THE MATRIX POWER FUNCTION AND ITS APPLICATION TO BLOCK CIPHER S-BOX CONSTRUCTION

Eligijus Sakalauskas and Kestutis Luksys

Department of Applied Mathematics
Kaunas University of Technology
Studentu str. 50-327a, LT-51368 Kaunas, Lithuania
{Eligijus.Sakalauskas; Kestutis.Luksys}@ktu.lt

Received March 2010; revised August 2011

Abstract. In this paper, we are continuing the security investigation of the new S-box function based on so-called matrix power function proposed earlier by the same authors. Security is considered against algebraic cryptanalysis. We claim that constructed S-box function is a candidate to be a one-way function since we showed that its inversion is related with one known NP-complete problem, namely with a solution of the system of multivariate quadratic (MQ) polynomial equations. This problem is known as an MQ problem. We show that the total break of the cipher providing key recovery having one or several plaintext-ciphertext pairs and using algebraic cryptanalysis corresponds to the solution of the large system of MQ polynomial equations. This system is underdefined and not sparse. Hence, we show that proposed S-box has a provable security property since the solution of such system of equations is intractable. The guess and determine attack also fails since the guessing and computing sets of solutions are too large to be solvable effectively due to the underdefined nature of MQ equations.

Keywords: Matrix power function, One-way function, S-box, Block cipher, MQ problem

1. Introduction. The matrix power function (MPF) and the S-box function (SBF) based on it are presented for the block cipher construction. These functions were firstly introduced in [1].

The MPF is based on matrix powering by other matrices. This function is some generalization of a discrete exponent function in cyclic groups by its expansion in matrix set. The security assumption of MPF is quite different from ordinary discrete exponent function since it does not rely on the difficulty of discrete logarithm problem (DLP).

The MPF can be interpreted as some matrix group action in some other set of matrices. All matrices are quadratic and have the same size. The idea to use rings (semirings) action problem in modules (or vectorial spaces) for the asymmetric cryptographic primitives’ construction can be found in [2]. However, the security of this approach is based on the difficulty of some generalized DLP. In the partial case this generalized DLP becomes the classical DLP in cyclic groups.

Due to Shor results cryptographic systems which rely on DLP can be compromised in quantum polynomial time. However, for the effective attacks quantum computers are needed which are still under some theoretical and practical obscurity. However, the ideas of developing quantum algorithms are still of great interest [3-5], and there are already available some first examples of quantum computers on the public market. The development of quantum computing is a threat to many traditional cryptosystems. It also concerns the algorithms based on elliptic curves cryptosystems [6,7].

This inspired us to use a rather different approach for the construction of cryptographic primitives and as a consequence we constructed the matrix power function. This function...
represents the certain matrix group action in some matrix set in particular way and is the composition of certain left and right matrix group action functions. We show that the security of the matrix power S-box is based on the complexity of solution of system of multivariate polynomial equations, which is known to be in NP-complete complexity class. Hence, we are making a conjecture that proposed S-box can be a candidate one-way function (OWF).

The preliminary security analysis of constructed S-box was considered in [8]. This analysis was based on the mathematical modelling of simplified S-box version since the analysis for the working example is infeasible due to too high dimensions. The algebraic degree, nonlinearity, differential uniformity, algebraic quadratic equations immunity and algebraic biaffine equations immunity characteristics were analyzed. These characteristics are not absolutely compatible with our construction since matrix power S-box realizes injective mapping. However, we have found that they are similar to those of ordinary power functions, like Gold, Kasami and Niho.

