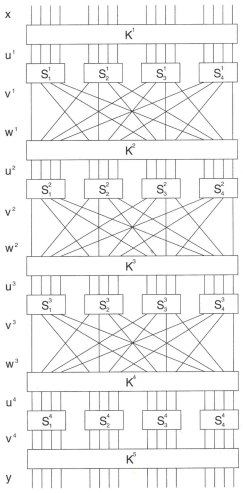
## Block ciphers

- Modern product ciphers incorporate a sequence of permutation and substitution operations.



## Substitution-permutation networks

- The game is to do this over and over again, substitution for confusion and permutation for defusion.
- A typical iterated cipher requires a round function and key schedule.






## Key schedules and round functions

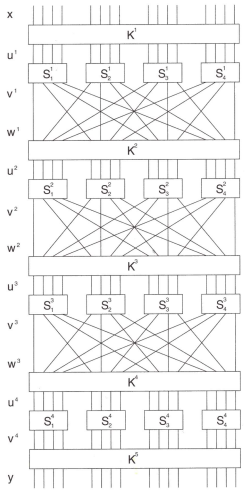
- Round keys, $K^1, \ldots, K^{Nr}$, are constructed from a random binary key, $K$, using some fixed, public algorithm.
- A round function, $g$, takes inputs $K^r$ and a current state $w^{r-1}$ and produces the next state, $w^r$.*

$$w^0 \leftarrow x$$
$$w^1 \leftarrow g(w^0, K^1)$$
$$w^2 \leftarrow g(w^1, K^2)$$
$$\ldots$$
$$w^{Nr-1} \leftarrow g(w^{Nr-2}, K^{Nr-1})$$
$$w^{Nr} \leftarrow g(w^{Nr-1}, K^{Nr})$$
$$y \leftarrow w^{Nr}$$

*The plaintext is the initial state, $w^0$.

## Substitution and permutation

- Plaintext and ciphertext are broken into binary sequences of length $\ell m$, the block length.
- A permutation $\pi_S: \{0, 1\}^\ell \to \{0, 1\}^\ell$, called an S-box, substitutes each set of $\ell$ bits for another.
- A permutation $\pi_P : \{1, \ldots, \ell m\} \to \{1, \ldots, \ell m\}$ mixes everything up.





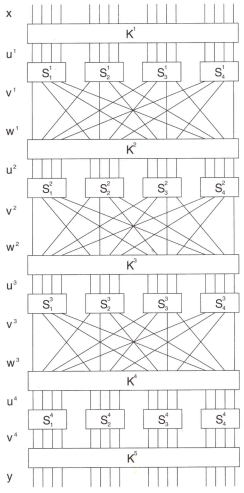

## In the example shown, . . .

- . . . the S-boxes are given by the substitutions:

| $z$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\pi_S(z)$ | E | 4 | D | 1 | 2 | F | B | 8 | 3 | A | 6 | C | 5 | 9 | 0 | 7 |

- . . . while the permutation is:

| $z$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\pi_P(z)$ | 1 | 5 | 9 | 13 | 2 | 6 | 10 | 14 | 3 | 7 | 11 | 15 | 4 | 8 | 12 | 16 |



## We still need a key schedule

- Given a 32-bit key $K = (k_1, \ldots, k_{31})$, define $K^r$, for $1 \le r \le 5$, to consist of 16 consecutive bits of $K$, beginning with $k_{4r-3}$.
- For K given by

  0011 1010 1001 0100 1101 0110 0011 1111

  the round keys are:

