ON THE CONSTRUCTION OF S- BOXES USING THE LEADERS AND FOLLOWERS METAHEURISTIC

Alejandro Freyre- Echevarría, Ismel Martínez- Díaz

University of Havana, Faculty of Math and Computer Science

a.freyre@estudiantes.matcom.uh.cu, ismel@matcom.uh.cu

RESUMEN

Los cifradores de bloque modernos enfrentan el peligro que suponen los ataques de canal colateral por consumo de potencia que tienen como objetivo principal las componentes no lineales conocidas como S - cajas. Una métrica teórica frente a este tipo de ataques es la propiedad denominada varianza del coeficiente de confusión. Mientras mayor sea el valor de esta métrica mejor será la resistencia teórica de la S- caja. En este trabajo se presenta el uso de la meta- heurística líderes y seguidores en la obtención de S- cajas con altos valores de la varianza del coeficiente de confusión.

ABSTRACT

Modern block ciphers are facing the threat of side- channel attacks by power leakage whose main target are the non- linear components known as S- boxes. A theoretical measure for the resistance of an S- box against this type of attacks is the confusion coefficient variance property. A higher value of this property represents a better theoretical resistance. In this work we use the leaders and followers meta- heuristic in order to achieve good confusion coefficient variance's valued S- boxes.

Keywords: S- box, confusion coefficient variance, Leaders and Followers

INTRODUCTION

In modern days, the information shared by users is a valuable resource. In order to protect the data of each user is necessary the use of symmetric cryptography and in a particular case the use of block ciphers. The most important non- linear component of block ciphers are S- boxes, vector boolean functions, which warranties the confusion during the encryption process [1]. Block ciphers are under the constant threat of correlation power attacks (CPA) that targets the power leakages measured by the attacker during the evaluation process of the S- box [2]. There are a set of properties to study the resistance associated to an S- box against the CPA attacks [3]. Some of those properties are theoretical and in general, their use is to design highly resilient S- boxes against CPA attacks.

One of those properties is the confusion coefficient variance (CCV) [4]. Picek *et al*. in [4] shows as higher the CCV the better resistance of an S- box against CPA attacks. Thus, higher CCV valued S- boxes are better on the construction of block ciphers. There are two other properties we use in this paper, they are the non- linearity (NL) [5] and differential uniformity (δ)[10, 11, 12]. The first measures the resistance of an S- box against linear attacks while the second refers to resilience against differential attacks.

From Tamayo- Vera in [7] we receive our motivation to use Leaders and Followers meta- heuristic. This investigation was made for the property of transparency order (TO) [8], being the first time the Leaders and Followers method was applied to combinatory and cryptographic problems. The results presented in [7] were very good for TO property, and is our interest to obtain those results but for the CCV property.

The Leaders and Followers meta- heuristic is an evolutionary algorithm[6], with the distinctive characteristic of using two populations the *leaders* and the *followers*. The *leaders* keep track of the best solutions found by the algorithm, while the *followers* are tasked with the search for new solutions. The new solutions generated within the followers are compared with their peers in the *followers* population. The *leaders* population is updated with solutions from the *followers* population only after enough search has been performed in the context of the *followers* [6].Figure below shows a pseudo- code of the method.

```
1:  L ← Initialize the leaders with n uniform random vectors.
2:  F ← Initialize the followers with n uniform random vectors.
3:  repeat
4:      for i ← 1, n do
5:          leader ← Pick a leader from L.
6:          follower ← Pick a follower from F.
7:          trial ← create_trial(leader, follower)
8:          if f(trial) < f(follower) then
9:              Substitute follower by trial in F.
10:         end if
11:     end for
12:     if median(f(F)) < median(f(L)) then
13:         L ← merge_populations(L, F)
14:         F ← Reinitialize the followers uniformly.
15:     end if
16: until The termination criterion is satisfied.
```

Fig. 1 – Pseudo- code of the Leaders and Followers method.

