computing@computingonline.net
www.computingonline.net

Print ISSN 1727-6209
On-line ISSN 2312-5381
International Journal of Computing

# OPPORTUNITIES TO MINIMIZE HARDWARE AND SOFTWARE COSTS FOR IMPLEMENTING BOOLEAN FUNCTIONS IN STREAM CIPHERS

**Alexandr Kuznetsov [1,2], Oleksandr Potii [1,2], Nikolay Poluyanenko [1,2], Serhii Ihnatenko [3], Igor Stelnyk [4], Danylo Mialkovsky [4]**

[1] V. N. Karazin Kharkiv National University, Svobody sq., 6, Kharkiv, 61022, Ukraine
[2] JSC "Institute of Information Technologies", Bakulin St., 12, Kharkiv, 61166, Ukraine,
kuznetsov@karazin.ua, potav@ua.fm, nlfsr01@gmail.com
[3] Security Service of Ukraine, Volodymyrska str., 33, Kyiv-1, 01601, Ukraine,
mongol_1979@ukr.net
[4] Department of Information Protection Administration of the State Service of Special Communication and Information Protection of Ukraine, Solomianska 13 str., Kyiv, 03680, Ukraine,
stelnik_i@i.ua, mdv@dsszzi.gov.ua

**Paper history:**
Received 14 January 2019
Received in revised form 25 November 2019
Accepted 20 December 2019
Available online 31 December 2019

**Keywords:**
stream ciphers;
pseudo-random sequence generators;
NLFSR;
nonlinear feedback shift register;
Boolean functions;
de Bruijn sequence;
linear complexity;
autocorrelation function;
cryptographic analysis;
nonlinear polynomials.

**Abstract:** Currently, nonlinear Boolean functions are actively investigated worldwide. However, many questions remain unanswered. The theory of nonlinear Boolean functions that are suitable for use in cryptographically strong algorithms is significantly incomplete. Despite the existence of numerous publications on these themes, many issues related to the interconnection of design characteristics affecting the generator's performance and its cryptographic characteristics still remain unsolved. The possibility of generating a special type of sequence, called de Bruijn sequence, at minimal hardware and software costs to implement nonlinear Boolean functions in stream encryption systems, is the main subject of this work. The paper presents the possible structure boundaries (algebraic degree of a Boolean function, the number of monomials in a function) of iterative de Bruijn sequence bitrate generators for various generated sequence characteristics, such as linear complexity and autocorrelation function. The profile of the linear complexity of the studied sequences is close to the expected value of the linear complexity, as well as for a truly random sequence.



## 1. INTRODUCTION

The analysis of the following modern stream cypher schemes: SNOW 2.0 [1], Decim [2], KCipher-2 [3], Sosemanuk [4], Grain [5], MICKEY 2.0 [6] and Trivium [7] shows that the main components are the iterative bitrate generators, as well as the function of complication that forms an output unit from several combinations of inner state bits.

Iterative bitrate stream generators are usually formed based on Linear Feedback Shift Register (LFSR). Their main function is to guarantee the uniqueness of the generator's inner state for a rather long period of time while it is working and make sure that good local statistical properties are provided. The registers that form the de Bruijn sequence meet these requirements. However, such sequence can also be generated by Nonlinear Feedback Shift Register (NLFSR). Applying NLFSR as iterative generators allows taking all the advantages of using LFSR, while omitting the main disadvantage of NLFSR – linearity.

Apart from that, when making crypto systems with LFSR new problems regarding the selection of nonlinear register arise. The LFSR theory is well studied [8], NLFSRs are much less researched that LFSR [9]. The first algorithm for building the smallest NLFSR of a given binary sequence was presented by Jansen in 1991 [10, 11]. Alternative algorithms were given in [12], as well as in [13], and [14].

The way to form a maximum period LFSRs is clear, their feedback functions correspond to primitive polynomials for $F_2$. Generally, it is unclear how to construct all NLFSR with the maximum period. The main method is to find such registers with the corresponding properties. Typically, all nonlinear feedback registers referenced in literature have a rather complex structure and consist of several structural elements. If LFSR is of a simple structure, then it is a small size register.

There are no general methods for designing maximum period NLFSR [8], [15]. The formation of a special class NLFSR with a maximum period was presented by Mykkeltveit and others in [16]. The work of E. Dubrova [15] contains an example of a Galois shift register with the size of $L=100$ cells that generates a sequence with a maximum period, but this sequence does not have the property of the de Bruijn sequence, i.e., some tuples of $L$ bits appear more than once in sequences.

At the same time, with the size increase in the register used, its constructive complexity increases. If it is possible to form the LFSR forming the de Bruijn sequence of the required size, then its structure will be so complex that its implementation in encryption systems, as an iterative generator, will be unacceptably resource-consuming.

The complexity of the design is directly related to the size and cost of the hardware implementation of the cipher that uses it, and in most cases affects its performance. If the structural complexity of a function is smaller, its circuit implementation is simpler. This is especially true for the so-called lightweight (or little-resources-consuming) cryptography.

