S-Box Reverse-Engineering
Boolean Functions, American/Russian Standards, and Butterflies

Léo Perrin
Based on joint works with Biryukov, Canteaut, Duval and Udovenko

June 6, 2018
CECC’18
Outline

1. Building Blocks for Symmetric Cryptography
2. Statistics and Skipjack
3. TU-Decomposition and Kuznyechik
4. The Butterfly Permutations and Functions
5. Conclusion
Outline

1. Building Blocks for Symmetric Cryptography
2. Statistics and Skipjack
3. TU-Decomposition and Kuznyechik
4. The Butterfly Permutations and Functions
5. Conclusion
Symmetric Cryptography

There are many symmetric algorithms! Hash functions, MACs...
Symmetric Cryptography

There are many **symmetric** algorithms! Hash functions, MACs...

**Definition (Block Cipher)**

- Input: \(n\)-bit block \(x\)
- Parameter: \(k\)-bit key \(\kappa\)
- Output: \(n\)-bit block \(E_\kappa(x)\)
- Symmetry: \(E\) and \(E^{-1}\) use the same \(\kappa\)
Symmetric Cryptography

There are many symmetric algorithms! Hash functions, MACs...

**Definition (Block Cipher)**

- Input: \( n \)-bit block \( x \)
- Parameter: \( k \)-bit key \( \kappa \)
- Output: \( n \)-bit block \( E_\kappa(x) \)
- Symmetry: \( E \) and \( E^{-1} \) use the same \( \kappa \)

**Properties needed:**

- Diffusion
- Confusion
- No cryptanalysis!
No Cryptanalysis?

Let us look at a typical cryptanalysis technique: the **differential attack.**
Differential Attacks

\[ 6 \text{ec}1067e5c5391\text{ae} \rightarrow \bigoplus \rightarrow 6 \text{ec}1067e5c5390\text{ae} \]

\[ a = 0000000000000000100 \]
Differential Attacks

$$6ec1067e5c5391ae \oplus a = 0000000000000000100$$

$$E_K$$

$$6ec1067e5c5390ae$$

$$E_K$$
Differential Attacks

\[ \text{6ec1067e5c5391ae} \quad \oplus \quad \text{6ec1067e5c5390ae} \]

\[ a = 00000000000000100 \]

\[ E_{\kappa} \]

\[ \text{0x7e6f661193739cea} \quad \rightarrow \quad \text{0x04d4595257eb06c8} \]
Differential Attacks

\[ E_k(6ec1067e5c5391ae) \oplus 6ec1067e5c5390ae \]

\[ a = 0000000000000100 \]

\[ E_k(0x7e6f661193739cea) \oplus 0x04d4595257eb06c8 \]

\[ b = 7abb3f43c4989a22 \]
Differential Attacks

\[ x \xrightarrow{E_{Kc}} E_{Kc}(x) \xrightarrow{\oplus} E_{Kc}(x \oplus a) \]

If there are many \( x \) such that \( E_{Kc}(x) = E_{Kc}(x \oplus a) = b \), then the cipher is not secure.
Differential Attacks

If there are many $x$ such that $E_{K}(x) \oplus E_{K}(x \oplus a) = b$, then the cipher is not secure.
Basic Block Cipher Structure

How do we build block ciphers that prevent such attacks (as well as others)?
How do we build block ciphers that prevent such attacks (as well as others)?

Such a block cipher iterates the round function above several times. $S$ is the substitution box (S-Box).
How do we build block ciphers that prevent such attacks (as well as others)?

Substitution-Permutation Network

Such a block cipher iterates the round function above several times. $S$ is the Substitution Box (S-Box).
The S-Box $\pi$ of the latest Russian standards, Kuznyechik (BC) and Streebog (HF).
Importance of the S-Box

If $S$ is such that

$$S(x) \oplus S(x \oplus a) = b$$

does not have many solutions $x$ for all $(a, b)$ then the cipher may be proved secure against differential attacks.
The S-Box (2/2)

**Importance of the S-Box**

If \( S \) is such that

\[
S(x) \oplus S(x \oplus a) = b
\]

does not have many solutions \( x \) for all \((a, b)\) then the cipher may be proved secure against differential attacks.

