Boolean Functions with Multiplicative Complexity 3 and 4

Çağdaş Çalık, Meltem Sönmez Turan, **René Peralta**

National Instutite of Standards and Technology (Gaithersburg MD, USA)

Boolean Functions and Applications, Florence, Italy 2019

Motivation

General Circuit Complexity Problem

Given a basis of Boolean gates, construct a circuit that computes a function that is optimal w.r.t. to some criteria.

Multiplicative Complexity (MC) of f, denoted $C_\wedge(f)$, is the minimum number of AND gates that is sufficient to evaluate f over the basis (AND, XOR, NOT).

- Relevant for side channel resistance, secure multi-party computation, cryptanalysis etc.

Some Properties of Multiplicative Complexity

- MC of a randomly selected n-variable Boolean function is at least $2^{n/2} - \mathcal{O}(n)$ with high probability [BPP00].
- MC of a function with degree d is at least $d - 1$ (degree bound).
- MC is affine invariant.
    - Boolean functions $f, g \in B_n$ are affine equivalent if there exists a transformation of the form $f(x) = g(Ax + a) + b \cdot x + c$, where $A \in GL(n, 2)$; $a, b \in \mathbb{F}_2^n$, and $c \in \mathbb{F}_2$.
    - The set of affine equivalent functions constitute an equivalence class denoted by $[f]$, where f is an arbitrary function from the class.
    - Affine equivalent Boolean functions have the same MC.

Enumeration by number of variables

MC distribution is known for up to 6-variables:

- $C_\wedge(f) \leq n - 1$ for $f \in B_n$, $n \leq 5$ [TP14],
- $C_\wedge(f) \leq 6$ for $f \in B_6$ [CTP18].

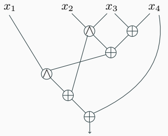
The method is infeasible for $n \geq 7$, due to the large number of affine equivalence classes and topologies.

Enumeration by multiplicative complexity

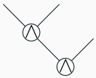
Exhaustively construct all Boolean topologies with 1,2, 3, . . . AND gates, and evaluate the topologies until a function from $[f]$ is generated.

- Topology: Abstraction of a Boolean circuit that shows the relations between AND gates

Boolean circuit



Topology



Boolean functions with MC 1 and 2

Boolean functions with MC 1 [FP02]

- Functions with MC 1 are affine equivalent to x_1x_2.
- The number of n-variable Boolean functions with MC 1 is $2\binom{2^n}{3}$.

Boolean functions with MC 2 [FTT17]

- Functions with MC 2 are affine equivalent to one of the functions from the set $\{x_1x_2x_3,\ x_1x_2x_3 + x_1x_4,\ x_1x_2 + x_3x_4\}$.
- The number of n-variable Boolean functions with MC 2 is

$$2^n(2^n-1)(2^n-2)(2^n-4)\left(\frac{2}{21}+\frac{2^n-8}{12}+\frac{2^n-8}{360}\right).$$

Boolean functions with MC 3 and 4

This work: Find exhaustive list of equivalence classes with MC 3 and 4.

Approach

Step 1. Construct Boolean circuits (topologies) with 3 and 4 AND gates.

Step 2. Evaluate the circuits to generate Boolean functions.

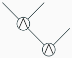
Step 3. Identify distinct affine equivalence classes with MC 3 and 4.

Constructing Topologies [CTP18]

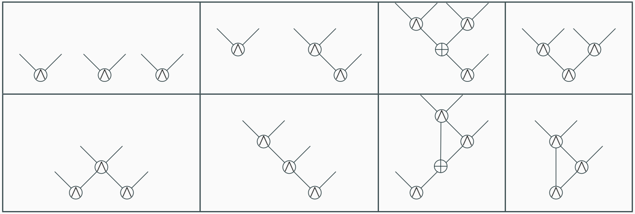
Topologies with 1 AND gate

Topologies with 2 AND gates

and



Topologies with 3 AND gates



Number of topologies with 4 AND gates is 84.

Boolean functions with MC 3 and 4

This work: Find exhaustive list of equivalence classes with MC 3 and 4.

Approach

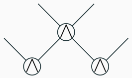
Step 1. Construct Boolean circuits (topologies) with 3 and 4 AND gates.

Step 2. Evaluate the circuits to generate Boolean functions.

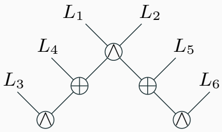
Step 3. Identify distinct affine equivalence classes with MC 3 and 4.

Evaluating Topologies to Generate Boolean Functions

- A topology with k AND gates can be supplied $2k$ linear function inputs $X = (L_1, \ldots, L_{2k})$.

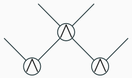


- Any affine transformation of the inputs $A(X) = (A(L_1), \ldots, A(L_{2k}))$ will produce a function from the same equivalence class. Hence, the inputs that are affine transformations of each other need not be considered.

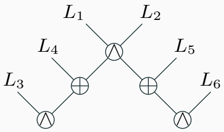


Evaluating Topologies to Generate Boolean Functions

- A topology with k AND gates can be supplied $2k$ linear function inputs $X = (L_1, \ldots, L_{2k})$.



