On Reverse-Engineering S-Boxes

Alex Biryukov[1], Léo Perrin[1], Aleksei Udovenko[1]

[1]SnT, University of Luxembourg

https://www.cryptolux.org

March 28, 2017

CryptoAction Symposium 2017

S-Box?

An S-Box is a small non-linear function mapping m bits to n usually specified via its look-up table.

S-Box?

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.

Example

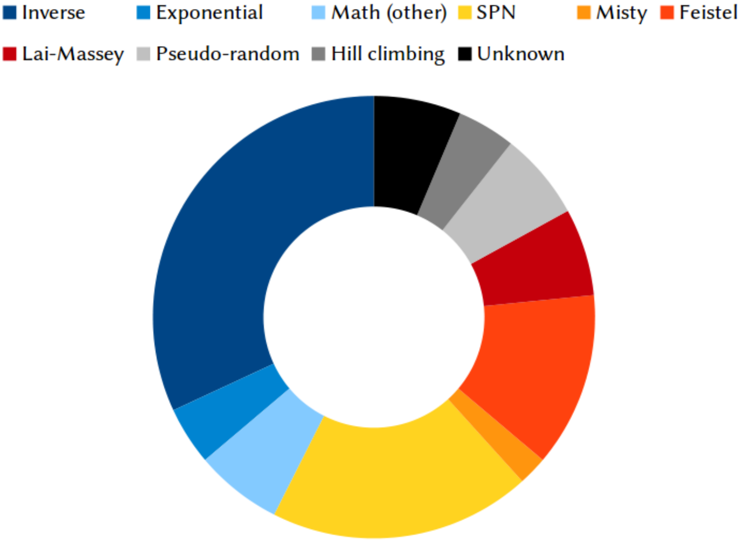
π' = (252, 238, 221, 17, 207, 110, 49, 22, 251, 196, 250, 218, 35, 197, 4, 77, 233,
119, 240, 219, 147, 46, 153, 186, 23, 54, 241. 187, 20, 205, 95, 193, 249, 24, 101,
90, 226, 92, 239, 33, 129, 28, 60, 66, 139, 1, 142, 79, 5, 132, 2, 174, 227, 106, 143,
160, 6, 11, 237, 152, 127, 212, 211, 31, 235, 52, 44, 81, 234, 200, 72, 171, 242, 42,
104, 162, 253, 58, 206, 204, 181, 112, 14, 86, 8, 12, 118, 18, 191, 114, 19, 71, 156,
183, 93, 135, 21, 161, 150, 41, 16, 123, 154, 199, 243, 145, 120, 111, 157, 158, 178,
177, 50, 117, 25, 61, 255, 53, 138, 126, 109, 84, 198, 128, 195, 189, 13, 87, 223,
245, 36, 169, 62, 168, 67, 201, 215, 121, 214, 246, 124, 34, 185, 3, 224, 15, 236,
222, 122, 148, 176, 188, 220, 232, 40, 80, 78, 51, 10, 74, 167, 151, 96, 115, 30, 0,
98, 68, 26, 184, 56, 130, 100, 159, 38, 65, 173, 69, 70, 146, 39, 94, 85, 47, 140, 163,
165, 125, 105, 213, 149, 59, 7, 88, 179, 64, 134, 172, 29, 247, 48, 55, 107, 228, 136,
217, 231, 137, 225, 27, 131, 73, 76, 63, 248, 254, 141, 83, 170, 144, 202, 216, 133,
97, 32, 113, 103, 164, 45, 43, 9, 91, 203, 155, 37, 208, 190, 229, 108, 82, 89, 166,
116, 210, 230, 244, 180, 192, 209, 102, 175, 194, 57, 75, 99, 182).

Screen capture from [GOST, 2015].

