MILP Modeling for (Large) S-boxes to Optimize Probability of Differential Characteristics

Ahmed Abdelkhalek[1], Yu Sasaki[2], **Yosuke Todo**[2], Mohamed Tolba[1], and Amr M. Youssef[1]

1:Concordia University, 2: NTT

Talk @ FSE2018, 5 March 2018

Copyright©2018 NTT corp. All Rights Reserved.

MILP for Differential Cryptanalysis

- Mouha et al. at Inscrypt 2011:

Problem of finding optimal differential trail → (convert) → Optimization problem in MILP

- Advantage:
  - The task of cryptographers is light.
  - We have several off-the-shelf solvers, where the speed of solving is dramatically improved recently.

Copyright©2018 NTT corp. All Rights Reserved.

Summary of Our Results

Previous consensus

1. **4-bit S-box is possible, but 8-bit is difficult.**
2. **# of active S-box is possible, but the probability is difficult.**

New Observations

1. **The algorithm to find minimized CNF of Boolean function enables us to evaluate 8-bit S-boxes.**
2. **Indicator constraints enables us to evaluate probability.**

Applications

- SKINNY-128: the max diff prob reaches 2^{-128} with 14 rounds (prev. 15 rounds)
- AES-round based Func from FSE2016: improved the max probability of diff trail

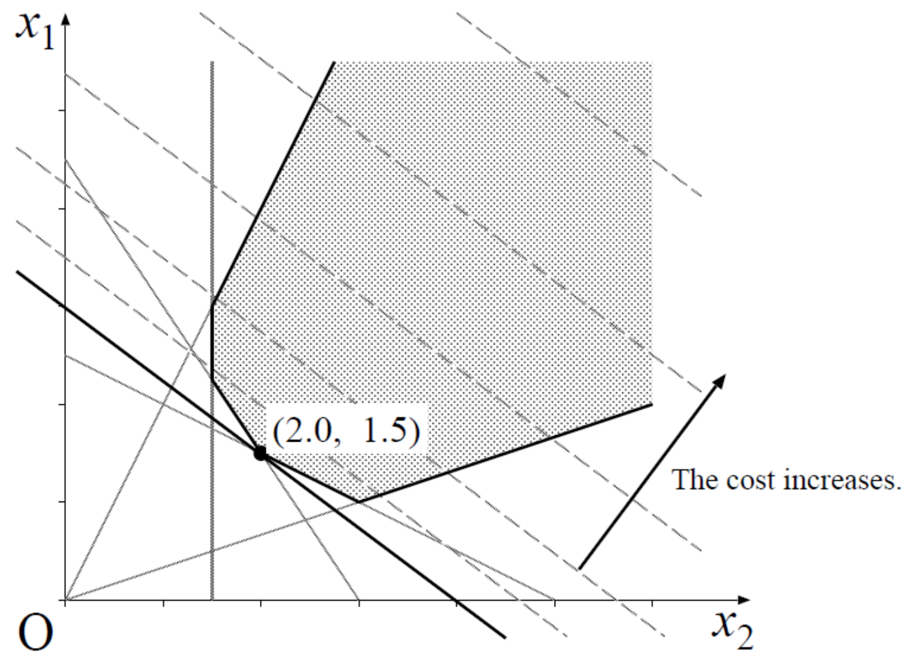
NTT
Copyright©2018 NTT corp. All Rights Reserved.

Mixed Integer Linear Programming (MILP)

- Optimize objective function within the solution range satisfying all the constraints.