In academic papers presenting new block ciphers, the choice of \( S \) is carefully explained.
S-Box Design

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

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

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

Grøstl...
iScream...
Khazad...
### S-Box Reverse-Engineering

- **AES S-Box**
- **Inverse (other)**
- **Exponential**
- **Math (other)**
- **SPN**
- **Misty**
- **Feistel**
- **Lai-Massey**
- **Pseudo-random**
- **Hill climbing**
- **Unknown**
S-Box Reverse-Engineering

- AES S-Box
- Inverse (other)
- Exponential
- Math (other)
- SPN
- Misty
- Feistel
- Lai-Massey
- Pseudo-random
- Hill climbing
- Unknown
Motivation (1/3)

A malicious designer can easily hide a structure in an S-Box.
Motivation (1/3)

A malicious designer can easily hide a structure in an S-Box.

To keep an advantage in implementation (WB crypto)...
A malicious designer can easily hide a structure in an S-Box.

To keep an advantage in implementation (WB crypto)...
... or an advantage in cryptanalysis (backdoor).
Definition (Kleptography)

The study of trapdoored cryptography is called kleptography (term introduced by Jung and Young).

S-Box based backdoors in the literature

Motivation (3/3)

Even without malicious intent, an unexpected structure can be a problem.

⇒ We need tools to reverse-engineer S-Boxes!
Outline

1. Building Blocks for Symmetric Cryptography
2. Statistics and Skipjack
3. TU-Decomposition and Kuznyechik
4. The Butterfly Permutations and Functions
5. Conclusion
Summary

We can recover parts of the design process of an S-Box using some statistics.

1. The two tables (basics of Boolean functions for cryptography)
2. A statistical tool based on the two tables
3. Application to NSA’s Skipjack
Let $S : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$ be an S-Box.
The Two Tables

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

**Definition (DDT)**

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

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

\[
\begin{bmatrix}
8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

The **LAT** of \( S \).

\[
\begin{bmatrix}
4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 2 & 0 & 0 & 2 & 0 & 2 \\
0 & 2 & 2 & 0 & 0 & 2 & 0 & 2 \\
0 & 2 & 0 & 2 & 0 & 2 & 0 & 2 \\
0 & 2 & 0 & 2 & 0 & 2 & 0 & 2 \\
0 & 2 & 0 & 2 & 0 & 2 & 0 & 2 \\
0 & 2 & 0 & 2 & 0 & 2 & 0 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]
Coefficient Distribution in the DDT

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

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

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

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

- Always even, $\geq 0$
- Typically between 0 and 16.
- Lower is better.
If an $n$-bit S-Box is bijective, then its LAT coefficients behave like independent and identically distributed random variables following this distribution:

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

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

- Always even, signed.
- Typically between -40 and 40.
- Lower absolute value is better.
### Looking Only at the Maximum

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>$\log_2 (\Pr [\max(\text{DDT}) \leq \delta])$</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>-0.006</td>
</tr>
<tr>
<td>12</td>
<td>-0.094</td>
</tr>
<tr>
<td>10</td>
<td>-1.329</td>
</tr>
<tr>
<td>8</td>
<td>-16.148</td>
</tr>
<tr>
<td>6</td>
<td>-164.466</td>
</tr>
<tr>
<td>4</td>
<td>-1359.530</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\ell$</th>
<th>$\log_2 (\Pr [\max(\text{LAT}) \leq \ell])$</th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
<td>-0.084</td>
</tr>
<tr>
<td>36</td>
<td>-0.302</td>
</tr>
<tr>
<td>34</td>
<td>-1.008</td>
</tr>
<tr>
<td>32</td>
<td>-3.160</td>
</tr>
<tr>
<td>30</td>
<td>-9.288</td>
</tr>
<tr>
<td>28</td>
<td>-25.623</td>
</tr>
<tr>
<td>26</td>
<td>-66.415</td>
</tr>
<tr>
<td>24</td>
<td>-161.900</td>
</tr>
<tr>
<td>22</td>
<td>-371.609</td>
</tr>
</tbody>
</table>

**DDT**

**LAT**

Probability that the maximum coefficient in the DDT/LAT of an 8-bit permutation is at most equal to a certain threshold.
## Looking Only at the Maximum