  $K^1$ = 0011 1010 1001 0100

  $K^2$ = 1010 1001 0100 1101

  $K^3$ = 1001 0100 1101 0110

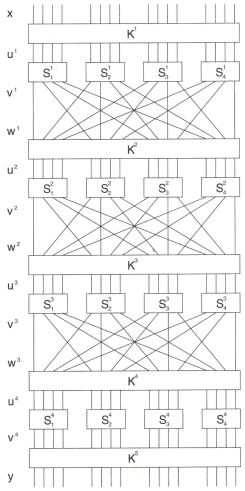
  $K^4$ = 0100 1101 0110 0011

  $K^5$ = 1101 0110 0011 1111



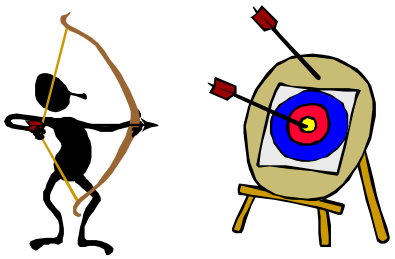
## For x = 0010 0110 1011 0111

$w^0 = 0010\ 0110\ 1011\ 0111$
$K^1 = 0011\ 1010\ 1001\ 0100$
$u^1 = 0001\ 1100\ 0010\ 0011$
$v^1 = 0100\ 0101\ 1101\ 0001$
$w^1 = 0010\ 1110\ 0000\ 0111$
$K^2 = 1010\ 1001\ 0100\ 1101$
$u^2 = 1000\ 0111\ 0100\ 1010$
$v^2 = 0011\ 1000\ 0010\ 0110$
$w^2 = 0100\ 0001\ 1011\ 1000$
$K^3 = 1001\ 0100\ 1101\ 0110$
$u^3 = 1101\ 0101\ 0110\ 1110$
$v^3 = 1001\ 1111\ 1011\ 0000$
$w^3 = 1110\ 0100\ 0110\ 1110$
$K^4 = 0100\ 1101\ 0110\ 0011$
$u^4 = 1010\ 1001\ 0000\ 1101$
$v^4 = 0110\ 1010\ 1110\ 1001$
$K^5 = 1101\ 0110\ 0011\ 1111,$
$y = 1011\ 1100\ 1101\ 0110$



## Linear cryptanalysis

- The object of linear cryptanalysis is to find a probabilistic linear relationship between subsets of plaintext and ciphertext bits*.
- The attacker computes XOR of relevant bits in relationship using various keys in order to find a key that yields a nonrandom distribution.



*Thus, this is known-plaintext attack.



## Before the details, we need . . .

- Suppose, $\mathbf{X}_1, \mathbf{X}_2, \ldots$ are independent random variables taking values from the set {0, 1} such that

  $$\Pr[\mathbf{X}_i = 0] = p_i \quad \text{and} \quad \Pr[\mathbf{X}_i = 1] = 1 - p_i$$

- The independence of $\mathbf{X}_i$ and $\mathbf{X}_j$ implies that

  $$\Pr[\mathbf{X}_i = 0, \mathbf{X}_j = 0] = p_i p_j$$
  $$\Pr[\mathbf{X}_i = 0, \mathbf{X}_j = 1] = p_i(1 - p_j)$$
  $$\Pr[\mathbf{X}_i = 1, \mathbf{X}_j = 0] = (1 - p_i) p_j$$
  $$\Pr[\mathbf{X}_i = 1, \mathbf{X}_j = 1] = (1 - p_i)(1 - p_j)$$

- We compute $\Pr[\mathbf{X}_i \oplus \mathbf{X}_j = 0]$ and $\Pr[\mathbf{X}_i \oplus \mathbf{X}_j = 1]$.

## A random variable's bias

- The bias of a random variable $\mathbf{X_i}$ is

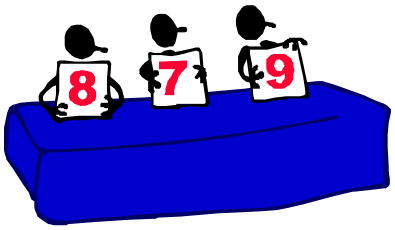
$$\varepsilon_i = p_i - \frac{1}{2}$$

- Observe that

$$-\frac{1}{2} \le \varepsilon_i \le \frac{1}{2}$$

$$\Pr[\mathbf{X_i} = 0] = \frac{1}{2} + \varepsilon_i$$

$$\Pr[\mathbf{X_i} = 1] = \frac{1}{2} - \varepsilon_i$$






## The piling-up lemma*

Lemma. Let $\varepsilon_{i_1,i_2,\ldots,i_k}$ denote the bias of the random variable $\mathbf{X}_{i_1} \oplus \mathbf{X}_{i_2} \oplus \ldots \oplus \mathbf{X}_{i_k}$. Then