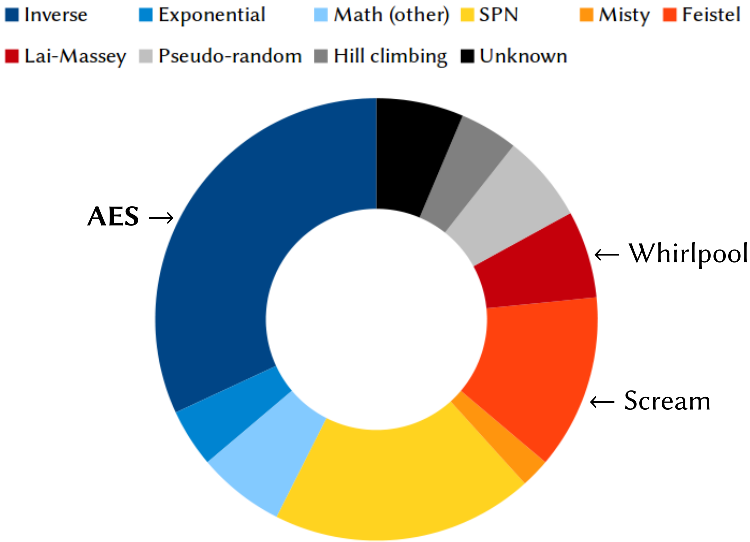
S-Box Design

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

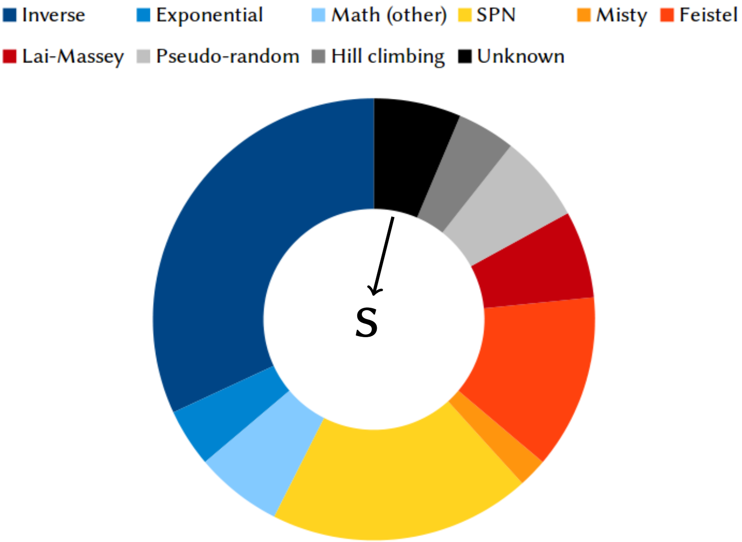
S-Box Design


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

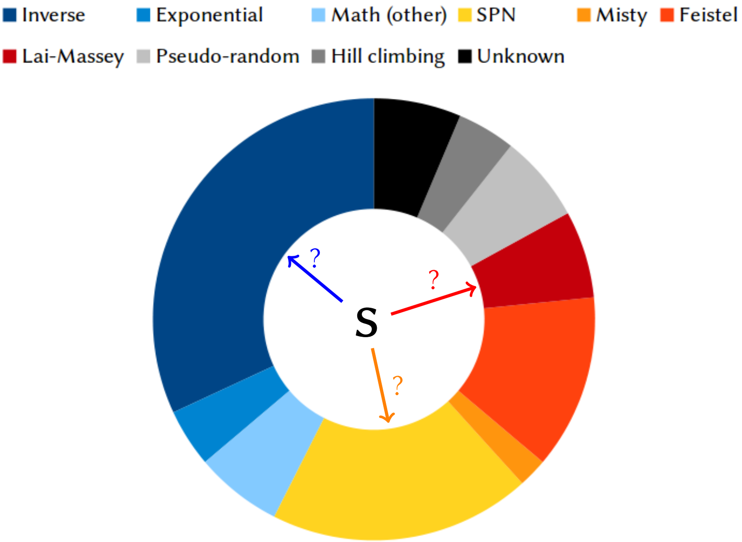
S-Box Design


Inverse
Exponential
Math (other)
SPN
Misty
Feistel
Lai-Massey
Pseudo-random
Hill climbing
Unknown
AES →
← Whirlpool
← Scream

S-Box Reverse-Engineering



S-Box Reverse-Engineering



Outline

Plan

The Two Tables

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

The Two Tables

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

Definition (DDT)

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

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

The Two Tables

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

Definition (DDT)

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

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

Definition (LAT)

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

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

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.

