Optimality and Beyond: The Case of 4×4 S-boxes

Stjepan Picek, Barış Ege,
Kostas Papagiannopoulos, Lejla Batina
Digital Security Group - ICIS
Radboud University Nijmegen

Domagoj Jakobović
Faculty of Electrical Engineering and Computing
University of Zagreb

***Abstract*—S-boxes with better transparency order are expected to have higher side-channel resistance. For 8×8 S-boxes this is not practical, considering the costs of lookup-table implementations and deterioration of many properties like nonlinearity or delta uniformity. However, if we concentrate on the 4×4 S-box size we can observe that it is possible to obtain S-boxes with better transparency order while maintaining proper "classical" properties. To prove this, we experiment with PRINCE and PRESENT S-boxes. We use various methods and show that evolutionary algorithms are also viable in obtaining the lowest known transparency order value for the nonlinearity value of 4. Next, we show that affine transformation changes the transparency order while keeping "classical" properties intact. By using this technique, it is possible to generate optimal S-boxes with improved DPA-related properties.**

I. Introduction

Besides more traditional linear [1] and differential cryptanalysis [2], the most practical attacks today against block ciphers belong to side-channel analysis (SCA) targeting actual implementations of cryptography in software or hardware. SCA relies on the physical leakages from the actual implementation and its efficiency is much greater than the one of linear or differential cryptanalysis [3]. There are various countermeasures such as hiding and masking schemes [4] that improve an algorithm's resistance to SCA. However, those countermeasures come with a substantial increase in cost due to larger memory requirements and the decrease in performance of the implemented algorithm. Many block ciphers are implemented either as Feistel network or as Substitution-Permutation network. In both of those design principles, S-boxes or Substitution boxes, are the most commonly used nonlinear element in a block cipher.

With the consideration of side-channel security, Prouff [5] defines the transparency order property that characterizes the resistance of S-boxes to SCA or more precisely to differential power analysis (DPA) [4]. Current methods to generate S-boxes with low transparency order values suffer from the same two major drawbacks: first being the deterioration of many properties related with linear and differential cryptanalysis of the algorithm (the most important of such properties are nonlinearity and delta uniformity) and second the fact that such new improved S-boxes can only be implemented as look-up tables (LUTs). For 8×8 S-boxes, such as the AES S-box, look up table (LUT) implementation is quite costly in terms of area. However, when 4×4 S-boxes are considered, LUTs are commonly used to implement these S-boxes. In addition, LUT implementations have to be protected against cache attacks [6].

Therefore, in this paper we concentrate on the bijective 4×4 S-boxes and their transparency orders. Furthermore, we are interested only in those S-boxes that retain the same nonlinearity level as the one in optimal S-boxes [7]. Leander and Poschmann defined optimal S-boxes as those of linearity 8, but since this is the same as having nonlinearity 4 we will continue to use nonlinearity property proposed in [8], instead of linearity. Finally, since many modern, widely-used 4×4 ciphers have S-boxes implemented as lookup tables, that also solves the second drawback as stated above.

A. Related Work

Previous works can be divided into several categories: first one relates to implementations of 4×4 S-boxes and in the second one several examples of transparency order property were investigated.

4×4 S-boxes. There is a plethora of cryptographic algorithms used today that have 4×4 S-boxes. In this paper we focus mainly on PRINCE algorithm [9] and PRESENT [10] algorithms. Leander and Poschmann classify all optimal 4×4 S-boxes [7]. Saarinen conducts exhaustive search of all bijective 4×4 S-boxes [11].

Transparency order. Mazumdar et al. construct rotation symmetric S-boxes with high nonlinearity and DPA resistance [12]. The same authors use constrained random search to find S-boxes with low transparency order and high nonlinearity [13]. Implementations of those S-boxes suggest that there is a significant increase in the necessary number of traces to perform DPA attack. Picek et al. use evolutionary algorithms to evolve 8×8 S-boxes with good transparency order values [14]. Anyhow, all investigations of transparency order property up to now were related with 8×8 S-boxes and our work is the first one to investigate the influence of transparency order property in 4×4 S-boxes. More details on our contributions are given below.

B. Our Contributions

Our first contribution is that, to the best of our knowledge, we are the first to investigate the influence of transparency order on 4×4 S-boxes. Furthermore, we use random search and genetic algorithms to find S-boxes with transparency order lower than the ones that can be usually found in modern ciphers. Next, we show that optimal S-boxes have different transparency order values, which implies that one should be careful even when searching among optimal ones. Third

This work was supported in part by the Technology Foundation STW (project 12624 - SIDES), The Netherlands Organization for Scientific Research NWO (project ProFIL 628.001.007) and the ICT COST action IC1204 TRUDEVICE.

