Conditional Linear Cryptanalysis

Stav Perle

Joint work with Eli Biham

Technion - Israel Institute of Technology
Cryptanalytic Techniques

- Differential and linear cryptanalysis are two major generic techniques for assessing the strength and vulnerabilities of block ciphers
- These techniques have various extensions which can improve their success in various cases
- Along with Davies’ attack, they are the best attacks against the Data Encryption Standard (DES)

<table>
<thead>
<tr>
<th>Technique</th>
<th>Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Differential Cryptanalysis</td>
<td>$2^{47}$</td>
</tr>
<tr>
<td>Linear Cryptanalysis</td>
<td>$2^{43}$</td>
</tr>
<tr>
<td>Improved Davies’ Attack</td>
<td>$2^{50}$</td>
</tr>
</tbody>
</table>

- Today, we will show a new extension that reduces this complexity further

<table>
<thead>
<tr>
<th>Technique</th>
<th>Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conditional Linear Cryptanalysis</td>
<td>$\leq2^{42}$</td>
</tr>
</tbody>
</table>
DES - Example of a Block Cipher

Plaintext (64 bits)

Key (56 bits)

Ciphertext (64 bits)
The F Function of DES

input (32 bits) → E → 48 bits

subkey (48 bits) → S1 → S2 → S3 → S4 → S5 → S6 → S7 → S8

output (32 bits)

Examples:

000000 → 0010
000001 → 1110
000010 → 1100
...
111110 → 0100
111111 → 0010
The Best Non-Trivial Approximation of S5

- It approximates the second bit of input to the XOR of the four output bits
  - In 12 cases: \( I_2 \oplus O_1 \oplus O_2 \oplus O_3 \oplus O_4 = 0 \)
  - In 52 cases: \( I_2 \oplus O_1 \oplus O_2 \oplus O_3 \oplus O_4 = 1 \)
Linear Cryptanalysis

- Linear Cryptanalysis uses statistical approximations that approximate parity of subsets of bits of the plaintext, ciphertext, and the subkeys.
  - E.g., (second bit of the plaintext) XOR (fifth bit of ciphertext) XOR (keys bits) = 0

- Each approximation has a probability, \( p \), to hold.
  - Which is the fraction of plaintexts whose encryption follow the approximation.
  - In random cases, the probability is expected to be \( \frac{1}{2} \), or close to \( \frac{1}{2} \).
Linear Cryptanalysis

- The ability to distinguish whether an approximation holds highly depends on the distance of the probability from $\frac{1}{2}$

- Let the bias be $\varepsilon = p - \frac{1}{2}$
  - Range: $-\frac{1}{2}$ to $+\frac{1}{2}$
  - The higher the (absolute value of the) bias, the easier to distinguish
  - $\varepsilon = 0$ means that the approximation is mostly useless
Linear Approximations

- A linear approximation is a tuple $(\lambda_P, \lambda_C, \lambda_K)$
  - $\lambda_P$ is a subset of bits of the plaintext
  - $\lambda_C$ is a subset of bits of the ciphertext
  - $\lambda_K$ is a subset of bits of the key (or the subkeys)

- The probability of the approximation is the probability that $P\lambda_P \oplus C\lambda_C \oplus K\lambda_K = 0$
Linear Approximations - Examples

- The best non-trivial approximation

\[ \lambda_p = 21\ 04\ 00\ 80\ 00\ 00\ 00\ 00_x \]

\[ 21\ 04\ 00\ 80_x \]

\[ 00\ 80\ 00_x \]

\[ P(00\ 00\ 00\ 00_x) \]

\[ F \]

\[ \lambda_c = 21\ 04\ 00\ 80\ 00\ 00\ 08\ 00_x \]

- It approximates the second bit of input to S5 to the XOR of the four output bits of S5

  - In 12 cases: \( P\lambda_P \oplus C\lambda_C \oplus K\lambda_K = 0 \)
  - In 52 cases: \( P\lambda_P \oplus C\lambda_C \oplus K\lambda_K = 1 \)

\[ \frac{1}{2} + \varepsilon = \frac{12}{64} \]

\[ \varepsilon = \frac{-20}{64} \]
This linear approximation has probability \( \frac{1}{2} + 2\left(\frac{-20}{64}\right)^2 = \frac{1}{2} + \frac{25}{128} \)
Given $\lambda=(\lambda_p, \lambda_c, \lambda_k)$, we know that $P\lambda_p \oplus C\lambda_c \oplus K\lambda_k = 0$ holds with probability $p = \frac{1}{2} + \varepsilon$

Given plaintext and the corresponding ciphertext, we can calculate the value of $P\lambda_p \oplus C\lambda_c$
Algorithm 1

- Given $\lambda = (\lambda_P, \lambda_C, \lambda_K)$, $\varepsilon(\lambda)$, and $N$ plaintexts and their ciphertexts, the algorithm counts the number $M$ of plaintexts satisfying

$$P\lambda_P \oplus C\lambda_C = 0$$

- Recall that $P\lambda_P \oplus C\lambda_C \oplus K\lambda_K = 0$ holds with probability $p = \frac{1}{2} + \varepsilon$

