On Reverse-Engineering S-Boxes

Alex Biryukov¹, Léo Perrin¹, Aleksei Udovenko¹

¹SnT, University of Luxembourg
https://www.cryptolux.org

March 28, 2017
CryptoAction Symposium 2017
An S-Box is a small non-linear function mapping $m$ bits to $n$ usually specified via its look-up table.
An S-Box is a small non-linear function mapping $m$ bits to $n$ usually specified via its look-up table.

- Typically, $m = n$, $n \in \{4, 8\}$
- Used by many block ciphers/hash functions/stream ciphers.
- Necessary for the wide trail strategy.

Screen capture from [GOST, 2015].
S-Box Design

- Inverse
- Exponential
- Math (other)
- SPN
- Misty
- Feistel
- Lai-Massey
- Pseudo-random
- Hill climbing
- Unknown
S-Box Design
S-Box Design

Introduction

AES → Whirlpool

AES → Scream

Biryukov, Perrin, Udovenko
On Reverse-Engineering S-Boxes 3 / 28
Introduction

S-Box Reverse-Engineering

- Inverse
- Exponential
- Math (other)
- SPN
- Misty
- Feistel
- Lai-Massey
- Pseudo-random
- Hill climbing
- Unknown

S
Introduction

S-Box Reverse-Engineering

- Inverse
- Exponential
- Math (other)
- SPN
- Misty
- Feistel
- Lai-Massey
- Pseudo-random
- Hill climbing
- Unknown
Outline

1. Introduction
2. Mathematical Background
3. Detailed Analysis of the Two Tables
4. TU-Decomposition
5. Conclusion
Plan

1. Introduction

2. Mathematical Background
   - The Two Tables
   - Coefficients Distribution

3. Detailed Analysis of the Two Tables

4. TU-Decomposition

5. Conclusion
The Two Tables

Let \( S : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n \) be an S-Box.
The Two Tables

Let $S : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be an S-Box.

**Definition (DDT)**

The *Difference Distribution Table* of $f$ is a matrix of size $2^n \times 2^n$ such that

$$
\text{DDT}[a, b] = \# \{ x \in \mathbb{F}_2^n \mid S(x \oplus a) \oplus S(x) = b \}.
$$
The Two Tables

Let $S : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be an S-Box.

**Definition (DDT)**

The *Difference Distribution Table* of $f$ is a matrix of size $2^n \times 2^n$ such that

$$\text{DDT}[a, b] = \# \{ x \in \mathbb{F}_2^n \mid S(x \oplus a) \oplus S(x) = b \}.$$

**Definition (LAT)**

The *Linear Approximations Table* of $S$ is a matrix of size $2^n \times 2^n$ such that

$$\text{LAT}[a, b] = \# \{ x \in \mathbb{F}_2^n \mid x \cdot a = S(x) \cdot b \} - 2^{n-1}.$$
Example

$S = [4, 2, 1, 6, 0, 5, 7, 3]$

**The DDT of $S$.**

$$
\begin{bmatrix}
8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 2 & 2 & 2 & 2 \\
0 & 0 & 0 & 0 & 2 & 2 & 2 & 2 \\
0 & 0 & 4 & 4 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 2 & 2 & 2 & 2 \\
0 & 4 & 4 & 0 & 0 & 0 & 0 & 0 \\
0 & 4 & 0 & 4 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 2 & 2 & 2 & 2
\end{bmatrix}
$$

**The LAT of $S$.**

$$
\begin{bmatrix}
4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 2 & 2 & 0 & 0 & 2 & -2 \\
0 & 2 & 2 & 0 & 0 & 2 & -2 & 0 \\
0 & 2 & 0 & 2 & 0 & -2 & 0 & 2 \\
0 & 2 & 0 & -2 & 0 & -2 & 0 & -2 \\
0 & -2 & 2 & 0 & 0 & -2 & -2 & 0 \\
0 & 0 & -2 & 2 & 0 & 0 & -2 & -2 \\
0 & 0 & 0 & 0 & -4 & 0 & 0 & 0
\end{bmatrix}
$$
If an $n$-bit S-Box is bijective, then its DDT coefficients behave like independent and identically distributed random variables following a Poisson distribution:

$$\Pr [\text{DDT}[a, b] = 2^z] = \frac{e^{-1/2}}{2^z z}.$$
Coefficient Distribution in the LAT

If an $n$-bit S-Box is bijective, then its LAT coefficients behave like independent and identically distributed random variables following this distribution:

$$\Pr \left[ \text{LAT}[a, b] = 2^z \right] = \frac{\binom{2^{n-1}}{2^{n-2+z}}}{\binom{2^n}{2^{n-1}}}.$$
Plan

