Linear Cryptanalysis

Nilanjan Datta

IAI, TCG CREST

tcg crest
Inventing Harmonious Future

Linear Cryptanalysis

- Consider a basic cipher: $C = M \oplus K$.
- Can you mount a key recovery attack?

Linear Cryptanalysis

- Consider a basic cipher: $C = M \oplus K$.
- Can you mount a key recovery attack?

Key Recovery Attack

- Make a query M_1. Say the ciphertext is C_1.
- Return $K = M_1 \oplus C_1$.

Linear Cryptanalysis

Consider a basic cipher of 4 bits:

$$
\begin{aligned}
C[1] &= M[1] \oplus M[2] \oplus K[1] \oplus K[2] \\
C[2] &= M[3] \oplus K[2] \oplus K[3] \\
C[3] &= M[1] \oplus M[3] \oplus K[3] \oplus K[4] \\
C[4] &= M[2] \oplus M[4] \oplus K[1] \oplus K[3]
\end{aligned}
$$

- Can you have a key recovery attack?
- What is the adversarial model?

Linear Cryptanalysis

Consider a basic cipher of 4 bits:

$$\begin{aligned}
C[1] &= M[1] \oplus M[2] \oplus K[1] \oplus K[2] \\
C[2] &= M[3] \oplus K[2] \oplus K[3] \\
C[3] &= M[1] \oplus M[3] \oplus K[3] \oplus K[4] \\
C[4] &= M[2] \oplus M[4] \oplus K[1] \oplus K[3]
\end{aligned}$$

- Guess a key say $K[1]$. Find $K[2], K[3], K[4]$. Complexity reduces from 2^4 to 2.
- Known Plaintext Attack is good enough to mount the attack.

Linear Cryptanalysis

Consider any cipher:

$$C = a \cdot M \oplus b \cdot K.$$

Generic Result

If ciphertext is a linear combination of the plaintext and the key, it is easy to mount

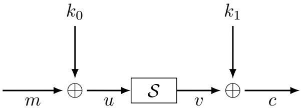
- key recovery attack,
- distinguishing attack.

What happens for non-linear functions?

Try to approximate a non-linear function by a linear function.

First Toy Cipher: Cipher1

$$c = S(m \oplus k_0) \oplus k_1$$



x	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
S(x)	F	E	B	C	6	D	7	8	0	3	9	A	4	2	1	5

Table: Sample S-Box

Can you recover the key here?

Main Idea: Linear Appoximation of the S-Box

Choose $\alpha = (1\ 0\ 0\ 1)$ and $\beta = (0\ 0\ 1\ 0)$.

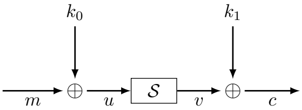
x	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
$S(x)$	F	E	B	C	6	D	7	8	0	3	9	A	4	2	1	5
$\alpha \cdot x$	0	1	0	1	0	1	0	1	1	0	1	0	1	0	1	0
$\beta \cdot S(x)$	1	1	1	0	1	0	1	0	0	1	0	1	0	1	0	0

Table: linear approximation of S-Box

Linear Approximation of the S-Box

$$\alpha \cdot x \oplus \beta \cdot S(x) = 1, \text{ with probability } 7/8.$$

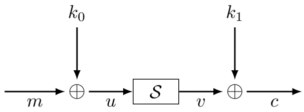
Main Idea: Linear Appoximation of the S-Box



Linear Approximation of the S-Box

- $u = m \oplus k_0$ with probability 1.
- $\alpha \cdot u \oplus \beta \cdot v = 1$ with probability $7/8$.
- $v = c \oplus k_1$ with probability 1.

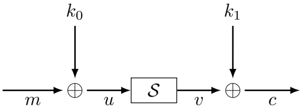
Main Idea: Linear Appoximation of the S-Box



Linear Approximation of the S-Box

- $\alpha \cdot u \oplus \beta \cdot v = 1$ with probability $7/8$.
- $\alpha \cdot (m \oplus k_0) \oplus \beta \cdot (c \oplus k_1) = 1$ with probability $7/8$.
- $\alpha \cdot k_0 \oplus \beta \cdot k_1 = \alpha \cdot m \oplus \beta \cdot c \oplus 1$ with probability $7/8$.

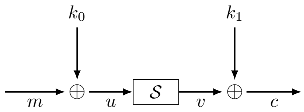
Main Idea: Linear Appoximation of the S-Box



Linear Approximation of the S-Box

- Key Recovery complexity reduces from 2^8 from 2^7.