The algorithm start with *leaders* and *followers* populations randomly selected from the search space, repeating its main loop until a defined termination criterion is satisfied. On every iteration of the main loop begins with a round of update to the followers, made by selecting a specified number of pair *(leader, follower)* to create a new trial child. If the child is better than its *follower* parent is, it will replace it in the *followers* population. For the final part of the main loop, it is checked if enough search has been performed in the context of the *followers* to enhance the *leaders*, and if it so merge the populations returning a new improved *leaders* population. Once the leaders are updated, the *followers* are restarted as uniformly sampled solutions.

EXPERIMENT AND RESULTS

We only test the method for the 8x8 bit S- box space. Like in [4] our fitness function *F* is taken as the sum of the values of non- linearity and confusion coefficient variance of analyzed S- box. Equation I represents the fitness function *F*.

$$F = NL + CCV \qquad (I)$$

We made a few changes from the method presented in [7]. Firstly, our method's stop criteria is reaching a fixed number of iterations (not function evaluations) given in the input. Like [7], we also modify the way to create a new trial. In addition to both parents; we took a value *x* within the interval $[0, 2^n - 1]$, being n the number of bits for each S- box input; making the new trial equal to the *follower* with the *follower* outputs, in the output of evaluating *follower*(*x*) and evaluating *leader*(*x*), swapped (see Fig. 1). In this new composition, the evaluation of *x* in each pair of *leader* and *follower* determine distinctively if the new trial S- box can improve the *follower* or not. Equation II represents the new trial creation:

$$trial = swap(follower,\ follower(x),\ leader(x)) \qquad (II)$$

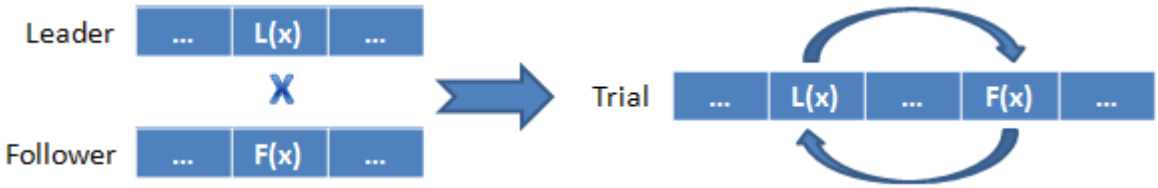

Fig. 1 – New trial creation using the function of equation II.

On every round of update to the *followers* within the main loop (see steps 4- 11 in [6]), we creates a new trial for every value of $x \in [0, 2^n - 1]$ selecting at random a *leader* and a *follower*. Like in the original method, if the evaluation of the trial is better than *follower* was, then we proceed to substitute the *follower* for the new solution. The merge criterion for the populations is satisfied if the mean of fitness function *F* on the set of *followers* is better than the mean of *F* on the *leaders*.

Tuning parameters to ensure a good result in the minimum possible time is always hard. Taking as reference the results of [7] over 500000 function evaluations (around 2000 of our iterations - main loop -), we decide to set the iterations to 2500 in order of finding good results in a good time.

For setting size to both populations, the *leaders* and the *followers,* we seek guidance in experiments presented in public literature. In [6] the author of Leaders and Followers suggest a population of the same size as the amount of components from an input of the search space domain. In our case results in having two populations of size 2^n. In [7], Tamayo- Vera, after a group of trials over sizes between 100 and 500, set the population size to 300. Picek et al. in [4] select a population size of 50. Finally, the author in [9] gets good results by setting the population size to 10 elements. We took the last reference to set out population size.

We run two independent test, making 2500 iterations over a population size of 10 elements for the *leaders* and the *followers*. Small population sizes warrantees a mayor rate of changes made over an S-box since it can be selected with better probability than from a larger set of S- boxes. However, we cannot assume the changes made will be always positive. We rely the majority of those changes become in upgrades, avoiding missed function evaluations as much as possible.

We establish a statistical comparison within the space of 8x8 bit S- boxes showing the behavior of 25000 randomly generated S- boxes and our experiments results (see Table 1.).

Method	Maximum NL	Minimum NL	Mean NL	Maximum CCV	Minimum CCV	Mean CCV	Generated S- Boxes
Random	98	76	92.7	0.249	0.076	0.124	25000
LaF	98	98	98	4.66	4.026	4.253	20

Table. 1– Statistical comparison on some S- boxes within the space 8x8 bits.