On the other hand, if searching initially for a LFSR with the given design features, in particular, ease of implementation, the register found may be vulnerable to certain types of attacks, to neutralize which one would need to introduce additional units into the algorithm and, as a result, increase the overall resource consumption of the entire scheme.

Thus, it is reasonable to ask about the possibility of optimizing the structure of the LFSR between the simplicity of the hardware/software implementation and correspondence with some of the specified properties of the generated sequence. The simplicity of implementation suggests the number of operations necessary and sufficient to calculate the next state of an iterative generator.

The work presents the obtained results that reflect the interrelation of the constructive characteristics of the LFSR (such as the maximum algebraic degree, the number of monomials) and some necessary cryptographic properties of the sequence it forms (autocorrelation and linear complexity).

## 2. RESULTS

### 2.1 BOOLEAN FUNCTIONS

Some definitions given below are used further in the paper:

$F_2$ – a finite field of two elements, 0 and 1. Operations in $F_2$ - multiplication and exclusive disjunction addition.

$V_L$ $L$ - dimensional vector space over a field $F_2$, $V_L=(F_2)^L$. Addition in $V_L$ space bitwise exclusive disjunction.

*A Boolean function of* $L$ *variables* is a mapping from $V_L$ to $F_2$. Extended Boolean functions are mappings from $V_L$ to $Z$ (set of integers). Even more common pseudo-Boolean functions are mappings from $V_L$ to $R$ (a set of real numbers).

The number of space $V_L$ vectors is equal to $2^L$, the number of all possible combinations $L$ of basis vectors with coefficients 0 and 1.

If $Z_2=\{0,1\}$, $Z_2^L$ will denote the set of all binary vectors $x=(x_1,\ldots,x_L)$ of the length $L$. It is assumed that all vectors are lexicographically ordered.

A random function of mapping from set $Z_2^L$ to set $Z_2$ is called a *Boolean function of* $L$ *variables*.

Each Boolean function of $L$ variables can be uniquely determined by the vector of its length $2^L$. For example, the functions $f$ and $g$ correspond to the vectors $(1001)$ and $(00010110)$. Further $f$ denotes the vector of $2^L$ length of the function $f$. It is assumed that the function arguments (i.e., the $L$-length vectors) are sorted out in lexicographical order.

The $\oplus$ denotes exclusive disjunction addition (XOR operation). It is known that each Boolean function can be uniquely defined by its *algebraic normal form* (ANF); the ANF is also known as the Zhegalkin polynomial.

ANF is an expression of a Boolean function:

$$f(x_1,\ldots x_L)=\bigoplus_{N\in P\{1,2,\ldots,L\}} a_N\prod_{i\in N}x_i\,, \qquad (1)$$

where $P\{1,2,\ldots L\}$ is the set of all subsets $\{1,2,\ldots L\}$ (Boolean), $a_N\in F_2$.

To calculate the ANF of a given function, there are simple algorithms (see, for example, [17]).

*The degree of a monomial* (a Boolean monomial) $x^N = \prod_{i \in N} x_i$ is defined as $|N|$ (the number of elements of the subset $N$).

The *algebraic degree* $\deg(f)$ or the *nonlinearity order* of a Boolean function $f$ is the number of variables in the longest addend (monomial) of its ANF. A Boolean function of the 1 degree is called *affine*. Its ANF is as follows: $f(x) = a_1 x_1 \oplus a_2 x_2 \oplus \cdots \oplus a_L x_L \oplus b = \langle a, x \rangle \oplus b$, where $b \in F_2, a \in V_L$. A function is called quadratic, cubic, etc., if its algebraic degree is 2, 3, etc., respectively.

## 2.2 DE BRUIJN SEQUENCE

Let $A = a_1, a_2, \ldots, a_{2^L-1}, a_{2^L}$, be a sequence of $2^L$ length from the elements of the $\{0,1\}$ alphabet.

Let $A = a_1, a_2, \ldots, a_{2^L-1}, a_{2^L}$ be a sequence of $2^L$ length from the elements of the $\{0,1\}$ alphabet. A is called a *de Bruijn sequence* of $L$ order if, among all tuples of length $L$: $L\,(a_1, a_2, \ldots, a_L)$, $(a_2, a_3, \ldots, a_{L+1})$, …, $(a_{2^L-L+1}, a_{2^L-L+2}, \ldots, a_{2^L})$, each of the possible tuples is present and occurs exactly once, i.e., there are various $2^L$ combinations over the $\{0,1\}$ alphabet [18].

Similar $2^L - 1$ length sequences without a tuple that contain only zeros are called *modified de Bruijn sequences*.

It is easy to notice that one can get a modified de Bruijn sequence from the de Bruijn sequence by leaving out one zero, as well as *vice versa* by adding one.

Similar sequences can also be constructed for the alphabet of $k$ elements. For example, 0011 and 002212011 are de Bruijn sequences of $L = 2$ order over alphabets $\{0,1\}$ and $\{0,1,2\}$ respectively. De Bruijn sequences are used in cryptography due to their proper statistical properties as indistinguishable, in a statistical sense, from truly random sequences.

