MATRIX POWER S-BOX ANALYSIS[1]

Kestutis Luksys, Petras Nefas

Abstract: Construction of symmetric cipher S-box based on matrix power function and dependant on key is analyzed. The matrix consisting of plain data bit strings is combined with three round key matrices using arithmetical addition and exponent operations. The matrix power means the matrix powered by other matrix. This operation is linked with two sound one-way functions: the discrete logarithm problem and decomposition problem. The latter is used in the infinite non-commutative group based public key cryptosystems. The mathematical description of proposed S-box in its nature possesses a good “confusion and diffusion” properties and contains variables “of a complex type” as was formulated by Shannon. Core properties of matrix power operation are formulated and proven. Some preliminary cryptographic characteristics of constructed S-box are calculated.

Keywords: Matrix power, symmetric encryption, S-box.

ACM Classification Keywords: E.3 Data Encryption, F.2.1 Numerical Algorithms and Problems.

Conference: The paper is selected from Sixth International Conference on Information Research and Applications – i.Tech 2008, Varna, Bulgaria, June-July 2008

Introduction

As it is known, the design criteria for the block ciphers as for other cryptographic systems are related with the known cryptoanalytic attacks. It is essential that after the new attack invention the old design criteria must be changed.

Traditional design criteria are oriented to the most powerful attacks such as linear and differential and were successfully satisfied for the several known ciphers, for example AES, Serpent, Camellia Misty/Kasumi etc. It was shown that the non-linearity properties of the inverse function in $GF(2^n)$ used as a single non-linear component in AES are close to optimality with respect to linear, differential and higher-order differential attacks [Canteaut and Videau, 2002].

But nevertheless it is shown that many known “optimal” ciphers have a very simple algebraic structure and are potentially vulnerable to the algebraic attack. This attack was declared in [Schaumuller-Bihl, 1983] and developed in [Courtois and Pieprzyk, 2002]. The vulnerability is related to S-box description by implicit input/output and key variables algebraic equations of polynomial type. For example the AES can be described by the system of multivariate quadratic equations in GF (2^8) for which the XL or XSL attack can be applied in principle. Then there is a principal opportunity to find the solution of these equations by some feasible algorithm that might be of sub-exponential time and recover the key from a few plaintext/ciphertext pairs.

The algebraic attack changes some old security postulates [Courtois, 2005]:

1. The complexity is no longer condemned to grow exponentially with the number of rounds.
2. The number of required plaintexts may be quite small (e.g. 1).
3. The wide trail strategy should have no impact whatsoever for the complexity of the attack.

Despite the fact that there are no practical results of breaking the entire AES by algebraic attack yet, it is sensible to build the new design methods possessing a higher resistance to algebraic attack. According to Courtois the design of ciphers will never be the same again and this is supported by the declared new ideas for the S-box

[1] Work supported by the Lithuanian State Science and Studies Foundation

construction laying on the sufficiently large random S-boxes to prevent all algebraic attacks one can think [Courtois et al., 2005].

In this paper we further discus so called matrix power operation introduced in [Sakalauskas and Luksys, 2007] for a matrix power S-box construction. Matrix power S-box is based on modular exponentiation over GF(2^n). This leads to some generalization of discrete logarithm problem (DLP) using a matrix group action problem over Galois field.

The idea to use the group or semigroup action problem in vectorial spaces for the asymmetric cryptographic primitives' construction can be found in [Monico, 2002]. We have generalized this approach and applied it to our S-box construction. As a result we have obtained some one way function (OWF) which is linked not only with a classical DLP but also with so called decomposition problem (DP), used in the asymmetric cryptosystems based on the hard problems in infinite non-commutative groups [Shpilrain and Ushakov, 2005]. The same kind of DP is used also in digital signature scheme and key agreement protocol construction [Sakalauskas, 2005] and [Sakalauskas et al., 2007].