In this paper we analyze the security of the S-box based on MPF against algebraic attack. This kind of attack was recently reapplied against the symmetric ciphers and especially against AES [9,10]. It is based on the construction of the system of algebraic equations relating the cipher’s input, secret key and output. If these equations could be solved then the secret key could be recovered and the cipher would be compromised having several (or even one) plaintext-ciphertext pairs. Since the relation between the input and output of AES cipher has a single non-linear function, then the system of algebraic equations, relating the input, key and output is considerably simple and it can be represented to be the system of multivariate quadratic (MQ) equations [9]. After this notice a lot of efforts were directed to analyze AES [11-13]. Despite these efforts, the attempts to create the ad-hoc methods for breaking the AES were unsuccessful. Nevertheless, the relative simplicity of systems of MQ equations is the remaining potential threat to AES. This stimulates further developments and there are made attempts to increase the complexity of AES by improving its S-box [14,15]. These improvements increase the algebraic complexity of the whole cipher together with better efficiency.

Due to algebraic cryptanalysis threat the design principles of symmetric ciphers have changed [16,17]. The increasing number of rounds does not protect against the algebraic attack in the same value as it protects against the traditional cryptanalytic attacks. Hence, the resistance against algebraic attack must be achieved by constructing S-boxes with as high as possible complexity level which is related to the S-box non-linearity. If S-box is highly complicated, then the number of rounds can be reduced to construct a secure block cipher.

Since we have proposed a considerable complicated S-box, the number of rounds of the block cipher using this S-box can be reduced significantly. Therefore, the investigation of this single S-box resistance to algebraic attack becomes sensible.

In Section 2, we formally introduce the MPF and SBF and define formal investigation of their resistance against algebraic attack. In Section 3, we introduce the basic algebraic structures and present the MPF construction on them. Some important properties of the MPF are proved. S-box construction for symmetric block cipher using the MPF is presented in Section 4. The resistance against algebraic attack is discussed in Section 5. Conclusions are given in Section 6.

2. The Definition of MPF and SBF. Formally, the matrix power function (MPF) $f$ is defined in the set of $m \times m$ matrices over the Galois field $GF(2^n)$ providing the mapping $f : GF(2^n)^{m^2} \to GF(2^n)^{m^2}$. The domain of $f$ is the arbitrary subset of $GF(2^n)^{m^2}$ consisting of matrices without zero entries and we denote it by $M$. We also assume
for simplicity that the range of \( f \) is the same, then \( f : M \to M \). According to our construction the MPF represents some matrix group \( M_G \) action in \( M \). Matrix group \( M_G \) is a set of \( m \times m \) matrices over the ring of integers \( \mathbb{Z}_{2^n-1} \). Hence, the matrix multiplication in \( M_G \) is defined and it consists of those matrices which have their inverses. Since \( M_G \) consists of \( m \times m \) matrices, then \( M_G \subset \mathbb{Z}_{2^n-1}^{m \times m} \). In order to define the MPF we must define the left and the right MPFs similar to the left and the right matrix multiplication operations. The left MPF \( f_{L^{-}} \) provides mapping \( f_{L^{-}} : M_G \times M \to M \) and the right MPF the mapping \( f_{R} : M \times M_G \to M \). Parameters \( L \) and \( R \) represent any matrices in matrix group \( M_G \) and reflect the fact that function \( f_{L^{-}} \) is defined by the left action of matrix \( L \) and \( f_{R} \) by the right action of matrix \( R \) in \( M \). In other words functions \( f_{L^{-}} \) and \( f_{R} \) can be interpreted also as left and right composition functions.

The MPF \( f \) we define as a composition \( \circ \) of functions \( f_{L^{-}} \) and \( f_{R} \), i.e.,

\[
f_{L,R} = f_{L^{-}} \circ f_{R}.
\]  

To define the S-box function (SBF) \( F \) firstly we must specify the S-box input/plaintext and output/ciphertext sets. Assuming that \( D \) is a plaintexts data set and \( C \) is a ciphertexts set, the set \( D \) is represented by \( m \times m \) matrices over \( GF(2^{n-1}) \), i.e., \( D = GF(2^{n-1})^{m^2} \) and \( C \subset GF(2^n)^{m^2} \). We are considering the SBF as performing autonomous ciphering operation, i.e., \( F : D \to C \). Hence, SBF is an injective mapping \( GF(2^{n-1})^{m^2} \to GF(2^n)^{m^2} \) guaranteeing the unique inverse mapping of \( F \) for decryption.