- The algorithm guesses that the parity of the key bits $K\lambda_K$ is

<table>
<thead>
<tr>
<th>$\varepsilon &gt; 0$</th>
<th>$\varepsilon &lt; 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M &gt; \frac{N}{2}$</td>
<td>0</td>
</tr>
<tr>
<td>$M &lt; \frac{N}{2}$</td>
<td>1</td>
</tr>
</tbody>
</table>

- This algorithm finds only one parity bit of the key

- The success rate of the algorithm grows as the number of plaintexts $N$ increases, and as the value of $|\varepsilon|$ increases

- For a high probability of success, $N \approx \frac{1}{\varepsilon^2}$ or higher
Matsui’s Best Approximation (14 rounds)

\[ \lambda_p = 21040080 \ 00008000_x \]

\[ \lambda_c = 00000000 \ 01040080_x \]
Matsui uses the best 14-round approximation with probability $\approx \frac{1}{2} - 2^{-20.75}$.

The attack requires about $2^{43}$ known plaintexts.
Conditional Linear Cryptanalysis

- Using conditions to discard data that reduces that bias
  - So the bias of the remaining data increase or decrease
- Conditions can be by any observable data available to the cryptanalyst
  - Plaintexts, ciphertexts, and formulae on them
- In Feistel ciphers we can compute various linear combinations of internal bits directly from the plaintext and ciphertext
A Case of Single Round

- The best non-trivial approximation

\[ \lambda_P = 21\ 04\ 00\ 80\ 00\ 00\ 00_2 \]

\[ \lambda_C = 21\ 04\ 00\ 80\ 00\ 08\ 00_2 \]

- It approximates the second bit of input to S5 to the XOR of the four output bits of S5
  - Probability \( \frac{1}{2} + \frac{-20}{64} \)
  - I.e., 12 cases with equality (parity 0 of the 5 bits), 52 cases with inequality (parity 1 of the 5 bits)
A Case of Single Round

- Conditioning on all the four output bits of S5 (16 cases) we get

\[
\begin{array}{cccccccccccccc}
0000 & 0001 & 0010 & 0011 & 0100 & 0101 & 0110 & 0111 & 1000 & 1001 & 1010 & 1011 & 1100 & 1101 & 1110 & 1111 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 2 & 0 & 2 & 0 & 3 & 1 & 2 & 0 \\
-0.5 & -0.5 & -0.5 & -0.5 & -0.25 & -0.5 & -0.25 & -0.5 & 0 & -0.5 & 0 & -0.5 & 0.25 & -0.25 & 0 & -0.5
\end{array}
\]

- Consider a condition on the LSB of the four output bits of S5 (a single bit)

<table>
<thead>
<tr>
<th>Condition</th>
<th>$P\lambda_p \oplus C\lambda_c \oplus K\lambda_K = 0$</th>
<th>$P\lambda_p \oplus C\lambda_c \oplus K\lambda_K = 1$</th>
<th>Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>none</td>
<td>12</td>
<td>52</td>
<td>-20/64</td>
</tr>
<tr>
<td>LSB=0</td>
<td>11</td>
<td>21</td>
<td>-5/32=-10/64</td>
</tr>
<tr>
<td>LSB=1</td>
<td>1</td>
<td>31</td>
<td>-15/32=-30/64</td>
</tr>
</tbody>
</table>
A Case of Single Round

- Scan from Adi Shamir’s CRYPTO’85 paper
  - He circled the values with an even parity of the four output bits

The 12 VS. the 52

- 1 vs. 31 for LSB=1, and 11 vs. 21 for LSB=0
Conditional Linear Cryptanalysis

- We can condition on the XOR of plaintext and ciphertext bits
  - even more than one bit at a time
- For example, on $P_L \oplus C_R = \bigoplus_{r \text{ is odd}} Y^r$
  - which is the XOR of the output of $F$ in all odd rounds
- Consider any one of these bits as a linear approximation
  - E.g., $P_{L,17} \oplus C_{R,17} = 0$
    - Equivalent to $Y^1_{17} \oplus Y^3_{17} = 0$
    - Such approximations are expected to have bias 0
- But they are very useful as conditions to other approximations
A Four-Round Example

- Consider four successive rounds taken from Matsui’s best linear approximation
- This approximation uses three active S boxes:
  - S5 on the first and third rounds, and
  - S1 on the fourth round
- Both odd rounds have the same active S box

\[
\begin{align*}
\lambda_p &= \begin{pmatrix} 21 & 04 & 00 & 80 & 00 & 00 & 00 & 00 \\ 00 & 00 & 80 & 00 & 00 & 00 & 00 & 00 \\ \end{pmatrix} \\
\lambda_c &= \begin{pmatrix} 00 & 00 & 00 & 00 & 00 & 01 & 04 & 00 \\ 00 & 00 & 80 & 00 & 00 & 00 & 00 & 00 \\ \end{pmatrix} \\
\end{align*}
\]
A Four-Round Example