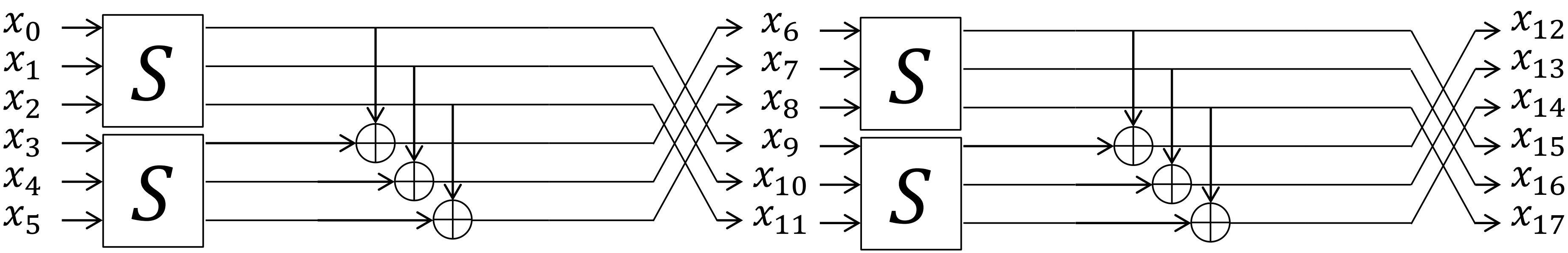
Minimize $50x_1 + 65x_2$

Constraints
$$\begin{cases} 3x_1 + 2x_2 \geq 9 \\ \frac{1}{15}x_1 + \frac{2}{15}x_2 \geq \frac{1}{3} \\ \frac{1}{6}x_1 \geq \frac{1}{4} \\ x_1 - 3x_2 \leq 0 \\ 2x_1 - x_2 \geq 0 \\ x_1 \geq 0 \\ x_2 \geq 0 \end{cases}$$



Copyright©2018 NTT corp. All Rights Reserved.

MILP Model for 3-Round Toy Cipher



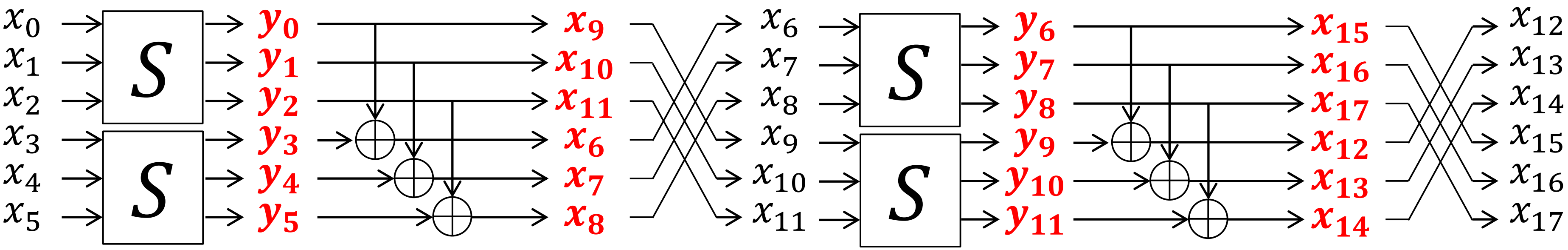
- 6-bit round function: 3-bit S-box, 3-bit xor, swap

To make the MILP model,
define a binary variable $x_i \in \{0,1\}$ for each round;

- $x_i = 0$ denotes the bit i has no difference
- $x_i = 1$ denotes the bit i has difference

Copyright©2018 NTT corp. All Rights Reserved.

Constraints for Linear Operations



- $a \oplus b = c$ can be modeled with 4 inequalities by removing each impossible (a, b, c).

$$(a, b, c) \neq (0,0,1) \implies a + b + (1 - c) \geq 1$$
$$(a, b, c) \neq (0,1,0) \implies a + (1 - b) + c \geq 1$$
$$(a, b, c) \neq (1,0,0) \implies (1 - a) + b + c \geq 1$$
$$(a, b, c) \neq (1,1,1) \implies (1 - a) + (1 - b) + (1 - c) \geq 1$$

Copyright©2018 NTT corp. All Rights Reserved.

Two Methods of Modeling S-box

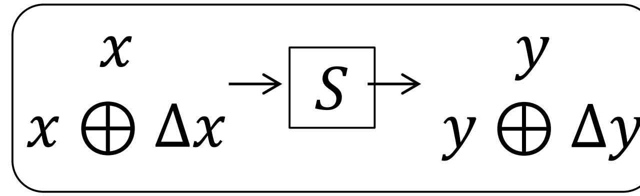
	H-representation of convex hull		**Logical condition model (Sun et al.)**	
tool	SAGE Math		N/A	
support alg	greedy	Sub MILP	greedy	Sub MILP
type	heuristic	optimal	heuristic	optimal
coefficients	any integer		{-1, 0, 1}	
#inequ.	small		large	
8-bit S-box	infeasible		?	

