New Method for Designing Binary 8×8 Matrices With Branch Number 5 Using Quasigroups

Vesna Dimitrova
Faculty of Computer Science and Engineering
Ss. Cyril and Methodius University, Skopje
Republic of N. Macedonia
email: vesna.dimitrova@finki.ukim.mk

Verica Bakeva
Faculty of Computer Science and Engineering
Ss. Cyril and Methodius University, Skopje
Republic of N. Macedonia
email: verica.bakeva@finki.ukim.mk

Zlatka Trajcheska
Faculty of Computer Science and Engineering
Ss. Cyril and Methodius University, Skopje
Republic of N. Macedonia
email: zlatka.trajcheska@gmail.com

***Abstract*—Quasigroups are algebraic structures, which are useful for application in cryptography and coding theory. Their specific properties and Boolean representations open a lot of scientific questions and new ideas for research. In this paper, we investigate the application of Boolean representation of quasigroups. We propose a new method for designing of binary matrices of order 8×8, which have the highest branch number.**

Keywords–quasigroup; Boolean functions; branch number; binary matrices.

I. INTRODUCTION

We start by providing definitions of quasigroups, their Boolean representation and their properties.

The groupoid $(Q, *)$, where $*$ is a binary operation, is called a quasigroup if it satisfies the following:

$$(\forall\, a, b \in Q)(\exists!\, x, y \in Q)(x * a = b \wedge a * y = b) \qquad (1)$$

meaning that the equations $x*a = b$ and $a*y = b$ have unique solutions for any $a, b \in Q$. These simple algebraic structures are suitable for application in cryptography, especially because of their large, exponentially growing number and their properties.

In this paper, we will consider the quasigroups of order 4. Their total number is 576, but not all of them are suitable for cryptographic purposes. Therefore, classifications of finite quasigroups are very important for choosing good quasigroups for designing cryptographic primitives. There are several classifications of quasigroups of order 4, for example, in [1][2].

Here, quasigroups are numbered according to their lexicographic ordering. Namely, this ordering is made such that each quasigroup is presented as an array of n^2 symbols, obtained by concatenation of the rows of the Latin square that represents the quasigroup operation. After that, the sorting of quasigroups is done by lexicographic ordering of the obtained arrays. Finally, a number is assigned to each quasigroup of the ordering starting with 1, and increasing by 1 sequentially until the last quasigroup is assigned a number.

Each quasigroup of order 2^n can be represented as a vector valued Boolean function [3][4]. It is done so that each element from the quasigroup $x \in Q$ can be represented as a binary vector $x = (x_1, x_2, ..., x_n) \in \{0,1\}^n$. In short, x is presented as a vector of n binary digits, which are its binary representation. Now, if we consider two elements from the quasigroup $x, y \in Q$ with their vector representations $x = (x_1, x_2, ..., x_n)$ and $y = (x_{n+1}, x_{n+2}, ..., x_{2n})$ the quasigroup operation can be presented as:

$$x * y \equiv f(x_1, ..., x_{2n}) = (f_1(x_1, ..., x_{2n}), ..., f_n(x_1, ..., x_{2n}))$$

where

$$f_i : \{0,1\}^{2n} \to \{0,1\}$$

are the components of the vector valued Boolean function f.

The quasigroups of order 4 are represented with pair of Boolean functions $(f_1(x_1, x_2, x_3, x_4), f_2(x_1, x_2, x_3, x_4))$.

A quasigroup $(Q, *)$ with Boolean representation

$$f(x_1, ..., x_{2n}) = (f_1(x_1, ..., x_{2n}), ..., f_n(x_1, ..., x_{2n})) \qquad (2)$$

is linear by Boolean representation if f_i is a linear Boolean polynomial for each $i = 1, 2, 3, ..., n$. The quasigroup is generally called nonlinear if there is at least one nonlinear f_i for $i = 1, 2, 3, ..., n$. The quasigroup is pure nonlinear if f_i is a nonlinear Boolean polynomial for each $i = 1, 2, 3, ..., n$.

A important property that we will strongly consider in the further text is the linearity by Boolean representation. The classification of order 4 by linearity was previously done in [5]. According to this, from 576 quasigroups of order 4, 144 are linear quasigroups and 432 are nonlinear quasigroups (144 of them are pure nonlinear).