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>$\log_2 \left( \Pr \left[ \max(\text{DDT}) \leq \delta \right] \right)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>-0.006</td>
</tr>
<tr>
<td>12</td>
<td>-0.094</td>
</tr>
<tr>
<td>10</td>
<td>-1.329</td>
</tr>
<tr>
<td>8</td>
<td>-16.148</td>
</tr>
<tr>
<td>6</td>
<td>-164.466</td>
</tr>
<tr>
<td>4</td>
<td>-1359.530</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\ell$</th>
<th>$\log_2 \left( \Pr \left[ \max(\text{LAT}) \leq \ell \right] \right)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
<td>-0.084</td>
</tr>
<tr>
<td>36</td>
<td>-0.302</td>
</tr>
<tr>
<td>34</td>
<td>-1.008</td>
</tr>
<tr>
<td>32</td>
<td>-3.160</td>
</tr>
<tr>
<td>30</td>
<td>-9.288</td>
</tr>
<tr>
<td>28</td>
<td>-25.623</td>
</tr>
<tr>
<td>26</td>
<td>-66.415</td>
</tr>
<tr>
<td>24</td>
<td>-161.900</td>
</tr>
<tr>
<td>22</td>
<td>-371.609</td>
</tr>
</tbody>
</table>

**DDT**

**LAT**

Probability that the maximum coefficient in the DDT/LAT of an 8-bit permutation is at most equal to a certain threshold.
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.
What is Skipjack? (2/2)

- Skipjack was supposed to be secret...

- ... but eventually published in 1998.

- Skipjack was to be used by the *Clipper Chip*,
What is Skipjack? (2/2)

- Skipjack was supposed to be secret...

- ... but eventually published in 1998.

- Skipjack was to be used by the Clipper Chip,

- It uses an $8 \times 8$ S-Box ($F$) specified only by its LUT.
Reverse-Engineering $F$

For Skipjack’s $F$, $\max(\text{LAT}) = 28$ and $\#28 = 3$. 
Reverse-Engineering F

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

![Graph showing probability distribution](image-url)
Reverse-Engineering $F$

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

![Graph showing probability distribution](image-url)
Reverse-Engineering $F$

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

\[
\Pr [\text{max(LAT)} = 28 \text{ and } \#28 \leq 3] \approx 2^{-55}
\]
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.
## Timeline

<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jun 98</td>
<td>Declassification of Skipjack</td>
</tr>
</tbody>
</table>
Timeline

1987  Initial design of Skipjack

Jul 93  “interim report” on Skipjack published by external cryptographers

Jun 98  Declassification of Skipjack
Timeline

1987  Initial design of Skipjack

Jul 93  “interim report” on Skipjack published by external cryptographers
Aug 95  Alleged “Skipjack” (actually not) is leaked to usenet
Sep 95  Schneier published his thoughts on “alleged Skipjack”, including the result of a FOIA request
Jun 98  Declassification of Skipjack
Timeline

1987  Initial design of Skipjack

Aug 92  The S-Box (“F-table”) of Skipjack is changed

Jul 93  “interim report” on Skipjack published by external cryptographers

Aug 95  Alleged “Skipjack” (actually not) is leaked to usenet

Sep 95  Schneier published his thoughts on “alleged Skipjack”, including the result of a FOIA request

Jun 98  Declassification of Skipjack
Timeline

1987  Initial design of Skipjack
Aug 90  (CRYPTO) Gilbert et al. use linear relations for key recovery (FEAL)
Aug 91  (CRYPTO) Attack against FEAL using linear relations between key, plaintext and ciphertext
May 92  (EUROCRYPT) Other attack against FEAL using linear relations between key, plaintext and ciphertext
Aug 92  The S-Box (“F-table”) of Skipjack is changed
Jul 93  “interim report” on Skipjack published by external cryptographers
Aug 95  Alleged “Skipjack” (actually not) is leaked to usenet
Sep 95  Schneier published his thoughts on “alleged Skipjack”, including the result of a FOIA request
Jun 98  Declassification of Skipjack
Conclusion on Skipjack

- AES S-Box
- Inverse (other)
- Exponential
- Math (other)
- SPN
- Misty
- Feistel
- Lai-Massey
- Pseudo-random
- Hill climbing
- Unknown
Conclusion on Skipjack

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

We can recover an actual decomposition using patterns in the LAT.