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

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:

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

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:

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

Plan

Looking Only at the Maximum

δ	$\log_2 (\Pr [\max(\mathcal{D}) \leq \delta])$
4	-1359.530
6	-164.466
8	-16.148
10	-1.329
12	-0.094
14	-0.006

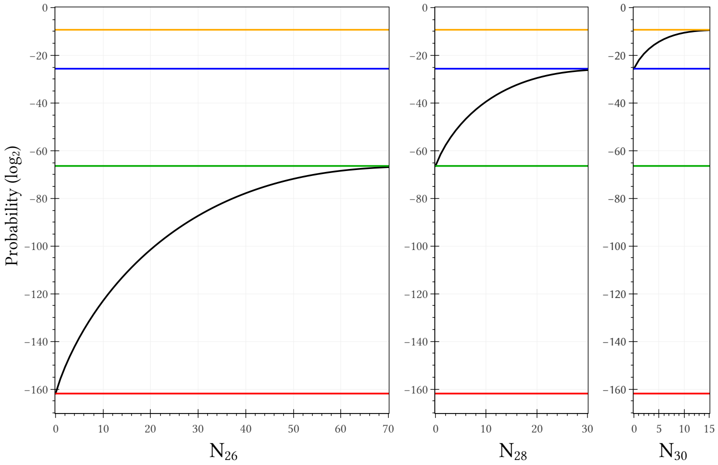
DDT

ℓ	$\log_2 (\Pr [\max(\mathcal{L}) \leq \ell])$
22	-371.609
24	-161.900
26	-66.415
28	-25.623
30	-9.288
32	-3.160
34	-1.008
36	-0.302
38	-0.084

LAT

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

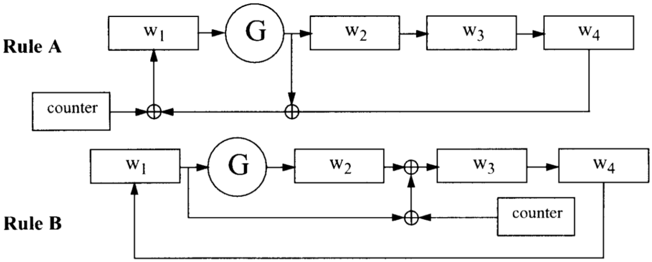
Taking Number of Maximum Values into Account



$\Pr[\max(\mathrm{LAT}) = 24]$, $\Pr[\max(\mathrm{LAT}) = 26]$, $\Pr[\max(\mathrm{LAT}) = 28]$, $\Pr[\max(\mathrm{LAT}) = 30]$

What is Skipjack? (1/2)

Type	Block cipher
Bloc	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)

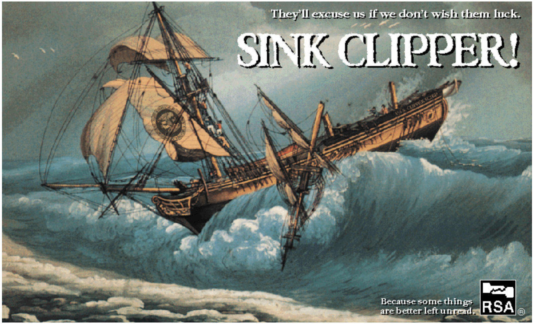
- Skipjack was supposed to be secret...
- ... but eventually published in 1998 [National Security Agency, 1998],
- It uses a 8×8 S-Box (F) specified only by its LUT,

What is Skipjack? (2/2)

- Skipjack was supposed to be secret...
- ... but eventually published in 1998 [National Security Agency, 1998],
- It uses a 8×8 S-Box (F) specified only by its LUT,
- 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 [National Security Agency, 1998],
- It uses a 8×8 S-Box (F) specified only by its LUT,
- Skipjack was to be used by the *Clipper Chip*.