Lets assume that binary matrices with suitable properties are also important for designing cryptographic primitives. There are several constructions of binary matrices with certain properties. In [6], the authors give an efficient way for generating circulant binary matrices with a prescribed number of ones which are invertible over $\mathbb{Z}_2$.

In [7], the authors investigate all binary matrices of order 8×8 and come to the conclusion that the Hamming weight of all matrices with branch number 5 varies from 33 to 44.

Our goal in this research is the construction of binary matrices of order 8×8 with linear and differential branch numbers 5 using Boolean representations of quasigroups, which have the maximal Hamming weight.

The rest of the paper is organized as follows. In Section II, we give definitions of linear and differential branch number. The new method for designing of branch number is given in Section III. Section IV presents the conclusions and ideas for future work.

 ISBN: 978-1-61208-732-0

II. BRANCH NUMBER

In this paper, we give a new method for constructing 8×8 nonsingular matrices with branch number 5. In [8], Kang has proved that the branch number of any 8×8 invertible binary matrix is less than or equal to 5, so the 8×8 binary matrices with branch number 5 are optimal. It is known that the diffusion layers of Camellia in [9] and E2, which are 8×8 binary matrices, have branch number 5. Kanda et al. in [10] found 10080 8×8 binary matrices with branch number 5 by using a searching algorithm, and for all candidate matrices, the total Hamming weight was 44 with 4 column (row) vectors with Hamming weight 6 and 4 column (row) vectors with Hamming weight 5.

At first, we give some definitions and principles in order to introduce the branch number of matrices.

Confusion in cryptography is a principle that indicates the lack of clarity in the relation between the plaintext and ciphertext. In the ciphers, this means that the key is not related to the ciphertext in a simple manner. It is usually made by substitution. Blocks that are used in ciphers for substitution are S-boxes. S-boxes are transformation units, which take m bits as input and give n bits as output. They are usually implemented with a lookup table [11][12].

Diffusion in cryptography means that, by changing of a single bit in the plaintext, approximately half of the bits in the ciphertext should be changed. It is usually implemented with a permutation of symbols.

Ciphers that have confusion and diffusion layer are called Substitution-Permutation Networks (SPNs). We are interested in the diffusion layer in order to apply quasigroups there.

Definition 1: [13] If a block cipher has n S-boxes in its structure, where each S-box has input and output of m bits, then the diffusion layer can be represented as:

$$A : (\{0,1\}^m)^n \longrightarrow (\{0,1\}^m)^n \tag{3}$$

or with this linear transformation:

$$A(x) = A \cdot x^T = \begin{bmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & ... & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \tag{4}$$

where $a_i \in \{0,1\}^m$, $x = (x_1, x_2, ..., x_n)$, $x_i \in \{0,1\}^m$, $i = 1, 2, ..., n$.

From this point on, we will only consider binary matrices and binary vectors.

The Hamming weight of a binary vector x is denoted by $wt(x)$ and represents the number of non-null components in x.

There are two different branch numbers – linear and differential. As their names indicate, one represents the resistance to the linear, and the other to the differential cryptanalysis.

Definition 2: Let A be a binary matrix of order $n \times n$.

$i)$ The linear branch number of A is defined by:

$$\beta_l(A) = \min\{wt(x)+wt(A^T \cdot x^T) | x \in \{0,1\}^n,\ x \neq 0\}. \tag{5}$$

$ii)$ The differential branch number of A is defined by:

$$\beta_d(A) = \min\{wt(x)+wt(A \cdot x^T) | x \in \{0,1\}^n,\ x \neq 0\}. \tag{6}$$

The design blocks in the ciphers should have good linear and differential properties, which means that the values of both branch numbers are high.

In our research, we consider only nonsingular binary matrices since the encryption and decryption are inverse processes.