1. Our target, the S-Box of Kuznyechik and Streebog
2. TU-decomposition: what is it and how to apply it to Kuznyechik
## Kuznyechik/Stribog

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stribog</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Type</strong></td>
<td>Hash function</td>
<td></td>
</tr>
<tr>
<td><strong>Publication</strong></td>
<td>2012</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Kuznyechik</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Type</strong></td>
<td>Block cipher</td>
<td></td>
</tr>
<tr>
<td><strong>Publication</strong></td>
<td>2015</td>
<td></td>
</tr>
</tbody>
</table>

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, $\pi$. 
## Kuznyechik/Striobog

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Striobog</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Type</strong></td>
<td>Hash function</td>
</tr>
<tr>
<td><strong>Publication</strong></td>
<td>2012</td>
</tr>
</tbody>
</table>

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Kuznyechik</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Type</strong></td>
<td>Block cipher</td>
</tr>
<tr>
<td><strong>Publication</strong></td>
<td>2015</td>
</tr>
</tbody>
</table>

### 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 $\eta$ (reordered columns)
The LAT of $\eta \circ \pi \circ \mu$
The TU-Decomposition

**Definition**

The **TU-decomposition** is a decomposition algorithm working against S-Boxes with vector spaces of zeroes in their LAT.

\[
\begin{align*}
T \text{ and } U & \text{ are mini-block ciphers; } \mu \text{ and } \eta \text{ are linear permutations.}
\end{align*}
\]
Final Decomposition Number 1

- Multiplication in $\mathbb{F}_{2^4}$
- $\alpha$ Linear permutation
- $\mathcal{I}$ Inversion in $\mathbb{F}_{2^4}$
- $\nu_0, \nu_1, \sigma$ $4 \times 4$ permutations
- $\phi$ $4 \times 4$ function
- $\omega$ Linear permutation
Hardware Performance

<table>
<thead>
<tr>
<th>Structure</th>
<th>Area ($\mu m^2$)</th>
<th>Delay (ns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive implementation</td>
<td>3889.6</td>
<td>362.52</td>
</tr>
<tr>
<td>Feistel-like</td>
<td>1534.7</td>
<td>61.53</td>
</tr>
<tr>
<td>Multiplications-first</td>
<td>1530.3</td>
<td>54.01</td>
</tr>
<tr>
<td>Feistel-like (with tweaked MUX)</td>
<td>1530.1</td>
<td>46.11</td>
</tr>
</tbody>
</table>
Conclusion for Kuznyechik/Stribog?

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

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

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

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

... or was it?

Belarussian inspiration

- The last standard of Belarus (BelT) uses an 8-bit S-box,
- somewhat similar to $\pi$...
Conclusion for Kuznyechik/Stribog?

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

... or was it?

Belarussian inspiration

- The last standard of Belarus (BelT) uses an 8-bit S-box,
- somewhat similar to $\pi$...
- ... based on a finite field exponential!
Final Decomposition Number 2 (!)

\[
\begin{array}{c|cccccccccccc}
0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & a & b & c & d & e & f \\
T_0 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & a & b & c & d & e & f \\
T_1 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & a & b & c & d & e & f \\
T_2 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & a & b & c & d & f & e \\
T_3 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & a & b & c & f & d & e \\
T_4 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & a & b & f & c & d & e \\
T_5 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & a & f & b & c & d & e \\
T_6 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & f & a & b & c & d & e \\
T_7 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & f & a & b & c & d & e \\
T_8 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & f & 8 & 9 & a & b & c & d & e \\
T_9 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & f & 7 & 8 & 9 & a & b & c & d & e \\
T_a & 0 & 1 & 2 & 3 & 4 & 5 & f & 6 & 7 & 8 & 9 & a & b & c & d & e & e \\
T_b & 0 & 1 & 2 & 3 & 4 & f & 5 & 6 & 7 & 8 & 9 & a & b & c & d & e \\
T_c & 0 & 1 & 2 & 3 & f & 4 & 5 & 6 & 7 & 8 & 9 & a & b & c & d & e \\
T_d & 0 & 1 & 2 & f & 3 & 4 & 5 & 6 & 7 & 8 & 9 & a & b & c & d & e \\
T_e & 0 & 1 & f & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & a & b & c & d & e \\
T_f & 0 & f & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & a & b & c & d & e \\
\end{array}
\]
Conclusion on Kuznyechik/Striбог

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

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

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

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

1. Building Blocks for Symmetric Cryptography
2. Statistics and Skipjack
3. TU-Decomposition and Kuznyechik
4. The Butterfly Permutations and Functions
5. Conclusion
Summary

We can obtain new mathematical results using reverse-engineering techniques.