The SBF \( F \) is a composition of some auxiliary function \( g_K \) and MPF \( f_{L,R} \) with both defined by matrix \( K \in \mathbb{Z}_{2^n-1}^{m^2} \) and matrices \( L, R \in M_G \) correspondingly. These matrices represent the S-box secret key by certain expansion of the cipher’s secret key \( K_0 \). Then the ciphering procedure for some plaintext \( D \) and corresponding ciphertext \( C \) symbolically can be expressed in the way

\[
C = F(D) = f_{L,R}(g_K(D)).
\]  

The security of the created S-box against algebraic attack relies on the complexity of \( F \) inversion in (2). This means that having several plaintext and ciphertext pairs \( (D, C) \) we can recover the S-box’s secret key matrices \( L, R \) and \( K \) by inverting \( F \), i.e., by finding \( F^{-1} \). The exact meaning of this statement we present below, but symbolically the breaking of the S-box can be expressed by the following equation recovering the secret keys \( (L, R, K) \)

\[
(L, R, K) = F^{-1}(D, C).
\]  

We will show that the complexity of algebraic attack based on the inversion of the SBF \( F \) does not rely on some kind of DLP problem (generalized or ordinary) since the orders of the Galois fields are considerably small. According to our notation it is rather linked with so-called generalized matrix decomposition problem (DP). To our knowledge the first idea to use the most general form of DP for the public key cryptography in abstract non-commuting groups is presented in [18]. One kind of DP is used in the asymmetric cryptosystems based on the hard problems in infinite non-commutative groups, e.g., braid groups [19]. The other kind of DP is also used in digital signature scheme construction and key agreement protocols [20-22].

We present below some preliminary complexity analysis of this problem. Obtained results show that the SBF can be considered as a candidate OWF, since its inversion corresponds to NP-complete problem. This is related with the hardness of its inversion. If it is true then the presented S-box construction must be not vulnerable against the main traditional and algebraic cryptanalytic attacks. Moreover, it is a very desirable property of any S-box or entire cipher to be an OWF since according to several fundamental theoretical results the existence of OWF implies the existence of pseudorandom number.
generators and vice versa [23,24]. Then the S-box output will be undistinguishable from the pseudorandom sequence. Therefore, the attempts are made to obtain functions as similar as possible to OWFs. The good tools for its design are to use the problems in NP-complete complexity class.

According to informal definition, if computing the direct value \( F(D) \) in (2) is performed easy by some deterministic algorithm in polynomial time but computing the inverse value \( F^{-1}(D, C) \) in (3) is hard, then the SBF \( F \) has a feature to be a candidate OWF. We will show that the first condition is satisfied since \( F(D) \) can be computed by fast (polynomial-time) algorithm using table based computations in small fields.

The main purpose of our analysis is the investigation of complexity of \( F \) inversion. We have found that the complexity of inversion of our SBF relies on the complexity of solution of system of multivariate polynomial (MP) equations, i.e., on the MP problem complexity. It is known that MP problems and even problems of solution of system of multivariate quadratic equations are in NP-complete class over any field [25,26]. We will show that the system of equations relating S-box input, key and output are MP equations of the third order over the ring \( \mathbb{Z}_{2^n-1} \). Hence, adversary must solve this problem to determine the S-box key matrices by having one or several plaintext-ciphertext pairs \((D, C)\).

3. The Construction of MPF. We give the exact meaning to the matrix action functions \( f_{L-} \), \( f_{-R} \) and \( f_{L,R} \) now. According to the definitions above for \( \forall L, R \in \mathbb{M}_G \) and \( \forall X \in \mathbb{M} \) there exist some \( Y, Z \in M \) such that \( f_{L-}(X) = Y \) and \( f_{-R}(X) = Z \). The elements of matrices \( L, X, R, Y \) and \( Z \) we denote by the indexed sets of its elements respectively, e.g., we denote matrix \( X \) by \( \{x_{ij}\} \).