Preliminaries

Let us define *m* x *m* matrices over $GF(2^n)$. The set of all those matrices over $GF(2^n)$ we denote as **M**. Plaintext and ciphertext data is represented in this set. We do not introduce any internal operations in the set **M**. For further considerations we are interested only in external operations performed in this set.

Let $\mathbf{M_G}$ be a group of *m* x *m* matrices over N_{2n-1} with the commonly defined matrix multiplication operation and matrix inverse. Keys' matrices should be chosen from $\mathbf{M_G}$.

Matrix group $\mathbf{M_G}$ left and right action operations in the set M are denoted by $\triangleright$ and $\triangleleft$ respectively.

In a formal way $\triangleright$ is a mapping $\triangleright : \mathbf{M_G} \times \mathbf{M} \to \mathbf{M}$ and $\triangleleft : \mathbf{M} \times \mathbf{M_G} \to \mathbf{M}$. Then $\forall L, R \in \mathbf{M_G}$ and $\forall X \in \mathbf{M}$ there exist some $Y, Z \in \mathbf{M}$ such that $L \triangleright X = Y$ and $X \triangleleft R = Z$.

The elements of matrices *L*, *X*, *R*, *Y and Z* we denote by the indexed set of its elements respectively, i.e. by $\{x_{ij}\}$ we denote matrix *X*.

We have chosen the following action operations which can be written for the matrix equation $L \triangleright X = Y$ elements

$$y_{ij} = \prod_{s=1}^{m} x_{sj}^{l_{is}} , \tag{1}$$

and for the matrix equation $X \triangleleft R = Z$ elements

$$z_{ij} = \prod_{t=1}^{m} x_{it}^{r_{tj}} . \tag{2}$$

The multiplication and power operations are performed using GF (2^n) arithmetic, i.e. modulo irreducible polynomial.

Example 1. To give a simple example, let us assume that all matrices have two rows and two columns, i.e. $m = 2$. In this case, matrix *Y* can be expressed in the following way

$$Y = L \triangleright X = \begin{pmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{pmatrix} \triangleright \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} x_{11}^{l_{11}} x_{21}^{l_{12}} & x_{12}^{l_{11}} x_{22}^{l_{12}} \\ x_{11}^{l_{21}} x_{21}^{l_{22}} & x_{12}^{l_{21}} x_{22}^{l_{22}} \end{pmatrix}.$$

Matrix *Z* can be expressed in the following way

$$Z = X \triangleleft R = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \triangleleft \begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{pmatrix} = \begin{pmatrix} x_{11}^{r_{11}} x_{12}^{r_{21}} & x_{11}^{r_{12}} x_{12}^{r_{22}} \\ x_{21}^{r_{11}} x_{22}^{r_{21}} & x_{21}^{r_{12}} x_{22}^{22} \end{pmatrix}.$$ □

Matrix power S-box

The S-box input data we denote by m x m matrix D with elements being binary strings in vector space F_2^{n-1}. Using the certain key expansion procedure we can generate the round keys for encryption: matrix K over F_2^{n-1} and matrices L, $R \in \mathbf{M_G}$. Input/output and key matrices are all of the same $m \times m$ size.

S-box transformations of input data D to ciphered output data C are performed as follows:

$$D + K + \mathbf{1} = X, \qquad \textbf{(3)}$$

$$L \rhd X \lhd R = C, \qquad \textbf{(4)}$$

where $D + K + \mathbf{1}$ denotes the ordinary arithmetical addition of matrices modulo 2^n; $\mathbf{1}$ is the matrix consisting of arithmetical unity elements in F_2^n. Combining (3) and (4) we obtain

$$L \rhd (D + K + \mathbf{1}) \lhd R = C. \qquad \textbf{(5)}$$

From (3) we obtain a matrix $X \in \mathbf{M}$ which does not contain zero elements, i.e. is without zero binary strings. This is necessary because of multiplications. If there would be at least one zero element, then ciphertext will be zero matrix.