Example 1: Let us calculate the branch number of the matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$. We consider all not-null binary vectors of order 2:

$$x_1 = (0,1), \quad x_2 = (1,0), \quad x_3 = (1,1).$$

Their Hamming weights are, respectively:

$$wt(x_1) = 1, \quad wt(x_2) = 1, \quad wt(x_3) = 2.$$

The products $y_i = A \cdot x_i^T$, $i = 1, 2, 3$ are

$$y_1 = (1,1), \quad y_2 = (0,1), \quad y_3 = (1,0),$$

whose Hamming weights are

$$wt(y_1) = 2, \quad wt(y_2) = 1, \quad wt(y_3) = 1,$$

respectively. The values of $\beta_i^{(d)} = wt(x_i) + wt(y_i)$, $i = 1, 2, 3$ are

$$\beta_1^{(d)} = 3, \quad \beta_2^{(d)} = 2, \quad \beta_3^{(d)} = 3.$$

The minimal value of $\beta_i^{(d)}$ $(i = 1, 2, 3)$ is the differential branch number, in this case $\beta_d(A) = 2$.

The linear branch number is calculated similarly.

III. CONSTRUCTING BINARY MATRICES 8×8 WITH BRANCH NUMBER 5

Our goal in this research is the construction of binary matrices of order 8×8 with linear and differential branch numbers 5 using Boolean representations of quasigroups. Namely, the maximal branch number for a binary matrix of order 8×8 is 5. These matrices are important in block ciphers and they are used in the design of a few ciphers in the lightweight cryptography, for example Camellia.

Before explaining the method of construction, we will divide the linear quasigroups into subclasses. Firstly, we choose a quasigroup that is linear by Boolean representation and take the algebraic normal forms of their Boolean functions. Since the quasigroup is linear by Boolean representation, the Boolean functions are also linear, i.e.,

$$f_j(x_1, x_2, x_3, x_4) = a_{j0} + a_{j1}x_1 + a_{j2}x_2 + a_{j3}x_3 + a_{j4}x_4,$$

where $a_{ji} \in \{0,1\}$, $i = 0, 1, 2, 3, 4$, $j = 1, 2$. Firstly, we discard the constants a_{j0} and obtain two linear polynomials of the Boolean representation

$$f_j(x_1, x_2, x_3, x_4) = a_{j1}x_1 + a_{j2}x_2 + a_{j3}x_3 + a_{j4}x_4,$$

for $j = 1, 2$. This way, the class of 144 linear quasigroups can be divided into 36 subclasses, each containing 4 quasigroups whose Boolean representations differ only by a constant $(const = 1)$:

- $Q_i(f_1, f_2)$
- $Q_j(f_1 + const, f_2)$
- $Q_k(f_1, f_2 + const)$
- $Q_l(f_1 + const, f_2 + const)$

In each subclass, all 4 quasigroups produce the same binary matrix. Therefore, the constants a_{j0} do not have influence on the final results and can be discarded. Further on, we will

use one representative quasigroup from each subclass. The lexicographic numbers of quasigroups in each subclass and corresponding representative are given in Table I.

TABLE I. SUBCLASSES OF LINEAR QUASIGROUPS AND REPRESENTATIVE.

No.	Quasigroups in the subclass	Representative
1	1, 172, 405, 576	1
2	4, 169, 408, 573	4
3	11, 189, 388, 566	11
4	14, 192, 385, 563	14
5	21, 179, 398, 556	21
6	24, 182, 395, 553	24
7	26, 147, 430, 551	26
8	27, 146, 431, 550	27
9	37, 163, 414, 540	37
10	40, 166, 411, 537	40
11	43, 157, 420, 534	43
12	46, 160, 417, 531	46
13	51, 246, 331, 526	51
14	54, 243, 334, 523	54
15	57, 259, 318, 520	57
16	60, 262, 315, 517	60
17	70, 252, 325, 507	70
18	71, 253, 324, 506	71
19	77, 272, 305, 500	77
20	80, 269, 308, 497	80
21	82, 284, 293, 495	82
22	83, 285, 292, 494	83
23	92, 274, 303, 485	92
24	93, 275, 302, 484	93
25	100, 197, 380, 477	100
26	101, 196, 381, 477	101
27	110, 212, 365, 467	110
28	111, 213, 364, 466	111
29	113, 203, 374, 464	113
30	116, 206, 371, 461	116
31	126, 223, 354, 451	126
32	127, 222, 355, 450	127
33	132, 234, 343, 445	132
34	133, 235, 342, 444	133
35	138, 228, 349, 439	138
36	139, 229, 348, 438	139