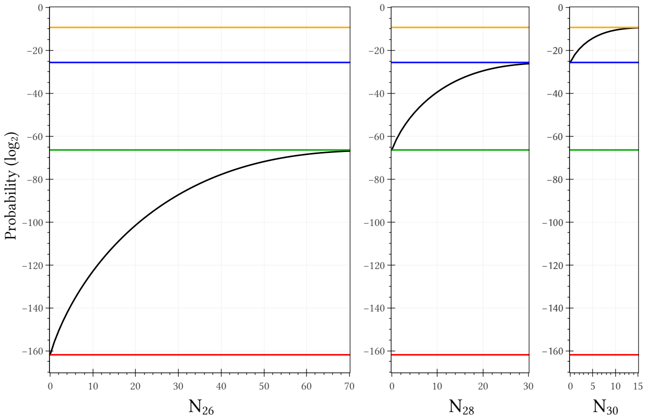


Reverse-Engineering F

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

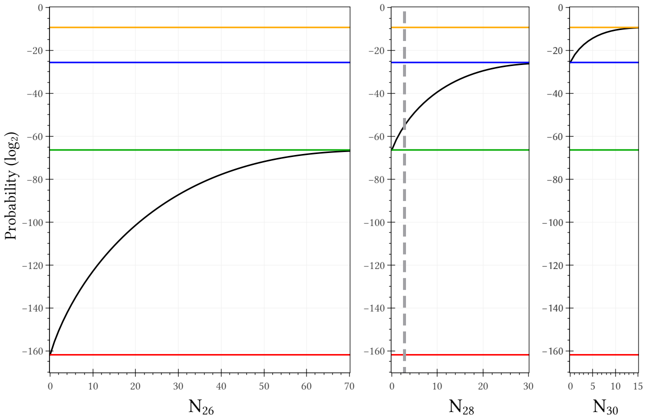
Reverse-Engineering F

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



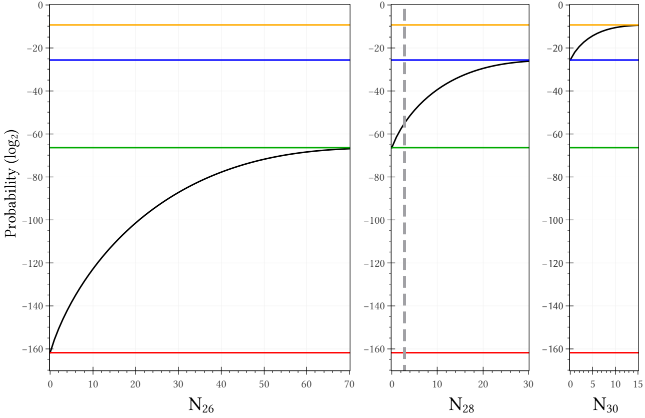
Reverse-Engineering F

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



Reverse-Engineering F

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



$$\Pr\left[\max(\text{LAT}) = 28 \text{ and } \#28 = 3\right] \approx 2^{-55}$$

What Can We Deduce?

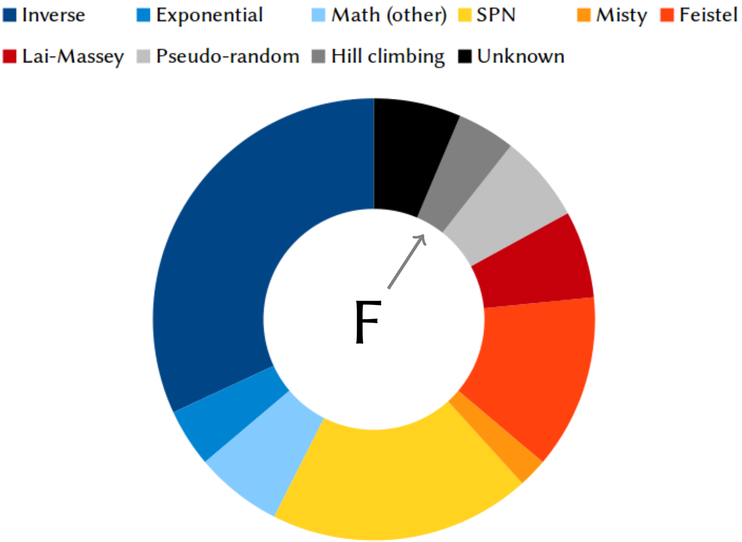
- F has not been picked uniformly at random.
- F has not been picked among a feasibly large set of random S-Boxes.
- Its linear properties were optimized (though poorly).

