Available online at www.sciencedirect.com
SCIENCE DIRECT®
Discrete Mathematics 306 (2006) 2016–2020
DISCRETE MATHEMATICS
www.elsevier.com/locate/disc

Note

On some probabilistic approximations for AES-like s-boxes

A.M. Youssef [a], S.E. Tavares[b], G. Gong[c]

[a] *Concordia Institute for Information Systems Engineering, Concordia University, Montreal, Quebec, Canada H3G 1M8*
[b] *Department of Electrical and Computer Engineering, Queen's University, Kingston, Ont., Canada, K7M 1B6*
[c] *Department of Electrical and Computer Engineering, University of Waterloo , Waterloo, Ont., Canada, N2L 3G1*

Received 30 June 2005; received in revised form 17 March 2006; accepted 28 March 2006
Available online 14 June 2006

Abstract

Several recently proposed block ciphers such as AES, Camellia, Shark, Square and Hierocrypt use s-boxes that are based on the inversion mapping over $GF(2^n)$. In order to hide the simple algebraic structure in this mapping, an affine transformation over F_2 is usually used after the output of the s-box. In some ciphers, an additional affine transformation is used before the input of the s-box as well. In this paper, we study the algebraic properties of a simple approximation in the form $s(x) = ax^{-1} + b, a, b \in GF(2^n)$ for such s-boxes. The implication of this result on the cryptanalysis of these ciphers remains an open problem.

Keywords: Finite fields; Cryptography; Monomial s-boxes; AES

1. Introduction

Differential [2] and linear cryptanalysis [11] are two of the most powerful attacks on iterative symmetric key block ciphers. The complexity of linear cryptanalysis depends on the size of the maximum entry in the linear approximation table of the Boolean functions used to construct the round function. Similarly, the complexity of differential cryptanalysis depends on the size of the largest entry of the XOR table of these Boolean functions. In [13], Nyberg suggested an s-box which is optimized towards these two criteria. Nyberg s-box is based on the inversion mapping

$$f(x) = x^{-1}, \quad x \in GF(2^n), \quad f(0) = 0.$$

The main disadvantage of this s-box is its simple algebraic description (by definition) over $GF(2^n)$ [7] which may enable some attacks such as the interpolation attacks [8,9]. In order to overcome this problem, this mapping was modified in a way that does not modify its resistance towards both linear and differential cryptanalysis while the overall s-box description becomes complex in $GF(2^n)$. The Nyberg s-box construction (with $n = 8$) was adopted in many block ciphers such as Shark [15], Square [3], AES [4–6,12], Camellia [1], and Hierocrypt [14]. Both AES and Camellia are of particular interest since AES is the current standard adopted by NIST and Camellia is included in the NESSIE (New European Schemes for Signatures, Integrity, and Encryption) portfolio of recommended cryptographic primitives.

E-mail addresses: youssef@ciise.concordia.ca (A.M. Youssef), tavares@ee.queensu.ca (S.E. Tavares), G.Gong@ece.uwaterloo.ca (G. Gong).

0012-365X/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.disc.2006.03.055

In order to hide the simple algebraic structure in this mapping, Shark, Square and AES use an affine transformation over F_2 after the output of the inversion mapping. In Camellia, an additional affine transformation is used before the input of the s-box as well. These affine transformations prove to be useful in preventing low-degree polynomial approximations in $GF(2^n)$. For example, using exhaustive search, we verified that the best third-degree polynomial approximation for the AES s-box holds with $p = 11/256$. Meanwhile, other simple (sparse) polynomial approximations may still prove to be useful for the cipher cryptanalysis. In this paper, we study the algebraic properties of a simple approximation in the form $s(x) = ax^{-1} + b$, $a, b \in GF(2^n)$ for the overall s-box. Our result applies to all s-boxes in the form $s(x) = L_1((L_2(x))^d)$ where L_1, L_2 are invertible affine transformations over $GF(2)$ and $gcd(n, d) = 1$.

2. Mathematical background and definitions

For a background about the general theory of finite fields, the reader is referred to [10]. Throughout this paper, we will use the hexadecimal notation to denote the field elements (e.g., let α denote the primitive element used to construct the finite field $GF(2^n)$, then the field element $b_{n-1}\alpha^{n-1} + b_{n-2}\alpha^{n-2} + \cdots + b_2\alpha^2 + b_1\alpha + b_0$, $b_i \in \{0, 1\}$, is represented by the hexadecimal number consisting of bits $(b_{n-1}b_{n-2}\cdots b_1b_0)$).