In 1894, S. Fly Santé Marie [19] and de Bruijn in 1946 [20] proved the existence of such sequences for any natural number $L$ for any alphabet of $k$ elements and showed that the number of different sequences $(B_n)$ is equal to

$$B_n = [(k-1)!]^{k^{L-1}} k^{k^{L-1}-L}, \quad (2)$$

(two sequences are considered different, if any of them cannot be obtained by cyclically shifting the other). For the case of Boolean functions, i.e., $k = 2$, the expression (2) takes the following form:

$$B_n = 2^{2^{L-1}-L}. \quad (3)$$

To assess the obtained results, Table 1 shows the quantitative values of the number of different de Bruijn sequences obtained in accordance with the expression (2)

$$B_n = [(k-1)!]^{k^{L-1}} k^{k^{L-1}-L}.$$

**Table 1. Number of different de Bruijn Sequences for given $L$ and $k$**

| $L$ | $k$ | | |
|---|---|---|---|
| | 2 | 3 | 4 |
| 2 | 1 | 24 | 20 736 |
| 3 | 2 | 373 248 | $\approx 1.8 \cdot 10^{20}$ |
| 4 | 16 | $\approx 1.2 \cdot 10^{19}$ | $\approx 8.4 \cdot 10^{85}$ |
| 5 | 2 048 | $\approx 4.4 \cdot 10^{60}$ | $\approx 2.1 \cdot 10^{350}$ |
| 6 | 67 108 864 | $\approx 1.7 \cdot 10^{186}$ | $\approx 5.3 \cdot 10^{1409}$ |
| 7 | $2^{57}$ | $\approx 1 \cdot 10^{1698}$ | $\approx 1.4 \cdot 10^{5649}$ |
| 8 | $2^{120}$ | $\approx 1 \cdot 10^{1698}$ | $6^{16384} \cdot 4^{16376}$ |
| 9 | $2^{247}$ | $\approx 1.4 \cdot 10^{5101}$ | $6^{65536} \cdot 4^{65527}$ |

As it can be seen, even with $L = 7$ and $k = 2$ the number of different de Bruijn sequences acquires the size that complicates processing (searching, storing and analysing) of the whole set.

It can be added that the number of primitive polynomials of the $L$ degree is equal to

$$\frac{\varphi(2^L - 1)}{L}, \quad (4)$$

where $\varphi$ is the Euler function.

The review [21] gives the most complete insight into the theory of de Bruijn sequences and the history of their use for various problems solving.

Those NLFSR, that implement those Boolean functions forming the modified de Bruijn sequences will be called *M-NLFSR*. If these functions are linear, then the corresponding registers will be called M-LFSR. In general, M-LFSR is a particular case of M-NLFSR.

The period of the de Bruijn sequence $(T_{PB})$ is determined by the size of the alphabet or, as it is also called, the basis of the de Bruijn sequence $k$, as

well as the digit capacity of the states (the number of memory cells of the iterative generator) – $L$:

$$T_{PB} = k^L. \quad (5)$$

And, correspondingly, the period of the modified de Bruijn sequence $(T)$ will be defined as:

$$T = k^L - 1. \quad (6)$$

## 2.3 MAXIMUM ALGEBRAIC ANF DEGREE OF DE BRUIJN SEQUENCES

The maximum algebraic degree of ANF for M-NLFSR with $k = 2$ is defined as

$$\deg(f) \le L - 2 \text{ (for } L > 2). \quad (7)$$

This statement was described by Golomb in the work (Theorem 4 [22]).

The distribution of the number of modified de Bruijn sequences was obtained depending on the non-linearity order of the M-NLFSR forming this sequence or, equivalently, on the maximum algebraic degree of the ANF of the forming polynomial. The resulting distribution for $k = 2$ is shown in Table 2.

**Table 2. Distribution of the M-NLFSR number depending on the order of nonlinearity $\deg(f)$ for $k = 2$**

| | Number of M-LFSR | Number of the M-LFSR 2nd Order | Number of the M-LFSR 3rd Order | Number of the M-LFSR 4th Order |
|---|---|---|---|---|
| 2 | 1 | – | – | – |
| 3 | 2 | – | – | – |
| 4 | 2 | 14 | – | – |
| 5 | 6 | 122 | 1,920 | – |
| 6 | 6 | 1,946 | 2,095,200 | 65,011,712 |
| 7 | 18 | 64,038 | *Unknown* | *Unknown* |
| 8 | 16 | 4,017,982 | *Unknown* | *Unknown* |
| 9 | 48 | 519,239,746 | *Unknown* | *Unknown* |

As it is seen, with an increase in the ANF algebraic degree, the number of polynomials that form the modified de Bruijn sequences and, consequently, the probability that a randomly chosen M-NLFSR has a more algebraic degree will increase exponentially.

## 2.4 THE NUMBER OF ANF M-NLFSR MONOMIALS

Let $\tau$ be the number of monomials in the recurrence relation defining the feedback for the LFSR. For all M-LFSR, $\tau$ is an even number. The distribution of the total number of monomials depending on $\tau$ for $4 \le L \le 6$ was published in [22]. The [22] also proved (Theorems 5, 6) that the minimum number of monomials in a polynomial corresponds to 2 (only for M-LFSR), and the maximum is calculated by the relation: $2^{L-1} - 2$ (except $L = 2$). The distribution has a Gaussian nature.