1. The big APN problem and its only known solution
2. Decomposing and generalizing this solution as butterflies
NSUCRYPTO (Olympiad in Cryptography)

Siberian Student’s Olympiad in Cryptography with International participation — 2014
Second round NSUCRYPTO November 17-24

Task 2. «An APN Permutation»

“Try to find an APN permutation on 8 variables or prove that it doesn’t exist.”

https://nsucrypto.nsu.ru/
The Big APN Problem

Definition (APN function)

A function $S : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$ is Almost Perfect Non-linear (APN) if

$$S(x \oplus a) \oplus S(x) = b$$

has 0 or 2 solutions for all $a \neq 0$ and for all $b$. 
The Big APN Problem

Definition (APN function)
A function $S : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$ is Almost Perfect Non-linear (APN) if

$$S(x \oplus a) \oplus S(x) = b$$

has 0 or 2 solutions for all $a \neq 0$ and for all $b$.

Big APN Problem
Are there APN permutations operating on $\mathbb{F}_2^n$ where $n$ is even?
Dillon et al.’s Permutation

Only One Known Solution!

For $n = 6$, Dillon et al. found an APN permutation.
Dillon et al.’s Permutation

Only One Known Solution!

For $n = 6$, Dillon et al. found an APN permutation.
Dillon et al.’s Permutation

Only One Known Solution!

For $n = 6$, Dillon et al. found an APN permutation.
Dillon et al.’s Permutation

Only One Known Solution!

For $n = 6$, Dillon et al. found an APN permutation.

It is possible to make a TU-decomposition!
On the Butterfly Structure

Definition (Open Butterfly $H^3_{\alpha,\beta}$)

This permutation is an open butterfly.
On the Butterfly Structure

Definition (Open Butterfly $H^3_{\alpha,\beta}$)
This permutation is an open butterfly.

Lemma
Dillon's permutation is affine-equivalent to $H^3_{w,1}$, where $Tr(w) = 0$. 

Closed Butterflies

**Definition (Closed butterfly $V^3_{\alpha,\beta}$)**

This quadratic function is a **closed butterfly**.
Closed Butterflies

Definition (Closed butterfly $V^3_{\alpha, \beta}$)
This quadratic function is a closed butterfly.

Lemma (Equivalence)
Open and closed butterflies with the same parameters are CCZ-equivalent.
Some Properties of Butterflies

**Theorem (Properties of butterflies)**

Let \( V^3_{\alpha, \beta} \) and \( H^3_{\alpha, \beta} \) be butterflies operating on \( 2n \) bits, \( n \) odd. Then:

- \( \deg \left( V^3_{\alpha, \beta} \right) = 2 \),
- if \( n = 3 \), \( \text{Tr} (\alpha) = 0 \) and \( \beta + \alpha^3 \in \{ \alpha, 1/\alpha \} \), then
  \[ \max(DDT) = 2, \quad \max(\mathcal{W}) = 2^{n+1} \quad \text{and} \quad \deg \left( H^3_{\alpha, \beta} \right) = n + 1 \]
- if \( \beta = (1 + \alpha)^3 \), then
  \[ \max(DDT) = 2^{n+1}, \quad \max(\mathcal{W}) = 2^{(3n+1)/2} \quad \text{and} \quad \deg \left( H^3_{\alpha, \beta} \right) = n \]
- otherwise,
  \[ \max(DDT) = 4, \quad \max(\mathcal{W}) = 2^{n+1} \quad \text{and} \quad \deg \left( H^3_{\alpha, \beta} \right) \in \{ n, n + 1 \} \]
  and \( \deg \left( H^3_{\alpha, \beta} \right) = n \) if and only if
  \[ 1 + \alpha \beta + \alpha^4 = (\beta + \alpha + \alpha^3)^2. \]
Outline

1. Building Blocks for Symmetric Cryptography
2. Statistics and Skipjack
3. TU-Decomposition and Kuznyechik
4. The Butterfly Permutations and Functions
5. Conclusion
Open Problem