Our Focus

Copyright©2018 NTT corp. All Rights Reserved.

Differential Distribution Table (DDT)

We compute the probability that Δx propagates to Δy for each $(\Delta x, \Delta y)$.



Input Difference (Δx)	Output Difference (Δy)							
	0x0	0x1	0x2	0x3	0x4	0x5	0x6	0x7
0x0	1	0	0	0	0	0	0	0
0x1	0	2^{-2}	2^{-2}	0	0	2^{-2}	2^{-2}	0
0x2	0	2^{-2}	2^{-2}	0	0	2^{-2}	2^{-2}	0
0x3	0	0	0	2^{-1}	0	0	0	2^{-1}
0x4	0	0	0	0	2^{-1}	0	0	2^{-1}
0x5	0	2^{-2}	2^{-2}	0	0	2^{-2}	2^{-2}	0
0x6	0	2^{-2}	2^{-2}	0	0	2^{-2}	2^{-2}	0
0x7	0	0	0	2^{-1}	2^{-1}	0	0	0

Copyright©2018 NTT corp. All Rights Reserved.

Truncated DDT ($*$-DDT)

- To count the # of active S-boxes, we only care whether each pattern is possible (non-zero probability) or impossible (zero probability). We call it "$*$-DDT".

Input Difference (Δx)	Output Difference (Δy)							
	0x0	0x1	0x2	0x3	0x4	0x5	0x6	0x7
0x0	**1**	0	0	0	0	0	0	0
0x1	0	**1**	**1**	0	0	**1**	**1**	0
0x2	0	**1**	**1**	0	0	**1**	**1**	0
0x3	0	0	0	**1**	0	0	0	**1**
0x4	0	0	0	0	**1**	0	0	**1**
0x5	0	**1**	**1**	0	0	**1**	**1**	0
0x6	0	**1**	**1**	0	0	**1**	**1**	0
0x7	0	0	0	**1**	**1**	0	0	0

NTT

Copyright©2018 NTT corp. All Rights Reserved.

Logical condition model

- We remove impossible propagations.

$$x_2 + x_1 + x_0 + y_2 + y_1 + (1 - y_0) \geq 1$$

Input Difference (Δx)	Output Difference (Δy)							
	0x0	0x1	0x2	0x3	0x4	0x5	0x6	0x7
0x0	**1**	0	0	0	0	0	0	0
0x1	0	**1**	**1**	0	0	**1**	**1**	0
0x2	0	**1**	**1**	0	0	**1**	**1**	0
0x3	0	0	0	**1**	0	0	0	**1**
0x4	0	0	0	0	**1**	0	0	**1**
0x5	0	**1**	**1**	0	0	**1**	**1**	0
0x6	0	**1**	**1**	0	0	**1**	**1**	0
0x7	0	0	0	**1**	**1**	0	0	0

Copyright©2018 NTT corp. All Rights Reserved.

Two Issues of the Previous S-box Model

1. The number of constraints for each S-box is exponential to the S-box size.
   - 4-bit to 4-bit S-box: feasible.
     - About 2^7 linear inequalities are required.
   - 8-bit to 8-bit S-box: difficult.
     - About 2^{15} linear inequalities are required.

2. Probability of differential transition is ignored.
   - An attempt was proposed by Sun et al. in 2014:
     - feasible only up to 4-bit to 4-bit S-box
     - Probability must be 2^{-x} where x is an integer.

NTT

Copyright©2018 NTT corp. All Rights Reserved.

New Method to Model ∗-DDT

Copyright©2018 NTT corp. All Rights Reserved.

New Method to Model (Large) $*$-DDT

- Core observation

Minimizing constraints for $*$-DDT

related ⟷

Finding the minimum product-of-sum representation of a Boolean function

a well-studied topic!!

NTT

Copyright©2018 NTT corp. All Rights Reserved.