Main Idea: Linear Appoximation of the S-Box



Interesting Observation

- If the probability of the linear approximation is $1/2$, you can not mount the attack.
- Goal: Find a linear approximation that has high deviation from $1/2$.

Combining Multiple Linear Approximations

Consider two random binary variables X_1 and X_2. Let $\Pr[X_1 = 0] = p_1$ and $\Pr[X_2 = 0] = p_2$

- Bias of X_i is defined by $p_i - 1/2$.

Combining Multiple Linear Approximations

Consider two random binary variables X_1 and X_2. Let $\Pr[X_1 = 0] = p_1$ and $\Pr[X_2 = 0] = p_2$

- Bias of X_i is defined by $p_i - 1/2$.
- If X_1 has bias ϵ_1 and X_2 has bias ϵ_2 and they are independent, what is the bias of $X_1 \oplus X_2$?

Combining Multiple Linear Approximations

Consider two random binary variables X_1 and X_2. Let $\Pr[X_1 = 0] = p_1$ and $\Pr[X_2 = 0] = p_2$

- Bias of X_i is defined by $p_i - 1/2$.
- If X_1 has bias ϵ_1 and X_2 has bias ϵ_2 and they are independent, what is the bias of $X_1 \oplus X_2$?
- Can you generalize it for any $X_1, \ldots, X_l$?

Combining Multiple Linear Approximations

Consider two random binary variables X_1 and X_2. Let $\Pr[X_1 = 0] = p_1$ and $\Pr[X_2 = 0] = p_2$

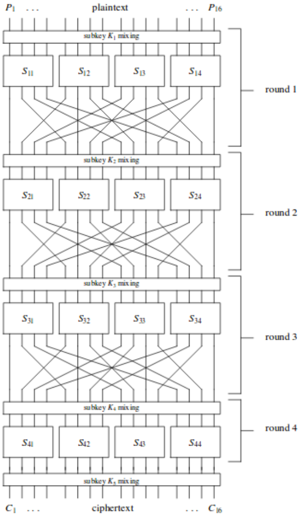
- Bias of X_i is defined by $p_i - 1/2$.
- If X_1 has bias ϵ_1 and X_2 has bias ϵ_2 and they are independent, what is the bias of $X_1 \oplus X_2$?
- Can you generalize it for any $X_1, \ldots, X_l$?

Piling-Up Lemma

If $\epsilon_{i_1,\ldots,i_l}$ denotes the bias of the random variable $X_{i_1} \oplus \cdots \oplus X_{i_l}$, then

$$\epsilon_{i_1,\ldots,i_l} = 2^{l-1} \prod_{j=1}^{l} \epsilon_{i_j}$$

Example of an Iterative SPN Block Cipher



Cipher4

- 16-bit Cipher
- Number of rounds: 4
- S-Box size: 4-bit

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

Table: S-Box

Examine Linear Pairs of the S-Box

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

Table: S-Box

X_1	X_2	X_3	X_4	Y_1	Y_2	Y_3	Y_4	$X_2 \oplus X_3$	$Y_1 \oplus Y_3 \oplus Y_4$	$X_1 \oplus X_4$	Y_2	$X_3 \oplus X_4$	$Y_1 \oplus Y_4$
0	0	0	0	1	1	1	0	0	0	0	1	0	1
0	0	0	1	0	1	0	0	0	0	1	1	1	0
0	0	1	0	1	1	0	1	1	0	0	1	1	0
0	0	1	1	0	0	0	1	1	1	1	0	0	1
0	1	0	0	0	0	1	0	1	1	0	0	0	0
0	1	0	1	1	1	1	1	1	1	1	1	1	0
0	1	1	0	1	0	1	1	0	1	0	0	1	0
0	1	1	1	1	0	0	0	0	1	1	0	0	1
1	0	0	0	0	0	1	1	0	0	1	0	0	1
1	0	0	1	1	0	1	0	0	0	0	0	1	1
1	0	1	0	0	1	1	0	1	1	1	1	1	0
1	0	1	1	1	1	0	0	1	1	0	1	0	1
1	1	0	0	0	1	0	1	1	1	1	1	0	1
1	1	0	1	1	0	0	1	1	0	0	0	1	0
1	1	1	0	0	0	0	0	0	0	1	0	1	0
1	1	1	1	0	1	1	1	0	0	0	1	0	1

Linear Approximation

Linear Approximation Table (LAT)

$2^n \times 2^n$ table to capture the linear approximation:

$$L_S(a, b) = |\{x \in \mathbb{F}_2^n : (a \cdot x) = (b \cdot S(x))\}| - 2^{n-1}.$$