- Conditioning on all the four XOR output bits of S5 (16 cases) we get

\[
\begin{array}{cccccccccccccccc}
0000 & 0001 & 0010 & 0011 & 0100 & 0101 & 0110 & 0111 & 1000 & 1001 & 1010 & 1011 & 1100 & 1101 & 1110 & 1111 \\
0.008 & 0 & 0.008 & 0 & 0.009 & 0 & 0.009 & 0 & 0.014 & 0 & 0.014 & 0 & 0.015 & 0 & 0.015 & 0 \\
\end{array}
\]

- Notice that this condition is based on the XOR of both odd rounds
  - Not just on one of them

- For applying Matsui’s Algorithm1 with our observation, we discard half of the known plaintexts, and use only the plaintexts in which the XOR of the LSB bits of S5 is zero
  - Their average bias is 0.0115
  - While the bias over all cases is 0.0057

- Using only these plaintexts increases the bias by a factor of two
A Four-Round Example

- We need a quarter of the data
  - Compared to a regular linear attack with the same approximation
  - But this is after we discard half of the given data that fails the condition
- We need half of the original data
  - We discard half of it, and get the required quarter
Conditional Linear Cryptanalysis of the Full DES

\[ \lambda_1 \]

\[ \lambda_p = 21 \ 04 \ 00 \ 80 \ 00 \ 00 \ 00 \]

\[ \lambda_c = 00 \ 00 \ 00 \ 00 \ 01 \ 04 \ 00 \ 80 \]

\[ P(00 \ 00 \ 00 \ 00) \]

\[ F \]

\[ 0 \]

\[ 0 \]

\[ 21 \ 04 \ 00 \ 80 \]

\[ P(00 \ 00 \ 00 \ 00) \]

\[ 00 \ 00 \ 00 \ 00 \ 00 \]

\[ F \]

\[ 0 \]

\[ 0 \]

\[ 21 \ 04 \ 00 \ 80 \]

\[ P(00 \ 00 \ 00 \ 00) \]
Conditional Linear Cryptanalysis of the Full DES

- Conditioning $\lambda_1$ on all the four XOR output bits of S5 (16 cases) we get

<table>
<thead>
<tr>
<th>Condition</th>
<th>Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>none</td>
<td>$\sim - 2^{-20.75}$</td>
</tr>
<tr>
<td>XOR LSB=0</td>
<td>$\sim - 2^{-21.74}$</td>
</tr>
<tr>
<td>XOR LSB=1</td>
<td>$\sim - 2^{-20.16}$</td>
</tr>
</tbody>
</table>
Conditional Linear Cryptanalysis of the Full DES

$\lambda_p = 21 \ 04 \ 00 \ 80 \ 00 \ 01 \ 10 \ 00_x$

$\lambda_1$

$\lambda_2$

$\lambda_c = 00 \ 00 \ 00 \ 00 \ 01 \ 04 \ 00 \ 80_x$
Conditional Linear Cryptanalysis of the Full DES

- Conditioning \( \lambda_1 \) on all the four XOR output bits of S5 (16 cases) we get

<table>
<thead>
<tr>
<th>0000</th>
<th>0001</th>
<th>0010</th>
<th>0011</th>
<th>0100</th>
<th>0101</th>
<th>0110</th>
<th>0111</th>
<th>1000</th>
<th>1001</th>
<th>1010</th>
<th>1011</th>
<th>1100</th>
<th>1101</th>
<th>1110</th>
<th>1111</th>
</tr>
</thead>
</table>

- Conditioning \( \lambda_2 \) on all the four XOR output bits of S5 (16 cases) we get

<table>
<thead>
<tr>
<th>0000</th>
<th>0001</th>
<th>0010</th>
<th>0011</th>
<th>0100</th>
<th>0101</th>
<th>0110</th>
<th>0111</th>
<th>1000</th>
<th>1001</th>
<th>1010</th>
<th>1011</th>
<th>1100</th>
<th>1101</th>
<th>1110</th>
<th>1111</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Condition</th>
<th>Bias ( \lambda_1 )</th>
<th>Bias ( \lambda_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>none</td>
<td>( \sim 2^{-20.75} )</td>
<td>( \sim 2^{-21.48} )</td>
</tr>
<tr>
<td>LSB=0</td>
<td>( \sim 2^{-21.74} )</td>
<td>( \sim 2^{-20.26} )</td>
</tr>
<tr>
<td>LSB=1</td>
<td>( \sim 2^{-20.16} )</td>
<td>( \sim 2^{-23.06} )</td>
</tr>
</tbody>
</table>
Success Probability by Complexity (#Ps&Time)
Known Plaintexts when Time is Fixed to $2^{50}$
Success Probability by Time for Various #KPs
Matsui’s Success Prob. by Time for Various #KPs
The End