Cryptographic uncertainness: some experiments on finite semifield based substitution boxes

I.F. Rúa[1] and E.F. Combarro[2]

Abstract Substitution boxes (S-boxes) are an important part of the design of block ciphers. They provide nonlinearity and so the security of the cipher depends strongly on them. Some block ciphers use S-boxes given by lookup tables (e.g., DES) where as others use S-boxes obtained from finite field operations (e.g., AES). As a generalization of the latter, finite semifields (i.e., finite nonassociative division rings) have been suggested as algebraic structures from which S-boxes with good cryptographic properties might be obtained. In this paper we present the results of experiments on the construction of S-boxes from finite semifields of orders 256 and 64, using the left and right inverses of these rings.
Keywords: Substitution box, S-box, block cipher, finite semifield

1 Introduction

> [...] a new science, called *Criptology*, arises. It has a field devoted to encryption (*Cryptography*) an another one to decryption (*Cryptanalysis*). Its origins are as old as humanity: remember the writing on a strip of parchment wrapped around a staff or Lacedaemonian 'scytale'; or the Caesar cipher consisting on a constant shifting of the letters of the alphabet.

These words are part of the opening lecture of the accademic year 1996-97 delivered by Pedro Gil at University of Oviedo [7]. The lecture, which was titled "The Mathematics of the uncertain", had a first part devoted to randomness, Probability and Statistics. The second part dealt with Information Theory and the mathematics of communication (it even has a third and final

Departamento de Matemáticas, Universidad de Oviedo, C/ Calvo Sotelo s/n 33007, +34 98510-3344 (Fax: 3352) Oviedo rua@uniovi.es · Departamento de Informática, Universidad de Oviedo, C/ Calvo Sotelo s/n 33007, +34 98510-9558 (Fax: 3382) Oviedo elias@aic.uniovi.es

part dedicated to fuzzy sets). It is difficult to understand modern Cryptography without a probabilistic point of view [8]. The first author to systematize this approach was Claude Shannon, the *father* of Information Theory. Apart from introducing the concepts of *entropy* and *information* in the context of communication in noiseless and noisy channels [19] (just as mentioned in Pedro Gil's lecture[1]), he considered a probabilistic model of *perfect secrecy* [20]. Following this idea, semantic security (which is, from a certain point of view, the theoretical notion of a secure cryptographic system) is founded on a probabilistic setting [9].

Block ciphers (which transform a block of bits of fixed size into another block of the same size with the help of a bit-key of also fixed, perhaps different, size) are a symmetric (i.e., private) key cryptographic primitive used in many other designs (e.g., cryptosystems, message authentication codes, hash functions,...) [13]. *Substitution boxes* (called *S-boxes*) are an important part of the design of block ciphers. They provide nonlinearity to the transformation and so the security of the cipher depends strongly on them. Some block ciphers use S-boxes given by lookup tables (e.g., DES) where as others use S-boxes obtained from finite field operations (e.g., AES) [21]. As a generalization of the latter, finite semifields (i.e., finite nonassociative division rings) have been suggested as algebraic structures from which S-boxes with good cryptographic properties might be obtained [5]. This is not the first time that nonassociative structures have been considered in a cryptographic setting (just recall, for instance, [6, 15, 22, 10]).

In this paper, following the path of [5], we present the results of experiments on the construction of S-boxes from finite semifields of order 256, using the left and right inverses of these rings. We process all finite semifields of such an order and rank 4 (and not only the 28 representatives up to isotopy considered in [5, Section 5.3]), and also all finite semifields of dimension 6 over $\mathbb{F}_2$ (as this is the biggest dimension for which all finite semifields of characteristic 2 have been classified). The paper is organized as follows: in section 2 basic notions of block ciphers (including properties of S-boxes) are reminded. Section 3 is devoted to finite semifields and their properties. Finally, in the last section we collect the results obtained from our computational experiments.

2 Block ciphers and substitution boxes

A *block cipher* is a deterministic cipher $E : \{0,1\}^b \times \{0,1\}^k \to \{0,1\}^b$ which transforms a block M of b bits of fixed size into another block C of the same size with the help of a key K of also fixed, perhaps different, size k [13].

[1] Incidentally, let us mention that we had the privilege of learning the basic aspects of Probability, Statistics and Information theory from Pedro himself, in two courses delivered at University of Oviedo some twenty years ago.

Well-known examples of block ciphers include the previous and the current NIST standards for encryption data: DES and AES [21]. For instance, in DES $b = 64, d = 56$, wheras in AES $b = 256, d \in \{128, 196, 256\}$. These ciphers are of utmost importance because, as pointed out in [2],

> Block ciphers are the "work horse" of practical cryptography: not only can they can be used to build a stream cipher, but they can be used to build ciphers with stronger security properties [...], as well as many other cryptographic primitives.