A similar distribution that takes into account $\deg(f)$ was obtained and is presented in Table 3.

Applying the principles of combinatorics, it can be indicated that the number of possible feedbacks for the NLFSR of $L$ cells and with a nonlinearity order $\deg(f)$ is determined by the correlation

$$n_L^{\deg(f)} = L \cdot \left( 1 + \sum_{i=1}^{\deg(f)-1} \frac{\prod_{j=1}^{i} (L-j)}{(1+i)!} \right). \quad (8)$$

**Table 3. Distribution of the ANF M-PCNOS number of monomials depending on the order of nonlinearity for $k = 2$**

| $\tau$ | Number of M-LFSR | Number of the M-LFSR 2nd Order | Number of the M-LFSR 3rd Order | Number of the M-LFSR 4th Order |
|---|---|---|---|---|
| $L = 2$ | | | | |
| 1 | 1 | – | – | – |
| $L = 3$ | | | | |
| 2 | 2 | – | – | – |
| $L = 4$ | | | | |
| 2 | 2 | – | – | – |
| 4 | – | 10 | – | – |
| 6 | – | 4 | – | – |
| $L = 5$ | | | | |
| 2 | 2 | – | – | – |
| 4 | 4 | 26 | 66 | – |
| 6 | – | 66 | 426 | – |
| 8 | – | 26 | 858 | – |
| 10 | – | 4 | 490 | – |
| 12 | – | – | 76 | – |
| 14 | – | – | 4 | – |
| $L = 6$ | | | | |
| 2 | 2 | – | – | – |
| 4 | 4 | 42 | 94 | 106 |
| 6 | – | 312 | 4,414 | 6,512 |
| 8 | – | 782 | 50,380 | 147,042 |
| 10 | – | 596 | 239,916 | 1,322,050 |
| 12 | – | 192 | 553,804 | 5,890,004 |

| | | | | |
|---|---|---|---|---|
| 14 | – | 22 | 658,398 | 14,115,280 |
| 16 | – | – | 420,692 | 19,139,124 |
| 18 | – | – | 141,894 | 15,146,272 |
| 20 | – | – | 23,808 | 7,057,286 |
| 22 | – | – | 1,742 | 1,892,112 |
| 24 | – | – | 58 | 274,994 |
| 26 | – | – | – | 20,294 |
| 28 | – | – | – | 628 |
| 30 | – | – | – | 8 |

As it can be seen in Table 4, the number of monomials for the studied M-NLFSR lies in the following range: $2 \le \tau \le \frac{2}{3} n_L^{\deg(f)}$, and the peak of the distribution equals approximately as follows: $\frac{1}{3} n_L^{\deg(f)}$.

## 2.5 AUTOCORRELATION FUNCTION

Let $S = (s(1), s(1), s(3), \ldots)$ be a periodic sequence with the $T$ period. The *autocorrelation function (ACF)* $S$ is an integer function $AC(t)$, defined as follows:

$$AC(t) = \frac{1}{T} \sum_{t=1}^{T} (2s(i) - 1)(2s(i+t) - 1) \qquad (9)$$

for $0 \le t \le T - 1$.

For example, Fig. 1 displays an ACF M-LFSR with $L = 4$ and, respectively, $T = 15$, determined by the recurrence relation $s_{(5+t)} = s_{(4+t)} + s_{(1+t)}$, and Figure 2 shows ACF M-NLFSR with $L = 4$ defined by the recurrence relation $s_{(5+t)} = s_{(4+t)} + s_{(3+t)} + s_{(2+t)} + s_{(1+t)} + s_{(4+t)} \cdot s_{(2+t)} + s_{(3+t)} \cdot s_{(2+t)}$.

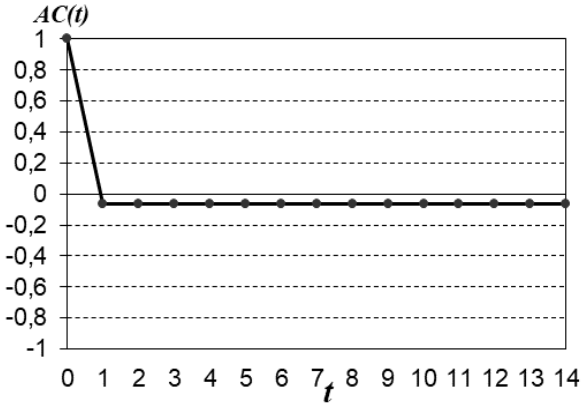


**Figure 1 – ACF for M-LFSR with $L = 4$**

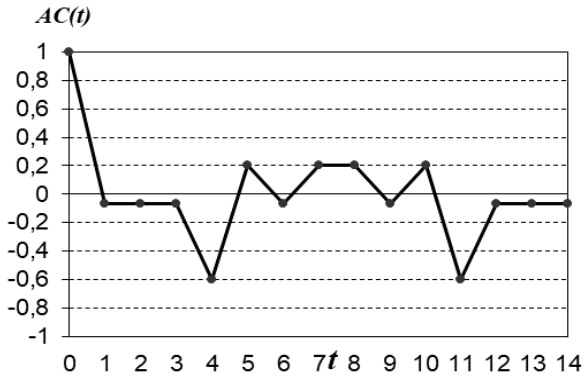


**Figure 2 – ACF for M-NLFSR with $L = 4$**

Analysing the obtained results, it can be argued that the appearance of all ACFs is symmetrical in relation to the middle of the graph (excluding a zero value, corresponding to the absence of a shift); only M-LFSR has a constant and minimal correlation with their shifted copies of the sequence.