The outcome results from both tests confirm that select a small population size is not a bad idea, because on every resulting set of *leaders* all the elements present values of confusion coefficient variance greater than 4.0 points. The two main differences between random generated S- boxes and our method are quality of CCV values and time.In matter of time, random generation is quite fast than applying Leaders and Followers with mean time of 24 hours in an Intel(R) Core(TM) i3- 2310 processor @ 2.10 GHz and 4GB of memory. In quality terms, results proves that Leaders and Followers is widely better than random generation. We always generate a set of S- boxes with high values of confusion coefficient variance and good non- linearity, while, as we show in *Table 1,* is very difficult to randomly generate S-boxes with confusion coefficient variance greater than 0.5, and a significant amount of them present low values of non- linearity.

Table 2 shows the values of non- linearity, CCV and differential uniformity of AES S- box, the best 8x8 S-box presented by Picek *et al.* in [4] and two of our results. In *Table 1*, the mean of CCV values is bigger than Picek's CCV value. Despite we could not achieve an almost optimal non- linearity like the Advanced Encryption Standard (AES) S- Box [12]; Table 1 shows that generated S- boxes presents a value of 98 like Picek's NL value. We also check the values of differential uniformity for all S- boxes, having two of the total generated with δ values greater than 12. Nevertheless, the average value for this metric among the evolved S- boxes is 12, including four S- boxes with δ value of 10.

S- box	Non- linearity	Confusion Coefficient Variance	Differential Uniformity
AES	112	0.11	4
Picek	98	4.06	12
S_1	98	4.66	12

S_2	98	4.22	10

Tab. 2 – Comparison of some S- boxes found using our method with AES S- box and Picek's S- box.

The table above shows that Leaders and Followers method is capable of produce a better set of results with a smaller population than used in [4]. From both tests made, 90% of the population is better than [4] and 20% of each sample shows improvement to the value of the differential uniformity.

CONCLUSIONS AND FUTURE WORK

In this work, we consider the study of the meta- heuristic technique known as Leaders and Followers applied to the search of 8x8 bit S- boxes with good values of confusion coefficient variance and non- linearity. We check that for small population sizes plus the new trial creation function (simpler than a crossover) the method is able to produce a set of good results in a relative short period. We hope on some modifications of the fitness function used in this paper lead to better results. Our future task is to continue the research of using Leaders and Followers for the improvement of S- boxes with the goal of reaching non- linearity values greater than 98 and make upgrades on another of their properties.

REFERENCES

[1] – Van Tilborg, H. C. and Jajodia, S. (2014). ***Encyclopedia of cryptography and security.*** Springer Science & Business Media.

[2] – Brier, E.; Clavier, C. and Olivier, F. (2004). ***Correlation power analysis with a leakage model.*** In International Workshop on Cryptographic Hardware Embedded Systems, pages 16- 29. Springer.

[3]– Stoffelen, K. (2015). ***Intrinsic side- channel analysis resistance and efficient masking.***Master's thesis, Radboud University.

[4] – Stjepan Picek, Kostas Papagiannopoulos, Barı¸s Ege, Lejla Batina, and Domagoj Jakobovic. ***Confused by confusion: Systematic evaluation of DPA resistance of various s- boxes***. In International Conference in Cryptology in India, pages 374–390. Springer, 2014.

[5] – Claude Carlet and Cunsheng Ding. ***Nonlinearities of s- boxes. Finite fields and their applications***, 13(1):121–135, 2007.

[6] – Gonzalez, Fernandez, Y., Chen, S. ***Leaders and Followers – A New Metaheuristic to Avoid the Bias of Accumulated Information***, 2015

[7] – Tamayo Vera, D. (2017). ***Algoritmos Heurísticos Híbridos para el diseño de S- cajas.***Master's thesis, Faculty of Mathematics and Computer Science, University of Havana.

[8] – Prouff, Emmanuel: ***DPA attacks and S- boxes***. In *International Workshop on Fast SoftwareEncryption*, pages 424–441. Springer, 2005.

[9] – Tesar P.: ***A New Method for Generating High Non- linearity S- Boxes***. Radioengineering - 2010. V. 19, NO. 1. - p. 23- 26.