978-1-4799-4112-4/14/$31.00 ©2014 IEEE

contribution is that we show that affine equivalent S-boxes can have different transparency order, which should be an important criteria in choosing proper S-boxes. Finally, we give a power analysis of a newly created S-box replacing the PRESENT S-box on a smartcard.

The remainder of this paper is organized as follows: In Section II we survey necessary information about cryptographic properties of S-boxes. Different S-boxes and their transparency orders are compared in Section III. Furthermore, we compare the results on transparency order obtained via random search and genetic algorithms. We also discuss implementations of the newly generated S-boxes and also their resistance to differential power analysis. Finally, in Section IV we conclude the paper.

II. Cryptographic Properties of S-boxes

Here we give necessary information about cryptographic properties of S-box that are of interest for this research, namely bijectivity, (non)linearity and δ-uniformity [7]. Besides those three properties, we also present a property that is related with DPA resistivity: transparency order (T_F).

The addition modulo 2 is denoted as " $\oplus$ ". The inner product of vectors $\bar{a}$ and $\bar{b}$ is denoted as $\bar{a}\cdot\bar{b}$ and equals $\bar{a}\cdot\bar{b} = \oplus_{i=1}^{n} a_i b_i$.
Function *F*, called S-box or vectorial Boolean function, of size (n, m) is defined as any mapping *F* from $\mathbb{F}_2^n$ to $\mathbb{F}_2^m$ [5]. When m equals 1 the function is called Boolean function. Boolean functions f_i, where $i \in \{1, ..., m\}$ are coordinate functions of *F* where every Boolean function has n variables. Hamming weight (HW) of a vector $\bar{a}$, where $\bar{a} \in \mathbb{F}_2^n$, is the number of non-zero positions in the vector.

An (n, m)-function is called balanced if it takes every value of $\mathbb{F}_2^m$ the same number 2^{n-m} of times [15]. Balanced (n, n)-functions are permutations on $\mathbb{F}_2^n$.

Nonlinearity N_F of an (n, m)-function *F* is equal to the minimum nonlinearity of all non-zero linear combinations $\bar{b}\cdot F$ of its coordinate functions f_i, where $\bar{b} \in \mathbb{F}_2^{n*}$ [3].

$$N_F = 2^{n-1} - \frac{1}{2} max_{\substack{\bar{a}\in\mathbb{F}_2^n \\ \bar{v}\in\mathbb{F}_2^{m*}}} |W_F(\bar{a}, \bar{v})|. \quad (1)$$

Here, $W_F(\bar{a}, \bar{v})$ represents Walsh transform of F [5].

$$W_F(\bar{a}, \bar{v}) = \sum_{\bar{x}\in\mathbb{F}_2^n} (-1)^{\bar{v}\cdot F(\bar{x})\oplus\bar{a}\cdot\bar{x}}. \quad (2)$$

Differential delta uniformity δ represents the largest value in the difference distribution table without counting the value 2^n in the first row and column position [2].

In 2005, Prouff introduced a new cryptographic property of S-boxes: transparency order [5] which can be defined for a (n, m)-function as follows.

$$T_F = max_{\bar{\beta}\in\mathbb{F}_2^m}(|m - 2HW(\bar{\beta})| - \frac{1}{2^{2n} - 2^n} \sum_{\bar{a}\in\mathbb{F}_2^{n*}} | \sum_{\substack{\bar{v}\in\mathbb{F}_2^m \\ HW(\bar{v})=1}} (-1)^{\bar{v}\cdot\bar{\beta}} W_{D_aF}(\bar{0}, \bar{v})|). \quad (3)$$

Here, W_{D_aF} represents Walsh transform of the derivative of F with respect to a vector $a \in \mathbb{F}_2^n$.

The higher the transparency order value is, the lower resistance an S-box exhibits to DPA. The worst transparency order is achieved when all coordinate function are f_i bent functions, and the best transparency order is achieved when *F* is an affine function [5]. In any (n, m)-function *F* transparency order is upper bounded by m. Carlet shows that some widely-used 8×8 S-boxes with very high nonlinearity have bad transparency orders [3].

III. S-Boxes and Transparency order

In this section we search for good S-boxes that have low transparency order values. By good S-boxes we mean primarily all optimal S-boxes [7]. In an effort to find S-boxes that have even lower transparency order values we also conduct experiments with random search and genetic algorithms.

