# Cryptanalysis of Modern Symmetric-Key Block Ciphers

[Based on "A Tutorial on Linear and Differential Cryptanalysis" by Howard Heys.]

Modern block ciphers (like DES and AES):
- proceed in rounds
- each round has its own round key or subkey
- the subkeys are computed from the master key by the key schedule

A simpler modern-type block cipher for now:
the substitution-permutation network
(similar to DES and AES but simplified structure)

# Substitution-Permutation Networks (SPN)

- consists of a number of rounds, each round (except the last), consists of XOR-ing the subkey (this is sometimes called key mixing), substitutions, and a permutation
- typically subkeys are derived from the master key but here they are randomly generated and unrelated

bit string XOR subkey, then do a substitution, then permute the bits

Let ℓ and m be positive integers. The block length of the cipher is ℓm.

We will use one substitution (also called an S-box)

$$\pi_S : \{0,1\}^{\ell} \to \{0,1\}^{\ell}$$

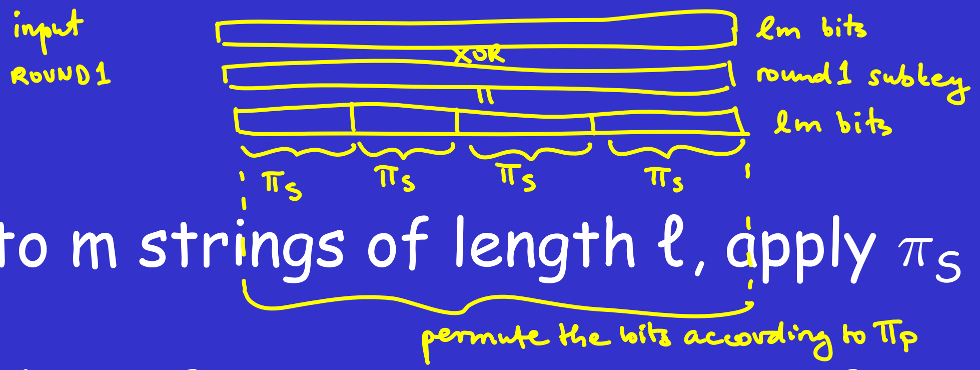
$\ell = 2$

$\pi_S$ : $(0,0) \to (1,0)$
$(0,1) \to (0,0)$
$(1,0) \to (1,1)$
$(1,1) \to (0,1)$

and one permutation

$$\pi_P : \{1, \dots, \ell m\} \to \{1, \dots, \ell m\}.$$

# Substitution-Permutation Networks (SPN)



In each round:
- XOR with the round key,
- split the current string into m strings of length ℓ, apply $\pi_S$ to each of these m strings
- if this is not the last round, perform permutation $\pi_P$; if it is the last round, XOR with the round key $K_{R+1}$ where R is the number of rounds

For example, if ℓ=2, m=3, $\pi_S$ and $\pi_P$ (see below), suppose the string before the round is 100011 and the round key is 100100 - what is the resulting string after this round ?

| | | | | |
|---|---|---|---|---|
| | 00 | 01 | 10 | 11 |
| x | 0 | 1 | 2 | 3 |
| $\pi_S(x)$ | 1 | 3 | 0 | 2 |

| | | | | | | |
|---|---|---|---|---|---|---|
| x | 1 | 2 | 3 | 4 | 5 | 6 |
| $\pi_P(x)$ | 6 | 4 | 2 | 1 | 3 | 5 |

input: 100011
subkey: 100100
after XOR: 00|01|11
after $\pi_S$: 01 11 10

after $\pi_P$:
1 2 3 4 5 6
1 1 1 1 0 0

# More on SPNs

- simple and very efficient, both in hardware and in software (assuming the S-boxes are not too large)
- decryption analogous to encryption (reverse each operation)
- very successful: DES and AES are variations on SPNs
- the first and last operations are XORing with subkeys (called whitening) – makes attacks harder

Figure 1 (Heys' tutorial): an example SPN that we will cryptanalyze

# Attacks on SPNs

- linear cryptanalysis and differential cryptanalysis
- both: known-plaintext, and they require a lot of plaintext-ciphertext pairs

**Linear cryptanalysis:**
Find a linear relationship between a subset of the plaintext bits and a subset of the ciphertext bits; this relationship should hold with probability bounded away from $\frac{1}{2}$ (the further away from $\frac{1}{2}$, the better). This probability, minus $\frac{1}{2}$, is called the probability bias.

Note:
In SPNs, all computations are linear, except for the S-boxes. Also, recall that linear cryptosystems are vulnerable to known-plaintext attacks.

