Circulant matrices over $\mathbb{F}_2$ and their use for construction of efficient linear transformations with high branch number

Stepan Davydov, Yuri Shkuratov

JSRPC "Kryptonite"

CTCrypt 2023
Table of contents

1. Introduction

2. Definitions and preliminaries

3. Linear transformations and their software implementation

4. Matrix decomposition into a sum of matrices $A_{a(x),f(x)}$

5. Decomposition of the circulant matrices over $\mathbb{F}_{2^s}$
Introduction
Linear transformations are used to construct block ciphers and hash functions.

High **branch numbers** of the linear transformation matrix and its transpose are needed to protect against **differential** and **linear** methods of cryptanalysis.
It is possible to construct MDS matrices using the following classes of matrices:

- **Cauchy** matrices (used in STREEBOG hash function);
- **Vandermonde** matrices;
- **recursive** (also named **serial**) matrices (used in PHOTON hash function, KUZNYECHIK block cipher);
- **Hadamard** matrices;
- **circulant** matrices (search methods, used in AES block cipher, SM4 block cipher, WHIRLPOOL hash function);
- etc.
Efficiency of using circulant matrices

We consider linear transformations, defined by multiplication in the ring $R = \mathbb{F}_2[x]/f(x)$.

Advantages of our approach:

✓ Software implementation is reduced to small count of the processor instructions usage (thanks to the use $CLMUL$ instruction set).

✓ Software implementation requires small amount of memory, much less than LUT-tables.

This class generalize the class of circulant matrices over $\mathbb{F}_2$:

\[
\text{Circulant matrices over } \mathbb{F}_2 \subset \text{Matrices of multiplication in ring } R = \mathbb{F}_2[x]/f(x)
\]
Definitions and preliminaries
Basic definitions

Let $Q$ be a field $\mathbb{F}_{2^s}$.

**Definition**

The *weight* of $\vec{a} \in Q^m$, denoted $wt(\vec{a})$, is the number of nonzero coordinates of $\vec{a}$.

**Definition**

*Branch number* of matrix $A \in Q_{m,m}$ is the following number:

$$\tau(A) = \min_{\vec{a} \neq \vec{0}} [wt(\vec{a}) + wt(\vec{a}A)].$$

It is obviously that $\tau(A) \leq m + 1$ and $\tau(A) = \tau(A^{-1})$, if $A^{-1}$ exists.

**Definition**

If $\tau(A) = m + 1$, $A$ is *Maximum Distance Separable* (further *MDS*) matrix.
Definition

Let $P = \mathbb{F}_2$ be the field of two elements, $Q = (P[x]/g(x), +, \cdot)$ and $g(x)$ be irreducible polynomial of degree $s$ over $P$. Let $B_{m \times m}$ be a matrix over $Q$, which transforms vectors from $Q^m$. Since elements of $Q$ are row vectors over $P$, it is possible to consider $B$ as linear transformation of row vectors of length $n = ms$ over $P$ and there exist corresponding matrix $A_{n \times n}$ over $P$.

*In such case we said:* matrix $A = A(B, g(x))$ *implements* linear transformation $B$ *on binary vectors.*
Basic definitions

Let $P$ be a field $\mathbb{F}_2$ and $\vec{a} \in P^{ms}$. We split $\vec{a}$ into $s$-subvectors: subvector $\vec{a}(i, s)$ with number $i$ is subvector of length $s$ equal to

$$(a_{i(s+1)}, a_{i(s+2)}, \ldots, a_{is}), \quad i \in \{0, \ldots, m-1\}.$$

Then

$$\vec{a} = (\vec{a}(m-1, s), \ldots, \vec{a}(0, s)).$$

**Definition**

$s$-weight of vector $\vec{a} \in P^{ms}$, denoted $wt_s(\vec{a})$, is the number of nonzero $s$-subvectors of vector $\vec{a}$.

**Definition**

Branch number on $s$-subvectors of matrix $A \in P_{ms,ms}$ is the following number:

$$\tau_s(A) = \min_{\vec{a} \in P^{ms} \setminus \{0\}} [wt_s(\vec{a}) + wt_s(\vec{a}A)].$$
**Remark**

Let \( f(x) \) be polynomial of degree \( n \) over, \( P_n[x] = P[x]/f(x) \) be the polynomial ring over \( P \) with addition and multiplication modulo \( f(x) \). Note that \( P_n[x] \) is vector space of dimension \( n \) over \( P \). There exist isomorphic mapping between \( P^n \) and \( P_n[x] \):

\[
\varphi(a_{n-1}, \ldots, a_1, a_0) = a_{n-1}x^{n-1} + \ldots + a_1x + a_0
\]

