A Genetic Algorithm for Evolving Plateaued Cryptographic Boolean Functions

TPNC 2015 - December 15-16 - Mieres

Luca Mariot, Alberto Leporati

Dipartimento di Informatica, Sistemistica e Comunicazione
Università degli Studi Milano - Bicocca
luca.mariot@disco.unimib.it, alberto.leporati@unimib.it

December 15, 2015

Boolean Functions - Basic Definitions

Boolean function: a mapping $f : \mathbb{F}_2^n \to \mathbb{F}_2$, where $\mathbb{F}_2 = \{0, 1\}$

Truth table representation:

(x_1, x_2, x_3)	000	100	010	110	001	101	011	111
$f(x_1, x_2, x_3)$	0	1	1	1	1	0	0	0

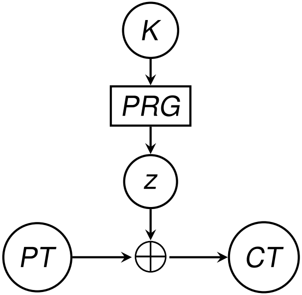
$$\Downarrow$$

$$\Omega_f = (0, 1, 1, 1, 1, 0, 0, 0)$$

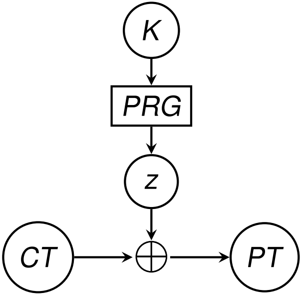
Algebraic Normal Form representation:

$$f(x_1, x_2, x_3) = x_1 \cdot x_2 \oplus x_1 \oplus x_2 \oplus x_3$$

Vernam Stream Cipher



(a) Encryption

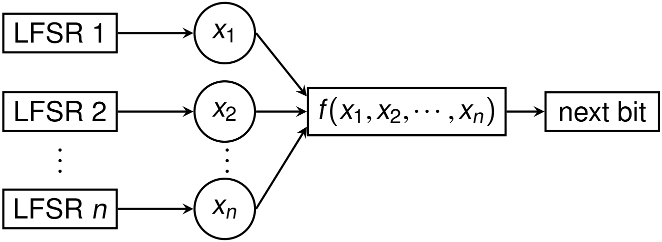


(b) Decryption

- K: secret key
- PRG: Pseudorandom Generator
- z: keystream
- $\bigoplus$: bitwise XOR
- PT: Plaintext
- CT: Ciphertext

An Example of PRG: The Combiner Model

- Function $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ combines the outputs of n Linear Feedback Shift Registers (LFSRs)



- Security of the model $\Leftrightarrow$ cryptographic properties of f

Walsh Transform

$$\hat{F}(\omega) = \sum_{x \in \mathbb{F}_2^n} \hat{f}(x) \cdot (-1)^{\omega \cdot x}$$

- $\hat{f}(x) = (-1)^{f(x)}$
- $\omega \cdot x = \omega_1 \cdot x_1 \oplus \cdots \oplus \omega_n \cdot x_n$
- Walsh Spectrum $\mathcal{S}_f = (\hat{F}(\underline{0}), \cdots, \hat{F}(\underline{1}))$
- Spectral Radius $W_M(f)$: maximum absolute value in $\mathcal{S}_f$

$$\Omega_f = (0,1,1,1,1,0,0,0)$$

$$\Downarrow \hat{F}$$

$$\mathcal{S}_f = (0,0,0,0,-4,4,4,4)$$

$$\Downarrow$$

$$W_M(f) = 4$$

Cryptographic Properties (1/3)

- Balancedness: Half of the truth table is composed of ones ($\Leftrightarrow \hat{F}(0) = 0$)

$$\Omega_f = (0,1,1,1,1,0,0,0) \;\Rightarrow\; \text{4 ones} \;\Rightarrow\; \text{BALANCED}$$

- Algebraic Degree: Degree of the ANF

$$f(x_1,x_2,x_3) = x_1 \cdot x_2 \oplus x_1 \oplus x_2 \oplus x_3 \;\Rightarrow\; deg(f) = 2$$

Cryptographic Properties (2/3)