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



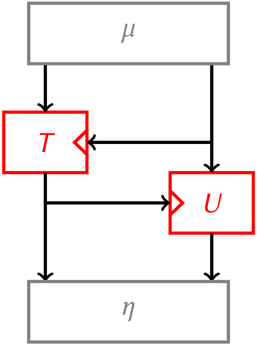
Plan

TU-Decomposition in a Nutshell

1. Identify linear patterns in zeroes of LAT;

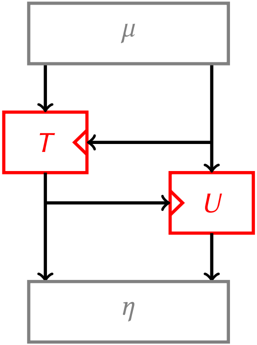
TU-Decomposition in a Nutshell

1. Identify linear patterns in zeroes of LAT;
2. Deduce linear layers μ, η such that π is decomposed as in right picture;



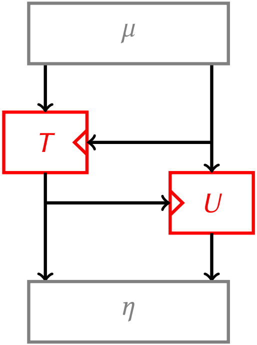
TU-Decomposition in a Nutshell

1. Identify linear patterns in zeroes of LAT;
2. Deduce linear layers μ, η such that π 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 μ, η such that π is decomposed as in right picture;
3. Decompose U, T;
4. Put it all together.



Kuznyechik/Stribog

Stribog

Type Hash function

Publication [GOST, 2012]

Kuznyechik

Type Block cipher

Publication [GOST, 2015]

Kuznyechik/Stribog

Stribog

Type	Hash function
Publication	[GOST, 2012]

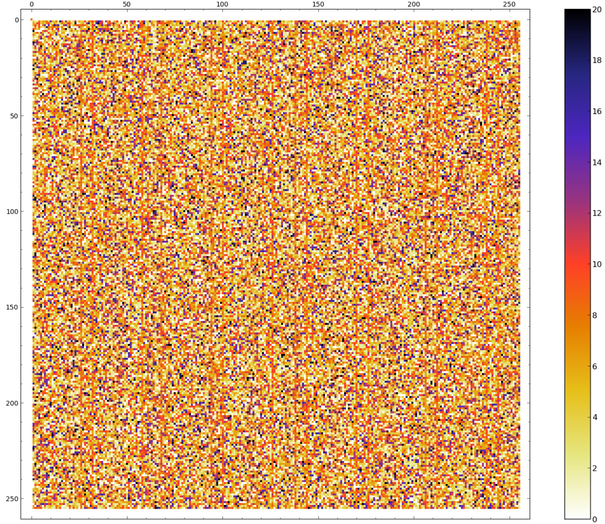
Kuznyechik

Type	Block cipher
Publication	[GOST, 2015]

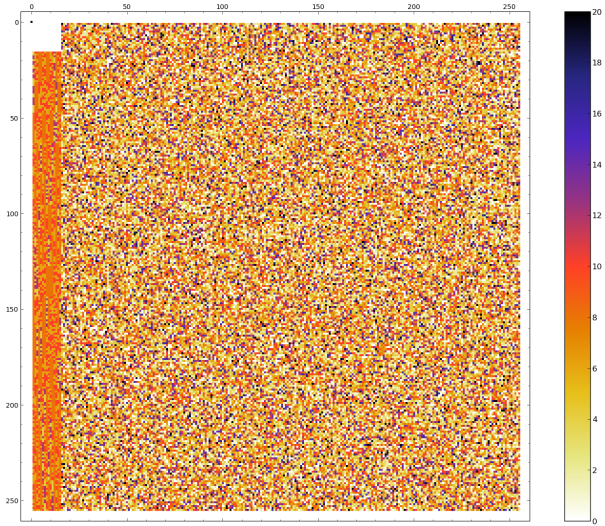
Common ground

- Both are standard symmetric primitives in Russia.
- Both were designed by the FSB (TC26).
- Both use the same 8×8 S-Box, π.