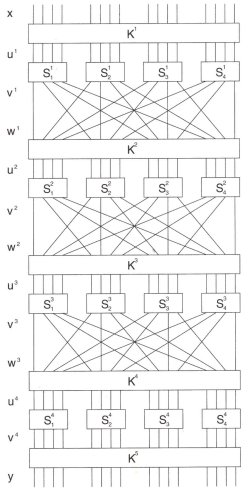
$$\varepsilon_{i_1,i_2,\ldots,i_k} = 2^{k-1} \prod_{j=1}^{k} \varepsilon_{i_j}$$

Corollary. Let $\varepsilon_{i_1,i_2,\ldots,i_k}$ denote the bias of the random variable $\mathbf{X}_{i_1} \oplus \mathbf{X}_{i_2} \oplus \ldots \oplus \mathbf{X}_{i_k}$. Suppose that $\varepsilon_{i_j} = 0$ for some *j*, then $\varepsilon_{i_1,i_2,\ldots,i_k} = 0$.

*Proof by induction on *k*.

# Linear approximations of S-boxes

- Consider an S-box $\pi_S$: $\{0, 1\}^m \rightarrow \{0, 1\}^n$.
- Assume input chosen uniformly at random from $\{0, 1\}^m$ .*
- Similarly, each output co-ordinate $y_j$ defines a random variable $\mathbf{Y}_j$ taking values 0 and 1.




*Thus, each input coordinate $x_i$ defines a random variable $\mathbf{X}_i$ taking on values 0 and 1 and these $\mathbf{X}_i$ are independent with zero biases.



# In our example, . . .

- . . . the permutation $\pi_S$: $\{0, 1\}^4 \rightarrow \{0, 1\}^4$, is given by

| $z$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\pi_S(z)$ | E | 4 | D | 1 | 2 | F | B | 8 | 3 | A | 6 | C | 5 | 9 | 0 | 7 |

- The random variable $\mathbf{X}_1 \oplus \mathbf{X}_4 \oplus \mathbf{Y}_2$ is unbiased.

| $\mathbf{X}_1$ | $\mathbf{X}_2$ | $\mathbf{X}_3$ | $\mathbf{X}_4$ | $\mathbf{Y}_1$ | $\mathbf{Y}_2$ | $\mathbf{Y}_3$ | $\mathbf{Y}_4$ | |
|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | |
| 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | ← |
| 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | |
| 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | |
| 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | ← |
| 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | ← |
| 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | ← |
| 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | |
| 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | |
| 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | ← |
| 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | ← |
| 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | |
| 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | ← |
| 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | ← |
| 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | |
| 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | |

# On the other hand, . . .

- . . . the permutation $\pi_S$: $\{0, 1\}^4 \rightarrow \{0, 1\}^4$, is still given by

| $z$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| $\pi_S(z)$ | $E$ | 4 | $D$ | 1 | 2 | $F$ | $B$ | 8 | 3 | $A$ |

| 9 | $A$ | $B$ | $C$ | $D$ | $E$ | $F$ |
|---|---|---|---|---|---|---|
| $A$ | 6 | $C$ | 5 | 9 | 0 | 7 |

- The random variable $\mathbf{X}_3 \oplus \mathbf{X}_4 \oplus \mathbf{Y}_1 \oplus \mathbf{Y}_4$ is has a bias of -3/8.