A. Optimal S-boxes

There are in total 16! bijective 4×4 S-boxes which is approximately 2^{44} options to search from. Leander and Poschmann define optimal S-boxes as those that are bijective, have linearity equal to 8 and δ-uniformity equal to 4. By using some clever shortcuts they found that all optimal S-boxes belong to 16 classes, i.e. all optimal S-boxes are affine equivalent to one of those classes [7]. For two S-boxes S_1 and S_2 to be equivalent, following equation needs to hold:

$$S_2(x) = B(S_1(A(x) + a)) + b \quad (4)$$

where A and B are invertible 4×4 matrices and $a, b \in \mathbb{F}_2^4$.

Transparency order values for 16 class representatives are given in Table I and we can see there is a difference between some of the values. For the transparency order, there are 3 values that are seen for the class representatives. This can immediately suggest that there is a difference in the DPA resistance between optimal S-boxes.

TABLE I. Transparency orders of optimal S-boxes class representatives

Optimal S-box	PRINCE notation	T_F
$G_0, G_4, G_5, G_6, G_8, G_{10}, G_{14}$	S_1, S_2, S_3	3.467
$G_1, G_3, G_9, G_{11}, G_{15}$	S_0, S_5	3.533
G_2, G_{12}, G_{13}	S_6, S_7	3.6

B. PRINCE Suitable S-boxes

Out of the 16 optimal classes, authors of PRINCE state that only 8 are suitable for PRINCE [9]. Out of the 8 representatives, there are only 3 with the best transparency order value of 3.467. However, on the basis of PRINCE S-boxes, we can reach one other, far more interesting fact.

In the PRINCE algorithm, authors chose class S_7 (or class G_{13} of optimal S-boxes) to derive the actual S-box. Class representatives are not suitable for the actual implementations since they are chosen according to the lexicographical order. Therefore, class representatives have fixed points, which is not acceptable for implementations. However, since an affine

TABLE II. AFFINE TRANSFORMATIONS

Number	Transformation
1	S(x) + c
2	S(B(x) + c)
3	(A(S(B(x) + c)) + d
4	(A(S(B(x) + c) + d)

transformation does not change linear and differential properties of S-boxes [7] one can apply that transformation to the class representative to remove fixed points and retain all good properties of optimal S-boxes. Therefore PRINCE uses an affine transformation to obtain the S-box that is used in the specification [9].

After running the analysis on that S-box, we can see that traditional differential and linear properties remain unchanged but, transparency order value is 3.4. It should be noted that the class representative has a transparency order value of 3.6. Based on this fact we can reach two conclusions. First one is that it matters (in terms of side-channel resistance) from which class an S-box is chosen from. Second one is that affine transformations change the transparency order, therefore changing DPA resistance of an algorithm using a certain S-box.

C. Is Affine Equivalence Also a DPA Equivalence

Based on the results of previous section, we give the following conjecture.

Conjecture 1: S-boxes that are affine equivalent (and therefore equivalent with respect to linear and differential properties) are not equivalent in terms of DPA resistance.

To experimentally support that the conjecture holds, it is enough to compare the transparency order values for the PRINCE case. Since we establish that an affine transformation can change transparency order values, next question is how difficult is to change it. To do this, we experiment with four affine transformations as listed in Table II.

Here, $c, d \in \mathbb{F}_2^4$ are constants, $+$ represents XOR operation and A and B are invertible matrices. For each of 8 class representatives we test with all 4 affine transformations for 10^6 times which gives in total 32 million generated S-boxes. Only when applying at least two multiplications with invertible matrices it results in the change of transparency order values. Therefore, only transformations 3 and 4 change transparency order where those values go from 3.4 to 3.73. If we check remaining 8 classes that are optimal, but not a good choice for PRINCE algorithm, we find S-boxes that have transparency order from 3.2 to 3.73 with affine transformations 3 and 4.

D. Random search

To better understand possible levels of transparency order values, we relax the criteria that the S-boxes need to satisfy and we look into all bijective 4×4 S-boxes. As a method to generate such S-boxes we employ random search (RS). The distribution of the random S-boxes values is shown in Table III.

We can observe from the results that it is possible to use random search to find S-boxes that have lower transparency orders than the ones in the class representatives. It is also important to observe that some of the best random S-boxes have transparency order of 3.26 and δ-uniformity and nonlinearity equal to 4, which means that those S-boxes belong to some of 16 optimal classes of S-boxes.