In addition, the value of the ACF varies significantly depending on the selected M-NLFSR in the formation of the resulting sequence. From a practical point of view, it is clear that the less the sequence correlates with its own shifted copy, the less it is subjected to attacks that use the given vulnerability. Besides, it is important that there is no correlation for any of the shifts.

Let's introduce the value of the ACF which will characterize its maximum value (taken in modulo) for all $1 \le t \le T - 1$ that will be denoted as $AC_{\max}$. The value $AC_{\max}$ is restrained from the bottom by the expression $AC_{\max} \ge 1/T$, achieved only when M-LFSRs are used as feedback functions.

The generalized results obtained for the values $AC_{\max}$ are given in Tables 4-7.

**Table 4. Distribution of the ACF maximum value for $L = 3 (T = 7)$**

| $AC_{\max}$ | Number of M-NLFSR | Number of M-LFSR |
|---|---|---|
| $1/T = 0.14286$ | 2 | 2 |

**Table 5. Distribution of the ACF maximum value for $L = 4 (T = 15)$**

| $AC_{\max}$ | Number of M-NLFSR | Number of M-LFSR | Number of the M-NLFSR 2nd Order |
|---|---|---|---|
| $1/T = 0.06667$ | 2 | 2 | – |
| $5/T = 0.33333$ | 10 | – | 10 |
| $9/T = 0.60000$ | 4 | – | 4 |

**Table 6. Distribution of the ACF maximum value for $L=5\left(T=31\right)$**

| $AC_{\max}$ | Number of M-NLFSR | Number of the M-NLFSR 2nd Order | Number of the M-NLFSR 3rd Order |
|---|---|---|---|
| $1/T=0.03226$ | 6 | – | – |
| $5/T=0.16129$ | | – | 56 |
| $7/T=0.22581$ | 154 | 8 | 146 |
| $9/T=0.29032$ | 1156 | 78 | 1078 |
| $11/T=0.35484$ | 170 | 12 | 158 |
| $13/T=0.41935$ | 382 | 14 | 368 |
| $15/T=0.48387$ | 8 | – | 8 |
| $17/T=0.54839$ | 110 | 10 | 100 |
| $21/T=0.67742$ | 6 | – | 6 |

**Table 7. Distribution of the ACF maximum value for $L=6\left(T=63\right)$**

| $AC_{\max}$ | Number of the M-NLFSR 2nd Order | Number of the M-NLFSR 3rd Order | Number of the M-NLFSR 4th Order |
|---|---|---|---|
| $5/T=0.07937$ | – | – | 12 |
| $7/T=0.11111$ | – | 10 | 468 |
| $9/T=0.14286$ | 10 | 12,772 | 391, 484 |
| $11/T=0.17460$ | 44 | 49,556 | 1,534,900 |
| $13/T=0.20635$ | 212 | 351,690 | 10,906,882 |
| $15/T=0.23810$ | 222 | 274,756 | 8,522,742 |
| $17/T=0.26984$ | 920 | 766,838 | 23,851,328 |
| $19/T=0.30159$ | 128 | 167,810 | 5,202,124 |
| $21/T=0.33333$ | 242 | 269,832 | 8,344,838 |
| $23/T=0.36508$ | 32 | 30,922 | 955,434 |
| $25/T=0.39683$ | 70 | 143,930 | 4,450 510 |
| $27/T=0.42857$ | 4 | 3,336 | 100,228 |
| $29/T=0.46032$ | 12 | 10,684 | 346,834 |
| $31/T=0.49206$ | – | 174 | 5,480 |
| $33/T=0.52381$ | 50 | 12,216 | 377,528 |
| $35/T=0.55556$ | – | 8 | 116 |
| $37/T=0.58730$ | – | 140 | 4,656 |
| $41/T=0.65079$ | – | 526 | 16,146 |
| $45/T=0.71429$ | – | – | 2 |

As it can be seen, all the studied sequences formed by non-linear registers are inferior in their characteristics to the ACF sequences formed by linear registers.

The best value for the studied M-NLFSRs corresponds to $AC_{\max}=5/T$ and is achieved in the presence of the maximum order of nonlinearity.

The nature of the distribution is approximately the same for any order of nonlinearity.

## 2.6 LINEAR COMPLEXITY

The *linear complexity* ($Li$) of a pseudo-random sequence is called the shortest shift register with the help of which this periodic sequence is formed, provided that the first $Li$ values of the sequence are the initial fillings of the register.

