КРИПТОНИТ

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

Stepan Davydov, Yuri Shkuratov

JSRPC "Kryptonite"

CTCrypt 2023

Table of contents

Introduction

Branch number importance

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 cryptanalisys.

MDS matrices over $\mathbb{F}_{2^s}$

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 then LUT-tables.

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

Circulant matrices over $\mathbb{F}_2$ $\subset$ *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

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

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:

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

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 $P_n[x]$:

$$\varphi(a_{n-1}, ..., a_1, a_0) = a_{n-1}x^{n-1} + ... + a_1x + a_0$$

Further we will equate row vectors of length n with corresponding polynomials from $P_n[x]$.

Linear transformations and their software implementation

String operations as processor instructions

Let P be a field $\mathbb{F}_2$. We consider the following operations on bit strings, which are implemented on computers as a processor instructions:

1. *XOR*$(\vec{\alpha}, \vec{\beta})$ is bitwise addition of strings modulo 2.
2. *AND*$(\vec{\alpha}, \vec{\beta})$ is bitwise conjunction of strings.
3. *OR*$(\vec{\alpha}, \vec{\beta})$ is bitwise disjunction of strings.
4. *SHFT*$(\vec{\alpha})$ is left (right) shift of the string by i positions with zero padding.
5. *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 $2n$.

Multiplication by an element of the ring

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:

$$\hat{a}_{f(x)} : h(x) \to h(x)a(x) \bmod f(x), \; h(x) \in R$$

The linear transformation matrix has the form:

$$A_{a(x),f(x)} = \begin{pmatrix} \hat{a}_{f(x)}(x^{n-1}) \\ \ldots \\ \hat{a}_{f(x)}(x^i) \\ \ldots \\ \hat{a}_{f(x)}(x) \\ \hat{a}_{f(x)}(1) \end{pmatrix} = \begin{pmatrix} a(x) \cdot x^{n-1} \bmod f(x) \\ \ldots \\ a(x) \cdot x^i \bmod f(x) \\ \ldots \\ a(x) \cdot x \bmod f(x) \\ a(x) \end{pmatrix}$$

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

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

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 $A_{a(x),x^n+1}$:

Matrix size \ s-subvector size	4-bit	6-bit	8-bit
4×4	5 (MDS)	5 (MDS)	5 (MDS)
6×6	6	6	6
8×8	7	-	8
16×16	12	-	-

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

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

Let $A \in P_{n\times n}$, $f(x) = x^n + f_{n-1}x^{n-1} + \ldots + f_1 x + 1$ be polynomial over P, $a_i(x)$ be polynomials over P of degree less than n, $i \in \overline{1,t}$.

We consider the following *decomposition*:

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

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

Remark

Multiplication by matrices D_i is implemented by instruction AND, by matrices A_i – according Statement 1. Sum is implemented by instruction XOR.

Number of summands in matrix decomposition

Since $f(0) = 1$, there exist matrix $A^{-1}_{x,f(x)}$:

$$A^{-1}_{x,f(x)} = \begin{pmatrix} x^{n-2} \\ \dots \\ x^{i-1} \\ \dots \\ 1 \\ x^{-1} \mod f(x) \end{pmatrix}$$

Definition

Let $Rev_{f(x)} : P_{n,n} \to P_{n,n}$ be transformation, which result on matrix A is matrix B such as every row $\vec{B_i} = \overrightarrow{A_i} \cdot A^{-i}_{x,f(x)}$.

Theorem

The minimum number of summands t in the decomposition of matrix A is equal to rank of the matrix $B = Rev_{f(x)}(A)$.

Probabilistic relations in matrix rows

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

$$\overrightarrow{\Omega_j} = (\overrightarrow{A_j} \parallel 0) + (0 \parallel \overrightarrow{A_{j+1}}), j \in \overline{0, n-2}$$

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 Θ: 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 Θ matrix A is *random matrix* defined by its decomposition.

Probabilistic relations in matrix rows

Theorem

Let probability space Θ be defined, A be $n \times n$ random matrix defined by the decomposition with t summands. Then for matrix A any $\overrightarrow{\Omega_j}$ equals $\vec{f}$ with probability:

$$\Pr(\overrightarrow{\Omega_j} = \vec{f}) \geq \frac{2^t - 1}{2^{2t+1}},$$

where $\vec{f}$ is vector of coefficients of the polynomial $f(x)$.

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

$Diag_{m\times m}(0x\alpha) = diag_{8m\times 8m}(\alpha_7, ..., \alpha_0, ..., \alpha_7, ..., \alpha_0)$

Some examples

Example 1 (*Whirlpool*)

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

$$W = Circ_{2^8}(0x01, 0x04, 0x01, 0x08, 0x05, 0x02, 0x09, 0x01).$$

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

$$A(W, g(x)) = Circ_2(0x01, 0x04, 0x01, 0x08, 0x05, 0x02, 0x09, 0x01)+$$
$$+ Diag(0x20)Circ_2(0x00, 0x00, 0x00, 0x08, 0xe8, 0x00, 0x08, 0xe8)+$$
$$+ Diag(0x40)Circ_2(0x00, 0x04, 0x74, 0x08, 0xec, 0x74, 0x08, 0xe8)+$$
$$+ Diag(0x80)Circ_2(0x00, 0x04, 0x74, 0x08, 0xec, 0x76, 0x32, 0xe8).$$

Some examples

Example 2 *(Alternative to Whirlpool matrix)*

Let $g(x) = x^8 + x^4 + x^3 + x^2 + 1$. Then the matrix $V = Circ_{2^8}(0x01, 0x02, 0x03, 0x05, 0x04, 0x03, 0x07, 0x07)$ is also 8×8 MDS circulant matrix over $\mathbb{F}_{2^8}$ and there exist matrix $A(V, g(x))$ decomposition, which consists of *three* summands:

$$A(V, g(x)) = Circ_2(0x01, 0x02, 0x03, 0x05, 0x04, 0x03, 0x07, 0x07) +$$
$$+ Diag(0x40) Circ_2(0x74, 0x00, 0x00, 0x04, 0x70, 0x74, 0x04, 0x70) +$$
$$+ Diag(0x80) Circ_2(0x4e, 0x02, 0x38, 0x3e, 0x70, 0x76, 0x3c, 0x48).$$

Some examples

Example 3 *(AES)*

Matrix $A(L, g(x))$ is used in the linear transformation of *AES* block cipher, where $L = Circ_{2^8}(0x03, 0x01, 0x01, 0x02)$ is 4×4 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:

$$A(L, g(x)) = Circ_2(0x03, 0x01, 0x01, 0x02) + \\ + Diag(0x80)Circ_2(0x34, 0x36, 0x00, 0x02).$$

Example 4

There exist 4×4 MDS matrix on $8 - subvectors$ over $\mathbb{F}_2$:

$$L' = Circ_2(0x01, 0x04, 0x04, 0x05).$$

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

Statement 4

Let *decomposition*

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

where $D_i = diag_{n \times n}(d_{i,n-1}, ..., d_{i,0}),\ d_{i,j} \in \{0, 1\},\ A_i = A_{a_i(x), x^n+1}$

holds for matrix A and polynomial $x^n + 1$. Then multiplication by matrix A can be implenented by t instructions AND, t instructions $CLMUL$ and $2t - 1$ instructions XOR.

КРИПТОНИТ

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