1. Introduction

2. Mathematical Background

3. Detailed Analysis of the Two Tables
   - Maximum Values in the Tables
   - Application to Skipjack

4. TU-Decomposition

5. Conclusion
Looking Only at the Maximum

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>$\log_2 (\Pr [\max(D) \leq \delta])$</th>
<th>$\ell$</th>
<th>$\log_2 (\Pr [\max(L) \leq \ell])$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>-1359.530</td>
<td>22</td>
<td>-371.609</td>
</tr>
<tr>
<td>6</td>
<td>-164.466</td>
<td>24</td>
<td>-161.900</td>
</tr>
<tr>
<td>8</td>
<td>-16.148</td>
<td>26</td>
<td>-66.415</td>
</tr>
<tr>
<td>10</td>
<td>-1.329</td>
<td>28</td>
<td>-25.623</td>
</tr>
<tr>
<td>12</td>
<td>-0.094</td>
<td>30</td>
<td>-9.288</td>
</tr>
<tr>
<td>14</td>
<td>-0.006</td>
<td>32</td>
<td>-3.160</td>
</tr>
<tr>
<td></td>
<td></td>
<td>34</td>
<td>-1.008</td>
</tr>
<tr>
<td></td>
<td></td>
<td>36</td>
<td>-0.302</td>
</tr>
<tr>
<td></td>
<td></td>
<td>38</td>
<td>-0.084</td>
</tr>
</tbody>
</table>

**DDT**

Probability that the maximum coefficient in the DDT/LAT of an 8-bit permutation is at most equal to a certain threshold.
Pr [max(LAT) = 24], Pr [max(LAT) = 26], Pr [max(LAT) = 28], Pr [max(LAT) = 30]
What is Skipjack? (1/2)

- **Type**: Block cipher
- **Block**: 64 bits
- **Key**: 80 bits
- **Authors**: NSA
- **Publication**: 1998
What is Skipjack? (2/2)

- Skipjack was supposed to be secret...
- ... but eventually published in 1998 [National Security Agency, 1998],
What is Skipjack? (2/2)

- Skipjack was supposed to be secret...
- ... but eventually published in 1998 [National Security Agency, 1998],
- It uses a $8 \times 8$ S-Box ($F$) specified only by its LUT,
Skipjack was supposed to be secret...

... but eventually published in 1998 [National Security Agency, 1998],

It uses a $8 \times 8$ S-Box ($F$) specified only by its LUT,

Skipjack was to be used by the *Clipper Chip*.
Skipjack was supposed to be secret...

... but eventually published in 1998 [National Security Agency, 1998],

It uses a $8 \times 8$ S-Box ($F$) specified only by its LUT,

Skipjack was to be used by the Clipper Chip.
Reverse-Engineering $F$

For Skipjack, $\max(LAT) = 28$ and $\#28 = 3$. 

Reverse-Engineering $F$

For Skipjack, $\text{max}(\text{LAT}) = 28$ and $\#28 = 3$. 

![Graphs showing probability distributions for different $N$ values: $N_{26}$, $N_{28}$, $N_{30}$, with probability values on the y-axis and $N$ values on the x-axis.](image)
Reverse-Engineering $F$

For Skipjack, $\max(LAT) = 28$ and $\#28 = 3.$
Reverse-Engineering F

For Skipjack, \( \text{max}(\text{LAT}) = 28 \) and \( \#28 = 3 \).

\[
\Pr \left[ \text{max}(\text{LAT}) = 28 \text{ and } \#28 = 3 \right] \approx 2^{-55}
\]
What Can We Deduce?

- $F$ has not been picked uniformly at random.
- $F$ has not been picked among a feasibly large set of random S-Boxes.
- Its linear properties were optimized (though poorly).
What Can We Deduce?

- $F$ has not been picked uniformly at random.
- $F$ has not been picked among a feasibly large set of random S-Boxes.
- Its linear properties were optimized (though poorly).

The S-Box of Skipjack was built using a dedicated algorithm.
Conclusion for Skipjack

![Pie chart showing various categories: Inverse, Exponential, Math (other), SPN, Misty, Feistel, Lai-Massey, Pseudo-random, Hill climbing, Unknown.]