Definition 1. A polynomial having the special form

$$L(x) = \sum_{i=0}^{t} \beta_i x^{2^i} \tag{1}$$

with coefficients β_i from $GF(2^n)$ is called a linearized polynomial over $GF(2^n)$.

Lemma 1. *There is a* $1-1$ *correspondence between the set of invertible linear transformations over* F_2^n *and the set of linearized polynomials over* $GF(2^n)$ [10].

Definition 2. Let S be a subgroup of $GF(2^n)$. A coset of S is a subset of $GF(2^n)$ whose elements can be expressed as $x + S = \{s + S, s \in S\}$.

Lemma 2. *The distinct cosets of a subgroup S in a group G are disjoint.*

Lemma 3. *The zeroes of* $L(x)$ *form a subspace of* $GF(2^n)$.

The following lemma [16] illustrates the effect of applying a linear transformation to the output coordinates of f on the coefficients of its corresponding polynomial.

Lemma 4. *Let* $F(x_1, \ldots, x_n) = (f_1(x), \ldots, f_n(x))$ *be the Boolean function corresponding to the polynomial function* $f(x) = x^d$ over $GF(2^n)$. *Let* $G(x)$ *be the Boolean mapping obtained by applying a linear transformation to the output coordinates of* $f(x_1, \ldots, x_n)$. *Then the polynomial function corresponding to G can be expressed as*

$$g(x) = \sum_{i=0}^{n-1} c_i x^{d2^i}, \quad c_i \in GF(2^n). \tag{2}$$

Proof. The proof follows directly by applying Lemma 1. □

It is straight forward to show that if the linear transformation is replaced by an affine one, then $g(x)$ can be expressed as

$$g(x) = \sum_{i=0}^{n-1} c_i x^{d2^i} + c_n, \quad c_i \in GF(2^n).$$

Example 1. Using Lagrange interpolation over $GF(2^8)$, where $GF(2^8)$ is defined by the irreducible polynomial $\alpha^8+\alpha^4+\alpha^3+\alpha+1$, the AES s-box can be expressed as

$$s(x) = 63 + 05x^{254\times1} + 09x^{254\times2} + f9x^{254\times4} + 25x^{254\times8} + f4x^{254\times16} + 01x^{254\times32} + b5x^{254\times64} + 8fx^{254\times128}.$$

By noting that for $x \in GF(2^n)$, $x^a = x^{a \bmod 2^n-1}$, then we have

$$s(x) = 63 + 05x^{254} + 09x^{253} + f9x^{251} + 25x^{247} + f4x^{239} + 01x^{223} + b5x^{191} + 8fx^{127}.$$

3. Monomial approximations for AES-like s-boxes

In this section, we present some algebraic properties for the affine transformations used in AES-like s-boxes. We also show some simple monomial approximations for these s-boxes.

Theorem 1. *Let $s(x)=L_1((L_2(x))^d)$ where L_1, L_2 are affine transformations over $GF(2)$ and $gcd(n,d)=1$. Then $\exists$ $a, \{b_1, b_2, \ldots, b_{2^{n-m}}\}$ for which the approximation $s(x)=ax^d+b_i$ holds with probability $p=2^m/2^n$ for some integer $m \geqslant 1$. Moreover, $\bigcup_{b_i}\{x|ax^d+b_i=x^d\}$ spans the whole vector space of $GF(2^n)$.*

Proof. The effect of the input transformation, L_2, can be moved to the output by changing the basis in which we perform the finite field computations. Thus, using the new basis, the above s-box is equivalent to another s-box in the form $s_1(x)=L_3(x^d)$, $L_3=L_1L_2^{-1}$. From Lemma 4, $s_1(x)$ can be expressed as $s_1(x)=\sum_{i=0}^{n-1}c_ix^{d2^i}$. Hence, the number of times in which the approximation $s_1(x)=ax^d$ holds is equivalent to the number of zeroes of the linearized polynomial $(a+c_0)x^d+\sum_{i=1}^{n-1}c_ix^{d2^i}$. From Lemma 3, this number is in the form 2^m for some integer $m \in \{0,\ldots,n\}$.Let X_1 denote the subgroup for which the approximation $s_1(x)=L_3(x^d)=ax^d$ holds. Thus $\forall x \in X_1$ and $\forall k_i \in GF(2^n)$ such that $(x+k_i) \in X_1$, we have $L_3(x)=ax \Rightarrow L_3(x+k_i)=a(x+k_i) \Rightarrow L_3(x)=ax+ak_i+L_3(k_i)=ax+b_i$ $\Rightarrow b_i=ak_i+L_3(k_i)$. The second part of the theorem follows from Lemma 2 and by noting that if $|X_1|=2^m$, then we have 2^{n-m} different choices of k_i such that $x+k_i$ belong to a different coset. In order to prove that $m \geq 1$, we note that $x=0$ is always a zero of the polynomial $(a+c_0)x^d+\sum_{i=1}^{n-1}c_ix^{d2^i}$. For $a=\sum_{i=0}^{n-1}c_i$, we have another zero at $x=1$ and hence $m \geq 1$. □