We can write now the implicit formula for an element c_{ij}:

$$c_{ij} = \prod_{t=1}^{m}\prod_{s=1}^{m} x_{st}^{l_{is}\cdot r_{tj}} = \prod_{t=1}^{m}\prod_{s=1}^{m}(d_{st} + k_{st} + 1)^{l_{is}r_{tj}}, \qquad \textbf{(6)}$$

where 1 is a bit string corresponding to arithmetical unit in F_2^n.

Since $\mathbf{M_G}$ is a group of matrices, then there exists the inverse matrix R^{-1} such that $RR^{-1} = R^{-1}R = I$, where I is the identity matrix.

Decryption operation can be written similarly to (5):

$$L^{-1} \rhd C \lhd R^{-1} - K - \mathbf{1} = D. \qquad \textbf{(7)}$$

Resulting matrix of inverse S-box can be expressed like this:

$$d_{ij} = \prod_{t=1}^{m}\prod_{s=1}^{m} c_{st}^{l'_{is}\cdot r'_{tl}} - k_{ij} - 1, \qquad \textbf{(8)}$$

where $\{l'_{ij}\} = L^{-1}$ and $\{r'_{ij}\} = R^{-1}$. Thus, we have to be able to calculate inverse matrices of L and R keys for decryption. Key matrix K remains the same, only during decryption ordinary subtraction is used instead of addition.

For the validity of the last equations the left-right action operations must satisfy the following properties:

1. The action operations must be associative, i.e.

$$L_2 \rhd (L_1 \rhd \mathbf{X}) = (L_2L_1) \rhd \mathbf{X}, \qquad \textbf{(9a)}$$

$$(\mathbf{X} \lhd R_1) \lhd R_2 = \mathbf{X} \lhd (R_1R_2). \qquad \textbf{(9b)}$$

2. The action operations are both left and right invertible, i.e.

$$L^{-1} \rhd (L \rhd X) = (L^{-1}L) \rhd X = I \rhd X = X, \qquad \textbf{(10a)}$$

$$(X \lhd R^{-1}) \lhd R = X \lhd (R^{-1}R) = X \lhd I = X. \qquad \textbf{(10b)}$$

Theorem 1. The action operations are associative.

Proof. Let us consider encryption and decryption scheme with plaintext matrix X, key matrixes L and R, their inverse matrices L^{-1} and R^{-1} and cipher text matrix C. According to (4) and (7), following relations should be true:

$$L \rhd X \lhd R = C,$$

$$L^{-1} \rhd C \lhd R^{-1} = X.$$

For the simplicity, we omit matrix *K* here and consider that matrix *X* has no zero elements. This does not affect generality because matrix *K* is added before matrix power operation and subtracted after, in case of inverse S-box.

Then plaintext matrix *X* can be expressed following way:

$$X = L^{-1} \triangleright C \triangleleft R^{-1} = L^{-1} \triangleright \left(L \triangleright X \triangleleft R\right) \triangleleft R^{-1} = L^{-1} \triangleright \left\{ \prod_{t=1}^{m} \prod_{s=1}^{m} x_{st}^{l_{is} \cdot r_{tj}} \right\} \triangleleft R^{-1} = \left\{ \prod_{v=1}^{m} \prod_{u=1}^{m} \left(\prod_{t=1}^{m} \prod_{s=1}^{m} x_{st}^{l_{us} \cdot r_{tv}} \right)^{l'_{iu} \cdot r'_{vj}} \right\} =$$