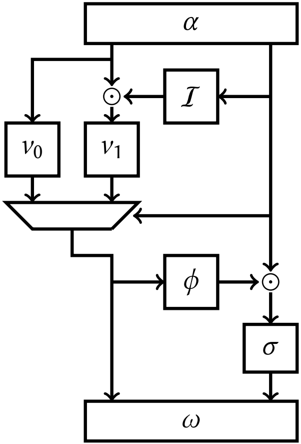
The LAT of π



The LAT of $\eta \circ \pi \circ \mu$

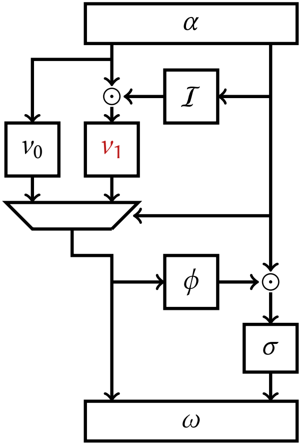


Final Decomposition Number 1



- $\odot$ Multiplication in $\mathbb{F}_{2^4}$
- α Linear permutation
- $\mathcal{I}$ Inversion in $\mathbb{F}_{2^4}$
- ν_0, ν_1, σ 4×4 permutations
- ϕ 4×4 function
- ω Linear permutation

Final Decomposition Number 1



- $\odot$ Multiplication in $\mathbb{F}_{2^4}$
- α Linear permutation
- $\mathcal{I}$ Inversion in $\mathbb{F}_{2^4}$
- ν_0, ν_1, σ 4×4 permutations
- ϕ 4×4 function
- ω Linear permutation

$$P[\nu_1(x \oplus \texttt{0x9}) \oplus \nu_1(x) = \texttt{0x2}] = \mathbf{1}$$

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/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 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 π...
- ... based on a finite field exponential!

Conclusion for Kuznyechik/Stribog?

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

... or was it?

Belarussian inspiration

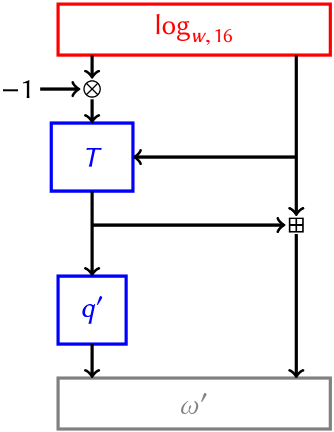
- The last standard of Belarus [STB 34.101.31-2011, 2011] uses an 8-bit S-box,
- somewhat similar to π...
- ... based on a finite field exponential!

Exponential in π

$$\pi \circ \exp$$

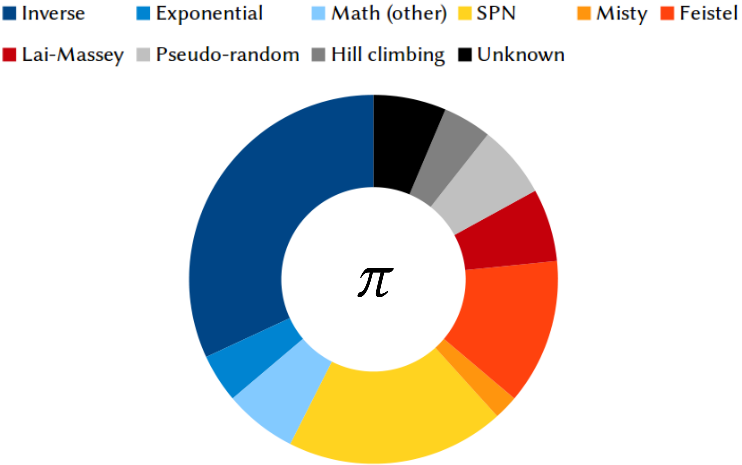
has $\max(DDT) = 128$ ($\Pr < 2^{-340}$) and a TU-decomposition!

Final Decomposition Number 2 (!)