Further we will equate row vectors of length \( n \) with corresponding polynomials from \( P_n[x] \).
Linear transformations and their software implementation
Let $P$ be a field $\mathbb{F}_2$. We consider the following operations on bit strings, which are implemented on computers as processor instructions:

1. $\text{XOR}(\vec{\alpha},\vec{\beta})$ is bitwise addition of strings modulo 2.
2. $\text{AND}(\vec{\alpha},\vec{\beta})$ is bitwise conjunction of strings.
3. $\text{OR}(\vec{\alpha},\vec{\beta})$ is bitwise disjunction of strings.
4. $\text{SHFT}(\vec{\alpha})$ is left (right) shift of the string by $i$ positions with zero padding.
5. $\text{CLMUL}(\vec{\alpha},\vec{\beta})$ is multiplication of binary strings of length $n$ as polynomials of degree $n - 1$ over $P$. The result is a string of length $2n$. 

Definition

Let $f(x)$ be a polynomial of degree $n$ over $P$. Linear transformation, which corresponds to multiplication by an element $a(x)$ of the ring $R = P[x]/f(x)$, is the following transformation:

$$\hat{a}_f(x) : h(x) \rightarrow h(x)a(x) \mod f(x), \ h(x) \in R$$

The linear transformation matrix has the form:

$$A_{a(x), f(x)} = \begin{pmatrix} \hat{a}_f(x)(x^{n-1}) \\ \vdots \\ \hat{a}_f(x)(x^i) \\ \vdots \\ \hat{a}_f(x)(x) \\ \hat{a}_f(x)(1) \end{pmatrix} = \begin{pmatrix} a(x) \cdot x^{n-1} \mod f(x) \\ \vdots \\ a(x) \cdot x^i \mod f(x) \\ \vdots \\ a(x) \cdot x \mod f(x) \\ a(x) \end{pmatrix}$$
## Implementation of linear transformation

### Statement 1

Let $f(x) = x^n + f_{n-1}x^{n-1} + ... + f_0 = x^n + \overline{f(x)}$ be a polynomial of degree $n$ over $P$, $a(x)$ be a polynomial of degree less than $n$ over $P$. Then the following statements are true for the transformation $\hat{a} = \hat{a}_f(x)$:

1. If $\deg \overline{f(x)} \leq n/2$, then transformation $\hat{a}$ can be implemented in 5 processor instructions: 3 $\text{CLMUL}$ + 2 $\text{XOR}$.

2. If $\deg \overline{f(x)} + \deg \overline{a(x)} \leq n$, then transformation $\hat{a}$ can be implemented in 3 processor instructions: 2 $\text{CLMUL}$ + 1 $\text{XOR}$.

3. If $\deg \overline{f(x)} = 0$, then transformation $\hat{a}$ can be implemented in 2 processor instructions: 1 $\text{CLMUL}$ + 1 $\text{XOR}$.

4. To implement the transformation $\hat{a}$, it is necessary to store the polynomials $a(x)$ and $\overline{f(x)}$ in memory in cases 1-2, and only the polynomial $a(x)$ in case 3.
Features of circulant matrices implementation

The circulant matrix looks like this:

\[ C_{n \times n} = \text{Circ}(c_{n-1}, ..., c_0) = \begin{pmatrix} c_0 & c_{n-1} & \ldots & c_2 & c_1 \\ c_1 & c_0 & \ldots & c_3 & c_2 \\ \vdots \\ c_{n-2} & c_{n-3} & \ldots & c_0 & c_{n-1} \\ c_{n-1} & c_{n-2} & \ldots & c_1 & c_0 \end{pmatrix} \]

Statement 2

Let \( f(x) = x^n + 1 \) be a polynomial over \( P \), \( \hat{a} = \hat{a}_{f(x)} \). Then:

1. Matrix of the linear transformation \( \hat{a} \) is circulant matrix over \( P \).
2. Branch numbers on \( s \)-subvectors of the matrices \( A \) and \( A^T \) are the same.
3. If \( n \) is even and the transformation \( \hat{a} \) is an involution, then for any \( s \geq 1 \) the branch number on \( s \)-subvectors of matrix \( A_{a(x), f(x)} \) does not exceed 4.
Transformations with the following maximum branch numbers on \(s\)-subvectors have been founded by enumeration on computers among transformations of the form \(A_{a(x), x^n+1}\):