The evaluation of linear complexity is one of the main parameters of the system. Any sequence that can be generated by an automaton (linear or nonlinear) over a finite field has finite linear complexity. Thus, it is possible to construct an algorithm by which the linear complexity of any sequence is determined, regardless of how it is generated, and knowledge of the structure of the circuit that forms the initial sequence is not necessary.

The most common mean to calculate the linear complexity is the Berlekamp-Massey algorithm, the essence of which is described in detail in [23, 24]. Thus, the large linear complexity of the formed sequence is a necessary (but not sufficient) condition for the practical durability of pseudo-random sequence generators.

Ideally, the linear complexity should be close to or equal to the period of the sequence [25-34]. Linear complexity was obtained by Golomb in [35]. M-NLFSR in most cases have the maximum linear complexity $Li_{\max}=2^L-2$ or one close to it [36-41]. The results of the distribution of linear complexity were obtained taking into consideration the order of nonlinearity.

Tables 8–10 show the distribution of linear complexity for the entire set of polynomials that form de Bruijn sequence sized as $L=2,\ldots,6$, and also the distribution depending on the order of nonlinearity of the register-forming sequence. Tables 11-13 show the distribution of linear complexity for the M-LFSR 2nd order and M-NLFSR 2nd order, if $7\le L\le 9$ .

**Table 8. Distribution of the linear complexity of the de Bruijn sequences for $L=4$**

| $Li$ | Number of M-NLFSR | Number of M-LFSR | Number of the M-NLFSR 2nd Order |
|---|---|---|---|
| 4 | 2 | 6 | – |
| 12 | 4 | – | 4 |
| 14 | 10 | – | 10 |

**Table 9. Distribution of the linear complexity of the de Bruijn sequences for $L = 5$**

| $Li$ | Number of M-NLFSR | Number of M-LFSR | Number of the M-NLFSR 2nd Order | Number of the M-NLFSR 3rd Order |
|---|---|---|---|---|
| 5 | 6 | 6 | – | – |
| 15 | 10 | – | – | 10 |
| 20 | 4 | – | – | 4 |
| 25 | 306 | – | 20 | 286 |
| 30 | 1 722 | | 102 | 1,620 |

As it can be seen from the obtained results, the M-NLFSRs overwhelming majority have the maximum linear complexity or it differs from the maximum by several units. The nature of the distribution is maintained for any value of $\deg(f)$ (See also [42, 43]). The profile of the linear complexity of the studied sequences is close to the expected value of the linear complexity, as well as for a truly random sequence.

**Table 10. The distribution of the linear complexity of the de Bruijn sequences for $L = 6$**

| $Li$ | Number of M-NLFSR | Number of M-LFSR | Number of the M-NLFSR 2nd Order | Number of the M-NLFSR 3rd Order |
|---|---|---|---|---|
| 6 | 6 | – | – | – |
| 27 | 10 | – | – | 10 |
| 30 | 8 | – | – | 8 |
| 32 | 12 | – | – | 12 |
| 33 | 8 | – | – | 8 |
| 35 | 62 | – | – | 62 |
| 36 | 152 | – | 10 | 142 |
| 38 | 478 | – | 14 | 464 |
| 39 | 1,036 | – | 48 | 988 |
| 41 | 3,572 | – | 106 | 3,466 |
| 42 | 6,100 | – | 200 | 5,900 |
| 44 | 17,240 | – | 536 | 16,704 |
| 45 | 28,702 | 4 | 936 | 27,762 |
| 47 | 86,056 | – | 2,650 | 83,406 |
| 48 | 134,290 | 4 | 4,184 | 130,102 |
| 50 | 401,102 | 8 | 12,692 | 388,402 |
| 51 | 453,734 | 20 | 14,184 | 439,530 |
| 53 | 1,364,978 | 48 | 43,184 | 1,321,746 |
| 54 | 1,819,148 | 68 | 56,930 | 1,762,150 |
| 56 | 5,453,680 | 158 | 171,298 | 5,282,224 |
| 57 | 3,190,982 | 126 | 100,724 | 3,090,132 |
| 59 | 9,557,084 | 256 | 297,988 | 9,258,840 |
| 60 | 11,148,860 | 338 | 347,518 | 10,801,004 |
| 62 | 33,441,564 | 916 | 1,041,998 | 32,398,650 |

**Table 11. Distribution of the linear complexity of the de Bruijn sequences for M-NLFSR with $\deg(f) \le 2$, if $L = 7$**

| $Li$ | Number of M-LFSR | Number of the M-LFSR 2nd order |
|---|---|---|
| 7 | 18 | – |
| 105 | – | 22 |
| 112 | – | 594 |
| 119 | – | 8,044 |
| 126 | – | 55,378 |

**Table 12. Distribution of the linear complexity of the de Bruijn sequences for M-NLFSR with $\deg(f \le 2)$, if $L = 8$**

| $Li$ | Number of M-LFSR | Number of the M-NLFSR 2nd Order |
|---|---|---|
| 8 | 16 | – |
| 208 | – | 2 |
| 214 | – | 2 |
| 218 | – | 2 |
| 220 | – | 6 |
| 222 | – | 18 |
| 224 | – | 26 |
| 226 | – | 96 |
| 228 | – | 204 |
| 230 | – | 726 |
| 232 | – | 1,020 |
| 234 | – | 2,914 |
| 236 | – | 5,954 |
| 238 | – | 18,168 |
| 240 | – | 17,578 |
| 242 | – | 52,538 |
| 244 | – | 96,554 |
| 246 | – | 289,222 |
| 248 | – | 147,584 |
| 250 | – | 440,762 |
| 252 | – | 738,236 |
| 254 | – | 2,206,370 |