- Nonlinearity: Hamming distance of f from affine functions (⇔ functions of degree 1)

$$n = 3,\ W_M(f) = 4 \ \Rightarrow nl(f) = 2^{-1}(2^n - W_M(f)) = 2$$

- m-Resiliency: $\hat{F}(\omega) = 0$ for all ω having at most m ones

$$\mathcal{S}_f = (0,0,0,0,-4,4,4,4) \ \Rightarrow \ \hat{F}(0,0,1) = -4 \neq 0$$

$$\Rightarrow f \text{ is NOT 1-resilient}$$

Cryptographic Properties (3/3)

- $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ with profile (n, m, d, nl) should:
  - be balanced
  - be resilient of high order m
  - have high algebraic degree d
  - have high nonlinearity nl
- Trade-offs:
  - *Siegenthaler's bound*: $d \leq n - m - 1$ [Siegenthaler84]
  - *Tarannikov's bound*: $Nl \leq 2^{n-1} - 2^{m+1}$ [Tarannikov00]

Search for Cryptographic Boolean Functions

- For $n > 5$, exhaustive search is unfeasible
- Evolutionary search offers a promising way to optimize cryptographic boolean functions
- Usual approach: directly search the space of boolean
- Complementary approach: Spectral Inversion

Spectral Inversion [Clark04] (1/2)

- Applying the Inverse Walsh Transform to a generic spectrum yields a pseudoboolean function $f : \mathbb{F}_2^n \to \mathbb{R}$

$$\mathcal{S}_f = (0, -4, -2, 2, 2, 4, 4, -2)$$

$$\Downarrow \hat{F}^{-1}$$

$$\Omega_{\hat{f}} = (0, 0, 0, -1, 0, -1, 2)$$

- New objective: minimize the deviation of Walsh spectra which satisfy the desired cryptographic constraints

Spectral Inversion [Clark04] (2/2)

Heuristic techniques proposed for this optimization problem:

- Clark et al. [Clark04]: Simulated Annealing (SA)
- Our work: Genetic Algorithms (GA)

Plateaued Functions [Zhang99]

- Our GA evolves spectra of plateaued functions
- A (pseudo)boolean function f is plateaued if its Walsh spectrum takes only three values: $-W_M(f)$, 0 and $+W_M(f)$

$$\mathcal{S}_f = (0,0,0,0,-4,4,4,4) \Rightarrow \text{plateaued}$$

- Motivations:
  - Simple combinatorial representation of candidate solutions, determined by a single parameter $r \geq n/2$
  - Plateaued functions reach both Siegenthaler's and Tarannikov's bounds

Chromosome Encoding

- Resiliency Constraint: ignore positions with at most m ones

x	<u>000</u>	<u>100</u>	<u>010</u>	110	<u>001</u>	101	011	111
$\mathcal{S}_f$	0	0	0	−4	0	4	4	4

- The chromosome c is the permutation of the spectrum in the positions with more than m ones:

x	110	101	011	111
c	−4	4	4	4

- The multiplicities of 0, $-W_M(f)$ and $+W_M(f)$ in the permutation depend on plateau index r

Fitness Function

- Given $\hat{f} : \mathbb{F}_2^n \to \mathbb{R}$, the nearest boolean function $\hat{b} : \mathbb{F}_2^n \to \mathbb{F}_2$ is defined for all $x \in \mathbb{F}_2^n$ as:

$$\hat{b}(x) = \begin{cases} +1 & , \text{if } \hat{f}(x) > 0 \\ -1 & , \text{if } \hat{f}(x) < 0 \\ +1 \text{ or } -1 \text{ (chosen randomly)} & , \text{if } \hat{f}(x) = 0 \end{cases}$$

- Objective function proposed in [Clark04]:

$$obj(f) = \sum_{x \in \mathbb{F}_2^n} (\hat{f}(x) - \hat{b}(x))^2$$

- Fitness function maximised by our GA: $fit(f) = -obj(f)$

Genetic Operators (1/2)

- Crossover between two Walsh spectra p_1, p_2 must preserve the multiplicities of $-W_M(f)$, 0 and $+W_M(f)$
- Idea: use counters to keep track of the multiplicities [Millan98]

Genetic Operators (2/2)