$$= \left\{ \prod_{v=1}^{m} \prod_{u=1}^{m} \prod_{t=1}^{m} \prod_{s=1}^{m} x_{st}^{l_{us} \cdot r_{tv} \cdot l'_{iu} \cdot r'_{vj}} \right\} = \left\{ \prod_{t=1}^{m} \prod_{s=1}^{m} \prod_{v=1}^{m} \prod_{u=1}^{m} x_{st}^{l_{us} \cdot r_{tv} \cdot l'_{iu} \cdot r'_{vj}} \right\} = \left\{ \prod_{t=1}^{m} \prod_{s=1}^{m} x_{st}^{\sum_{v=1}^{m} \sum_{u=1}^{m} l_{us} \cdot r_{tv} \cdot l'_{iu} \cdot r'_{vj}} \right\} =$$

$$= \left\{ \prod_{t=1}^{m} \prod_{s=1}^{m} x_{st}^{\sum_{u=1}^{m} l_{us} \cdot l'_{iu} \cdot \sum_{v=1}^{m} r_{tv} \cdot r'_{vj}} \right\} = \left\{ \prod_{t=1}^{m} \prod_{s=1}^{m} x_{st}^{l''_{ik} \cdot r''_{tj}} \right\} = L'' \triangleright X \triangleleft R'' = \left(L^{-1}L\right) \triangleright X \triangleleft \left(RR^{-1}\right)$$

Thus we receive that

$$L^{-1} \triangleright \left(L \triangleright X \triangleleft R\right) \triangleleft R^{-1} = \left(L^{-1}L\right) \triangleright X \triangleleft \left(RR^{-1}\right). \qquad \textbf{(11)}$$

We used ordinary features of power function and matrix multiplication, so this equation holds for any matrices. □

Lemma 1. Defined matrix power operation has the following property: if key matrices are identities, then resulting matrix of ciphertext is equal to the matrix of plaintext.

Proof. Any element of ciphertext matrix can be written following way:

$$c_{ij} = t_{ij}^{l_{ii} \cdot rjj} \cdot \prod_{u=1}^{m} \prod_{\substack{v=1 \\ v \neq j \,\&\, u \neq i}}^{m} t_{vu}^{r_{iv} l_{uj}} .$$

If key matrices are $L = R = I$, then we have that $l_{ii} = r_{jj} = 1$ for any *i* and *j* from1 to *m*, and $l_{ij} = r_{jj} = 0$, if $i \neq j$:

$$c_{ij} = t_{ij}^{1 \cdot 1} \cdot \prod_{u=1}^{m} \prod_{\substack{v=1 \\ v \neq j \,\&\, u \neq i}}^{m} t_{vu}^{0} = t_{ij} . \qquad \square$$

Theorem 2. The action operations are both left and right invertible.

Proof. According to (11) and Lemma 1 we obtain:

$$L^{-1} \triangleright \left(L \triangleright X \triangleleft R\right) \triangleleft R^{-1} = \left(L^{-1}L\right) \triangleright X \triangleleft \left(RR^{-1}\right) = I \triangleright X \triangleleft I = X \,. \qquad \textbf{(12)} \ \square$$

Key matrices

Key matrices *L* and *R* should be invertible in order that defined matrix power S-box would be bijective, i.e. would have inverse S-box. This is obvious from the (12). Decryption of the ciphertext can be done only with L^{-1} and R^{-1} matrices. If *L* or *R* did not have inverse, then matrix power S-box is surjective and inverse S-box does not exist.

Matrices *L* and *R* are chosen from group $\mathbf{M_G}$, therefore they have inverse matrices. The problem is, how to construct group $\mathbf{M_G}$ that it would be large enough and brute force attacks would be useless.

One of the methods is to use the certain non-commutative group representation in the set of matrix group *GL*(*m*, *GF*(2^n)). The non-commutative group is presented by finite sets of generators and relations. Then it is required to construct representation matrices and their inverses for each initial group generator [Sakalauskas and Luksys, 2007].

Other method is to generate random matrices and to check if they are invertible. We have used this method to evaluate key space and matrix power S-box security properties.

For the analysis, we have chosen 3 x 3 matrices (m = 3) and n = 7. In this case key matrices L and R are over N_{127} and data matrices are over $GF(2^7)$. Irreducible polynomial was $x^7 + x + 1$.