A common design of block ciphers is that of iterated ciphers, where a round function is used repeatedly r times to process the block of bits M using a set of round keys obtained from the master key K with the help of an auxiliary key schedule algorithm (e.g., in DES $r = 14$, in AES $r \in \{10, 12, 14\}$). In these ciphers, the ultimate transformation of the block M depends on the round function F. Traditionally, the function F can be of Feistel type (such as in DES) or a Substitution-Permutation Network (such as in AES) [21]. In either case, both use substitution boxes in the design of F.

A *substitution box* (called *S-box*) is a fixed boolean function $S : \{0,1\}^n \to \{0,1\}^m$, where the parameters n, m depend on the actual cipher considered (for instance, in DES $n = 6, m = 4$, in AES $n = m = 8$). S-boxes are a core part of the design of block ciphers as they provide nonlinearity to the transformation. The security of the cipher (e.g., robustness against differential or linear attacks) depends strongly on them. Some block ciphers use S-boxes given by lookup tables (e.g., DES) where as others use S-boxes obtained from finite field operations [21]. For instance, AES S-boxes identify the set $\{0,1\}^8$ with the Galois field $\mathbb{F}_{2^8}$ of 256 elements (multiplication is taken modulo the polynomial $x^8+x^4+x^3+x+1$) and before applying an $\mathbb{F}_2-$affine transformation, the input element is changed into its multiplicative inverse in $\mathbb{F}_{2^8}$ (the zero element is replicated).

Different properties of an S-box can be introduced in order to determine its cryptographic utility, and so multiple criteria can be found in the literature (e.g., [14, 18]). In this paper we study properties #1, #3 and #4 in [5] for S-boxes of sizes 256 and 64. Namely, we identify the sets $\{0,1\}^8$ and $\{0,1\}^6$ with $\mathbb{F}_2^8$ and $\mathbb{F}_2^6$, and consider

1. Bijectivity: $n = m = 8$ (alt. $n = m = 6$), and the S-box must be bijective.
3. Non-linearity: the linear invariant λ_S is defined as

$$\lambda_S = \max\{\ |-2^{n-1} + \#\{x \in \mathbb{F}_2^n\ :\ (a|x) = (b|S(x))\}|\ :\ a, b \in \mathbb{F}_2^n, b \neq \mathbf{0}\}$$

   where $(a|x)$ denotes the usual inner product in $\mathbb{F}_2^n$, $n = 8$ (alt. $n = 6$).
4. The differential invariant δ_S is equal to

$$\delta_S = \max\{\ \#\{x \in \mathbb{F}_2^n\ :\ S(x) \oplus S(a \oplus x) = b\}\ :\ a, b \in \mathbb{F}_2^n, a \neq \mathbf{0}\}$$

   where $a \oplus x$ denotes bitwise addition mod 2, and $n = 8$ (alt. $n = 6$).

With respect to these properties AES S-boxes are optimal in the sense that they are bijective, have minimal non-linearity $\lambda_{AES} = 16$, and minimal differential invariant $\delta_{AES} = 4$ among non-APN functions [5]. Also, $\lambda_{\mathbb{F}_{64}} = 8$ and $\delta_{\mathbb{F}_{64}} = 4$.

3 Finite semifields

In this section we collect definitions and facts on finite semifields [11, 4]. A finite nonassociative ring D is called *finite semifield*, if the set of nonzero elements D^* is closed under the product, and it has an identity element. In such a case D^* is a multiplicative loop. That is, there exists an element $e \in D^*$ (the identity of D) such that $ex = xe = x$, for all $x \in D$ and, for all $a, b \in D^*$, the equation $ax = b$ (resp. $xa = b$) has a unique solution. Let us emphasize that these *left* and *right inverses* might be different elements of the finite semifield. This is an important fact apparently obviated in [5] and [6].

Finite semifields are nonassociative finite division rings and, apart from finite fields, *proper* finite semifields exist. The characteristic of a finite semifield D is a prime number p, and D is a finite-dimensional algebra over $\mathbb{F}_q$ ($q = p^c$) of dimension d, for some $c, d \in \mathbb{N}$, so that the order of D is $|D| = q^d$. Moreover, $\mathbb{F}_q$ can be chosen to be contained in the associative-commutative center $Z(D)$ of D. In this paper we will be interested in finite semifields of order 256, i.e., of dimension 8 over its center $Z(D) = \mathbb{F}_2$ or of rank 4 (i.e., of dimension 4 over $\mathbb{F}_4 \subseteq Z(D)$). The finite field $\mathbb{F}_{256}$ is included in the latter case. Also, we will be interested in semifields of order 64, i.e., 8-dimensional over $\mathbb{F}_{64}$. E.g., the Galois field $\mathbb{F}_{2^8}$.