Linearity

Maximum value in the LAT (non-zero appoximation):

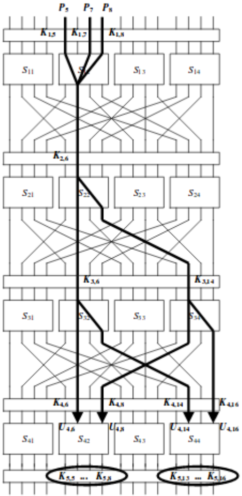
$$L_S = |max_{a,b \neq 0} L_S(a, b)|.$$

High Propagation Ratios for Linear Approximation Table (LAT) for the S-Box

Input Sum \ Output Sum	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

$\text{Bias}_{[1011 \rightarrow 0100]} = \frac{1}{4}, \quad \text{Bias}_{[0100 \rightarrow 0101]} = -\frac{1}{4}$

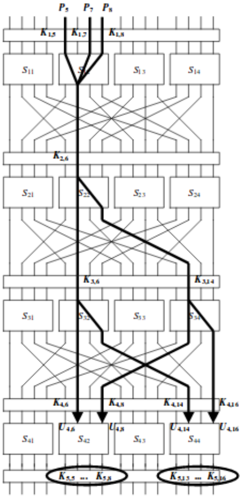
Linear Trail for the SPN



Computing Biases in the Propagation for the S-Boxes

- Bias of 1011 $\xrightarrow{S_2^1}$ 0100 is $\frac{1}{4}$
- Bias of 0100 $\xrightarrow{S_2^2}$ 0101 is $-\frac{1}{4}$
- Bias of 0100 $\xrightarrow{S_2^3}$ 0101 is $-\frac{1}{4}$
- Bias of 0100 $\xrightarrow{S_4^3}$ 0101 is $-\frac{1}{4}$

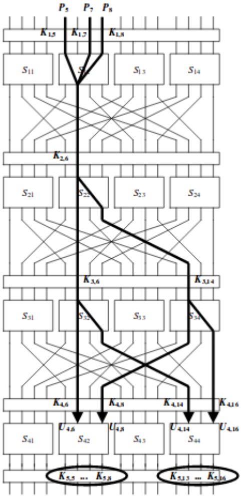
Linear Trail for the SPN



Linear Approximation for the First Round:

- $V_6^1 = U_5^1 \oplus U_7^1 \oplus U_8^1$ with bias $1/4$.
- $V_6^1 = (P_5 \oplus K_5^1) \oplus (P_7 \oplus K_7^1) \oplus (P_8 \oplus K_8^1)$ with bias $1/4$.

Linear Trail for the SPN



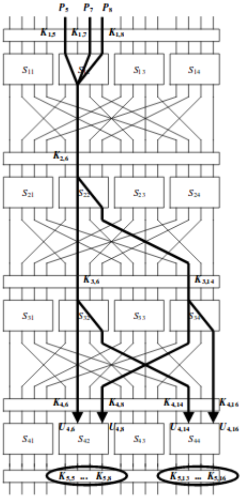
Linear Approximation for the First Round:

- $V_6^1 = U_5^1 \oplus U_7^1 \oplus U_8^1$ with bias $1/4$.
- $V_6^1 = (P_5 \oplus K_5^1) \oplus (P_7 \oplus K_7^1) \oplus (P_8 \oplus K_8^1)$ with bias $1/4$.

Linear Approximation for the Second Round:

- $V_6^2 \oplus V_8^2 = U_6^2$ with bias $-1/4$.
- $V_6^2 \oplus V_8^2 = V_6^1 \oplus K_6^2$ with bias $-1/4$.

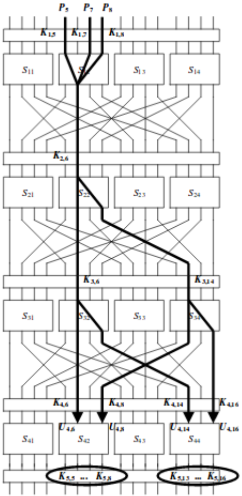
Linear Trail for the SPN



Linear Approximation upto Second Round (Piling-up Lemma):

- $V_6^2 \oplus V_8^2 \oplus P_5 \oplus K_5^1 \oplus P_7 \oplus K_7^1 \oplus P_8 \oplus K_8^1 \oplus K_6^2 = 0$ has bias $-1/8$.