**Cellular Message Encryption Algorithm**

From Wikipedia, the free encyclopedia

In **cryptography**, the **Cellular Message Encryption Algorithm** (CMEA) is a **block cipher** which was used for securing **mobile phones** in the **United States**. CMEA is one of four cryptographic primitives specified in a **Telecommunications Industry Association** (TIA) standard, and is designed to **encrypt** the control channel, rather than the voice data. In 1997, a group of cryptographers published attacks on the **cipher** showing it had several weaknesses which give it a trivial effective strength of a 24-bit to 32-bit cipher.[1]

<table>
<thead>
<tr>
<th><strong>CMEA</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General</strong></td>
</tr>
<tr>
<td><strong>Designers</strong></td>
</tr>
<tr>
<td><strong>First published</strong></td>
</tr>
<tr>
<td><strong>Cipher detail</strong></td>
</tr>
<tr>
<td><strong>Key sizes</strong></td>
</tr>
<tr>
<td><strong>Block sizes</strong></td>
</tr>
<tr>
<td><strong>Rounds</strong></td>
</tr>
</tbody>
</table>
### Cellular Message Encryption Algorithm

From Wikipedia, the free encyclopedia

In cryptography, the **Cellular Message Encryption Algorithm** (CMEA) is a block cipher which was used for securing mobile phones in the United States. CMEA is one of four cryptographic primitives specified in a Telecommunications Industry Association (TIA) standard, and is designed to encrypt the control channel, rather than the voice data. In 1997, a group of cryptographers published attacks on the cipher showing it had several weaknesses which give it a trivial effective strength of a 24-bit to 32-bit cipher.[1]

<table>
<thead>
<tr>
<th>CMEA</th>
<th>General</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Designers</strong></td>
<td>James A. Reeds III</td>
</tr>
<tr>
<td><strong>First published</strong></td>
<td>1991</td>
</tr>
<tr>
<td><strong>Cipher detail</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Key sizes</strong></td>
<td>64 bits</td>
</tr>
<tr>
<td><strong>Block sizes</strong></td>
<td>16-64 bits</td>
</tr>
<tr>
<td><strong>Rounds</strong></td>
<td>3</td>
</tr>
</tbody>
</table>

**A hidden structure!**

CMEA uses an 8-bit (non-bijective) S-Box... With a TU-decomposition!