- Any affine transformation of the inputs $A(X) = (A(L_1), \ldots, A(L_{2k}))$ will produce a function from the same equivalence class. Hence, the inputs that are affine transformations of each other need not be considered.



Warning: One topology can correspond to multiple equivalence classes of functions.

Dimension of a Boolean function

The following functions are affine equivalent:

$$x_1 x_2$$

$$x_1 x_3 + x_1 x_4 + x_2 x_3 + x_2 x_4$$

Affine transformations can eliminate variables.

It is easier to work on smaller number of variables.

Dimension of a Boolean function

The following functions are affine equivalent:

$$x_1x_2$$

$$x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4$$

Affine transformations can eliminate variables.

It is easier to work on smaller number of variables.

Definition. Let L_f be the number of input variables that appear in the *algebraic normal form* (ANF) of a Boolean function f. The dimension of f is the smallest number of variables that appear in the ANF among the functions that are affine equivalent to f:

$$dim(f) = \min_{g \in [f]} L_g.$$

Autocorrelation, Linear Structures, and Dimension

The autocorrelation of a Boolean function f at $\alpha \in \mathbb{F}_2^n$ is

$$R_f(\alpha) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x)+f(x+\alpha)}.$$

The autocorrelation spectrum of f is the vector $[R_f(0), \ldots, R_f(2^n - 1)]$.

$\alpha \in \mathbb{F}_2^n$ is a linear structure[1] of f if $f(x) + f(x+\alpha)$ is constant.

The linearity dimension of f is equal to

$$d_l(f) = \log_2 \#\{|R_f(\alpha)| = 2^n, \alpha \in \mathbb{F}_2^n\}.$$

Observation: the dimension of an n-variable Boolean function is:

$$dim(f) = n - d_l(f).$$

[1] X. Lai, Additive and Linear Structures of Cryptographic Functions, FSE'94.

A New MC Lower Bound based on Dimension

Theorem
For $f \in B_n$, $C_\wedge(f) \geq \lceil \dim(f)/2 \rceil$.

Sketch of the proof.

1. Let $C_\wedge(f) = k$, consider a circuit implementing f with k AND gates.
2. The topology with k AND gates has $2k$ linear function inputs.
3. The rank of $2k$ linear functions can be at most $2k$.
4. Any set of $2k$ linear functions on $n > 2k$ variables can be affine transformed to functions having at most $2k$ variables.
5. Therefore, $dim(f) \leq 2k$, which implies $C_\wedge(f) \geq \lceil dim(f)/2 \rceil$.

Example. Let $f = \Sigma_4^8 = x_1x_2x_3x_4 + \ldots + x_5x_6x_7x_8$. According to the degree bound, $C_\wedge(f) \geq 3$. By dimension bound, $C_\wedge(f) \geq 8/2 = 4$.

Boolean functions with MC 3 and 4

This work: Find exhaustive list of equivalence classes with MC 3 and 4.

Approach

Step 1. Construct Boolean circuits (topologies) with 3 and 4 AND gates.

Step 2. Evaluate the circuits to generate Boolean functions.

Step 3. Identify distinct affine equivalence classes with MC 3 and 4.

Affine Equivalence Classes with MC 3

Dimension 4:

$x_1x_2x_3x_4$
$x_1x_2 + x_1x_2x_3x_4$
$x_2x_3 + x_1x_4 + x_1x_2x_3x_4$

Dimension 5:

$x_3x_4 + x_1x_5 + x_1x_2x_5 + x_1x_2x_3x_4$	$x_3x_4 + x_1x_3x_4 + x_1x_2x_5$
$x_2x_4 + x_1x_5 + x_1x_2x_3$	$x_4x_5 + x_1x_2x_3$
$x_1x_2x_5 + x_1x_2x_3x_4$	$x_1x_3x_4 + x_1x_2x_5$
$x_2x_3x_5 + x_1x_4x_5 + x_1x_2x_3x_4$	$x_3x_5 + x_1x_2x_5 + x_1x_2x_3x_4$
$x_1x_3 + x_1x_2x_5 + x_1x_2x_3x_4$	$x_3x_4 + x_1x_2x_5 + x_1x_2x_3x_4$
$x_1x_5 + x_1x_2x_3x_4$	$x_2x_3 + x_1x_5 + x_1x_2x_3x_4$
$x_2x_3 + x_2x_3x_5 + x_1x_4x_5 + x_1x_2x_3x_4$	$x_1x_5 + x_1x_2x_5 + x_1x_2x_3x_4$

Dimension 6:

$x_3x_4 + x_2x_5 + x_1x_6$	$x_1x_6 + x_1x_3x_4 + x_1x_2x_5$
$x_3x_4 + x_1x_6 + x_1x_3x_4 + x_1x_2x_5$	$x_4x_5 + x_1x_6 + x_1x_2x_3$
$x_1x_6 + x_1x_2x_5 + x_1x_2x_3x_4$	$x_5x_6 + x_3x_4x_5 + x_1x_2x_6 + x_1x_2x_3x_4$
$x_3x_4 + x_1x_6 + x_1x_2x_5 + x_1x_2x_3x_4$	