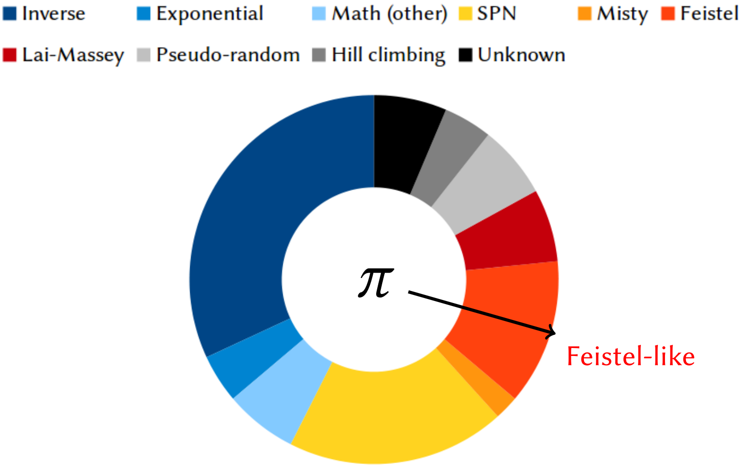


	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
T_0	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
T_1	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
T_2	0	1	2	3	4	5	6	7	8	9	a	b	c	d	f	e
T_3	0	1	2	3	4	5	6	7	8	9	a	b	c	f	d	e
T_4	0	1	2	3	4	5	6	7	8	9	a	b	f	c	d	e
T_5	0	1	2	3	4	5	6	7	8	9	a	f	b	c	d	e
T_6	0	1	2	3	4	5	6	7	8	9	f	a	b	c	d	e
T_7	0	1	2	3	4	5	6	7	8	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

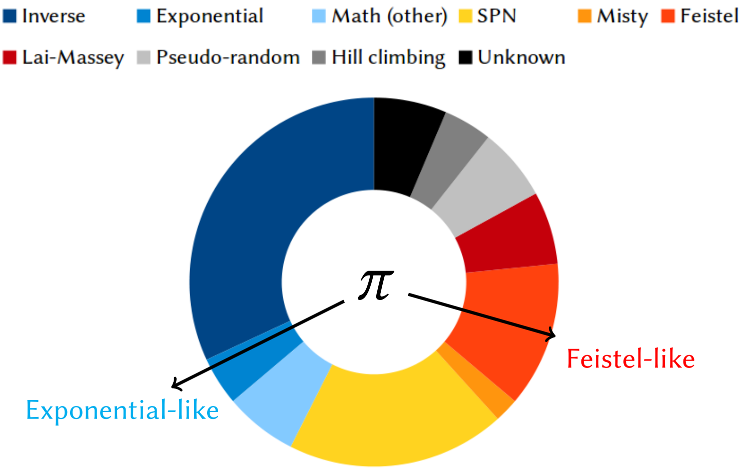
Conclusion on Kuznyechik/Stribog



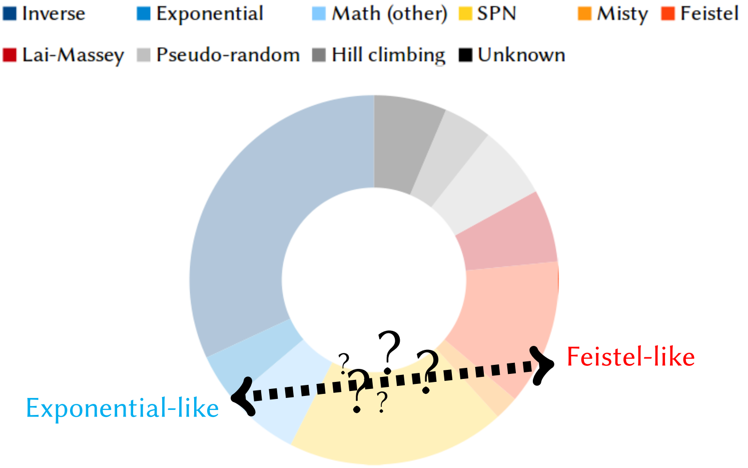
Conclusion on Kuznyechik/Stribog



Conclusion on Kuznyechik/Stribog



Conclusion on Kuznyechik/Stribog



Plan

For More Information (1/2)

Theoretical background + S-Box of Skipjack

Biryukov, A. and Perrin, L. (2015). On Reverse-Engineering S-Boxes with Hidden Design Criteria or Structure.
In *Advances in Cryptology – CRYPTO 2015*, pages 116–140

