Journal of Computer Science 5 (12): 1091-1094, 2009
ISSN 1549-3636
© 2009 Science Publications

New Directions in Cryptanalysis of Block Ciphers

Davood RezaeiPour and Mohamad Rushdan Md Said
Institute for Mathematical Research, University Putra Malaysia,
43400 UPM Serdang, Selangor Darul Ehsan, Malaysia

Abstract: Problem statement: The algebraic expression of the Advanced Encryption Standard (AES) RIJNDAEL S-box involved only 9 terms. The selected mapping for RIJNDAEL S-box has a simple algebraic expression. This enables algebraic manipulations which can be used to mount interpolation attack. **Approach:** The interpolation attack was introduced as a cryptanalytic attack against block ciphers. This attack is useful for cryptanalysis using simple algebraic functions as S-boxes. **Results:** In this study, we presented an improved AES S-box with good properties to improve the complexity of AES S-box algebraic expression with terms increasing to 255. **Conclusion:** The improved S-box is resistant against interpolation attack. We can develop the derivatives of interpolation attack using the estimations of S-box with less nonlinearity.

Key words: Block cipher, AES, S-box, interpolation attack, Lagrange interpolation formula

INTRODUCTION

The interpolation attack is a technique for attacking block ciphers built from simple algebraic functions. A block cipher algorithm may not include any algebraic property that can be efficiently distinguishable, since an interpolation attack can be applied to such a block cipher which leads to the leakage of information about the secret key.

This mathematical property has effective implications using a block cipher with a fixed secret key. If the ciphertext is described as a polynomial -with unknown coefficients-of the plaintext, and if the degree of this polynomial is sufficiently low, then a limited number of plaintext-ciphertext pairs is capable to completely determine the encryption function[1]. Constructing this polynomial will not immediately yield the key. Actually this is a polynomial that emulates the encryption function. It produces valid ciphertexts from given plaintexts.

It can be applied by constructing an implicit polynomial expression involving parts of the plaintext and the ciphertext.

Now, we can check the polynomial against another value that was not used in the construction to test it. If the polynomial produces the correct result, then we have guessed the key bits. This allows the cryptanalyst to encrypt and decrypt data for the unknown key-without doing any key-recovery.

In this article, we first describe the main parts of AES (RIJNDAEL) which consists of the individual transformations and AES S-box. We will introduce the interpolation attack with considering of the points of weakness and strength in AES S-box. Finally, we will discuss the manner of doing interpolation attack using the different representations of AES S-box.

MATERIALS AND METHODS

AES cryptosystem (RIJNDAEL cipher): The RIJNDAEL cipher, designed by Daemen and Rijmen[2] in 1998, is a successor of SQUARE. It was submitted to the US National Institute of Standards and Technology (NIST) in response to an open call for 128 bit block ciphers. It was, together with 14 other candidates, extensively evaluated during two years, before NIST announced in 2000 that RIJNDAEL would replace DES and become the new AES. Just as its predecessor SQUARE, RIJNDAEL was specifically designed to resist differential and linear cryptanalysis.

In RIJNDAEL cipher, the individual transformations SubBytes, ShiftRows, MixColumns, and AddRoundKey process the state[3]. The SubBytes transformation is a non-linear byte substitution that operates independently on each byte of the state using a substitution table (S-box). AES S-box is presented in hexadecimal form in Fig. 1.

Actually, S-box is non-linear substitution table which used in several byte substitution transformations

Corresponding Author: Davood RezaeiPour, Institute for Mathematical Research, University Putra Malaysia,
43400 UPM Serdang, Selangor Darul Ehsan, Malaysia Tel: +601-73507481 Fax: +603-89423789

and in the Key Expansion routine to perform a one-for-one substitution of a byte value. This S-box is invertible and constructed by composing two transformations:

- Take the multiplicative inverse in the finite field GF(2^8)
- Apply the following affine transformation over GF(2):

$$b'_i = b_i \oplus b_{(i+4) \bmod 8} \oplus b_{(i+5) \bmod 8} \oplus b_{(i+6) \bmod 8} \oplus b_{(i+7) \bmod 8} \oplus c_i \quad (1)$$

for $0 \le i \le 8$, where b_i and c_j are the i^{th} bit of the b and c, respectively.