Key matrices L and R have nine elements and each element can be randomly chosen from 127 values. Thus we can generate $127^9 \approx 2^{62.9}$ distinct matrices. But that would be all possible variants, including those matrices, which do not have inverse. We have generated 2^{34} matrices and 0.969% of those matrices were not invertible. Therefore rough estimate of matrix power key space would be around 2^{63} (for two key matrices).

Table 1. Cryptographic characteristics of 500 000 random invertible matrix power S-boxes

Group No.	Bijective	Algebraic degree	Nonlinearity	*k-uniform*	Algebraic quadratic equations immunity	Algebraic biaffine equations immunity	Percentage, %
1.	T	4	56	2	$2^{19,53}$	-	10,9
2.	T	6	54	2	$2^{15,63}$	$2^{12,68}$	5,5
3.	T	5	44	4	$2^{19,53}$	$2^{19,65}$	3,5
4.	T	5	44	4	$2^{19,53}$	$2^{19,44}$	1,8
5.	T	5	44	4	$2^{19,53}$	$2^{19,23}$	0,2
6.	T	5	44	6	$2^{13,84}$	2^{12}	10,9
7.	T	4	56	2	$2^{10,75}$	$2^{19,65}$	3,5
8.	T	4	56	2	$2^{10,75}$	$2^{19,44}$	1,8
9.	T	4	56	2	$2^{10,75}$	$2^{19,23}$	0,2
10.	T	4	56	2	$2^{11,72}$	2^{12}	11,0
11.	T	3	56	2	$2^{19,53}$	2^{999}	11,0
12.	T	3	44	4	$2^{19,53}$	$2^{19,65}$	3,6
13.	T	3	44	4	$2^{19,53}$	$2^{19,44}$	1,8
14.	T	3	44	4	$2^{19,53}$	$2^{19,23}$	0,2
15.	T	3	44	6	$2^{13,84}$	2^{12}	11,0
16.	T	2	56	2	$2^{10,75}$	$2^{19,65}$	3,5
17.	T	2	56	2	$2^{10,75}$	$2^{19,44}$	1,8
18.	T	2	56	2	$2^{10,75}$	$2^{19,23}$	0,2
19.	T	2	56	2	$2^{11,72}$	2^{12}	10,9
20.	T	1	0	128	$2^{7,814}$	$2^{6,34}$	5,5
21.	F	7	0	2	$2^{7,17}$	$2^{6,34}$	1,5
22.	F	7	0	2	$2^{7,135}$	$2^{6,294}$	0,1

It is very difficult to evaluate cryptographic characteristics of S-box with 54 bit input and 63 bit output. Therefore we have made two simplifications for the security analysis. First of all we did not do key addition (3). If data matrix had zero element (-s) that matrix was left unchanged. This let us to analyze S-box with equal input and output size. For the second simplification, we fixed all input matrix elements except one and analyzed only one particular element of the output matrix. This led us to the analysis of the S-box with input and output size of 7 bits.

We have chosen to evaluate five cryptographic characteristics: algebraic degree [Meier et al., 2004], nonlinearity, differential coefficient *k-uniform*, algebraic quadratic equations immunity and algebraic biaffine equations immunity [Courtois et al., 2005]. Algebraic immunity is calculated as follows:

$$\Gamma = \left(\frac{t}{n}\right)^{\left\lceil \frac{t}{r} \right\rceil},$$

where t is number of terms in algebraic normal form (ANF) of Boole function, n – number of variables, r – number of biaffine or quadratic equations.

We have generated 500 000 random invertible matrix power S-boxes. Analysis results are shown in Table 1.