Isomorphism of finite semifields is defined as usual for algebras, and the classification of finite semifields up to isomorphism can be naturally considered. Because of the connections to finite geometries [1], the following notion must be considered. An *isotopy* between two finite semifields D_1 and D_2 is a triple (F, G, H) of bijective $\mathbb{F}_q-$linear maps $D_1 \to D_2$ such that $H(ab) = F(a)G(b)$, for all $a, b \in D_1$. Clearly, any isomorphism between two semifields is an isotopy, but the converse is not necessarily true. From any finite semifield D, a projective plane $\mathcal{P}(D)$ can be constructed [11]. Theorem 6 in [1] shows that isotopy of finite semifields is the algebraic translation of the isomorphism between the corresponding projective planes.

By [11][Theorem 5.2.1], up to six projective planes can be constructed from a given finite semifield D using the transformations of the group S_3. Actually, S_3 acts on the set of semifield planes of a given order producing, for each semifield D, its *Knuth orbit* [11]. So, the classification of finite semifields can be reduced to the classification of the corresponding Knuth orbits.

In the particular case of semifields of order 256 and rank 4, i.e., with center containing $\mathbb{F}_4$, a computer-assisted classification was presented in [3].

A total amount of 28 Knuth classes were obtained. The actual number of semifields is much bigger. Namely, the number of isotopy classes is 51 and the number of nonisomorphic finite semifields containing $\mathbb{F}_4$ is 75939 (these numbers were obtained with the techniques describe in [3]). Unfortunately, a complete classification of finite semifields of order 256 has not been achieved (not even of order 128 [17]). Moreover, it is even unknown how many of them might there exist (the number must be clearly much bigger than those 75939 containing $\mathbb{F}_4$ in the center).

The biggest dimension for which all finite semifields of characteristic 2 have been classified is 6 [16]. There are 80 Knuth orbits of such an order containing 322 isotopy classes for a total amount of 376971 semifields.

4 Some experiments on finite semifield based S-boxes

Inspired by the S-boxes of AES, the authors propose in [5] the construction of S-boxes from the multiplicative structure of finite semifields. Namely, they suggest "using the inverse function" [5, Section 5.3]. As was noticed in the previous section, a distinction between left and right inverse is needed when dealing with (noncommutative) finite semifields. So, given a finite semifield D of order 256 (alt. 64) and identity e, we have considered the two following S-boxes:

$$\begin{array}{rl} S_r: & D \to D \\ & a \neq 0 \to b \text{ s.t. } ab = e \\ & 0 \to 0 \end{array} \qquad \begin{array}{rl} S_l: & D \to D \\ & a \neq 0 \to b \text{ s.t. } ba = e \\ & 0 \to 0 \end{array}$$

It is clear that, when D is commutative (in particular, if S is the Galois field $\mathbb{F}_{2^8}$), both S-boxes coincide. It is also evident that, because D is a finite semifield, the bijectivity property holds in both cases. In order to compute the linear $\lambda_{S_r}, \lambda_{S_l}$ and differential $\delta_{S_r}, \delta_{S_l}$ invariants we identify the elements of D with those of the set $\mathbb{F}_2^8$ (alt. $\mathbb{F}_2^6$). This can be straightforwardly done as the representation of finite semifields introduced in [3] is exactly that one. Moreover, in Table 2 of such a paper it is contained a complete description of all finite semifields of order 256 and rank 4, i.e., and center containing the finite field $\mathbb{F}_4$ [3, Section 4.2], up to Knuth orbit. These are the semifields also considered in [5, Section 5.3], where it is claimed that

> We thus have also tried to construct S-boxes based on all these 28 semifields up to isotopy, by using the inverse function.

It appears that the authors have only consider the 28 representatives in their construction and at most one of the two possible "inverse function". As it was said in the previous section, up to isomorphy, the actual number of finite semifields or order 256 with center containing $\mathbb{F}_4$ is much bigger. So, we have used the computational machinery described in [3, Section 3] to generate all

those finite semifields. For each one of them, we have explicitly constructed the aforementioned S-boxes S_r and S_l. Since we are interested in S-boxes with "good" cryptographic properties, we have taken as a reference the invariants for the AES S-box ($\lambda_{AES} = 16, \delta_{AES} = 4$). Let us remark the following fact.