| $\mathbf{X}_1$ | $\mathbf{X}_2$ | $\mathbf{X}_3$ | $\mathbf{X}_4$ | $\mathbf{Y}_1$ | $\mathbf{Y}_2$ | $\mathbf{Y}_3$ | $\mathbf{Y}_4$ |
|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 |
| 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 |
| 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 |
| 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 |
| 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 |
| 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 |
| 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 |
| 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 |
| 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 |
| 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 |



# Linear approximation table $N_L(a, b)$

| $a$ \ $b$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 16 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |
| 1 | 8 | 8 | 6 | 6 | 8 | 8 | 6 | 14 | 10 | 10 | 8 | 8 | 10 | 10 | 8 | 8 |
| 2 | 8 | 8 | 6 | 6 | 8 | 8 | 6 | 6 | 8 | 8 | 10 | 10 | 8 | 8 | 2 | 10 |
| 3 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 10 | 2 | 6 | 6 | 10 | 10 | 6 | 6 |
| 4 | 8 | 10 | 8 | 6 | 6 | 4 | 6 | 8 | 8 | 6 | 8 | 10 | 10 | 4 | 10 | 8 |
| 5 | 8 | 6 | 6 | 8 | 6 | 8 | 12 | 10 | 6 | 8 | 4 | 10 | 8 | 6 | 6 | 8 |
| 6 | 8 | 10 | 6 | 12 | 10 | 8 | 8 | 10 | 8 | 6 | 10 | 12 | 6 | 8 | 8 | 6 |
| 7 | 8 | 6 | 8 | 10 | 10 | 4 | 10 | 8 | 6 | 8 | 10 | 8 | 12 | 10 | 8 | 10 |
| 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 6 | 10 | 10 | 6 | 10 | 6 | 6 | 2 |
| 9 | 8 | 8 | 6 | 6 | 8 | 8 | 6 | 6 | 4 | 8 | 6 | 10 | 8 | 12 | 10 | 6 |
| A | 8 | 12 | 6 | 10 | 4 | 8 | 10 | 6 | 10 | 10 | 8 | 8 | 10 | 10 | 8 | 8 |
| B | 8 | 12 | 8 | 4 | 12 | 8 | 12 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |
| C | 8 | 6 | 12 | 6 | 6 | 8 | 10 | 8 | 10 | 8 | 10 | 12 | 8 | 10 | 8 | 6 |
| D | 8 | 10 | 10 | 8 | 6 | 12 | 8 | 10 | 4 | 6 | 10 | 8 | 10 | 8 | 8 | 10 |
| E | 8 | 10 | 10 | 8 | 6 | 4 | 8 | 10 | 6 | 8 | 8 | 6 | 4 | 10 | 6 | 8 |
| F | 8 | 6 | 4 | 6 | 6 | 8 | 10 | 8 | 8 | 6 | 12 | 6 | 6 | 8 | 10 | 8 |

*Bias of the binary 8-tuple: $\varepsilon(a, b) = \Pr(a,b) - 1/2 = N_L(a,b)/16 - 1/2$.

## A linear attack on an SPN

- We find a linear approximation of S-boxes incorporating four active S-boxes:

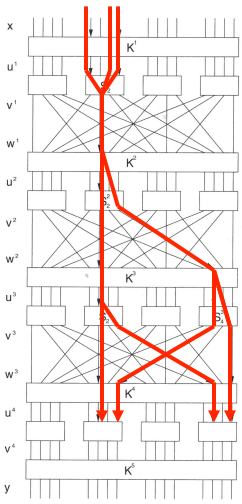
$S_2^1$ : $\mathbf{T}_1 = \mathbf{U}_5^1 \oplus \mathbf{U}_7^1 \oplus \mathbf{U}_8^1 \oplus \mathbf{V}_6^1$ has bias 1/4

$S_2^2$ : $\mathbf{T}_2 = \mathbf{U}_6^2 \oplus \mathbf{V}_6^2 \oplus \mathbf{V}_8^2$ has bias -1/4