Example 2. Let $s(x)=L(x^{-1})$, $x \in GF(2^4)$ where $GF(2^4)$ is defined by the irreducible polynomial x^4+x+1, and L is given by

$$L=\begin{bmatrix}1&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0\end{bmatrix}.$$

Then we have

$$\begin{aligned} s(x) &= 9x^{-1} \quad \text{if } x \in \{0,1,4,5,a,b,e,f\},\\ &= 9x^{-1}+5 \quad \text{if } x \in \{2,3,6,7,8,9,c,d\}.\end{aligned}$$

Let $L_{\hat{p}}$ denote the number of invertible linear transformations for which the *best* approximation of the form $L(x)=ax$ holds with probability $\hat{p}$, i.e.,

$$\hat{p}=\max_{a\in GF(2^n)}\frac{|\{x|L(x^{-1})=ax^{-1}\}|}{2^n}.$$

Finding an expression for the distribution $L_{\hat{p}}$ over all randomly selected linear transformations seem to be a hard combinatorial problem. Table 1 shows the distribution of $L_{\hat{p}}$ for all the invertible linear transformations over F_2 for

Table 1
$L_{\hat{p}}$ distribution for $n = 4$

$\hat{p}$	$\frac{2}{16}$	$\frac{4}{16}$	$\frac{8}{16}$	1
$L_{\hat{p}}$	30	18540	1575	15

Table 2
$L_{\hat{p}}$ distribution for 10,000 randomly selected linear transformations, $n = 8$

$\hat{p}$	$\frac{4}{256}$	$\frac{8}{256}$	$\frac{16}{256}$
$L_{\hat{p}}$	2271	7446	283

Table 3
$\{a, b_i\}$-pairs for the AES s-box

a	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}	b_{11}	b_{12}	b_{13}	b_{14}	b_{15}	b_{16}
5f	e	15	23	38	4e	55	63	78	8e	95	a3	b8	ce	d5	e3	f8
ae	17	1a	20	2d	54	59	63	6e	91	9c	a6	ab	d2	df	e5	e8
1f	b	11	25	3f	4d	57	63	79	87	9d	a9	b3	c1	db	ef	f5

Table 4
$\{a, b_i\}$-pairs for the Shark and Square s-box

a	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	b_{10}	b_{11}	b_{12}	b_{13}	b_{14}	b_{15}	b_{16}
ac	6	15	22	31	46	55	62	71	86	95	a2	b1	c6	d5	e2	f1

$n = 4$. Table 2 shows the distribution of $L_{\hat{p}}$ for 10000 randomly selected invertible linear transformations over F_2 for $n = 8$.

Using the irreducible polynomial $x^8 + x^4 + x^3 + x + 1$, Table 3 shows the different $\{a, b_i\}$ pairs for which the approximation $s(x) = ax^{-1} + b_i$ holds for the AES s-box. Thus for AES, we have $\hat{p}_{\text{AES}} = 16/256$.

Similarly, using the irreducible polynomial $\alpha^8 + \alpha^7 + \alpha^6 + \alpha^5 + \alpha^4 + \alpha^2 + 1$, Table 4 shows the corresponding pairs of the Shark (and Square) s-box. Thus we also have $\hat{p}_{\text{Shark}} = \hat{p}_{\text{Square}} = 16/256$. Since many randomly selected linear transformations have $\hat{p} = 4/256 < \hat{p}_{\text{AES}} = 16/256$, one can conclude that the linear transformation of these ciphers performs worse than randomly selected linear transformations in hiding approximations of the above form.

Similar approximations were found for the Camellia s-boxes. For example, for $S1$, $S2$, $S3$, $\hat{p} = 8/256$. However, Camellia uses four s-boxes. The linear transformation at the input of S_4 is different from that at the input of S_1, S_2, S_3 which prevents us from finding one set of basis for which all our approximations hold. When applying our approach to the Hierocrypt s-box, we have $\hat{p} = 8/256$.