Proposition 1. *$\lambda_{S_l} = \lambda_{S_r}$, for any finite semifield D of order 2^n.*

Proof. For all $x, y \in D$, we have that $y = S_l(x)$ iff $x = S_r(x)$. Therefore, for all nonzero $a, b \in \mathbb{F}_2^n$:

$$\#\{x \in \mathbb{F}_2^n \ : \ (a|x) = (b|S_l(x))\} = \#\{y \in \mathbb{F}_2^n \ : \ (b|y) = (a|S_r(y))\}$$

On the other hand, since $(\mathbf{0}|x) = 0$, for all $x \in \mathbb{F}_2^n$, and because the maps S_l and S_r are bijections:

$$\#\{x \in \mathbb{F}_2^n \ : \ 0 = (c|S_l(x))\} = \#\{y \in \mathbb{F}_2^n \ : \ 0 = (c|S_r(y))\}$$

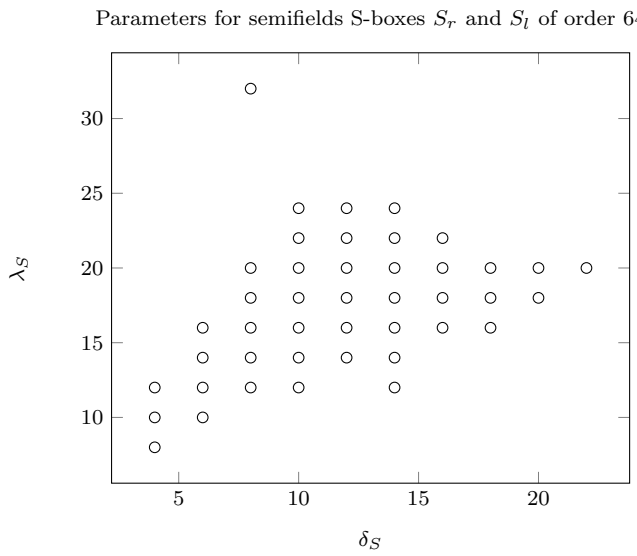
for all $\mathbf{0} \neq c \in \mathbb{F}_2^n$. Hence,

$$\lambda_{S_l} = \max\{ \ | - 2^{n-1} + \#\{x \in \mathbb{F}_2^n \ : \ (a|x) = (b|S_l(x))\}| \ : \ a, b \in \mathbb{F}_2^n, b \neq \mathbf{0}\}$$

$$= \max\{ \ | - 2^{n-1} + \#\{y \in \mathbb{F}_2^n \ : \ (b|y) = (a|S_r(x))\}| \ : \ a, b \in \mathbb{F}_2^n, a \neq \mathbf{0}\} = \lambda_{S_r}$$

Our computations show that none of the generated S-boxes had a pair of invariants matching those of the finite field $\mathbb{F}_{2^8}$. So, no S-box with "good" cryptographic properties was obtained from the constructions S_r or S_l on semifields of order 256 containing $\mathbb{F}_4$ in the center. Let us mention, for the record, that the linear and differential parameters might be different for isotopic non-isomorphic finite semifields. This means that these parameters are not isotopy invariants, such as the center or nuclei sizes [12]. So, for instance, a full computation of the linear and differential parameters for finite semifields isotopic to Semifield #II of [3, Table 2], shows that we can find parameters $(\lambda_{S_r}, \delta_{S_r}) = (38, 12), (38, 14), (36, 10), (34, 10), \ldots$.

The construction of S-boxes S_r and S_l was also applied to all finite semifields of order 64. Remember that the parameters of the finite field of such an order are $\lambda_{\mathbb{F}_{64}} = 8$ and $\delta_{\mathbb{F}_{64}} = 4$. The computational results show that there are some proper semifields with $\delta_{S_l} = 4$. Namely, semifields falling in Knuth orbits #IV,V,VIII,X. Among these, only 6 proper semifields in Knuth orbit #V share the pair $(\lambda_{S_r}, \delta_{S_r}) = (8, 4)$ with the finite field $\mathbb{F}_{64}$. We have plotted in the following graph all pairs $(\lambda_{S_r}, \delta_{S_r})$ and $(\lambda_{S_l}, \delta_{S_l})$ found in our study.



We finish this short note by showing one of the S-boxes with the same parameters of the finite field S-box $S_{\mathbb{F}_{64}}$, but constructed from left inverses in a finite semifield of order 64.