We have grouped all generated S-boxes into 22 groups according their characteristics. Two groups of S-boxes are not bijective, i.e. the representation of one element of input matrix into one output matrix element is not bijective, but the whole matrix power S-box remains invertible. Group 20th represents S-boxes which performs a linear transformation. Characteristics of groups 1–19 are similar to those of ordinary power functions, like Gold, Kasami, Niho etc. [Cheon and Lee, 2004].

These are just preliminary results and further analysis of matrix power S-box should be done.

Conclusion

In this paper we have analyzed key depended S-box based on introduced matrix power operation. We have formulated and proven core properties of this operation.

Some preliminary cryptographic characteristics of constructed S-box are calculated. Characteristics of simplified version of matrix power S-box are similar to those of ordinary power functions, like Gold, Kasami, Niho etc.

Bibliography

[Canteaut and Videau, 2002] A. Canteaut and M. Videau. Degree of composition of highly nonlinear functions and applications to higher order differential cryptanalysis. Advances in Cryptology Eurocrypt'2002, Springer Verlag 2002.

[Cheon and Lee, 2004] J.H. Cheon, D.H. Lee. Resistance of S-boxes against Algebraic Attacks. Fast Software Encryp¬tion, LNCS 3017, pp. 83-94, Springer-Verlag, 2004.

[Courtois, 2005] N.T. Courtois. General Principles of Algebraic Attacks and New Design Criteria for Cipher Components. Advanced Encryption Standard – AES, LNCS 3373, pp. 67-83, 2005.

[Courtois et al., 2005] N.T. Courtois, B. Debraize and E. Garrido. On exact algebraic [non-]immunity of S-boxes based on power functions. Cryptology ePrint Archive: Report, no. 203 (2005), http://eprint.iacr.org/2005/203.

[Courtois and Pieprzyk, 2002] N.T. Courtois and J. Pieprzyk. Cryptanalysis of Block Ciphers with Overdefined Systems of Equations. Proceedings of Asiacrypt'2002, LNCS 2501, pp. 267-287, Springer-Verlag, 2002.

[Meier et al., 2004] W.Meier, E.Pasalic and C.Carlet. Algebraic Attacks and Decomposition of Boolean Functions. Advances in Cryptology - EUROCRYPT 2004, LNCS 3027, Springer Berlin / Heidelberg, 2004, pp. 474-491.

[Monico, 2002] C. Monico. Semirings and Semigroup actions in Public–Key Cryptography. PhD. thesis, University of Notre Dame, May 2002.

[Sakalauskas and Luksys, 2007] E.Sakalauskas and K.Luksys, Matrix Power S-Box Construction. Cryptology ePrint Archive: Report, no. 214 (2007), http://eprint.iacr.org/2007/214.

[Sakalauskas et al., 2007] E. Sakalauskas, P. Tvarijonas and A. Raulinaitis. Key Agreement Protocol (KAP) Using Conjugacy and Discrete Logarithm Problems in Group Representation Level, Informatica, Vol. 18, No. 1, 2007, pp. 115-124.

[Sakalauskas, 2005] E. Sakalauskas. One Digital Signature Scheme in Semimodule over Semiring, Informatica, vol. 16, no. 3, 2005, pp. 383-394.

[Schaumuller-Bichl, 1983] Schaumuller-Bichl. Cryptanalysis of the Data Encryption Standard by the Method of Formal Coding. Advances in Cryptology EUROCRYPT-1982, LNCS 149, Springer-Verlag, 1983, pp. 235-255.

[Shpilrain and Ushakov, 2005] V. Shpilrain and A. Ushakov. A new key exchange protocol based on the decomposition problem. Cryptology ePrint Archive: Report, no. 447 (2005), http://eprint.iacr.org/2005/447.

Authors' Information

Kestutis Luksys – PhD student, Kaunas University of Technology, Studentu st. 50-327A, Kaunas 51368, Lithuania; e-mail: kestutis.luksys@ktu.lt.

Petras Nefas – Dr., Head of Division, GRC of State Security Department, Lithuania;
e-mail: petras.nefas@gmail.com.