$*$-DDT to Product-of-Sum Representation

- Define a $2n$-bit to 1-bit Boolean function that outputs 1 only if the propagation is possible.
- Let us consider the product-of-sum (resp. Conjunctive Normal Form)
- CNF of an example 3-bit S-box

$$
\begin{aligned}
& f(x_2, x_1, x_0, y_2, y_1, y_0) \\
& = (x_2 \vee x_1 \vee x_0 \vee y_2 \vee y_1 \vee \overline{y_0}) \wedge (x_2 \vee x_1 \vee x_0 \vee y_2 \vee \overline{y_1} \vee y_0) \\
& \quad \wedge (x_2 \vee x_1 \vee x_0 \vee y_2 \vee \overline{y_1} \vee \overline{y_0}) \wedge (x_2 \vee x_1 \vee x_0 \vee \overline{y_2} \vee y_1 \vee y_0) \wedge \\
& \quad \cdots \\
& \quad \wedge (\overline{x_2} \vee \overline{x_1} \vee \overline{x_0} \vee \overline{y_2} \vee \overline{y_1} \vee y_0) \wedge (\overline{x_2} \vee \overline{x_1} \vee \overline{x_0} \vee \overline{y_2} \vee \overline{y_1} \vee \overline{y_0})
\end{aligned}
$$

Innovative R&D by NTT

NTT

Copyright©2018 NTT corp. All Rights Reserved.

CNF and Linear Inequalities

- Such Boolean function must be 1.

$$f(x_2, x_1, x_0, y_2, y_1, y_0)$$
$$= (x_2 \vee x_1 \vee x_0 \vee y_2 \vee y_1 \vee \overline{y_0}) \wedge (x_2 \vee x_1 \vee x_0 \vee y_2 \vee \overline{y_1} \vee y_0)$$
$$\wedge (x_2 \vee x_1 \vee x_0 \vee y_2 \vee \overline{y_1} \vee \overline{y_0}) \wedge (x_2 \vee x_1 \vee x_0 \vee \overline{y_2} \vee y_1 \vee y_0) \wedge$$
$$\cdots$$
$$\wedge (\overline{x_2} \vee \overline{x_1} \vee \overline{x_0} \vee \overline{y_2} \vee \overline{y_1} \vee y_0) \wedge (\overline{x_2} \vee \overline{x_1} \vee \overline{x_0} \vee \overline{y_2} \vee \overline{y_1} \vee \overline{y_0})$$

- In other words, all sum representations in the Boolean function must be 1.
- Convert the sum repr. to linear inequality.
- $x_2 \vee x_1 \vee x_0 \vee y_2 \vee y_1 \vee \overline{y_0} = 1$
  $\Longrightarrow x_2 + x_1 + x_0 + y_2 + y_1 + (1 - y_0) \geq 1$

NTT

Copyright©2018 NTT corp. All Rights Reserved.

Minimization of Boolean functions

- Our goal is to find minimized representations.

41 terms

$$f(x_2, x_1, x_0, y_2, y_1, y_0)$$
$$= (x_2 \vee x_1 \vee x_0 \vee y_2 \vee y_1 \vee \overline{y_0}) \wedge (x_2 \vee x_1 \vee x_0 \vee y_2 \vee \overline{y_1} \vee y_0)$$
$$\wedge (x_2 \vee x_1 \vee x_0 \vee y_2 \vee \overline{y_1} \vee \overline{y_0}) \wedge (x_2 \vee x_1 \vee x_0 \vee \overline{y_2} \vee y_1 \vee y_0) \wedge$$
$$\cdots$$
$$\wedge (\overline{x_2} \vee \overline{x_1} \vee \overline{x_0} \vee \overline{y_2} \vee \overline{y_1} \vee y_0) \wedge (\overline{x_2} \vee \overline{x_1} \vee \overline{x_0} \vee \overline{y_2} \vee \overline{y_1} \vee \overline{y_0})$$