S_l	0	1	2	3	4	5	6	7
0	00	40	73	24	30	45	62	27
1	41	70	55	47	05	03	46	32
2	15	37	31	11	17	66	74	06
3	72	34	57	02	10	35	14	64
4	44	77	43	67	71	36	53	25
5	20	21	13	56	33	54	01	61
6	60	50	26	12	75	16	76	65
7	23	52	51	22	42	63	07	04

Table 1 An S-Box with minimal linear and differential parameters (constructed from a proper semifield of order 64)

Conclusion

We have explicitly constructed S-boxes from proper finite semifields of orders 256 and 64, and computed their linear and differential parameters. The results in the case of 256 elements are not satisfactory, since none of these S-boxes have the same minimal invariants as those of the AES S-box. This is not surprising since only rank 4 semifields of such an order where analyzed, as this was the only subclass of semifields of order 256 for which a complete

classification has been achieved so far. On the other hand, the case of order 64 semifields (for which a full classification is known) is more promising. Some S-boxes have been constructed with the same parameters of those obtained from the Galois field $\mathbb{F}_{64}$.

Acknowledgements I.F. Rúa is partially supported by MINECO-13-MTM2013-45588-C3-1-P, and Principado de Asturias Grant GRUPIN14-142. E.F. Combarro is partially supported by MINECO-16-TEC2015-67387-C4-3-R

References

1. Albert, A.A. (1960) Finite division algebras and finite planes. In: Proceedings of Symposia in Applied Mathematics 10:53–70.
2. Boneh, D. and Shoup, V. (2015) A Graduate Course in Applied Cryptography. `https://crypto.stanford.edu/~dabo/cryptobook/draft_0_2.pdf`
3. Combarro, E.F. and Rúa, I.F. and Ranilla, J. (2011) New advances in the computational exploration of semifields. International Journal of Computer Mathematics 88(9): 1990–2000.
4. Cordero, M. and Wene, G.P. (1999) A survey of finite semifields. Discrete Mathematics 208/209: 125–137.
5. Dumas, J-G. and Orfila, J-B. (2014) Generating S-Boxes from Semifields Pseudo-extensions. `arXiv:1411.2503`
6. Figueroa, R. and Salzberg, P.M. and Shiue, P.J-S. (1994) A family of cryptosystems based on combinatorial properties of finite geometries. In: Contemporary Mathematics 168: 63–68.
7. Gil Álvarez, P. (1996) Las matemáticas de lo incierto. Servicio de publicaciones. Universidad de Oviedo.
8. Goldreich, O. (2001) Foundations of Cryptography. Cambridge University Press.
9. Goldwasser, S. and Micali, S. (1984) Probabilistic Encryption. Journal of Computer and System Science 28(2): 270–299.
10. Kalka, A. (2012) Non-associative public-key cryptography. `arXiv:1210.8270`
11. Knuth, D.E. (1965) Finite semifields and projective planes. Journal of Algebra 2: 182–217.
12. Lavrauw M. and Polverino O. (2011) Finite semifields and Galois geometry. In: De Beule J., Storme L. (eds.) Current Research Topics in Galois Geometry. NOVA Academic Publishers. ISBN 978-1-61209-523-3.
13. Menezes, A.J. and van Oorschot, P.C. and Vanstone, S.A. (1996) Hanndbook of Applied Cryptography. CRC Press.
14. Mister, S. and Adams, C. (1996). Practical S-Box Design. In: Selected areas in cryptography. `doi=10.1.1.40.7715`
15. Rúa, I.F. (2004) Anillos no asociativos en codificación y criptografía. PhD Thesis, University of Oviedo, Oviedo.
16. Rúa, I.F. and Combarro, E.F. and Ranilla, J. (2009) Classification of semifields of order 64. Journal of Algebra 322: 4011–4029.
17. Rúa, I.F. and Combarro, E.F. and Ranilla, J. (2012) Determination of division algebras with 243 elements. Finite Fields and Their Applications 18: 1148–1155.
18. Saarinen, M-J.O. (2012) Cryptographic Analysis of All 4 4-Bit S-Boxes. In: Selected Areas in Cryptography, Lecture Notes in Computer Science 7118: 118–133.
19. Shannon, C. (1948) A Mathematical Theory of Communication. Bell System Technical Journal 27(3): 379–423.

20. Shannon, C. (1949) Communication Theory of Secrecy Systems. Bell System Technical Journal 28(4): 656–715.
21. Stinson D.R. (2006) Cryptography Theorey and Practice (3rd edition). Chapman & Hall/CRC.
22. Malekian, E. and Zakerolhosseini, A. (2010) A non-associative lattice-based public key cryptosystem. Security and Communication Networks 5(2): 145–163.