Then we define the left matrix action function \( f_{L-}(X) = Y \) which can be written for the elements of matrix \( Y = \{y_{ij}\} \) in the way

\[
y_{ij} = \prod_{s=1}^{m} x_{sij}^{l_{ij}}, \quad (4)
\]

and the right matrix action function \( f_{-R}(X) = Z \) for the elements of matrix \( Z = \{z_{ij}\} \) respectively

\[
z_{ij} = \prod_{t=1}^{m} x_{jit}^{r_{ij}}. \quad (5)
\]

Since the elements \( l_{is} \) and \( r_{tj} \) are in \( \mathbb{Z}_{2^n-1} \), i.e., they are integers, and \( x_{ij} \) are in \( GF(2^n) \) the powering and multiplication operations are performed using \( GF(2^n) \) field arithmetic.

Since both the left and the right action operations consist of matrix \( X \) elements powered by the entries of other matrices \( L \) and \( R \) in \( \mathbb{M}_G \), we have called them the left and the right matrix power functions (MPFs) and denoted them as \( f_{L-} \) and \( f_{-R} \) correspondingly. As it is seen these functions are uniquely defined by matrices \( L \) and \( R \) and provide mappings from \( \mathbb{M} \) to \( \mathbb{M} \). According to this mathematical meaning we can give symbolical notation of \( f_{L-} \) and \( f_{-R} \) as left and right matrix power functions:

\[
Y = f_{L-}(X) = L X, \quad (6)
\]

\[
Z = f_{-R}(X) = X^R. \quad (7)
\]

Hence, \( f_{L-} \) corresponds to matrix \( X \) left powered by matrix \( L \) and \( f_{-R} \) to \( X \) right powered by matrix \( R \). To give a simple example, let us assume that all matrices have two rows and two columns, i.e., \( m = 2 \). Then matrix \( Y \) can be expressed in the following way

\[
Y = \begin{pmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} l_{11} x_{11} + l_{12} x_{12} \\ l_{21} x_{21} + l_{22} x_{22} \end{pmatrix} = \begin{pmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}.
\]
and matrix $Z$ has the form

$$Z = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{pmatrix} = \begin{pmatrix} x_{11}^r r_{11} & x_{12}^r r_{12} \\ x_{21}^r r_{21} & x_{22}^r r_{22} \end{pmatrix}. $$

If matrix $X$ has no zero elements, then it is clearly seen that there will be no zeros in matrices $Y$ and $Z$.

We denote the multiplication of any two matrices $P$, $S$ in $\mathbf{M}_G$ by $P \cdot S$.

**Definition 3.1.** Operations $\cdot$ and $f_{L-}$ (respectively $\cdot$ and $f_{-R}$) are associative if

$$L_1 \left(L_2 X\right) = L_1 \cdot L_2 X, \quad (X R_1) R_2 = X R_1 \cdot R_2. $$

(8) (9)

**Theorem 3.1.** Operations $\cdot$ and $f_{L-}$ (respectively $\cdot$ and $f_{-R}$) are associative.

**Proof:** Let us denote the elements of matrices $L_1$ and $L_2$ by \{l'_{ij}\} and \{l''_{ij}\} respectively.

Then according to (4), following relations are true:

$$L_1 \left(L_2 X\right) = L_1 \left(\prod_{s=1}^{m} x_{ss}^{l''_{sij}}\right) = L_1 \left(\prod_{s=1}^{m} \left(\prod_{t=1}^{m} x_{ts}^{l'_{tsj}}\right)\right) = \prod_{s=1}^{m} \prod_{t=1}^{m} x_{st}^{l_{tsij}} = L_1 \cdot L_2 X, $$

since $\sum_{t=1}^{m} l'_{ts} l''_{st}$ corresponds to the element in the $i$th row and $s$th column of matrix $L_1 \cdot L_2$. Similarly, it can be shown that $\cdot$ and $f_{-R}$ are associative.