TABLE III. DISTRIBUTION OF RANDOM S-BOXES VALUES

Property	Max	Min	Mean	Std. dev.
Nonlinearity	4	0	2.105	0.694
Transparency order	3.73	3	3.467	0.099

E. Genetic Algorithm

In the next step of the search we tried to go even lower with the transparency order values but remain on the nonlinearity level of 4 (so, here we relax δ-uniformity criterion). To be able to do that, we use a genetic algorithm (GA) where the goal was to evolve balanced bijective S-boxes with high nonlinearity and low transparency order. Our fitness function is the sum of nonlinearity (N_F) and transparency order (T_F) properties. Since the transparency order value should be as low as possible, we subtract the value obtained from the upper bound value for transparency order. The fitness function represents definition of the problem to solve with evolutionary algorithm:

$$fitness = N_F + (m - T_F) \tag{5}$$

where we want to maximize the fitness function value. For a detailed explanation about genetic algorithms we refer reader to [16].

To represent the problem we use a permutation representation where an S-box is represented with decimal values between 0 and 15 where each of those values is one entry for S-box lookup table. Parameters for the evolutionary algorithm are as follows: the size of (n, m)-function is 4×4, number of independent runs for each evolutionary experiment is 30 and the population size is 100. Tournament size in steady-state tournament selection is equal to 3. Mutation probability is set to 0.3 per individual. The evolution process lasts until the stopping criterion is fulfilled, here the stopping criterion is a certain number of generations without improvement of the best solution.

The best solution in respect to the transparency order that we were able to find with GA has transparency order value of 3.2 and nonlinearity of 4, but with δ-uniformity of 8. GA also found solutions with transparency order of 3.26 and nonlinearity and δ-uniformity 4 (as in random search) but with GA we also succeeded in further lowering the transparency order value. Although this improvement in transparency order is relatively small, it is still worth mentioning that it is possible to obtain better results than with random search. The improvement also came at a price since δ-uniformity property deteriorated. After adding the δ-uniformity property to the fitness equation, we were able to find S-boxes that have nonlinearity and δ-uniformity 4 and transparency order 3.2 which is the same as with affine transformations.

In Table IV we give hexadecimal values for the best created S-boxes for random search and genetic algorithms.

TABLE IV. S-BOXES OBTAINED WITH RS AND GA

Alg.	x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
RS	S(x)	C	A	1	9	B	8	0	5	E	2	7	D	3	6	4	F
GA	S(x)	6	5	9	D	C	0	7	F	A	B	2	4	E	3	1	8

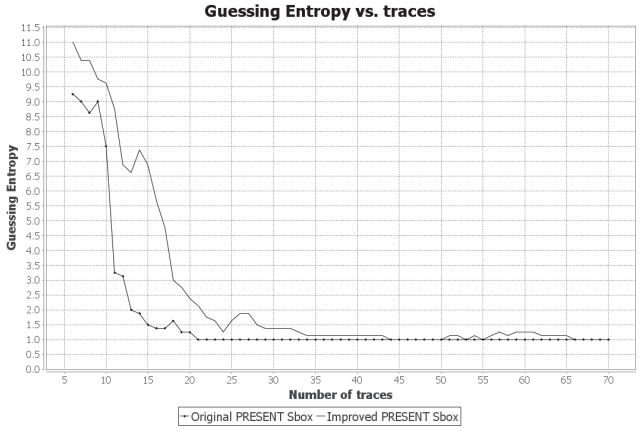

Fig. 1. Guessing entropy for the original and improved S-box.

F. Preliminary Side Channel Experiments

We implemented two versions of the PRESENT cipher on an ATmega163 smartcard, one using the improved "S-box 2" and another with the original PRESENT S-box. We acquired traces and performed CPA on the first cipher round in order to recover the 64-bit round key. To quantify the effect of a different S-box we used the guessing entropy metric, as defined by Standaert et al. [17]. Specifically, we focused on the top 10 key-byte candidates of the CPA attack and we measured the success order. In order to observe the S-box effect on a *full-round* attack, we averaged the guessing entropy of the 8 key bytes that constitute the PRESENT 64-bit round key. The averaged entropy demonstrates the workload required for round-key recovery, given the number of traces. The results are presented in Figure 1.

IV. CONCLUSION

In this work we consider the DPA resistance properties of 4×4 S-boxes. We show that there is a difference between 16 optimal classes of S-boxes in terms of properties related with DPA resistance. We believe this can serve as a more strict guideline when choosing between some of the optimal S-box classes. Furthermore, we show that an affine transformation changes the transparency order values, which can be important not only from theoretical perspective, but also from the practical one. We reiterate that with 4×4 S-boxes two main disadvantages when comparing with DPA resistant 8×8 S-boxes are not present. There is no deterioration of linear and differential properties nor it is impractical to implement them as lookup tables. Finally, our experiments show a practical improvement of DPA resistance of the newly generated S-box. Since the security does not degrade with using S-boxes with improved transparency order, we consider them as a good choice for future algorithms.