Minimize

13 terms

$$f(x_2, x_1, x_0, y_2, y_1, y_0)$$
$$= (x_2 \vee x_1 \vee y_2 \vee \overline{y_1}) \wedge (x_2 \vee x_1 \vee \overline{y_2} \vee y_1) \wedge (x_2 \vee \overline{x_1} \vee y_2 \vee y_1)$$
$$\wedge (\overline{x_2} \vee x_1 \vee y_2 \vee y_1) \wedge (x_2 \vee \overline{x_1} \vee \overline{y_2} \vee \overline{y_1}) \wedge (\overline{x_2} \vee x_1 \vee \overline{y_2} \vee \overline{y_1}) \wedge$$
$$\cdots$$
$$\wedge (x_3 \vee \overline{x_2} \vee \overline{x_1} \vee y_1) \wedge (\overline{x_3} \vee y_3 \vee y_2 \vee y_1)$$

Every term is converted into 1 linear constraint.

Copyright©2018 NTT corp. All Rights Reserved.

Minimization of Boolean functions

- This problem has well been studied in the area of logic circuits.
  - Quine-McCluskey (QM) algorithm
    - optimal but exponential
  - Espresso algorithm
    - heuristic but efficient

inequalities to represent ∗-DDT of 8-bit S-boxes

Structure	# non-zero entries	QM	Espresso
AES S-box	33150	-	8302
SKINNY-128 S-box	54067	372	376

Copyright©2018 NTT corp. All Rights Reserved.

New Methods to Evaluate Probability

Copyright©2018 NTT corp. All Rights Reserved.

Core Observation

- Separate DDT to multiple tables so that each table contains entries with the same probability.

$$pb\text{-DDT} \begin{cases} 1 & \text{if the entry in DDT has probability } pb \\ 0 & \text{otherwise} \end{cases}$$

- Use indicator constraints (with the big-M method) to activate only a single pb-DDT.

Copyright©2018 NTT corp. All Rights Reserved.

Input Difference (Δx)	Output Difference (Δy)							
	0x0	0x1	0x2	0x3	0x4	0x5	0x6	0x7
0x0	1	0	0	0	0	0	0	0
0x1	0	2^{-2}	2^{-2}	0	0	2^{-2}	2^{-2}	0
0x2	0	2^{-2}	2^{-2}	0	0	2^{-2}	2^{-2}	0
0x3	0	0	0	2^{-1}	0	0	0	2^{-1}
0x4	0	0	0	0	2^{-1}	0	0	2^{-1}
0x5	0	2^{-2}	2^{-2}	0	0	2^{-2}	2^{-2}	0
0x6	0	2^{-2}	2^{-2}	0	0	2^{-2}	2^{-2}	0
0x7	0	0	0	2^{-1}	2^{-1}	0	0	0

DDT

Copyright©2018 NTT corp. All Rights Reserved.

Input Difference (Δx)	Output Difference (Δy)							
	0x0	0x1	0x2	0x3	0x4	0x5	0x6	0x7
0x0	1	0	0	0	0	0	0	0
0x1	0	2^{-2}	2^{-2}	0	0	2^{-2}	2^{-2}	0
0x2	0	2^{-2}	2^{-2}	0	0	2^{-2}	2^{-2}	0
0x3	0	0	0	2^{-1}	0	0	0	2^{-1}
0x4	0	0	0	0	2^{-1}	0	0	2^{-1}
0x5	0	2^{-2}	2^{-2}	0	0	2^{-2}	2^{-2}	0
0x6	0	2^{-2}	2^{-2}	0	0	2^{-2}	2^{-2}	0
0x7	0	0	0	2^{-1}	2^{-1}	0	0	0

DDT

2^{-1}-DDT

Δx	Δy							
	0x0	0x1	0x2	0x3	0x4	0x5	0x6	0x7
0x0	0	0	0	0	0	0	0	0
0x1	0	0	0	0	0	0	0	0
0x2	0	0	0	0	0	0	0	0
0x3	0	0	0	1	0	0	0	1
0x4	0	0	0	0	1	0	0	1
0x5	0	0	0	0	0	0	0	0
0x6	0	0	0	0	0	0	0	0
0x7	0	0	0	1	1	0	0	0