**Theorem 3.2.** The composition of left and right matrix power functions $f_{L-}$ and $f_{-R}$ is commuting, i.e.,

$$f_{-R} \left(f_{L-} \left(X\right)\right) = f_{L} \left(f_{-R} \left(X\right)\right). $$

(10)

**Proof:** Using (4) and (5), we obtain

$$f_{-R} \left(f_{L-} \left(X\right)\right) = f_{-R} \left(\prod_{s=1}^{m} x_{ss}^{l_{sij}}\right) = \prod_{t=1}^{m} \left(\prod_{s=1}^{m} x_{ts}^{l_{tsj}}\right) = \prod_{t=1}^{m} \prod_{s=1}^{m} x_{st}^{l_{tsij}} = f_{L} \left(\prod_{t=1}^{m} x_{st}^{r_{tsj}}\right) = f_{L-} \left(f_{-R} \left(X\right)\right). $$

**Corollary 3.1.** The left and right matrix power functions are associative, i.e.,

$$(L X)^R = L \left(X^R\right) = L X^R. $$

(11)

Then the composition $f_{LR}$ of the left and the right functions $f_{L-}$ and $f_{-R}$ in (1) can be denoted as $f$ since left and right MPFs are commuting which is reflected in (11). The function $f$ is our proposed MPF which we used to construct the SBF.

**4. The S-Box Construction.** As it was mentioned above, the S-box is represented by S-box function (SBF) $F$. The SBF input is the plaintext $D$ in $GF(2^n)^m$ and the output is the ciphertext $C$ in $GF(2^n)^m$. They satisfy the formal relation (2). Let the symmetric secret key be some bit string $K_0$. Then using certain key expansion algorithm we can generate the key matrix $K$ over $\mathbb{Z}_{2^{n-1}}$ and matrices $L, R \in \mathbf{M}_G$. All matrices $D, C, K, L$ and $R$ are square $m \times m$ matrices. The SBF $F$ has the key matrices $K, L, R$ and performs the injective mapping $F : GF(2^{n-1})^m \rightarrow GF(2^n)^m$ thus yielding ciphertext matrix $C$. 


This function includes the auxiliary function \( g_K \) and matrix power function \( f \). The SBF is defined by the following relations

\[
\begin{align*}
g_K(D) &= D + K + 1 = X, \quad (12) \\
f(g_K(D)) &= f(X) = L X^R = C. \quad (13)
\end{align*}
\]

The addition operations in (12) are ordinary addition operations of matrices. It is the additions of entries of matrices but they are defined according to the addition rules in \( \mathbb{Z}_{2^n} \). \( \mathbf{1} \) is the matrix in \( \mathbb{Z}_{2^n}^m \) consisting of arithmetical unity elements in all its positions. Combining (12) and (13) the SBF \( F \) has the following formal expression:

\[
F(D) = L(D + K + 1)^R = C. \quad (14)
\]

Using the operations defined in (12) we obtain a matrix \( X \in \mathbf{M} \) which does not contain zero elements in \( GF(2^n) \). The smallest possible element of \( \{x_{ij}\} \) is 1 and the largest one can be coded in numerical form as \( 2^n - 1 \). This condition must be necessarily satisfied since the presence of at least one zero element among \( \{x_{ij}\} \) will cause to gain the zero ciphertext matrix \( \mathbf{C} \).

Hence, now we can write the following formula for the element \( c_{ij} \):

\[
\prod_{t=1}^{m} \prod_{s=1}^{m} (d_{st} + k_{st} + 1)^{l_{st}r_{tj}} = \prod_{t=1}^{m} \prod_{s=1}^{m} x_{st}^{l_{st}r_{tj}} = c_{ij}, \quad i, j = 1, 2, \ldots, m, \quad (15)
\]

where \( 1 \) is a unity in \( \mathbb{Z}_{2^n} \).

Since \( \mathbf{M}_G \) is the group of matrices, there exist the inverse matrices \( R^{-1} \) and \( L^{-1} \) such that \( R R^{-1} = R^{-1} R = \mathbf{I} = L L^{-1} = L^{-1} L \), where \( \mathbf{I} \) is the identity diagonal matrix in \( \mathbb{Z}_{2^n}^{m^2} \).