4. Conclusions

The main design objective of the AES-like s-boxes' affine transformations is to hide simple algebraic relations over $GF(2^n)$. In this paper, we showed that some simple approximations are still possible. While the implication of our results on the cryptanalysis of AES and other ciphers remains an open problem, special care should be taken when designing linear transformations that hide the $GF(2^n)$ structure within monomial-based s-boxes. For example, if we replace the affine transformation of the AES s-box (denote it by L_{AES}) with L' obtained by applying L_{AES} to itself, i.e., $L'(x) = L_{\text{AES}}(L_{\text{AES}}(x))$, then we can obtain approximations of the above form that hold with $\hat{p} = 64/256$.

One should also note that while the approximations introduced in this paper hold with a relatively small probability, the possibility of discovering a clever method for concatenating these approximations so that the overall round function may also have some simple (sparse) approximation over $GF(2^n)$ should not be excluded. If such approximations exist, then the overall cipher can be attacked using some of the recently proposed algebraic attacks.

Acknowledgment

The authors would like to thank the anonymous referees for their valuable comments. The first author is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Grant N00930.

References

[1] K. Aoki, T. Ichikawa, M. Kanda, M. Matsui, S. Moriai, J. Nakajimaand, T. Tokita, Camellia: a 128-bit block cipher suitable for multiple platforms—design and analysis, Proceedings of Seventh Annual International Workshop on Selected Areas in Cryptography, SAC'2000, Lecture Notes in Computer Science, vol. 2012, Springer, Berlin, 2001, pp. 39–56.
[2] E. Biham, A. Shamir, Differential cryptanalysis of DES-like cryptosystems, J. Cryptology 4 (1) (1991) 3–72.
[3] J. Daemen, L.R. Knudsen, V. Rijmen, The block cipher Square, Fast Software Encryption, Lecture Notes in Computer Scienc, vol. 1267, Springer, Berlin, 1997, pp. 149–165.
[4] J. Daemen, V. Rijmen, The Block Cipher Rijndael, Springer, Berlin, ISBN 3-540-42580-2, 2000.
[5] Federal Information Processing Standards Publication (FIPS 197), Advanced Encryption Standard (AES), November 26, 2001.
[6] N. Ferguson, R. Schroeppel, D. Whiting, A simple algebraic representation of Rijndael, Proceedings of the Eighth International Workshop on Selected Areas in Cryptography, SAC'2001, Lecture Notes in Computer Science, vol. 2259, 2001, pp. 103–111.
[7] G. Gong, S.W. Golomb, Transform domain analysis of DES, IEEE Trans. Inform. Theory IT-45 (6) (1999) 2065–2073.
[8] T. Jakobsen, Cryptanalysis of block ciphers with probabilistic non-linear relations of low degree, Proceedings of Crypto'99, Lecture Notes in Computer Science, vol. 1462, 1999, pp. 213–222.
[9] T. Jakobsen, L. Knudsen, The interpolation attack on block ciphers, Fast Software Encryption, Lecture Notes in Computer Science, vol. 1267, 1997, pp. 28–40.
[10] R. Lidl, H. Niederreiter, Finite fields, Encyclopedia of Mathematics and its Applications, vol. 20, Addison-Wesley, Reading, MA, 1983.
[11] M. Matsui, Linear cryptanalysis method for DES Cipher, Adv. Cryptology, Proceedings of Eurocrypt'93, Lecture Notes in Computer Sceince, vol. 765, Springer, Berlin, 1994, pp. 386–397.
[12] S. Murphy, M.J.B. Robshaw, Essential algebraic structure within the AES, Proceedings of Crypto 2002, Lecture Notes in Computer Science, vol. 2442, 2002, pp. 1–16.
[13] K. Nyberg, Differentially uniform mappings for cryptography, Proceedings of Eurocrypt'93, Lecture Notes in Computer Sceince, vol. 765, Springer, Berlin, 1994, pp. 55–64.
[14] K. Ohkuma, H. Muratani, F. Sano, S. Kawamura, The block cipher hierocrypt, Proceedings of Seventh Annual International Workshop on Selected Areas in Cryptography, SAC'2000, Lecture Notes in Computer Science, vol. 2012, Springer, Berlin, 2001, pp. 72–88.
[15] V. Rijmen, J. Daemen, B. Preneel, A. Bosselaers, E. De Win, The cipher Shark, in: D. Gollmann (Ed.), Fast Software Encryption, Lecture Notes in Computer Scienc, vol. 1039, Springer, Berlin, 1996, pp. 99–112.
[16] A. M. Youssef, G. Gong, On the interpolation attacks on block ciphers, Proceedings of Fast Software Encryption 2000, Lecture Notes in Computer Science, vol. 1978, 2001, pp. 109–120.