ACKNOWLEDGEMENTS

Authors would like to thank Gregor Leander from Ruhr University Bochum for advice and answers to our numerous questions.

REFERENCES

[1] M. Matsui and A. Yamagishi, "A new method for known plaintext attack of FEAL cipher," in *Proceedings of*, ser. EUROCRYPT'92. Berlin, Heidelberg: Springer-Verlag, 1993, pp. 81–91.

[2] 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.

[3] C. Carlet, "On highly nonlinear S-boxes and their inability to thwart DPA attacks," in *Proceedings of*, ser. INDOCRYPT'05. Berlin, Heidelberg: Springer-Verlag, 2005, pp. 49–62.

[4] S. Mangard, E. Oswald, and T. Popp, *Power Analysis Attacks: Revealing the Secrets of Smart Cards (Advances in Information Security)*. Secaucus, NJ, US1A: Springer-Verlag New York, Inc., 2007.

[5] E. Prouff, "DPA Attacks and S-Boxes," in *12th International Workshop, FSE 2005, Paris, France*, ser. Lecture Notes in Computer Science, vol. 3557. Springer, 2005, pp. 424–441.

[6] D. J. Bernstein, "Cache-timing attacks on AES," 2004. [Online]. Available: http://cr.yp.to/papers.html#cachetiming

[7] G. Leander and A. Poschmann, "On the Classification of 4 Bit S-Boxes," in *Arithmetic of Finite Fields*, ser. Lecture Notes in Computer Science, C. Carlet and B. Sunar, Eds. Springer Berlin Heidelberg, 2007, vol. 4547, pp. 159–176.

[8] A. Braeken, "Cryptographic Properties of Boolean Functions and S-Boxes," Ph.D. dissertation, Katholieke Universiteit Leuven, 2006.

[9] J. Borghoff, A. Canteaut, T. Gneysu, E. Kavun, M. Knezevic, L. Knudsen, G. Leander, V. Nikov, C. Paar, C. Rechberger, P. Rombouts, S. Thomsen, and T. Yaln, "PRINCE : A Low-Latency Block Cipher for Pervasive Computing Applications," in *Advances in Cryptology: ASIACRYPT 2012*, ser. Lecture Notes in Computer Science, X. Wang and K. Sako, Eds. Springer Berlin Heidelberg, 2012, vol. 7658, pp. 208–225.

[10] A. Bogdanov, L. R. Knudsen, G. Leander, C. Paar, A. Poschmann, M. J. Robshaw, Y. Seurin, and C. Vikkelsoe, "PRESENT: An Ultra-Lightweight Block Cipher," in *Proceedings of*, ser. CHES '07. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 450–466.

[11] M.-J. Saarinen, "Cryptographic Analysis of All 4 x 4-Bit S-Boxes," in *Selected Areas in Cryptography*, ser. Lecture Notes in Computer Science, A. Miri and S. Vaudenay, Eds. Springer Berlin Heidelberg, 2012, vol. 7118, pp. 118–133.

[12] B. Mazumdar, D. Mukhopadhyay, and I. Sengupta, "Design and implementation of rotation symmetric S-boxes with high nonlinearity and high DPA resilience," in *Hardware-Oriented Security and Trust (HOST), 2013 IEEE International Symposium on*, 2013, pp. 87–92.

[13] B. Mazumdar, D. Mukhopadhay, and I. Sengupta, "Constrained Search for a Class of Good Bijective S-Boxes with Improved DPA Resistivity," *Information Forensics and Security, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2013.

[14] S. Picek, B. Ege, L. Batina, D. Jakobovic, L. Chmielewski, and M. Golub, "On Using Genetic Algorithms for Intrinsic Side-channel Resistance: The Case of AES S-box," in *Proceedings of the First Workshop on Cryptography and Security in Computing Systems*, ser. CS2 '14. New York, NY, USA: ACM, 2014, pp. 13–18.

[15] 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.

[16] A. E. Eiben and J. E. Smith, *Introduction to Evolutionary Computing*. Springer-Verlag, Berlin Heidelberg New York, USA, 2003.

[17] F.-X. Standaert, T. G. Malkin, and M. Yung, "A unified framework for the analysis of side-channel key recovery attacks," in *Proceedings of the 28th Annual International Conference on Advances in Cryptology: The Theory and Applications of Cryptographic Techniques*, ser. EUROCRYPT '09. Berlin, Heidelberg: Springer-Verlag, 2009, pp. 443–461.