Decryption operation can be written in a similar formal way as in (14):

\[
L^{-1} C^{R^{-1}} - K - 1 = D. \quad (16)
\]

The elements of input-plaintext data matrix \( D = \{d_{ij}\} \) can be expressed in the form

\[
\prod_{t=1}^{m} \prod_{s=1}^{m} x_{st}^{l_{st}r_{tj}'} - k_{ij} - 1 = d_{ij}, \quad i, j = 1, 2, \ldots, m, \quad (17)
\]

where \( \{l_{ij}'\} = L^{-1} \) and \( \{r_{ij}'\} = R^{-1} \). Thus, for decryption we have to be able to calculate the inverse key matrices of \( L \) and \( R \). The key matrix \( K \) remains the same but in decryption the subtraction is used instead of addition.

The decryption is valid since the left-right action operations satisfy the condition of Theorem 3.1 and matrices \( L \) and \( R \) are in matrix group \( \mathbf{M}_G \); thus they have their inverses. Then formally

\[
F'(C) = L^{-1} C^{R^{-1}} - K - 1 = L^{-1} (L(D + K + 1)^R)^{R^{-1}} - K - 1
\]

\[
= L^{-1} (D + K + 1)^{R R^{-1}} - K - 1 = D + K + 1 - K - 1 = D.
\]

We would like to point out that all arithmetical operations for the encryption/decryption are performed in relatively small fields and with small integers or short bit strings. Hence, they can be carried out using look-up tables.

5. Analysis of the S-Box Function’s One-Way Properties. As it is seen from (15) the calculation of the SBF’s direct value \( F(D) \) using the table multiplication and powering operations in field \( GF(2^n) \) are performed in polynomial time algorithm with respect to parameters \( m \) and \( n \), i.e., it would take \( m^4 \) multiplications and \( m^2 \) sums in the ring, \( m^4 \) exponentiations and \( (m^2 - 1)m^2 \) multiplications in the field or \( 3m^4 \) operations in total. Separating MPF part to the left and the right functions the total number of operations
can be reduced to $4m^3 - m^2$ operations. Thus the asymptotic computation time of SBF is $O(m^3n)$. Hence, function $F$ satisfies the first OWF condition defined in Section 2.

To estimate the complexity of $F$ inversion we must consider the complexity of system (15) solution to find the unknown key matrices $K$, $L$ and $R$ with given several (or even one) instance of plaintext-ciphertext pairs $(D, C)$. Since we do not know how to deal with this system of $m^2$ equations directly, we must transform it to more convenient form. Let us fix a generator $\alpha$ of multiplication group in $GF(2^n)$ which can be expressed as $n - 1$ degree polynomial over the field $\mathbb{Z}_2$

$$\alpha = 0 \cdot x^{n-1} + 0 \cdot x^{n-2} + \ldots + 0 \cdot x^2 + 1 \cdot x + 0 \cdot x^0.$$ (18)

Then we can take a discrete logarithm with the base $\alpha$ of both sides of each equation of (15). This yields the following system of $m^2$ equations for one plaintext-ciphertext pair

$$\sum_{t=1}^{m} \sum_{s=1}^{m} l_{is} r_{tj} \log_\alpha(d_{st} + k_{st} + 1) = \log_\alpha c_{ij}, \quad i, j = 1, 2, \ldots, m.$$ (19)

As we have mentioned above for the effective practical implementation we choose the sufficiently small field $GF(2^n)$, e.g., $n = 8$ or similar value. Then DLP can be solved effectively. Moreover, the $\log_\alpha c_{ij}$ can be computed even more effectively using the discrete logarithm table with only $2^{16}$ entries. This means that security of our S-box does not rely on the difficulty of DLP.