2^{-2}-DDT

Δx	Δy							
	0x0	0x1	0x2	0x3	0x4	0x5	0x6	0x7
0x0	0	0	0	0	0	0	0	0
0x1	0	1	1	0	0	1	1	0
0x2	0	1	1	0	0	1	1	0
0x3	0	0	0	0	0	0	0	0
0x4	0	0	0	0	0	0	0	0
0x5	0	1	1	0	0	1	1	0
0x6	0	1	1	0	0	1	1	0
0x7	0	0	0	0	0	0	0	0

Innovative R&D by NTT

Flag Variables

2^{-1}-DDT

Δx	Δy							
	0x0	0x1	0x2	0x3	0x4	0x5	0x6	0x7
0x0	0	0	0	0	0	0	0	0
0x1	0	0	0	0	0	0	0	0
0x2	0	0	0	0	0	0	0	0
0x3	0	0	0	1	0	0	0	1
0x4	0	0	0	0	1	0	0	1
0x5	0	0	0	0	0	0	0	0
0x6	0	0	0	0	0	0	0	0
0x7	0	0	0	1	1	0	0	0

$Q_1 \in \{0, 1\}$

2^{-2}-DDT

Δx	Δy							
	0x0	0x1	0x2	0x3	0x4	0x5	0x6	0x7
0x0	0	0	0	0	0	0	0	0
0x1	0	1	1	0	0	1	1	0
0x2	0	1	1	0	0	1	1	0
0x3	0	0	0	0	0	0	0	0
0x4	0	0	0	0	0	0	0	0
0x5	0	1	1	0	0	1	1	0
0x6	0	1	1	0	0	1	1	0
0x7	0	0	0	0	0	0	0	0

$Q_2 \in \{0, 1\}$

- $Q_1 + Q_2$ is 1 if the S-box is active.
- Linear inequalities for 2^{-i}-DDT are activated If $Q_i = 1$.
  - Such constraint is called ***indicator constraint*** in MILP.

Copyright©2018 NTT corp. All Rights Reserved.

Experimental Data for pb-DDT

Structure		Num. of zero entries	QM	Espresso
AES S-box	2^{-7}	33406	-	8241
	2^{-6}	65281	-	350
SKINNY-128 S-box	2^{-7}	62848	206	208
	2^{-6}	60530	275	283
	$2^{-5.4}$	65472	33	34
	2^{-5}	62708	234	239
	$2^{-4.4}$	65458	42	52
	2^{-4}	64884	147	159
	$2^{-3.7}$	65534	15	15
	$2^{-3.4}$	65518	24	28
	$2^{-3.2}$	65534	15	15
	2^{-3}	65435	62	67
	$2^{-2.7}$	65534	16	16
	$2^{-2.4}$	65532	17	17
	2^{-2}	65513	37	40

Copyright©2018 NTT corp. All Rights Reserved.

Application to Skinny

Copyright©2018 NTT corp. All Rights Reserved.

SKINNY

- Proposed at CRYPTO2016 by Beierle et al.
- Tweakable block cipher supporting n-bit block and n-, $2n$-, and $3n$-bit tweakey, where $n \in \{64,128\}$.

- In this talk, we focus our attention on the single-key analysis of SKINNY-128.

Copyright©2018 NTT corp. All Rights Reserved.

Search Results

Rounds	9	10	11	12	13	14
LB [BJK+16]	2^{-82}	2^{-92}	2^{-102}	2^{-110}	2^{-116}	2^{-122}
Tight bound	2^{-86}	2^{-96}	2^{-104}	2^{-112}	$\mathbf{2^{-123}}$	$\mathbf{\leq 2^{-128}}$

- To achieve these results, we use several heuristic strategy.
  - We first get better upper bound heuristically.
  - Then we get the tight bound by using the knowledge of the upper bound.

Please refer to our paper in detail.

NTT

Copyright©2018 NTT corp. All Rights Reserved.

Concluding Remarks

- New MILP model
  - QM and Espresso for modeling $*$-DDT.
  - pb-DDT and indicator constraints.
- Applications
  - Improved diff resistance of SKINNY-128
  - Evaluated prob of AES-round based function.
- MILP can be applied to 8-bit Sboxes!!

Thank you for your attention!!

Copyright©2018 NTT corp. All Rights Reserved.