Number of Boolean functions with MC 3

The number of n-variable Boolean functions with MC 3 is

$$2^{n-4}\prod_{i=0}^{3}\frac{2^n-2^i}{2^4-2^i}s_4+2^{n-5}\prod_{i=0}^{4}\frac{2^n-2^i}{2^5-2^i}s_5+2^{n-4}\prod_{i=0}^{5}\frac{2^n-2^i}{2^6-2^i}s_6,$$

where

$$\begin{aligned} s_4 &= 32768, \\ s_5 &= 1576479744, \\ s_6 &= 183894007808. \end{aligned}$$

Affine Equivalence Classes with MC 4

After evaluating 84 topologies with 4 AND gates, we obtained

- 26 classes with dimension 5,
- 888 classes with dimension 6,
- 321 classes with dimension 7,
- 42 classes with dimension 8.

Complete list is available at:

`https://github.com/usnistgov/Circuits/tree/master/data/mc_dim`

Conclusion

- Provided a new lower bound for the MC of Boolean functions based on their dimension.
- Identified all equivalence classes with MC 3 (24 classes) and MC 4 (1277 classes).
- Ongoing. The identification of classes with MC 5 is still in progress.

MC	dimension										
	2	3	4	5	6	7	8	9	10	11	12	Total
1	1											1
2		1	2									3
3			3	14	7							24
4				26	888	321	42					1277
5					148483	*	*	*	575			*
6					931	*	*	*	*	*	*	*

Table 1: The Distribution of Classes w.r.t MC and Dimension.

More Information

NIST Circuit Complexity Project Webpage:

`https://csrc.nist.gov/Projects/Circuit-Complexity`

GitHubLink:

`https://github.com/usnistgov/Circuits/`

Contact email:

`circuit_complexity@nist.gov`

References

[BPP00] J. Boyar, R. Peralta, and D. Pochuev, “On the multiplicative complexity of Boolean functions over the basis $(\wedge, \oplus, 1)$, Theoretical Computer Science, vol. 235, no. 1, pp. 43 – 57, 2000.

[CTP18] Ç. Çalık, M. Sönmez Turan, R. Peralta, The Multiplicative Complexity of 6-variable Boolean Functions, Cryptography and Communications 2018.

[FP02] M. J. Fischer and R. Peralta. Counting Predicates of Conjunctive Complexity One. Yale Technical Report 1222, February 2002.

[FTT17] M. G. Find, D. Smith-Tone, M. Sönmez Turan, The Number of Boolean Functions with Multiplicative Complexity 2, International Journal of Information and Coding Theory, 2017.

[Lai94] X. Lai, Additive and Linear Structures of Cryptographic Functions, FSE 1994, LNCS 1008, Springer-Verlag, pp. 75–85, 1994.

[Nyb92] K. Nyberg, On the Construction of Highly Nonlinear Permutations, Eurocrypt’92.

[TP14] M. Sönmez Turan and R. Peralta. The Multiplicative Complexity of Boolean functions on Four and Five Variables. LightSec 2014, Turkey.

Computing the Autocorrelation Spectrum

Wiener-Khintchine Theorem
The autocorrelation spectrum and the Walsh spectrum of a Boolean function are related in the following way:

$$\left[R_f(0), \ldots, R_f(2^n - 1)\right] = \frac{1}{2^n}\left[W_f^2(0), \ldots, W_f^2(2^n - 1)\right] H_n,$$

where H_n is the Sylvester-Hadamard matrix of order 2^n.

Computing the autocorrelation spectrum of $f \in \mathcal{B}_n$ can be carried out as follows:

1. Compute the Walsh spectrum of f using Fast Walsh Transform.
2. Take the squares of Walsh spectrum entries.
3. Apply another Fast Walsh Transform to the resulting sequence.
4. Divide each entry by 2^n.

The complexity of computing the autocorrelation spectrum is $\mathcal{O}(n2^n)$

Autocorrelation-ANF Relationship

Any Boolean function can be expressed in the form

$$f(x) = x_i g_1(x) + g_2(x),$$

where the functions $g_1(x)$ and $g_2(x)$ do not depend on x_i. Let $\alpha_i \in \mathbb{F}_2^n = e_i$, i.e., $w_H(\alpha_i) = 1$. Then,

$$\begin{aligned} R_f(\alpha_i) &= f(x) + f(x+\alpha_i) \\ &= [x_i g_1(x) + g_2(x)] + [(x_i+1) g_1(x) + g_2(x)] \\ &= g_1(x) \end{aligned}$$

If $|R_f(\alpha_i)| = 2^n$ implies $g_1(x)$ is constant. Also,

- If $g_1(x) = 0$ then x_i does not appear in the ANF.
- If $g_1(x) = 1$ then $f(x) = x_i + g_2(x)$, i.e., x_i appears as a linear term.
- Conclusion. f is either independent of x_i or can be transformed to a function that is independent of x_i.