In Matrix form, the affine transformation element of the S-box can be written as:

$$\begin{bmatrix} b'_0 \\ b'_1 \\ b'_2 \\ b'_3 \\ b'_4 \\ b'_5 \\ b'_6 \\ b'_7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \\ b_7 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

The design principle for the RIJNDAEL S-box is influenced by linear and differential cryptanalysis and also interpolation attacks. The designers considered these criteria:

- Invertibility
- Minimization of the largest non-trivial correlation between linear combinations of input bits and linear combination of output bits
- Minimization of the largest non-trivial value in the XOR table
- Complexity of its algebraic expression in GF(2^8)
- Simplicity of description

The affine transformation (1) does not affect the properties with respect to the first 3 criteria, but if properly chosen, allows the S-box to satisfy the 4rth criterion.

We have chosen an affine mapping which has a very simple algebraic expression. It can be seen as modular polynomial multiplication followed by an addition:

$$b(x) = (x^7 + x^6 + x^2 + x) + a(x)$$
$$(x^7 + x^6 + x^5 + x^4 + 1) \bmod x^8 + 1$$

x \ y	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
0	63	7c	77	7b	f2	6b	6f	c5	30	01	67	2b	fe	d7	ab	76
1	ca	82	c9	7d	fa	59	47	f0	ad	d4	a2	af	9c	a4	72	c0
2	b7	fd	93	26	36	3f	f7	cc	34	a5	e5	f1	71	d8	31	15
3	04	c7	23	c3	18	96	05	9a	07	12	80	e2	eb	27	b2	75
4	09	83	2c	1a	1b	6e	5a	a0	52	3b	d6	b3	29	e3	2f	84
5	53	d1	00	ed	20	fc	b1	5b	6a	cb	be	39	4a	4c	58	cf
6	d0	ef	aa	fb	43	4d	33	85	45	f9	02	7f	50	3c	9f	a8
7	51	a3	40	8f	92	9d	38	f5	bc	b6	da	21	10	ff	f3	d2
8	cd	0c	13	ec	5f	97	44	17	c4	a7	7e	3d	64	5d	19	73
9	60	81	4f	dc	22	2a	90	88	46	ee	b8	14	de	5e	0b	db
a	e0	32	3a	0a	49	06	24	5c	c2	d3	ac	62	91	95	e4	79
b	e7	c8	37	6d	8d	d5	4e	a9	6c	56	f4	ea	65	7a	ae	08
c	ba	78	25	2e	1c	a6	b4	c6	e8	dd	74	1f	4b	bd	8b	8a
d	70	3e	b5	66	48	03	f6	0e	61	35	57	b9	86	c1	1d	9e
e	e1	f8	98	11	69	d9	8e	94	9b	1e	87	e9	ce	55	28	df
f	8c	a1	89	0d	bf	e6	42	68	41	99	2d	0f	b0	54	bb	16

Fig. 1: S-box: Substitution values for the byte xy (in hexadecimal format)

The modulus has been chosen as the simplest modulus possible. The multiplication polynomial is selected from the set of polynomials coprime to themodulus as the one with the simplest description. The constant is selected such that S-box has no fixed points $(S-box(a) = a)$ and no "opposite fixed points" $(S-box(a) = \overline{a})$.

Interpolation attack: The interpolation attacks depend only on the number of S-boxes and number of rounds in the cipher. This attack is independent of the sizes of the S-boxes.

Based on the following theorem, Jakobsen and Knudsen[4] introduced the interpolation attack in 1997.

Theorem 1: Let R be a field. Given 2n elements $x_1, x_2, \ldots, x_n \in R$, $y_1, y_2, \ldots, y_n \in R$, where the x_is are distinct. Define:

$$f(x) = \sum_{i=1}^{n} y_i \prod_{1 \le j \le n, j \ne i} \frac{x - x_j}{x_i - x_j} \quad (2)$$

Then f(x) is the only polynomial over R of degree at most n-1 such that $f(x_j) = y_j$ for $1 \le i \le n$. This equation is known as the Lagrange interpolation formula.

Based on this theorem, in the cipher algorithm, every ciphertext is describable as polynomial inclusive of plaintext, which its coefficients are the specific functions of the key. It means that the ciphertext can be interpolated by a polynomial in the plaintext and key variables, i.e., by Lagrange interpolation.

If the message length be m, and the describer polynomial of the cipher consists of nonzero coefficients $\{n \mid n < 2^m\}$, then the interpolation attack is done, with having n plaintexts and corresponding ciphertexts.