# Linear Approximations of S-boxes

4-bit inputs (ℓ = 4)

The S-box from Figure 1:

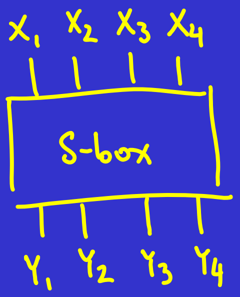
0000 0001 1000 1111

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

0011

Understanding the table: ℓ=4, the possible 4-bit strings are given in HEX.

Let $X_1$, $X_2$, $X_3$, $X_4$ be random variables for the input bits (independent, uniform), and let $Y_1$, $Y_2$, $Y_3$, $Y_4$ be random variables for the output bits.



# Linear Approximations of S-boxes

The S-box from Figure 1:

$x_2 = x_3 = 0$ ← equation (*) holds for this input

1001

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

1010

$Y_1 = 1$

$Y_3 = 1$

$Y_4 = 0$

Consider the linear equation:

$X_2 \oplus X_3 \oplus Y_1 \oplus Y_3 \oplus Y_4 = 0$, or, equivalently $X_2 \oplus X_3 = Y_1 \oplus Y_3 \oplus Y_4$. (*)

This equation holds for 12 or the 16 possible input values $X_1$, $X_2$, $X_3$, $X_4$. What is the probability bias of this equation ?

the equation holds w. prob. $\frac{12}{16}$ → the bias is $\frac{12}{16} - \frac{1}{2} = \frac{1}{4}$

# Linear Approximations of S-boxes

The S-box from Figure 1:

$\underline{0}00\underline{0}$

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

1110

N

Consider the linear equation:

$$X_1 \oplus X_4 = Y_2$$

What is the probability bias of this equation ? 0

(do yourself)

# Linear Approximations of S-boxes

The S-box from Figure 1:

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

Consider the linear equation:

$$X_3 \oplus X_4 = Y_1 \oplus Y_4$$

What is the probability bias of this equation ?

$-\frac{3}{8}$

$\uparrow$

negative

Coming back from the next slide:

$a_1 = a_2 = 0$ $\quad$ $b_1 = b_4 = 1$

$a_3 = a_4 = 1$ $\quad$ $b_2 = b_3 = 0$

# Linear Approximations of S-boxes

The S-box from Figure 1:

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

We can compute the probability biases for all linear equations relating the $X_i$'s and the $Y_i$'s. I.e. for any $a_i, b_i \in \{0,1\}$, we can compute the bias of the equation

$$a_1X_1 \oplus a_2X_2 \oplus a_3X_3 \oplus a_4X_4 = b_1Y_1 \oplus b_2Y_2 \oplus b_3Y_3 \oplus b_4Y_4.$$

$2^8$ such equations

See Tables 3 and 4 in Heys's tutorial.

Next task: combining the linear approximations of the S-boxes to get a linear approximation of the entire SPN.

# Piling-up Lemma

We will combine S-box approximations… What happens to the biases ?

**Piling-up Lemma:**
For k independent random variables $X_1, X_2, \ldots, X_k$ where $X_i=0$ has bias $\epsilon_i$, the equation $X_1 \oplus \ldots \oplus X_k=0$ has bias $2^{k-1} \Pi_{i=1,\ldots,k} \epsilon_i$.

Note: lemma by Matsui, inventor of linear cryptanalysis

Proving the lemma for k=2:

$Pr(X_1=1) = p_1$ $\qquad \varepsilon_1 = p_1 - \frac{1}{2}$

$Pr(X_2=1) = p_2$ $\qquad \varepsilon_2 = p_2 - \frac{1}{2}$

want: $Pr(X_1 \oplus X_2 = 0) = Pr(X_1=0 \text{ and } X_2=0) + Pr(X_1=1 \text{ and } X_2=1) \overset{\text{independent } X_1, X_2}{=}$

$$= Pr(X_1=0) \cdot Pr(X_2=0) + Pr(X_1=1) \cdot Pr(X_2=1)$$

$$= (1-p_1) \cdot (1-p_2) + p_1 p_2 = \left(\frac{1}{2}-\varepsilon_1\right) \cdot \left(\frac{1}{2}-\varepsilon_2\right) + \left(\varepsilon_1+\frac{1}{2}\right)\left(\varepsilon_2+\frac{1}{2}\right)$$

$$= \frac{1}{4} - \frac{\varepsilon_1}{2} - \frac{\varepsilon_2}{2} + \varepsilon_1\varepsilon_2 + \frac{1}{4} + \frac{\varepsilon_1}{2} + \frac{\varepsilon_2}{2} + \varepsilon_1\varepsilon_2 = \frac{1}{2} + 2\varepsilon_1\varepsilon_2$$

hence,
the bias of $X_1 \oplus X_2 = 0$
is $\frac{1}{2} + 2\varepsilon_1\varepsilon_2 - \frac{1}{2} = 2\varepsilon_1\varepsilon_2$

# Piling-up Lemma

We will combine S-box approximations… What happens to the biases ?

**Piling-up Lemma:**
For k independent random variables $X_1, X_2, \ldots, X_k$ where $X_i=0$ has bias $\epsilon_i$, the equation $X_1 \oplus \ldots \oplus X_k=0$ has bias $2^{k-1} \Pi_{i=1,\ldots,k} \epsilon_i$.

Note: lemma by Matsui, inventor of linear cryptanalysis

Give a simple example that shows that the assumption that the $X_i$'s are independent is necessary.

$X_1 = X_2$ $\quad$ $X_1 \oplus X_2 = 0$ always true : $\quad$ $Pr(X_1 \oplus X_2 = 0) = 1$

bias : $\frac{1}{2}$

not $2 \cdot 0 \cdot 0$

# Linear Approximation for the Cipher

Recall the SPN from Figure 1 (also see Figure 3; we do not do the last round on this slide).

Our approximation will involve S-boxes $S_{12}$, $S_{22}$, $S_{32}$, and $S_{34}$. We call them the active S-boxes.

We will use the following approximations of these S-boxes:

$S_{12}$: $X_1 \oplus X_3 \oplus X_4 = Y_2$ bias $\frac{1}{4}$
$S_{22}$: $X_2 = Y_2 \oplus Y_4$ bias $-\frac{1}{4}$
$S_{32}$: $X_2 = Y_2 \oplus Y_4$ bias $-\frac{1}{4}$
$S_{34}$: $X_2 = Y_2 \oplus Y_4$ bias $-\frac{1}{4}$

# Linear Approximation for the Cipher

plaintext

Let $P_i$ be the random variable for the i-th plaintext bit, let $U_{r,i}$ be the random variable for the i-th input bit to the round r S-boxes, let $V_{r,i}$ be the random variable for the i-th output bit of the round r S-boxes, and let $K_{r,i}$ be the i-th bit of the r-th subkey.

subkeys

Let $T_1, T_2, T_3, T_4$ be random variables such that

$T_1 = U_{1,5} \oplus U_{1,7} \oplus U_{1,8} \oplus V_{1,6}$ 1/4 bias

$T_2 = U_{2,6} \oplus V_{2,6} \oplus V_{2,8}$ -1/4 bias

$T_3 = U_{3,6} \oplus V_{3,6} \oplus V_{3,8}$ -1/4 bias

$T_4 = U_{3,14} \oplus V_{3,14} \oplus V_{3,16}$ -1/4 bias

What are the biases of $T_i=0$ for $i \in \{1,2,3,4\}$ ?

# Linear Approximation for the Cipher

Let $P_i$ be the random variable for the i-th plaintext bit, let $U_{r,i}$ be the random variable for the i-th input bit to the round r S-boxes, let $V_{r,i}$ be the random variable for the i-th output bit of the round r S-boxes, and let $K_{r,i}$ be the i-th bit of the r-th subkey.

Let $T_1, T_2, T_3, T_4$ be random variables such that

$$T_1 = U_{1,5} \oplus U_{1,7} \oplus U_{1,8} \oplus V_{1,6}$$
$$T_2 = U_{2,6} \oplus V_{2,6} \oplus V_{2,8}$$
$$T_3 = U_{3,6} \oplus V_{3,6} \oplus V_{3,8}$$
$$T_4 = U_{3,14} \oplus V_{3,14} \oplus V_{3,16}$$

Note: the $T_i$'s are not independent but pretending that they are indep. works well in practice.

# Linear Approximation for the Cipher

Let $P_i$ be the random variable for the i-th plaintext bit, let $U_{r,i}$ be the random variable for the i-th input bit to the round r S-boxes, let $V_{r,i}$ be the random variable for the i-th output bit of the round r S-boxes, and let $K_{r,i}$ be the i-th bit of the r-th subkey.

Let $T_1, T_2, T_3, T_4$ be random variables such that

$$T_1 = U_{1,5} \oplus U_{1,7} \oplus U_{1,8} \oplus V_{1,6}$$
$$T_2 = U_{2,6} \oplus V_{2,6} \oplus V_{2,8}$$
$$T_3 = U_{3,6} \oplus V_{3,6} \oplus V_{3,8}$$
$$T_4 = U_{3,14} \oplus V_{3,14} \oplus V_{3,16}$$

Applying the Piling-up Lemma:
what is the bias of $T_1 \oplus T_2 \oplus T_3 \oplus T_4 = 0$ ?

$$2^{4-1} \cdot \frac{1}{4} \cdot \left(-\frac{1}{4}\right)^3 = -\frac{1}{32} = -0.03125$$

# Linear Approximation for the Cipher

Expressing $T_1 \oplus T_2 \oplus T_3 \oplus T_4$ as the XOR of plaintext bits, subkey bits, and bits of the input (straightforward but tedious):

3 plaintext bits

4 bits computed via S-box approx. (just before the final S-box comput.)

$$T_1 \oplus T_2 \oplus T_3 \oplus T_4 = \underline{P_5 \oplus P_7 \oplus P_8} \oplus \underline{U_{4,6} \oplus U_{4,8} \oplus U_{4,14} \oplus U_{4,16}} \oplus K_{1,5} \oplus K_{1,7} \oplus K_{1,8} \oplus K_{2,6} \oplus K_{3,6} \oplus K_{3,14} \oplus K_{4,6} \oplus K_{4,8} \oplus K_{4,14} \oplus K_{4,16}$$

→ fixed, indep. of the plaintext

For fixed key bits, their XOR-sum is <u>either 0 or 1</u>. Then the bias of

$$P_5 \oplus P_7 \oplus P_8 \oplus U_{4,6} \oplus U_{4,8} \oplus U_{4,14} \oplus U_{4,16} = 0$$

← if the XOR of the key-bits is 0, then this bias is $-\frac{1}{32}$

is either -1/32 or 1/32.

# Extracting Key Bits

Recall: we are performing a known-plaintext attack, and we assume that we have a large pool of plaintext-ciphertext pairs (all encrypted with the same key).

How to use our linear approximation to determine a part of subkey $K_5$ ?

We will partially decrypt each ciphertext, and see if our linear approximation

$$P_5 \oplus P_7 \oplus P_8 \oplus U_{4,6} \oplus U_{4,8} \oplus U_{4,14} \oplus U_{4,16} = 0$$

holds or not.

# Extracting Key Bits

In particular, we will go through all possible $2^8$ possibilities for the subkey bits $K_{5,5}$, $K_{5,6}$, $K_{5,7}$, $K_{5,8}$, $K_{5,13}$, $K_{5,14}$, $K_{5,15}$, $K_{5,16}$.

For each candidate subkey, compute the bias of

$$P_5 \oplus P_7 \oplus P_8 \oplus U_{4,6} \oplus U_{4,8} \oplus U_{4,14} \oplus U_{4,16} = 0$$

(described on the next slide).

We are looking for a subkey for which the bias is the closest to 1/32 or -1/32.

# Extracting Key Bits

How to compute the bias for a specific candidate subkey ? For each plaintext-ciphertext pair, partially decrypt the ciphertext (in our case, XOR with the candidate subkey, then invert the two S-boxes to get $U_{4,5}$, $U_{4,6}$, $U_{4,7}$, $U_{4,8}$, $U_{4,13}$, $U_{4,14}$, $U_{4,15}$, $U_{4,16}$), then compute the value of

$$P_5 \oplus P_7 \oplus P_8 \oplus U_{4,6} \oplus U_{4,8} \oplus U_{4,14} \oplus U_{4,16}. = 0 \quad (*)$$

Determine the fraction of plaintext-ciphertext pairs for which this value is 0, subtract $\frac{1}{2}$ to get the bias (see Table 5).

```
bestbias = undef.
for all $2^8$ subkey possib. for $K_{5,5} \dots K_{5,8}$ , $K_{5,13} \dots K_{5,16}$
    count = 0
    for all plaintext/ ciphertext pairs (suppose there are b such pairs)
        XOR the key with the ciphertext to get $V_{4,5} \dots V_{4,8}$ , $V_{4,13} \dots V_{4,16}$
        run the S-boxes backward to get $U_{4,5} \dots U_{4,8}$ , $U_{4,13} \dots U_{4,16}$
        check if the equation (*) holds → if yes, count++
    bias = count / b - 1/2  ←  choose the subkey for which |bias| is the closest to $\frac{1}{32}$
```

# Extracting Key Bits

How many plaintext-ciphertext pairs do we need ?
If the bias is $\epsilon$ (for us $|\epsilon|=1/32$), we need about $c\epsilon^{-2}$ pairs for some "small" constant c. For our example c=8 is sufficient.

How many pairs do we need for our example ?

$$c \cdot \frac{1}{\varepsilon^2} = c \cdot (32)^2 = 8 \cdot 32^2$$

**Questions:**

- What are some disadvantages of linear cryptanalysis ?

- How can you make your SPN more secure against linear cryptanalysis ?