**What is its actual structure?**
Conclusion

1 Cryptographers use mathematics but mathematicians could also use crypto!
Conclusion

1. Cryptographers use mathematics but mathematicians could also use crypto!

2. If you design a cipher, justify every step of your design.
Conclusion

1. Cryptographers use mathematics but mathematicians could also use crypto!

2. If you design a cipher, justify every step of your design.

3. If you choose a cipher, demand a full design explanation.
## The Last S-Box

<table>
<thead>
<tr>
<th>14</th>
<th>11</th>
<th>60</th>
<th>6d</th>
<th>e9</th>
<th>10</th>
<th>e3</th>
<th>2</th>
<th>b</th>
<th>90</th>
<th>d</th>
<th>17</th>
<th>c5</th>
<th>b0</th>
<th>9f</th>
<th>c5</th>
</tr>
</thead>
<tbody>
<tr>
<td>d8</td>
<td>da</td>
<td>be</td>
<td>22</td>
<td>8</td>
<td>f3</td>
<td>4</td>
<td>a9</td>
<td>fe</td>
<td>f3</td>
<td>f5</td>
<td>fc</td>
<td>bc</td>
<td>30</td>
<td>be</td>
<td>26</td>
</tr>
<tr>
<td>bb</td>
<td>88</td>
<td>85</td>
<td>46</td>
<td>f4</td>
<td>2e</td>
<td>e</td>
<td>fd</td>
<td>76</td>
<td>fe</td>
<td>b0</td>
<td>11</td>
<td>4e</td>
<td>de</td>
<td>35</td>
<td>bb</td>
</tr>
<tr>
<td>30</td>
<td>4b</td>
<td>30</td>
<td>d6</td>
<td>dd</td>
<td>df</td>
<td>df</td>
<td>d4</td>
<td>90</td>
<td>7a</td>
<td>d8</td>
<td>8c</td>
<td>6a</td>
<td>89</td>
<td>30</td>
<td>39</td>
</tr>
<tr>
<td>e9</td>
<td>1</td>
<td>da</td>
<td>d2</td>
<td>85</td>
<td>87</td>
<td>d3</td>
<td>d4</td>
<td>ba</td>
<td>2b</td>
<td>d4</td>
<td>9f</td>
<td>9c</td>
<td>38</td>
<td>8c</td>
<td>55</td>
</tr>
<tr>
<td>d3</td>
<td>86</td>
<td>bb</td>
<td>db</td>
<td>ec</td>
<td>e0</td>
<td>46</td>
<td>48</td>
<td>bf</td>
<td>46</td>
<td>1b</td>
<td>1c</td>
<td>d7</td>
<td>d9</td>
<td>1b</td>
<td>e0</td>
</tr>
<tr>
<td>23</td>
<td>d4</td>
<td>d7</td>
<td>7f</td>
<td>16</td>
<td>3f</td>
<td>3</td>
<td>3</td>
<td>44</td>
<td>c3</td>
<td>59</td>
<td>10</td>
<td>2a</td>
<td>da</td>
<td>ed</td>
<td>e9</td>
</tr>
<tr>
<td>8e</td>
<td>d8</td>
<td>d1</td>
<td>db</td>
<td>cb</td>
<td>cb</td>
<td>c3</td>
<td>c7</td>
<td>38</td>
<td>22</td>
<td>34</td>
<td>3d</td>
<td>db</td>
<td>85</td>
<td>23</td>
<td>7c</td>
</tr>
<tr>
<td>24</td>
<td>d1</td>
<td>d8</td>
<td>2e</td>
<td>fc</td>
<td>44</td>
<td>8</td>
<td>38</td>
<td>c8</td>
<td>c7</td>
<td>39</td>
<td>4c</td>
<td>5f</td>
<td>56</td>
<td>2a</td>
<td>cf</td>
</tr>
<tr>
<td>d0</td>
<td>e9</td>
<td>d2</td>
<td>68</td>
<td>e4</td>
<td>e3</td>
<td>e9</td>
<td>13</td>
<td>e2</td>
<td>c</td>
<td>97</td>
<td>e4</td>
<td>60</td>
<td>29</td>
<td>d7</td>
<td>9b</td>
</tr>
<tr>
<td>d9</td>
<td>16</td>
<td>24</td>
<td>94</td>
<td>b3</td>
<td>e3</td>
<td>4c</td>
<td>4c</td>
<td>4f</td>
<td>39</td>
<td>e0</td>
<td>4b</td>
<td>bc</td>
<td>2c</td>
<td>d3</td>
<td>94</td>
</tr>
<tr>
<td>81</td>
<td>96</td>
<td>93</td>
<td>84</td>
<td>91</td>
<td>d0</td>
<td>2e</td>
<td>d6</td>
<td>d2</td>
<td>2b</td>
<td>78</td>
<td>ef</td>
<td>d6</td>
<td>9e</td>
<td>7b</td>
<td>72</td>
</tr>
<tr>
<td>ad</td>
<td>c4</td>
<td>68</td>
<td>92</td>
<td>7a</td>
<td>d2</td>
<td>5</td>
<td>2b</td>
<td>1e</td>
<td>d0</td>
<td>dc</td>
<td>b1</td>
<td>22</td>
<td>3f</td>
<td>c3</td>
<td>c3</td>
</tr>
<tr>
<td>88</td>
<td>b1</td>
<td>8d</td>
<td>b5</td>
<td>e3</td>
<td>4e</td>
<td>d7</td>
<td>81</td>
<td>3</td>
<td>15</td>
<td>17</td>
<td>25</td>
<td>4e</td>
<td>65</td>
<td>88</td>
<td>4e</td>
</tr>
<tr>
<td>e4</td>
<td>3b</td>
<td>81</td>
<td>81</td>
<td>fa</td>
<td>1</td>
<td>1d</td>
<td>4</td>
<td>22</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td>27</td>
<td>68</td>
<td>27</td>
<td>2e</td>
</tr>
<tr>
<td>3b</td>
<td>83</td>
<td>c7</td>
<td>cc</td>
<td>25</td>
<td>9b</td>
<td>d8</td>
<td>d5</td>
<td>1c</td>
<td>1f</td>
<td>e5</td>
<td>59</td>
<td>7f</td>
<td>3f</td>
<td>3f</td>
<td>ef</td>
</tr>
</tbody>
</table>