**Table 13. Distribution of the linear complexity of the de Bruijn sequences for M-NLFSR with $\deg(f) \le 2$, if $L = 9$**

| $Li$ | Number of M-LFSR | Number of the M-NLFSR 2nd order |
|---|---|---|
| 9 | 48 | |
| 456 | – | 2 |
| 459 | – | 2 |
| 462 | – | 6 |
| 465 | – | 34 |
| 468 | – | 30 |
| 471 | – | 504 |

| | | |
|---|---|---|
| 474 | – | 1,866 |
| 477 | – | 1,386 |
| 480 | – | 21,112 |
| 483 | – | 74,458 |
| 486 | – | 42,542 |
| 489 | – | 599,644 |
| 492 | – | 2,097,832 |
| 495 | – | 795,706 |
| 498 | – | 11,163,082 |
| 501 | – | 39,051,092 |
| 504 | – | 7,268,192 |
| 507 | – | 101,824,464 |
| 510 | – | 356,297,792 |

## 3. CONCLUSION

The Boolean functions that form the de Bruijn sequence are efficient as iterative generators in stream ciphers due to their proper local statistical characteristics, the maximum period of the generated sequence, and the simplicity of implementation.

However, the use of nonlinearity in Boolean functions leads to an increase in design characteristics.

Thus, the majority of M-NLFSRs have the number of monomials equal to 1/3 of the maximum value in their structure. For example, for $L = 32$, most M- NLFSRs will have approximately $10^9$ of the items, and each item will contain the product of up to 30 different values, that from the point of view of practical implementation for streaming encryption systems is unacceptable. However, as it is shown in the work, there are M- NLFSRs with a minimum number of feedback coefficients.

All studied sequences formed by non-linear registers are inferior as of ACF characteristics to sequences formed by linear registers.

The best value for the studied M-NLFSR equals $AC_{\max} = 5/T$ (for M-LFSR all $AC_{\max} = 1/T$) and is achieved with the maximum order of nonlinearity. The nature of the distribution is approximately the same for any order of nonlinearity. However, given the relatively uniform distribution of $AC_{\max}$, it can be assumed that the randomly taken Boolean function forming the de Bruijn sequence will have a low $AC_{\max}$.

Unlike the M-LFSR, the overwhelming majority of the M-NLFSRs with $\deg(f) \geq 2$ have a maximum, or a few units different from the maximum, linear complexity as $Li_{\max} = 2^L - 2$. The nature of the distribution is maintained for any $\deg(f) \leq 2$. At the same time, the profile of linear complexity is close to the mathematically expected, as well as for truly random sequences.

Thus, the search for constructively simple NLFSRs forming the de Bruijn sequence, as well as the optimization of the structure to match the necessary cryptographic properties is a complex task that requires further study.

## 4. REFERENCES


[1] M. Schafheutle, *A First Report on the Stream Cipher SNOW*. [Online]. Available at: http://www.cryptonessie.org

[2] C. Berbain, O. Billet, A. Canteaut, N. Courtois, B. Debraize, H. Gilbert, L. Goubin, A. Gouget, L. Granboulan, C. Lauradoux, M. Minier, T. Pornin, and H. Sibert, Decim – A New Stream Cipher for Hardware applications, *in ECRYPT Stream Cipher Project Report 2005/004.* [Online]. Available at: http://www.ecrypt.eu.org

[3] S. Kiyomoto, T. Tanaka, and K. Sakurai, "A word-oriented stream cipher using clock control," *Workshop Record of SASC 2007, pp.260–274, January 2007* [Online]. Available at: https://www.cosic.esat.kuleuven.be/ecrypt/stream/papersdir/2007/029.pdf

[4] *The eSTREAM Project – eSTREAM Phase 3. SOSEMANUK (Portfolio Profile 1)*. [Online]. Available at: http://www.ecrypt.eu.org

[5] *The eSTREAM Project – eSTREAM Phase 3. Grain (Portfolio Profile 2).* [Online]. Available at: http://www.ecrypt.eu.org/stream/grainpf.html

[6] *The eSTREAM Project – eSTREAM Phase 3. MICKEY (Portfolio Profile 2).* [Online]. Available at: http://www.ecrypt.eu.org/stream/mickeypf.html

[7] *The eSTREAM Project – eSTREAM Phase 3. Trivium (Portfolio Profile 2).* [Online]. Available: http://www.ecrypt.eu.org/stream/triviumpf.html

[8] P. Dabrowski, G. Łabuzek, T. Rachwalik, J. Szmidt, "Searching for nonlinear feedback shift registers with parallel computing," 2013. [Online]. Available att: https://eprint.iacr.org/2013/542.pdf

[9] H. Fredricksen, "A survey of full length nonlinear shift register cycle algorithms," *SIAM Review*, vol. 24, no. 2, pp. 195-221, 1982.