$S_2^3$ : $\mathbf{T}_3 = \mathbf{U}_6^3 \oplus \mathbf{V}_6^3 \oplus \mathbf{V}_8^3$ has bias -1/4

$S_4^3$ : $\mathbf{T}_4 = \mathbf{U}_{14}^3 \oplus \mathbf{V}_{14}^3 \oplus \mathbf{V}_{16}^3$ has bias -1/4

- Assuming independences of $T_j$, piling up lemma implies $\mathbf{T}_1 \oplus \mathbf{T}_2 \oplus \mathbf{T}_3 \oplus \mathbf{T}_4$ has bias -1/32.

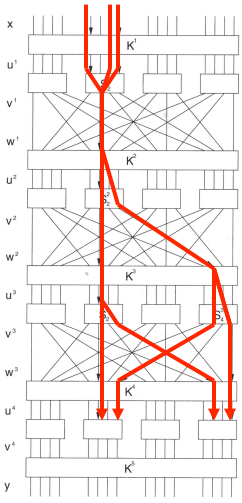




## Canceling "intermediate" variables

- The XOR of the $\mathbf{T}_i$ can be expressed in terms of plaintext bits, bits of $U^4$, and key bits.

$$\mathbf{T}_1 = \mathbf{U}_5^1 \oplus \mathbf{U}_7^1 \oplus \mathbf{U}_8^1 \oplus \mathbf{V}_6^1$$
$$= \mathbf{X}_5 \oplus \mathbf{K}_5^1 \oplus \mathbf{X}_7 \oplus \mathbf{K}_7^1 \oplus \mathbf{X}_8 \oplus \mathbf{K}_8^1 \oplus \mathbf{V}_6^1$$
$$\mathbf{T}_2 = \mathbf{U}_6^2 \oplus \mathbf{V}_6^2 \oplus \mathbf{V}_8^2$$
$$= \mathbf{V}_6^1 \oplus \mathbf{K}_6^2 \oplus \mathbf{V}_6^2 \oplus \mathbf{V}_8^2$$
$$\mathbf{T}_3 = \mathbf{U}_6^3 \oplus \mathbf{V}_6^3 \oplus \mathbf{V}_8^3$$
$$= \mathbf{V}_6^2 \oplus \mathbf{K}_6^3 \oplus \mathbf{V}_6^3 \oplus \mathbf{V}_8^3$$
$$\mathbf{T}_4 = \mathbf{U}_{14}^3 \oplus \mathbf{V}_{14}^3 \oplus \mathbf{V}_{16}^3$$
$$= \mathbf{V}_8^2 \oplus \mathbf{K}_{14}^3 \oplus \mathbf{V}_{14}^3 \oplus \mathbf{V}_{16}^3$$



## Plaintext, bits of $u^4$ and keybits

- $\mathbf{T}_1 \oplus \mathbf{T}_2 \oplus \mathbf{T}_3 \oplus \mathbf{T}_4 =$

  $\mathbf{X}_5 \oplus \mathbf{X}_7 \oplus \mathbf{X}_8 \oplus \mathbf{V}_6^3 \oplus \mathbf{V}_8^3 \oplus \mathbf{V}_{14}^3 \oplus \mathbf{V}_{16}^3$

  $\oplus \mathbf{K}_5^1 \oplus \mathbf{K}_7^1 \oplus \mathbf{K}_8^1 \oplus \mathbf{K}_6^2 \oplus \mathbf{K}_6^3 \oplus \mathbf{K}_{14}^3$

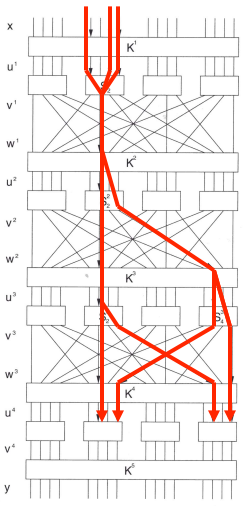
- Next, replace the $\mathbf{V}_i^3$ by expressions involving $\mathbf{U}_i^4$.