Biryukov, Perrin, Udovenko

On Reverse-Engineering S-Boxes
Plan

1. Introduction
2. Mathematical Background
3. Detailed Analysis of the Two Tables
4. TU-Decomposition
   - Principle
   - Results on Kuznyechik/Streebog
5. Conclusion
TU-Decomposition in a Nutshell

1. Identify linear patterns in zeroes of LAT;

2. Deduce linear layers $\mu$, $\eta$ such that $\pi$ is decomposed as in right picture;

3. Decompose $U$, $T$;

4. Put it all together.
TU- Decomposition in a Nutshell

1. Identify linear patterns in zeroes of LAT;
2. Deduce linear layers $\mu$, $\eta$ such that $\pi$ is decomposed as in right picture;
TU-Decomposition in a Nutshell

1. Identify linear patterns in zeroes of LAT;

2. Deduce linear layers $\mu, \eta$ such that $\pi$ is decomposed as in right picture;

3. Decompose $U, T;$
TU-Decomposition in a Nutshell

1. Identify linear patterns in zeroes of LAT;
2. Deduce linear layers $\mu, \eta$ such that $\pi$ is decomposed as in right picture;
3. Decompose $U, T$;
4. Put it all together.
## Kuznyechik/Stribog

<table>
<thead>
<tr>
<th><strong>Stribog</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type</strong></td>
<td>Hash function</td>
</tr>
<tr>
<td><strong>Publication</strong></td>
<td>[GOST, 2012]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Kuznyechik</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type</strong></td>
<td>Block cipher</td>
</tr>
<tr>
<td><strong>Publication</strong></td>
<td>[GOST, 2015]</td>
</tr>
</tbody>
</table>

Both are standard symmetric primitives in Russia. Both were designed by the FSB (TC26). Both use the same $8 \times 8$ S-Box, $\pi$. 

Biryukov, Perrin, Udovenko

On Reverse-Engineering S-Boxes
Kuznyechik/Stribog

---

**Stribog**

- **Type**: Hash function
- **Publication**: [GOST, 2012]

**Kuznyechik**

- **Type**: Block cipher
- **Publication**: [GOST, 2015]

---

**Common ground**

- Both are standard symmetric primitives in Russia.
- Both were designed by the FSB (TC26).
- Both use the same $8 \times 8$ S-Box, $\pi$. 
The LAT of $\pi$
The LAT of $\eta \circ \pi \circ \mu$
Final Decomposition Number 1

- Multiplication in $\mathbb{F}_2^4$
- $\alpha$ Linear permutation
- $I$ Inversion in $\mathbb{F}_2^4$
- $\nu_0, \nu_1, \sigma$ 4 × 4 permutations
- $\phi$ 4 × 4 function
- $\omega$ Linear permutation
Final Decomposition Number 1

\[ P[\nu_1(x \oplus 0x9) \oplus \nu_1(x) = 0x2] = 1 \]

- \( \alpha \): Linear permutation
- \( I \): Inversion in \( \mathbb{F}_{2^4} \)
- \( \nu_0, \nu_1, \sigma \): 4 \times 4 permutations
- \( \phi \): 4 \times 4 function
- \( \omega \): Linear permutation

\( \odot \): Multiplication in \( \mathbb{F}_{2^4} \)
Conclusion for Kuznyechik/Stribog?

The Russian S-Box was built like a strange Feistel...
Conclusion for Kuznyechik/StriBog?

The Russian S-Box was built like a strange Feistel...

... or was it?
Conclusion for Kuznyechik/Striobg?

The Russian S-Box was built like a strange Feistel...

... or was it?

Belarussian inspiration

- The last standard of Belarus [STB 34.101.31-2011, 2011] uses an 8-bit S-box,
- somewhat similar to π...
Conclusion for Kuznyechik/Stribog?

The Russian S-Box was built like a strange Feistel...

... or was it?

Belarussian inspiration

- The last standard of Belarus [STB 34.101.31-2011, 2011] uses an 8-bit S-box,
- somewhat similar to $\pi$...
- ... based on a finite field exponential!
Conclusion for Kuznyechik/Strigob?

The Russian S-Box was built like a strange Feistel...

... or was it?

Belarussian inspiration

- The last standard of Belarus \([\text{STB 34.101.31-2011, 2011}]\) uses an 8-bit S-box,
- somewhat similar to \(\pi\)...
- ... based on a finite field exponential!

Exponential in \(\pi\)

\[ \pi \circ \exp \]

has \(\text{max}(DDT) = 128\ (\text{Pr} < 2^{-340})\) and a TU-decomposition!
Final Decomposition Number 2 (!)