<table>
<thead>
<tr>
<th>Matrix size</th>
<th>(s)-subvector size</th>
<th>4-bit</th>
<th>6-bit</th>
<th>8-bit</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 \times 4</td>
<td>4-bit</td>
<td>5 (MDS)</td>
<td>5 (MDS)</td>
<td>5 (MDS)</td>
</tr>
<tr>
<td>6 \times 6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>8 \times 8</td>
<td>7</td>
<td>-</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>16 \times 16</td>
<td>12</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Matrix decomposition into a sum of matrices $A_{a(x), f(x)}$
Let $A \in P_{n \times n}$, $f(x) = x^n + f_{n-1}x^{n-1} + \ldots + f_1x + 1$ be polynomial over $P$, $a_i(x)$ be polynomials over $P$ of degree less than $n$, $i \in 1, t$.

We consider the following decomposition:

$$A = \sum_{i=1}^{t} D_i A_i,$$

where $D_i = \text{diag}_{n \times n}(d_{i,n-1}, \ldots, d_{i,0})$, $d_{i,j} \in \{0, 1\}$, $A_i = A_{a_i(x), f(x)}$.

**Remark**

Multiplication by matrices $D_i$ is implemented by instruction $\text{AND}$, by matrices $A_i$ – according Statement 1. Sum is implemented by instruction $\text{XOR}$.
Since \( f(0) = 1 \), there exist matrix \( A_{x,f(x)}^{-1} \):

\[
A_{x,f(x)}^{-1} = \begin{pmatrix}
\chi^{n-2} \\
\vdots \\
\chi^{i-1} \\
\vdots \\
1 \\
(x^{-1} \mod f(x))
\end{pmatrix}
\]

**Definition**

Let \( \text{Rev}_{f(x)} : P_{n,n} \to P_{n,n} \) be transformation, which result on matrix \( A \) is matrix \( B \) such as every row \( \vec{B}_i = \vec{A}_i \cdot A_{x,f(x)}^{-i} \).

**Theorem**

The minimum number of summands \( t \) in the decomposition of matrix \( A \) is equal to \textit{rank} of the matrix \( B = \text{Rev}_{f(x)}(A) \).
Probabilistic relations in matrix rows

Let $A \in P_{n \times n}$. We consider the set of the vectors:

$$\overrightarrow{\Omega}_j = (\overrightarrow{A}_j \parallel 0) + (0 \parallel \overrightarrow{A}_{j+1}), \quad j \in 0, n - 2$$

of length $n + 1$ over $P$. Due to the decomposition of matrix $A$ we obtain vector that $\overrightarrow{\Omega}_j$ is equal to:

$$\overrightarrow{\Omega}_j = \sum_{i=1}^{t} d_{i,j}(\overrightarrow{A}_{i,j} \parallel 0) + \sum_{i=1}^{t} d_{i,j+1}(0 \parallel \overrightarrow{A}_{i,j+1})$$

**Probability space** $\Theta$: let all $d_{i,j}$ and all coefficients of the polynomials $a_i(x)$ be mutually independent random variables with a uniform distribution on $P = \mathbb{F}_2$.

In case of probability space $\Theta$ matrix $A$ is **random matrix** defined by its decomposition.
<table>
<thead>
<tr>
<th><strong>Theorem</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Let probability space $\Theta$ be defined, $A$ be $n \times n$ random matrix defined by the decomposition with $t$ summands. Then for matrix $A$ any $\Omega_j$ equals $\vec{f}$ with probability:</td>
</tr>
<tr>
<td>$\Pr(\Omega_j = \vec{f}) \geq \frac{2^t - 1}{2^{2t+1}}$,</td>
</tr>
<tr>
<td>where $\vec{f}$ is vector of coefficients of the polynomial $f(x)$.</td>
</tr>
</tbody>
</table>
Decomposition of the circulant matrices over $\mathbb{F}_{2^s}$
Decomposition of the circulant matrices over $\mathbb{F}_{2^s}$

Let $P$ be a field $\mathbb{F}_2$ and $Q = (P[x]/g(x), +, \cdot)$ with some irreducible polynomial $g(x)$ of degree $s$ over $P$, $Q \cong \mathbb{F}_{2^s}$, $f(x) = x^n + 1$.