Let us explain the method of construction of the binary matrices. We choose two linear quasigroups by Boolean representation from different subclasses. Then, we take the corresponding linear polynomials based on the algebraic normal form of the four (two by two) Boolean functions that represent the chosen quasigroups. Let us denote these linear polynomials provided from the first quasigroup by $Q_1 : f_1$ and $Q_1 : f_2$, and from the second quasigroup by $Q_2 : f_1$ and $Q_2 : f_2$. Firstly, we fill two matrices A_1 and A_2 of order 8×4 and after that we form the matrix A (of order 8×8) as concatenation by rows. In the following, we give the method of construction.

- For each not-null a_{1i} in $Q_1 : f_1$, the cell in the row 1 and column i in A_1 is filled with 1, and for each null a_{1i} the corresponding cell is filled with 0 ($i \in \{1, 2, 3, 4\}$). Then, we fill the third row in A_1 by shifting the bits from the first row on left with offset 1, the fifth row by shifting the bits from the third row on left with offset 1 and the seventh row by shifting the bits from the fifth row on left also with offset 1.

- For each not-null a_{1i} in $Q_2 : f_1$, the cell in the row 2 and column i in A_1 is filled with 1, and for each null a_i the same cell is filled with 0 ($i \in \{1, 2, 3, 4\}$). Then, we fill the fourth row by shifting the bits from the second row on left with offset 1, the sixth row by shifting the bits from the fourth row on left with offset 1 and the eight row by shifting the bits from the sixth row on left also with offset 1.

- The same method is applied for construction of matrix A_2 using the polynomials $Q_1 : f_2$ (for odd rows) and $Q_2 : f_2$ (for even rows).

- We form the matrix $A = [A_1|A_2]$.

This construction will be denoted as $Q_1f_1 - Q_1f_2 - Q_2f_1 - Q_2f_2$, referring to the order of the Boolean functions that are used in the construction. A graphical presentation of the explained construction using the quasigroups with lexicographic numbers 4 and 14 is given in Figure 1.

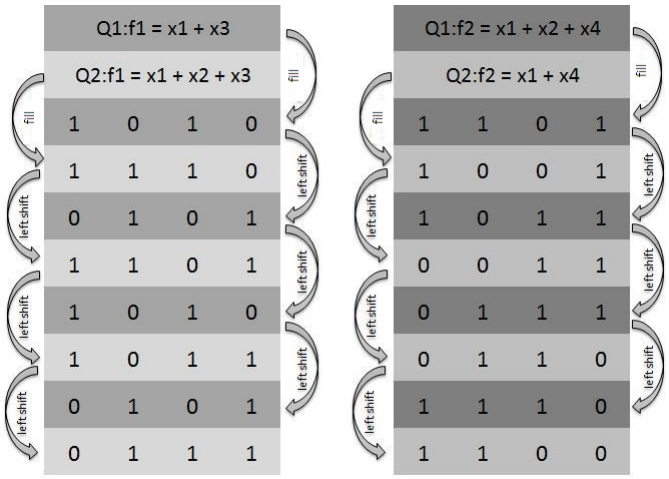

Figure 1. An example of binary matrix constructed using the Boolean function obtained from quasigroups 4 and 14.

Using this method, we obtain 384 nonsingular matrices such that:

- 160 matrices have branch number 3
- 192 matrices have branch number 4 and
- 32 matrices have branch number 5.

All 32 matrices with branch number 5 obtained by our method have the Hamming weight of 44, which is the maximal weight according to [7].

In our investigation, we consider a similar construction where the right shift is used instead of the left shift. The results were very similar and the distribution of matrices based on their branch number was the same. Also, the construction $Q_1f_1 - Q_2f_1 - Q_1f_2 - Q_2f_2$ was analyzed, either using left and right shift, and the results also were the same.

The complete results can be found in [14].