$log_{w, 16}$

$-1$ 

$q'$

$\omega'$

| $T_0$ | 0 1 2 3 4 5 6 7 8 9 a b c d e f |
| $T_1$ | 0 1 2 3 4 5 6 7 8 9 a b c d e f |
| $T_2$ | 0 1 2 3 4 5 6 7 8 9 a b c d f e |
| $T_3$ | 0 1 2 3 4 5 6 7 8 9 a b c f d e |
| $T_4$ | 0 1 2 3 4 5 6 7 8 9 a b f c d e |
| $T_5$ | 0 1 2 3 4 5 6 7 8 9 a f b c d e |
| $T_6$ | 0 1 2 3 4 5 6 7 8 9 f a b c d e |
| $T_7$ | 0 1 2 3 4 5 6 7 8 9 f 9 a b c d e |
| $T_8$ | 0 1 2 3 4 5 6 7 f 8 9 a b c d e |
| $T_9$ | 0 1 2 3 4 5 6 f 7 8 9 a b c d e |
| $T_a$ | 0 1 2 3 4 5 f 6 7 8 9 a b c d e |
| $T_b$ | 0 1 2 3 4 f 5 6 7 8 9 a b c d e |
| $T_c$ | 0 1 2 3 f 4 5 6 7 8 9 a b c d e |
| $T_d$ | 0 1 2 f 3 4 5 6 7 8 9 a b c d e |
| $T_e$ | 0 1 f 2 3 4 5 6 7 8 9 a b c d e |
| $T_f$ | 0 f 1 2 3 4 5 6 7 8 9 a b c d e |
Conclusion on Kuznyechik/Striog

Inverse  Exponential  Math (other)  SPN  Misty  Feistel  Lai-Massey  Pseudo-random  Hill climbing  Unknown
Conclusion on Kuznyechik/Stribog

- **Inverse**
- **Exponential**
- **Math (other)**
- **SPN**
- **Misty**
- **Feistel**
- **Lai-Massey**
- **Pseudo-random**
- **Hill climbing**
- **Unknown**

π

Feistel-like
Conclusion on Kuznyechik/Stribog

- Inverse
- Exponential
- Math (other)
- SPN
- Misty
- Feistel
- Lai-Massey
- Pseudo-random
- Hill climbing
- Unknown

π

Feistel-like

Exponential-like
Conclusion on Kuznyechik/Stribog

![Diagram showing categories of cryptographic techniques: Inverse, Exponential, Math (other), SPN, Misty, Feistel, Lai-Massey, Pseudo-random, Hill climbing, Unknown. The categories are color-coded and arranged in a circular graph with Exponential-like and Feistel-like techniques highlighted.]
Plan

1. Introduction
2. Mathematical Background
3. Detailed Analysis of the Two Tables
4. TU-Decomposition
5. Conclusion
For More Information (1/2)

Theoretical background + S-Box of Skipjack


S-Box of Stribog/Kuznechik (Feistel)


S-Box of Stribog/Kuznechik (Exponential)

*IACR Transactions on Symmetric Cryptology*, 2016(2):99–124
APN Permutation


Online

1. https://eprint.iacr.org/2015/976 (Skipjack)
2. https://eprint.iacr.org/2016/071 (Stribog/Kuznechik 1)
Conclusion

- We can recover *a lot* from an LUT
- white-box crypto is all the hardest,
- we can use cryptanalysis to discover new math results,
- secret services’ algorithms are all the more suspicious!
Conclusion

- We can recover a lot from an LUT
- white-box crypto is all the hardest,
- we can use cryptanalysis to discover new math results,
- secret services’ algorithms are all the more suspicious!

Nothing-up-my-sleeve

Always justify your constants!
Open Positions @ uni.lu

- post-doc in real-world crypto/blockchain/ privacy
- post-doc in lightweight crypto and side-channel attacks (FDISC project)
- PhDs in applied crypto (PRIDE project)

https://www.cryptolux.org/index.php/Home
Open Positions @ uni.lu

- post-doc in real-world crypto/blockchain/ privacy
- post-doc in lightweight crypto and side-channel attacks (FDISC project)
- PhDs in applied crypto (PRIDE project)

https://www.cryptolux.org/index.php/Home

Thank you!
Appendix

Details About Skipjack

![Graph showing the number of occurrences of different absolute values of coefficients in the LAT for various methods: Random permutation, Skipjack's F, Imitating F, Improving R to the max, and Improving (L/2, N_{1/2}).]
Bibliography I


SKIPJACK and KEA Algorithm Specifications.

Exponential S-boxes: a link between the S-boxes of BelT and Kuznyechik/Streebog.