Statement 3

Let $C = C_{m \times m}$ be circulant matrix over $Q$, $n = ms$ and matrix $A_{n \times n} = A(C, g(x))$ implements corresponding $C$ transformation on binary vectors of length $n$. Then:

1. There exist decomposition for matrix $A$ and polynomial $x^n + 1$, which consists of no more than $s$ summands.

2. If binary representation of any element of matrix $C$ contains $s - k$ zeros in most significant bits, then there exist decomposition for matrix $A$, which consists of no more $k$ summands.
Matrix decomposition into a sum of matrices $A_{a(x),f(x)}$

**Definition**

Let $\alpha$ be byte, $\alpha = (\alpha_7, ..., \alpha_0)$, then

$\text{Diag}_{m \times m}(0x\alpha) = \text{diag}_{8m \times 8m}(\alpha_7, ..., \alpha_0, ..., \alpha_7, ..., \alpha_0)$
Example 1 (*Whirlpool*)

Matrix $A(W, g(x))$ is used in the linear transformation of *Whirlpool* hash function, where $W$ is $8 \times 8$ MDS circulant matrix over $\mathbb{F}_{2^8}$ and $g(x) = x^8 + x^4 + x^3 + x^2 + 1$.

$$W = \text{Circ}_{2^8}(0x01, 0x04, 0x01, 0x08, 0x05, 0x02, 0x09, 0x01).$$

Matrix $A(W, g(x))$ decomposition consists of *four* summands:

$$A(W, g(x)) = \text{Circ}_2(0x01, 0x04, 0x01, 0x08, 0x05, 0x02, 0x09, 0x01) +$$

$$+ \text{Diag}(0x20)\text{Circ}_2(0x00, 0x00, 0x00, 0x08, 0xe8, 0x00, 0x08, 0xe8) +$$

$$+ \text{Diag}(0x40)\text{Circ}_2(0x00, 0x04, 0x74, 0x08, 0xec, 0x74, 0x08, 0xe8) +$$

$$+ \text{Diag}(0x80)\text{Circ}_2(0x00, 0x04, 0x74, 0x08, 0xec, 0x76, 0x32, 0xe8).$$
Some examples

Example 2 (Alternative to Whirlpool matrix)

Let $g(x) = x^8 + x^4 + x^3 + x^2 + 1$. Then the matrix

$V = \text{Circ}_{2^8}(0x01, 0x02, 0x03, 0x05, 0x04, 0x03, 0x07, 0x07)$

is also an $8 \times 8$ MDS circulant matrix over $\mathbb{F}_{2^8}$ and there exist matrix $A(V, g(x))$ decomposition, which consists of three summands:

$A(V, g(x)) = \text{Circ}_2(0x01, 0x02, 0x03, 0x05, 0x04, 0x03, 0x07, 0x07) +
\text{Diag}(0x40)\text{Circ}_2(0x74, 0x00, 0x00, 0x04, 0x70, 0x74, 0x04, 0x70) +
\text{Diag}(0x80)\text{Circ}_2(0x4e, 0x02, 0x38, 0x3e, 0x70, 0x76, 0x3c, 0x48)$. 
Some examples

Example 3 (AES)
Matrix $A(L, g(x))$ is used in the linear transformation of AES block cipher, where $L = \text{Circ}_{28}(0x03, 0x01, 0x01, 0x02)$ is $4 \times 4$ MDS circulant matrix over $\mathbb{F}_{2^8}$, $g(x) = x^8 + x^4 + x^3 + x + 1$. Matrix $A(L, g(x))$ decomposition consists of two summands:

$$
A(L, g(x)) = \text{Circ}_2(0x03, 0x01, 0x01, 0x02) + \\
+ \text{Diag}(0x80)\text{Circ}_2(0x34, 0x36, 0x00, 0x02).
$$

Example 4
There exist $4 \times 4$ MDS matrix on $8$ – subvectors over $\mathbb{F}_2$:

$$
L' = \text{Circ}_2(0x01, 0x04, 0x04, 0x05).
$$
## Statement 4

Let decomposition

\[ A = \sum_{i=1}^{t} D_i A_i, \]

where \( D_i = \text{diag}_{n \times n}(d_{i,n-1}, \ldots, d_{i,0}), \) \( d_{i,j} \in \{0, 1\}, \) \( A_i = A_{a_i(x), x^{n+1}} \)

holds for matrix \( A \) and polynomial \( x^n + 1. \) Then multiplication by matrix \( A \) can be implemented by \( t \) instructions \( \text{AND}, \) \( t \) instructions \( \text{CLMUL} \) and \( 2t - 1 \) instructions \( \text{XOR}. \)
Thanks for your attention!

The report was prepared by:

- **Stepan Davydov:**
  JSRPC “Kryptonite”, s.davydov@kryptonite.ru
- **Yuri Shkuratov:**
  JSRPC “Kryptonite”, y.shkuratov@kryptonite.ru