Actually, if the number of terms in polynomial be less, then we can get the coefficients of polynomial, instead of the key variables. If the number of nonzero coefficients is n, then we can form an equations system by n equations and n unknowns, with having n plaintexts and corresponding ciphertexts. With solving of such system, we will find the coefficients and we will have a specific polynomial from input to output. Using this polynomial, we can recover the ciphertext without the knowledge about key.

The performing of interpolation attack over AES S-box: Using the interpolation attack, SHARK cryptosystem[5] was analyzed by Knudsen and Jakobsen[4]. This cryptosystem was designed by AES designers, whereas they had enough information about the interpolation attack. But this is not certain reason for resistance of SHARK against interpolation attack. In this cryptosystem, a carefully chosen S-box imposes most number of terms on the equations. Since in the polynomial representation of S-box, the all possible terms will be with hamming weights 7.With forming of the equation for one round cipher, we have:

$$S(x + k_1) + k_2 = y \quad (3)$$

x = Plaintext
y = Cipher text

which x and y are known but, $k_1$and k_2 are unknowns.

Using extension (3), we can find a polynomial in terms of x with 255 terms of degree 254, such that all possible powers of x appear in it. So, the interpolation attack is not possible. Since in the AES, S-box equation has the all possible terms with hamming weights 7, it can be seen that all terms appear in the representation of other rounds and the number of terms cannot be less than 2^m, so the interpolation attack is impossible even on one round.

Now, we can express this question: Is interpolation attack possible using S-box estimation? As an example, if we form the describer polynomial of one round using $S_{85}(x)$ estimation, then we will have 31 terms with nonzero coefficients instead of 255 terms, namely, we can get the coefficients with using 31 suitable texts instead of using 255 texts. Since the probability of truth for every pair is:

$$\frac{1}{3}(\cong \frac{86}{255})$$

We thus need (31×3 = 93) pairs of plaintext and ciphertext for solving of this probable equation, which is less than 255. The computational complexity of this attack is more than exhaustive key search attack, so it is not successful.

RESULTS AND DISCUSSION

Jakobsen and Knudsen presented interpolation attacks in[4] as a reaction to ciphers using algebraically constructed S-Boxes such as those proposed by Nyberg[6]. In fact, interpolation attacks were the first demonstration of successful polynomial-based algebraic attacks against block ciphers. Interpolation attacks work by expressing the relationship between the plaintext and ciphertext for a fixed key as either one or as a vector of polynomials.

If the degree of these polynomials is low enough, the coefficients of the polynomials can be interpolated from a number of plaintext/ciphertext pairs. A key-dependent equivalent of the encryption or the decryption algorithm has then been determined. In[4] upper bounds on the data complexity-the number of required pairs for known-plaintext interpolation attacks-are given for selected examples. In general, this number increases exponentially with the degree of the polynomial function describing the S-Box, the number of rounds and the number of elements in the internal state.

Since AES provides “full diffusion” after only two rounds, so it can be considered resistant against the interpolation attack.

CONCLUSION

We described the interpolation attack against AES cryptosystem which utilized from algebraic properties of AES. We also introduced the version of AES S-box which was resistant against interpolation attack. Finally we illustrated the new directions for the future research. We can develop the derivatives of interpolation attack using the estimations of S-box with less nonlinearity. Also, one can speed up the attacks using the Newton interpolation instead of Lagrange interpolation.

REFERENCES

1. C., Swenson, 2008. Modern Cryptanalysis. Wiley Publishing, Inc., Indiana, ISBN: 13: 978-0470135938, pp: 222.
2. Daemen, J. and V. Rijmen, 1999. AES proposal: rijndael, AES algorithm submission. http:// csrc.nist.gov/encryption/aes/aes-home.htm
3. National Institute of Standards and Technology, 2001. Advanced encryption standard, FIPS 197. http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf

4. Jakobsen, T. and L.R. Knudsen, 1997. The interpolation attack on block ciphers. Lecture Notes Comput. Sci., 1267: 28-40. DOI: 10.1007/BFb0052329
5. Rijmen, V., J. Daemen, B. Preneel, A. Bosselaers and E. De Win, 1996. The cipher SHARK. Lecture Notes Comput. Sci., 1039: 99-111. DOI: 10.1007/3-540-60865-6
6. Nyberg, K., 1993. Differentially uniform mappings for cryptography. Lecture Notes Comput. Sci., 765: 55-64. DOI: 10.1007/3-540-48285-7