The expression $\log_\alpha(d_{st} + k_{st} + 1)$ cannot be computed since it contains the unknown key matrix elements $k_{st}$. Then denoting $\log_\alpha(d_{st} + k_{st} + 1) = z_{st}$ and $\log_\alpha c_{ij} = h_{ij}$ we obtain the following system of $m^2$ multivariate polynomial equations of degree 3 of each monomial

$$\sum_{t=1}^{m} \sum_{s=1}^{m} l_{is} r_{tj} z_{st} = h_{ij}, \quad i, j = 1, 2, \ldots, m.$$ (20)

This system with $m^2$ equations has $3m^2$ unknowns $\{l_{is}\}, \{r_{tj}\}, \{z_{st}\}$ and $m^4$ cubic monomials. Due to Fermat theorem it is defined over the ring. Moreover, the system is not sparse and overdefined as in the case of AES cipher equations and at first look no special methods as XL and XSL algorithms can be applied efficiently [9,27].

If matrix $Z$ is found then the recovery of the key $K$ can be performed by computing

$$\alpha^{z_{st}} - d_{st} - 1 = k_{st}.$$ (21)

The solution of the system (20) totally breaks the S-box. Hence, we must investigate the complexity of solving system (20) to estimate the S-box security.

Since the system (20) is underdefined, then there might be more than one solution of key matrices $K$, $L$ and $R$. If the set of possible solutions has a big cardinality with respect to the total scan area then it is possible to try to guess some solutions applying brute force attack. However, this approach does not break the cipher since finding any matrices $L'$, $R'$ and $Z'$ helps nothing. To break the cipher we must find the actual key matrices. We are not able either to encrypt or decrypt the plaintext or ciphertext validly with non-original $L'$, $R'$ and $Z'$ since all mappings in encryption/decryption must be one-to-one. This attack is of complexity $O(2^{3m^3-n})$.

In another way the attacker can try to mount the guess and determine attack. If any two matrices of $Z$, $L$ and $R$ are guessed, then the system (20) becomes a system of linear equations and the third matrix can be calculated in $O(m^6n)$ time. The whole attack is of complexity $O(2^{3m^3-n} m^6n)$. If we fix parameter $n = 8$, i.e., elements of matrices are bytes, then computational limit of $2^{80}$ operations would be exceeded with $m \geq 3$.

The current MP equations of (20) can be transformed to corresponding MQ problem in polynomial time by introducing additional variables. There are several different ways
to do it. We have chosen the way that the number of new monomials and new equations would be minimal. By introducing new variable \( u_{sti} = l_{is}z_{st} \) we have an underdefined system of MQ equations over \( \mathbb{Z}_2^{n-1} \):

\[
\sum_{t=1}^{m} l_{is}z_{st} = u_{sti}, \quad i, j, s, t = 1, 2, \ldots, m.
\]

Then the number of additional variables is equal to the number of terms, i.e., \( m^3 \). The same number of auxiliary quadratic equations is added to the initial system. In total we have \( m^2 + m^3 \) equations with \( 3m^2 + m^3 \) unknowns, \( m^3 + m^4 \) quadratic terms and \( m^3 \) linear terms.

As it is known the general decision MQ problem is in NP-complete class over any field. We are making a conjecture that the same is valid for the general decision MP problem. So far, the universal algorithm suitable for solution of general MP and MQ problems is Grobner basis algorithm when number of unknowns and equations is the same [28]. It is widely believed that for the random generated MQ problem over field \( GF(2) \) Grobner basis algorithm is impractical if the number of equations and unknowns is more than 80 [29].

Our system of Equations (22) cannot be reckoned as random generated since it has some predetermined structure influenced by rigorous rules of matrix power function. However, if key matrices are randomly generated with the uniform distribution, we do not know how to employ this structure to facilitate the solution of system (20).

The size of the system (22) depends only on the chosen parameter \( m \). When \( m = 4 \), the MQ system has 80 equations with 112 unknowns, i.e., the solution of such system is intractable.