S-Box of Stribog/Kuznechik (Feistel)

Biryukov, A., Perrin, L., and Udovenko, A. (2016). Reverse-Engineering the S-Box of Streebog, Kuznyechik and STRIBOBr1.
In *Advances in Cryptology – EUROCRYPT 2016*, pages 372–402

S-Box of Stribog/Kuznechik (Exponential)

Perrin, L. and Udovenko, A. (2017). Exponential S-boxes: a link between the S-boxes of BelT and Kuznyechik/Streebog.
IACR Transactions on Symmetric Cryptology, 2016(2):99–124

For More Information (2/2)

APN Permutation

Perrin, L., Udovenko, A., and Biryukov, A. (2016). Cryptanalysis of a Theorem: Decomposing the Only Known Solution to the Big APN Problem.
In *Advances in Cryptology – CRYPTO 2016*, pages (93–122)

Online

1. https://eprint.iacr.org/2015/976 (Skipjack)
2. https://eprint.iacr.org/2016/071 (Stribog/Kuznechik 1)
3. https://eprint.iacr.org/2016/539 (6-bit APN)
4. http://tosc.iacr.org/index.php/ToSC/article/view/567/509 (Stribog/Kuznechik 2)

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`

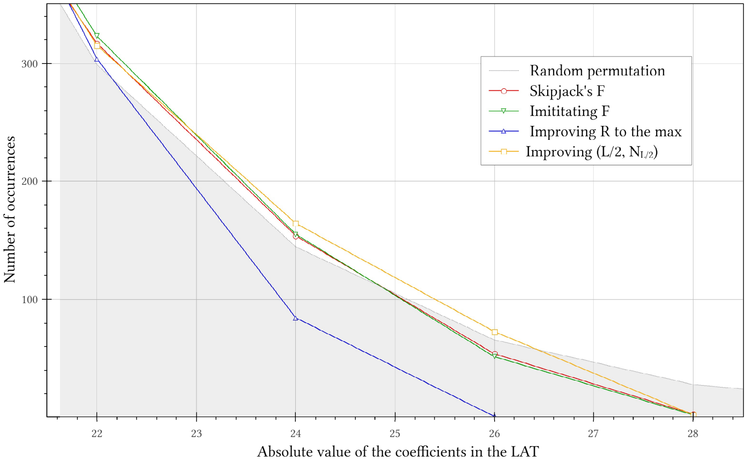
Open Positions @ uni.lu

- 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!

Details About Skipjack



Bibliography I

Biryukov, A. and Perrin, L. (2015).
On Reverse-Engineering S-Boxes with Hidden Design Criteria or Structure.
In *Advances in Cryptology – CRYPTO 2015*, pages 116–140.

Biryukov, A., Perrin, L., and Udovenko, A. (2016).
Reverse-Engineering the S-Box of Streebog, Kuznyechik and STRIBOBr1.
In *Advances in Cryptology – EUROCRYPT 2016*, pages 372–402.

GOST (2012).
Gost r 34.11-2012: Streebog hash function.
https://www.streebog.net/.

Bibliography II

GOST (2015).
(GOST R 34.12–2015) information technology – cryptographic data security – block ciphers.
http://tc26.ru/en/standard/gost/GOST_R_34_12_2015_ENG.pdf.

National Security Agency, N. S. A. (1998).
SKIPJACK and KEA Algorithm Specifications.

Perrin, L. and Udovenko, A. (2017).
Exponential S-boxes: a link between the S-boxes of BelT and Kuznyechik/Streebog.
IACR Transactions on Symmetric Cryptology, 2016(2):99–124.

Bibliography III

Perrin, L., Udovenko, A., and Biryukov, A. (2016).
Cryptanalysis of a Theorem: Decomposing the Only Known Solution to the Big APN Problem.
In *Advances in Cryptology – CRYPTO 2016*, pages (93–122).

STB 34.101.31-2011 (2011).
"Information technologies. Data protection. Cryptographic algorithms for encryption and integrity control.".
State Standard of Republic of Belarus (STB 34.101.31-2011).
http://apmi.bsu.by/assets/files/std/belt-spec27.pdf.