IV. CONCLUSION

Quasigroups are algebraic structures, which are useful for application in cryptography and coding theory. In this paper, using quasigroups of order 4, we propose a new way of constructing binary matrices of order 8×8 with branch number 5. These matrices are very useful for designing cryptographic primitives, especially in the field of lightweight cryptography.

Our research opens some questions that will be subjects for future investigation. Some of them are the following:

- Check if there is something specific and try to find a pattern in the matrices, which would lead to better

 ISBN: 978-1-61208-732-0

understanding of their design or simplifying their generation.

- Check all possible combinations of the constructions and try to find a theoretical dependency between them. It is very important to check why all constructions lead to the same results and if these results depend on the quasigroup properties in some way.

ACKNOWLEDGMENT

This research was partially supported by Faculty of Computer Science and Engineering at "Ss Cyril and Methodius" University in Skopje, Republic of N. Macedonia.

REFERENCES

[1] V. Dimitrova and S. Markovski, "Classification of Quasigroups by Image Patterns", Proceedings of 5th International Conference for Informatics and Information Technology, Bitola, Macedonia, pp. 152–160, 2007.

[2] V. A. Artamonov, S. Chakrabarti and S. K. Pal, "Characterization of polynomially complete quasigroups based on Latin squares for cryptographic transformations", Discrete Applied Mathematics, vol. 200, pp. 5–17, 2016.

[3] D. Gligoroski, V. Dimitrova and S. Markovski, "Quasigroups as Boolean functions, their equation systems and Groebner bases", In: Groebner Bases, Coding, and Cryptography, Springer 2009, pp. 415-420, 2009.

[4] I.S.M. Sala, "An algebraic description of Boolean functions", Proceedings of WCC07, pp. 343–349, 2007.

[5] S. Markovski, V. Bakeva, V. Dimitrova and A. Popovska-Mitrovikj, "Representation of algebraic structures by Boolean functions and its applications", In: D. Trajanov, V. Bakeva (eds.): ICT-Innovations 2017, Data-Driven Innovation, Communications in Computer and Information Science Series (CCIS) Vol.778, Springer, Cham, pp. 229–23, 2017.

[6] T. Fabsic, O. Grosek, K. Nemoga and P. Zajac, "On generating invertible circulant binary matrices with a prescribed number of ones", Cryptography and Communications, vol. 10, no. 1, pp. 159–175, 2018.

[7] Y. Gao and G. Guo, "Unified approach to construct 8x8 binary matrices with branch number 5", In: Proceedings of the 2010 First ACIS International Symposium on Cryptography, and Network Security, Data Mining and Knowledge Discovery, E-Commerce and Its Applications, and Embedded Systems, CDEE '10, IEEE Computer Society, Washington, DC, USA, pp. 413–416, 2010.

[8] J. Kang, "Practical and provable security against differential and linear cryptanalysis for substitution-permutation networks", ETRI journal, vol. 23, no. 4, pp. 158-167, 2001.

[9] K. Aoki et al., "Camellia: A 128-bit block cipher suitable for multiple platforms - design and analysis", In: D. R. Stinson, S. Tavares (eds.) SAC 2000. LNCS, vol. 2012, Springer, Heindelberg, pp. 39–56, 2001.

[10] M. Kanda, Y. Takashima, T. Matsumoto, K. Aoki and K. Ohta, "A strategy for constructing fast round functions with practical security against differential and linear cryptanalysis", Selected Areas in Cryptography, LNCS 1556, pp. 264–279, 1999.

[11] I. Vergili and M. Yucel, "Avalanche and bit independence properties for the ensembles of randomly chosen n×n S-boxes", Turkish Journal of Electrical Engineering and Computer Sciences, vol. 9, no. 2, pp. 161–176, 2001.

[12] D. Loebenberger and M. Nsken, "A family of 6-to-4-bit S-boxes with large linear branch number", Cryptology ePrint Archive, Report 2013/188, pp. 1–11, 2013.

[13] D. Kwon, S. H. Sung, J. H. Song and S. Park, "Design of block ciphers and coding theory", Trends in Mathematics", Information Center for Mathematical Sciences, vol. 8, no. 1, pp. 13–20, 2005.

[14] https://goo.gl/vbtvuM, Last access: 22.07.2019

 ISBN: 978-1-61208-732-0