  $\mathbf{V}_6^3 = \mathbf{U}_6^4 \oplus \mathbf{K}_6^4$

  $\mathbf{V}_8^3 = \mathbf{U}_{14}^4 \oplus \mathbf{K}_{14}^4$

  $\mathbf{V}_{14}^3 = \mathbf{U}_8^4 \oplus \mathbf{K}_8^4$

  $\mathbf{V}_{16}^3 = \mathbf{U}_{16}^4 \oplus \mathbf{K}_{16}^4$






## Selecting the biased random variable

- The result

  $\mathbf{X}_5 \oplus \mathbf{X}_7 \oplus \mathbf{X}_8 \oplus \mathbf{U}_6^4 \oplus \mathbf{U}_8^4 \oplus \mathbf{U}_{14}^4 \oplus \mathbf{U}_{16}^4$

  $\oplus \mathbf{K}_5^1 \oplus \mathbf{K}_7^1 \oplus \mathbf{K}_8^1 \oplus \mathbf{K}_6^2 \oplus \mathbf{K}_6^3 \oplus \mathbf{K}_{14}^3 \oplus \mathbf{K}_6^4 \oplus \mathbf{K}_8^4 \oplus \mathbf{K}_{14}^4 \oplus \mathbf{K}_{16}^4$

- If the keybits are fixed, then the random variable

  $\mathbf{K}_5^1 \oplus \mathbf{K}_7^1 \oplus \mathbf{K}_8^1 \oplus \mathbf{K}_6^2 \oplus \mathbf{K}_6^3 \oplus \mathbf{K}_{14}^3 \oplus \mathbf{K}_6^4 \oplus \mathbf{K}_8^4 \oplus \mathbf{K}_{14}^4 \oplus \mathbf{K}_{16}^4$

  has fixed value 0 or 1 and

  $\mathbf{X}_5 \oplus \mathbf{X}_7 \oplus \mathbf{X}_8 \oplus \mathbf{U}_6^4 \oplus \mathbf{U}_8^4 \oplus \mathbf{U}_{14}^4 \oplus \mathbf{U}_{16}^4$

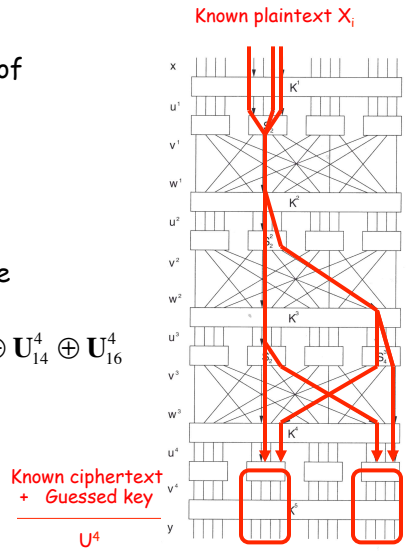
  has bias equal to ±1/32, where the sign depends on the values of the unknown key bits.

## Candidate subkeys

- We collect a large number of plaintext/ciphertext pairs.
- For each of the $2^8 = 256$ possible keys that are XORed with the 2nd and 4th S-boxes in the final row, we calculate the value of

$$\mathbf{X}_5 \oplus \mathbf{X}_7 \oplus \mathbf{X}_8 \oplus \mathbf{U}_6^4 \oplus \mathbf{U}_8^4 \oplus \mathbf{U}_{14}^4 \oplus \mathbf{U}_{16}^4$$

- The correct key should produce a bias of ±1/32.






## Success

- It is suggested that a linear attacked based on a linear approximation having bias equal to ε will be successful if the number of plaintext-ciphertext pairs is approximately $c\varepsilon^{-2}$, for a small constant c.