There are known several methods for solving underdefined systems of MQ equations in polynomial time [30]. However, these methods work effectively only when the number of unknowns significantly exceeds the number of equations and it cannot be applied in our case.

The systems (20) and (22) are derived in the case of one plaintext-ciphertext pair. By adding additional pair the number of equations in system (20) would double, but the number of unknowns would not double. For each additional different pair the number of unknowns \( \{l_{is}\}, \{r_{tj}\} \) is the same and only the number of \( \{z_{st}\} \) is different. Thus analyzing \( p \) plaintext-ciphertext pairs we obtain \( pm^2 \) equations with \( 2m^2 \) unknowns \( \{l_{is}\} \) and \( \{r_{tj}\} \), \( pm^2 \) unknowns \( \{z_{st}\} \) and \( pm^4 \) cubic monomials in (20). Converting this system to the system of MQ equations we would get \( p(m^2 + m^3) \) equations with \( 2m^2 + p(m^2 + m^3) \) unknowns. Therefore, no matter how many plaintext-ciphertext pairs we analyze, the MQ system remains underdefined and the difference between the number of unknowns and equations remains \( 2m^2 \). This fact adds extra security against the attacks based on the large number of available plaintext-ciphertext pairs.

According to these conjectures we can define matrix power S-box security parameters and their values. Since the complexity of solution of initial system of Equations (15) depends on the number of equations \( m^2 \) and Galois field size \( 2^n \) then the meaningful security parameters are \( m \) and \( n \).

As it was pointed out above the solution of MQ system becomes intractable when \( m \geq 4 \). When \( m = 4 \), parameter \( n \) must be equal to 3 or greater for S-box resistance against guess and determine attack. Thus security parameters should meet these bounds: \( m \geq 4 \) and \( n \geq 3 \).

The complexity of resultant MQ system with respect to \( m \) and \( n \) allows us to make a conjecture that constructed S-box can be considered as a good candidate to be an OWF. Despite the fact that the existence of OWF is not proved yet, the conjectured OWFs
based on NP-complete or conjectured NP-complete problems have a great interest since they provide an output which is not distinguishable from the pseudorandom sequence [23,24].

6. Conclusions. The security investigation against algebraic cryptanalysis of the S-box function based on matrix power function proposed earlier by the same authors is presented [1]. It is shown that constructed S-box function is a candidate to be a one-way function since its inversion is related with one known NP-complete problem, namely with a solution of system of multivariate polynomial equations, i.e., MP problem.

Algebraic cryptanalysis yields the system of MP equations relating $m \times m$ matrices of plaintext, keys and ciphertext. The derived system of equations showed that determination of the S-box key matrices having one or several plaintext-ciphertext pairs $(D, C)$ corresponds to the solutions of underdefined MP system of the third order equations over $\mathbb{Z}_{2^n-1}$. These features essentially worsen the solution of the obtained equations system since neither classical, nor special methods such as Grobner basis algorithm and XL, XSL methods cannot be applied efficiently.

Performed analysis allows us making a conjecture that the proposed S-box has a provable security property since its inversion is based on NP-complete problem and hence that matrix power S-box can be reckoned as a candidate one-way function. If the S-box function is close to one-way function then its output (ciphertext) is indistinguishable from the output obtained from pseudorandom sequence [23]. Hence, both the S-box and the entire cipher functions must be as close to OWF as possible.

The S-box security parameters are defined as being $n$ and $m$. It is assumed that their secure values are $n \geq 3$ and $m \geq 4$. Compliance with these limits ensures that brute force as well as guess and determine attacks are infeasible.

The practical implementation of the S-box realization requires the $4m^3 - m^2$ operations in $\mathbb{Z}_{2^n-1}$ and in $GF(2^n)$. The software realization is fast since all operations can be implemented with the look-up tables consisting of $2^{2n}$ entries. For example, if $m = 4$ and $n = 8$, then input data matrix is of 112 bits size and look-up tables consist of $2^{16}$ entries.

Acknowledgment. This work was partially supported by the Lithuanian State Science and Study Foundation. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES