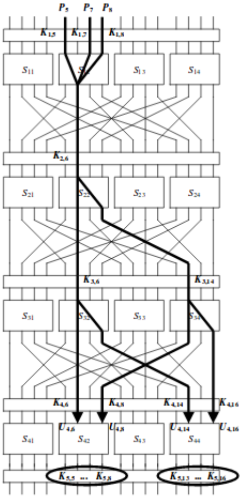
Linear Trail for the SPN



First Linear Approximation for the Third Round:

- $V_6^3 \oplus V_8^3 = U_6^3$ with bias $-1/4$.
- $V_6^3 \oplus V_8^3 = V_6^2 \oplus K_6^3$ with bias $-1/4$.

Linear Trail for the SPN



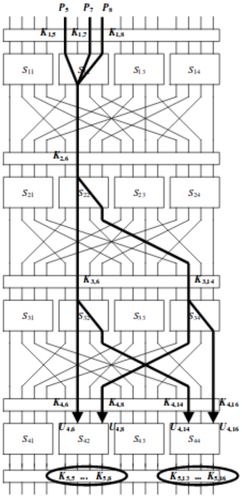
First Linear Approximation for the Third Round:

- $V_6^3 \oplus V_8^3 = U_6^3$ with bias $-1/4$.
- $V_6^3 \oplus V_8^3 = V_6^2 \oplus K_6^3$ with bias $-1/4$.

Second Linear Approximation for the Third Round:

- $V_{14}^3 \oplus V_{16}^3 = U_{14}^3$ with bias $-1/4$.
- $V_{14}^3 \oplus V_{16}^3 = V_8^2 \oplus K_{14}^3$ with bias $-1/4$.

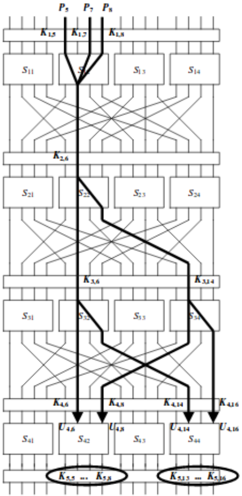
Linear Trail for the SPN



Linear Approximation for Third Round (Piling-up Lemma):

- $V_6^3 \oplus V_8^3 \oplus V_6^2 \oplus K_6^3 \oplus V_{14}^3 \oplus V_{16}^3 \oplus V_8^2 \oplus K_{14}^3 = 0$ has bias $1/8$.

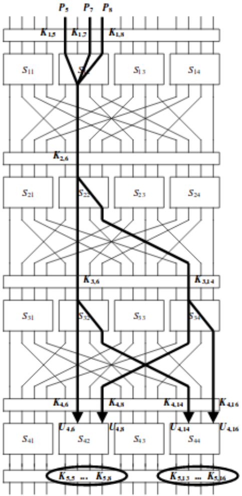
Linear Trail for the SPN



Linear Approximation Upto Third Round (Piling-up Lemma):

- $P_5 \oplus K_5^1 \oplus P_7 \oplus K_7^1 \oplus P_8 \oplus K_8^1 \oplus K_6^2 \oplus U_6^4 \oplus K_6^4 \oplus U_{14}^4 \oplus K_{14}^4 \oplus K_6^3 \oplus U_8^4 \oplus K_8^4 \oplus U_{16}^4 \oplus K_{16}^4 \oplus K_{14}^3 = 0$ has bias $-1/32$.

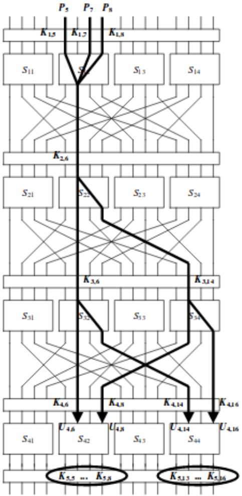
Linear Trail for the SPN



Linear Approximation Upto Third Round (Piling-up Lemma):

- $P_5 \oplus P_7 \oplus P_8 \oplus U_6^4 \oplus U_{14}^4 \oplus U_8^4 \oplus U_{16}^4 \oplus \Sigma_K = 0$ has bias $-1/32$.
- Since Σ_K is fixed (either 0 or 1), $P_5 \oplus P_7 \oplus P_8 \oplus U_6^4 \oplus U_{14}^4 \oplus U_8^4 \oplus U_{16}^4 = 0$ has bias of magnitude $1/32$.

Extracting Key-bits



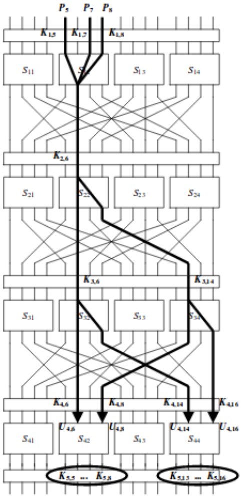
Objective

Extract bits from subkey K_5

Target partial sub-key bits

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

Extracting Key-bits



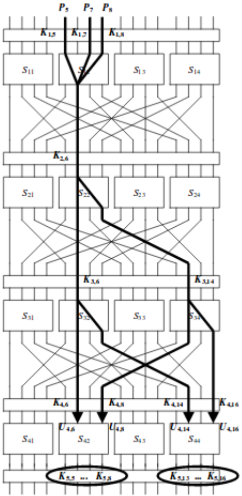
Objective

Extract bits from subkey K_5

Target partial sub-key bits

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

Extracting Key-bits



Towards Obtaining the partial key

- Collect 10000 (plaintext-ciphertext).
- For all possible values of the partial key:
  - Execute partial decryption to get U^4 values
  - $Count = \#$ the linear approximation holds
  - Compute the bias: $|bias| = |Count - 5000|/10000$

Extracting Key-bits

partial subkey $[K_{5,5}...K_{5,8}, K_{5,13}...K_{5,16}]$	| bias |	*partial subkey* $[K_{5,5}...K_{5,8}, K_{5,13}...K_{5,16}]$	| bias |
1 C	0.0031	2 A	0.0044
1 D	0.0078	2 B	0.0186
1 E	0.0071	2 C	0.0094
1 F	0.0170	2 D	0.0053
2 0	0.0025	2 E	0.0062
2 1	0.0220	2 F	0.0133
2 2	0.0211	3 0	0.0027
2 3	0.0064	3 1	0.0050
2 4	**0.0336**	3 2	0.0075
2 5	0.0106	3 3	0.0162
2 6	0.0096	3 4	0.0218
2 7	0.0074	3 5	0.0052
2 8	0.0224	3 6	0.0056
2 9	0.0054	3 7	0.0048

Report the partial sub-key with highest *prob* (here 0010 0100)

Estimation on the Number of Known (Plaintext,Ciphertext)

Active S-Boxes

S-Boxes involved in a linear characteristic

Find Linear Bias

γ: # Active S-Boxes

β_i: occurrence of the particular linear approximation in the i^{th} Active S-box of the characteristic

$$\mathrm{LB} = 2^{\gamma-1} \prod \beta_i,$$

- Number of Chosen (Plaintext,Ciphertext) Pair: $N_L = \frac{1}{\mathrm{LB}^2}$ (Result by Matsui)

How to Build Linear Cryptanalysis Resistant Cipher

Step 1: Calculate Minimum Number of Active S-Box (w) for round r

Use Mixed Integer Linear Programming (MILP)

How to Build Linear Cryptanalysis Resistant Cipher

Step 1: Calculate Minimum Number of Active S-Box (w) for round r

Use Mixed Integer Linear Programming (MILP)

Step 2: Find An (Trivial) Upper bound on the Linear Probability for round r

- Find Linear Characteristics (lc) of the S-Box (maximum propagation ratio)
- Compute $\mathrm{LB} = 2^{w-1}(\mathrm{lc})^w$

How to Build Linear Cryptanalysis Resistant Cipher

Step 1: Calculate Minimum Number of Active S-Box (w) for round r

Use Mixed Integer Linear Programming (MILP)

Step 2: Find An (Trivial) Upper bound on the Linear Probability for round r

- Find Linear Characteristics (lc) of the S-Box (maximum propagation ratio)
- Compute $\mathrm{LB} = 2^{w-1}(\mathrm{lc})^w$

Step 3: Estimate Number of Rounds r

Find r such that $\mathrm{LB}^2 \leq 2^{-n}$ (Recall number of Known Plaintext-Ciphertext Pairs)

Exercise

Given the following facts, find the minimum number of rounds for GIFT-64 to resist linear cryptanalysis:

- Linear bias of the S-Box is 2^{-2}.
- Number of active S-Boxes in the linear trail for any r rounds of GIFT-64 is $2r$.

References

- Howard Heys, *"A Tutorial on Linear and Differential Cryptanalysis"*
- Kazuo Sakiyama, Yu Sasaki and Yang Li, *"Security of Block Ciphers: From Algorithm Design to Hardware Implementation"*
- Douglas Stinson, *"Cryptography Theory and Practice"*

Thank You..!!!