- Mutation: swap two random positions in the chromosome with different values
- Selection operators adopted:
  - Roulette-Wheel (*RWS*)
  - Deterministic Tournament (*DTS*)

Experimental Settings

Common parameters:

- Number of variables $n = 6,7$ and plateau index $r = 4$

(n,m,d,nl)	$\lvert 0_{res}\rvert$	$\lvert 0_{add}\rvert$	$\lvert -W_M(f)\rvert$	$\lvert +W_M(f)\rvert$
$(6,2,3,24)$	22	26	6	10
$(7,2,4,56)$	29	35	28	36

GA-related parameters:

- Population size $N = 30$
- max generations $G = 500000$
- GA runs $R = 500$
- Crossover probability $p_\chi = 0.95$
- Mutation probability $p_\mu = 0.05$
- Tournament size $k = 3$

SA-related parameters:

- Inner loops $MaxIL = 3000$
- Moves in loop $MIL = 5000$
- SA runs $R = 500$
- Initial temperatures $T = 100, 1000$
- Cooling parameter: $\alpha = 0.95, 0.99$

Results

Statistics of the best solutions found by our GA and SA over $R = 500$ runs.

n	Stat	$GA(RWS)$	$GA(DTS)$	$SA(T_1, \alpha_1)$	$SA(T_2, \alpha_2)$
6	avg_o	14.08	13.02	19.01	19.03
	min_o	0	0	0	0
	max_o	16	16	28	28
	std_o	5.21	6.23	4.89	4.81
	$\#opt$	60	93	11	10
	avg_t	83.3	79.2	79.1	79.4
7	avg_o	53.44	52.6	45.09	44.85
	min_o	47	44	32	27
	max_o	58	59	63	57
	std_o	2.40	2.77	4.39	4.18
	$\#opt$	0	0	0	0
	avg_t	204.2	204.5	180.3	180.2

Conclusions

- Main contribution: Genetic Algorithm for evolving Walsh spectra of boolean functions by spectral inversion
- The GA focuses exclusively on plateaued functions, due to their good cryptographic properties
- Specialized crossover and mutation to preserve the multiplicities in the spectra
- For $n = 6$, our GA is more efficient than SA [Clark04] in generating plateaued boolean functions

Future Developments

- $n = 6$ is too low for practical cryptographic applications! (necessary at least $n = 13$ to avoid algebraic attacks)
- Our GA does not scale to higher number of variables
- Future experiments: combine our GA with local search technique of [Kavut07]
- Further improvements: different fitness functions, additional cryptographic properties, ...

References

- [Clark04] Clark, J.A., Jacob, J., Maitra, S., Stanica, P.: Almost Boolean Functions: The Design of Boolean Functions by Spectral Inversion. Comput. Intell. 20(3):450-462 (2004)
- [Kavut07] Kavut, S., Yucel, M.D., Maitra, S.: Construction of Resilient Functions by Concatenation of Boolean Functions Having Nonintersecting Walsh Spectra. In: Michon, J.-F., Valarcher, P., Yunès, J.-B. (eds.) BFCA '07, pp. 43–62. Universités de Rouen et du Havre (2007)
- [Millan98] Millan, W., Clark, A., Dawson, E.: Heuristic Design of Cryptographically Strong Balanced Boolean Functions. In: Nyberg, K. (ed.) EUROCRYPT '98. LNCS, vol. 1403, pp. 489-499. Springer, Heidelberg (1998)
- [Siegenthaler84] Siegenthaler, T.: Correlation-Immunity of Nonlinear Combining Functions for Cryptographic Applications. IEEE Trans. Inf. Theory 30(5), 776–780 (1984)
- [Tarannikov00] Tarannikov, Y.V.: On Resilient Boolean Functions with Maximum Possible Nonlinearity. In: Roy, B.K., Okamoto, E. (eds.) INDOCRYPT 2000. LNCS, vol. 1977, pp. 19-30. Springer, Heidelberg (2000)
- [Zhang99] Zheng, Y., Zhang, X.-M.: Plateaued Functions. In: Varadharajan, V., Mu, Y. (eds.) ICICS '99. LNCS, vol. 1726, pp. 284-300. Springer, Heidelberg (1999)