[10] – E. Biham and A. Shamir, ***Differential Cryptanalysis of DES- like Cryptosystems***, in Proceedings of, ser. CRYPTO '90. London, UK, UK: Springer- Verlag, 1991, pp. 2- 21.

[11] – Y. Crama and P. L. Hammer, ***Boolean Models and Methods in Mathematics, Computer Science, and Engineering***, 1st ed. New York, NY, USA: Cambridge University Press, 2010.

[12] – K. Nyberg, ***Perfect Nonlinear S- Boxes***, in Advances in Cryptology - EUROCRYPT '91, Workshop on the Theory and Application of Cryptographic Techniques, Brighton, UK, April 8- 11, 1991, Proceedings, ser. Lecture Notes in Computer Science, vol. 547. Springer, 1991, pp. 378- 386.

APENDIX A – SOME S- BOXES FOUNDWITH LEADERS AND FOLLOWERS

NL = 98, CCV = 4.662, δ = 12

CC 24 2F C4 44 F0 8A FC DD 32 47 80 A8 BE 41 97
55 4D 12 99 F5 65 A4 85 E0 73 D9 CB 6A 15 B3 78
FA 27 5B 1A E4 FB 21 E9 8D 0C 7D 0B 01 A3 8C 3F
89 4C 53 CD 74 91 DB 6D 2B 75 A0 33 C7 9D C5 82
14 B5 C6 BD 5C 06 DF B4 63 6F C0 39 77 52 E6 22
E5 25 E1 26 57 BC A1 9B 62 38 5F 69 11 8B D4 B9
4A 7F 42 CA 9E 1B 96 46 D0 D3 9C ED 7C 0A 7B 5A
72 81 37 AD 70 45 C1 E3 D5 1F A5 54 4B AF 1C 2E
D1 B6 C2 93 CE 71 98 88 58 1E A6 F6 2A 0E 7E 3B
36 20 F8 61 18 59 86 FD BF 3C AE 09 D8 F7 0D AA
B0 3A F2 79 1D A2 F9 E8 AC F4 90 B8 4F 95 8E C8
FF B2 EC 03 2C E7 28 5E 6C 50 8F 23 02 DC 87 F3
A7 C9 A9 64 2D EA 83 56 B1 10 6B DA 94 35 0F EB
68 3D 16 EF BA 84 FE 66 31 EE 60 D6 F1 07 17 08
51 43 5D D2 30 49 3E 7A 9F AB 9A 19 4E DE 05 13
E2 CF 04 76 BB 48 C3 40 00 67 92 D7 6E 29 B7 34

NL = 98, CCV = 4.216, δ = 10

8B 78 BB EF BF 7E 03 CA 76 95 6E DC 6B CB B8 F1
8A 23 BD FB F3 F5 D4 65 3E C7 7C B3 ED 8F AD BC
15 9F 08 83 0A 20 C9 B4 A4 B0 85 44 40 A0 19 87
35 5C 01 04 80 11 98 D3 26 51 92 2A 64 24 2E 32
DA 5E DD 2F F2 7D C5 F8 93 97 A7 F7 D7 9D 4D 90

62 33 D5 77 F4 CD CE 9B 71 8E FF 3F F9 FD 5D 36
3C B1 58 13 4A A1 46 53 56 4E 14 21 18 00 9A 1A
0E 29 41 1C 96 0D 0F 72 2D 4C 42 05 12 0B 3A EE
16 EB 82 09 C1 94 E9 5A E1 89 17 0C 86 E0 D0 69
E8 A9 45 06 84 30 34 E7 E4 D1 50 07 38 49 C3 C2
27 9C 6F FE DF DE 6D 4B 63 66 5F B9 3D D6 55 F0
39 2B F6 7B E5 BE A8 9E D9 E3 37 BA AE 67 57 EC
D2 43 25 60 61 8C B2 C6 2C B5 28 10 A2 02 1E 6C
91 1D 31 54 81 88 A3 C4 79 6A 48 68 C0 22 52 AC
A5 8D EA E6 1F 73 D8 74 E2 1B B6 DB FA 5B AB 70
59 A6 75 CF 4F 3B 47 AA 7A CC 7F AF B7 FC C8 99