[10] C.J. Jansen, *Investigations on Nonlinear Streamcipher Systems: Construction and Evaluation Methods*, Ph.D. Thesis, Technical University of Delft, 1989.

[11] C.J. Jansen, "The maximum order complexity of sequence ensembles," *Lecture Notes in Computer Science*, Adv. Cryptology-Eupocrypt'1991, Berlin, Germany, 1991, vol. 547, pp. 153-159.

[12] D. Linardatos, N. Kalouptsidis, "Synthesis of minimal cost nonlinear feedback shift registers," *Signal Process*., vol. 82, no. 2, pp. 157–176, 2002.

[13] P. Rizomiliotis, N. Kalouptsidis, "Results on the nonlinear span of binary sequences," *IEEE*

*Transactions on Information Theory*, vol. 51, no. 4, pp. 1555–5634, 2005.

[14] K. Limniotis, N. Kolokotronis, N. Kalouptsidis, "On the nonlinear complexity and Lempel-Ziv complexity of finite length sequences," *IEEE Transactions on Information Theory*, vol. 53, no. 11, pp. 4293–4302, 2007.

[15] E. Dubrova, "A scalable method for constructing Galois NLFSRs with period 2n-1 using cross-join pairs," *IEEE Transactions on Information Theor*y, vol. 59, issue 1, pp. 703-709, 2013.

[16] J. Mykkeltveit, M.-K. Siu, P. Tong, "On the cyclic structure of some nonlinear shift register sequences," *Inform. and Control*., vol. 43, pp. 202-215, 1979.

[17] C. Carlet, "Boolean functions for cryptography and error correcting codes," *in: Crama Y., Hammer P. L. (Eds.), Boolean Methods and Models, Cambridge University Press*. [Online]. Available at: http://www.rocq.inria.fr/secret/Claude.Carlet/chap-fcts-Bool.pdf

[18] D. Knuth, *The Art of Computer Programming, vol. II, Seminumerical Algorithms*, USA, Commonwealth of Massachusetts: Addison-Wesley, 1969, 634 p.

[19] M.C. Flye-Sainte, "Solution to question number 48," *l'Intermediaire des Mathematiciens*, vol. 1, pp. 107-110, 1894.

[20] N.G. de Bruijn, "A combitorial problem," *Nederl. Akad. Wetensch*., vol. 49, pp. 758-764, 1946.

[21] H. Fredricksen, "A survey of full length nonlinear shift register cycle algorithm," *SIAM Review*, vol. 24, issue 2, pp. 195–221, 1982.

[22] G.L. Mayhew, S.W. Golomb, "Characterizations of generators for modified de Bruijn sequences," *Advances in Applied Mathematics*, vol. 13, issue 4, pp. 454-461, 1992. [Online]. Available at: https://www.sciencedirect.com/science/article/pii/01968858 9290021N

[23] E. R. Berlekamp, *Algebraic Coding Theory*, McGraw-Hill, NY, 1968, 474 p.

[24] F.J. McWilliams, N.J. Sloane, *The Theory of Error-Correcting Codes*, North-Holland, 1978, 762 p.

[25] T. Kaida and J. Zheng, "On linear complexity of periodic sequences over extension fields from cyclic difference sets," *Proceedings of the 2015 Seventh International Workshop on Signal Design and its Applications in Communications (IWSDA)*, Bengaluru, 2015, pp. 15-18. DOI: 10.1109/IWSDA.2015. 7458392

[26] O. Kuznetsov, O. Potii, A. Perepelitsyn, D. Ivanenko, N. Poluyanenko, "Lightweight stream ciphers for green IT engineering," *in: Kharchenko V., Kondratenko Y., Kacprzyk J. (eds) Green IT Engineering: Social, Business and Industrial Applications. Studies in Systems, Decision and Control*, vol. 171, Springer, Cham, 2019, pp. 113-137. DOI: 10.1007/978-3-030-00253-4_6.

[27] K. Tsuchiya, C. Ogawa, Y. Nogami and S. Uehara, "Linear complexity of generalized NTU sequences," *Proceedings of the 2017 Eighth International Workshop on Signal Design and its Applications in Communications (IWSDA)*, Sapporo, 2017, pp. 74-78. DOI: 10.1109/IWSDA.2017.8095739

[28] A. Andrushkevych, Y. Gorbenko, O. Kuznetsov, R. Oliynykov, M. A. Rodinko, "A prospective lightweight block cipher for green IT engineering," *in: Kharchenko V., Kondratenko Y., Kacprzyk J. (eds) Green IT Engineering: Social, Business and Industrial Applications. Studies in Systems, Decision and Control*, vol. 171, Springer, Cham, 2019, pp. 95-112. DOI: 10.1007/978-3-030-00253-4_5.

[29] J. Szmidt and P. Dąbrowski, "The construction of nonlinear feedback shift registers of small orders," *Proceedings of the 2015 International Conference on Military Communications and Information Systems (ICMCIS)*, Cracow, 2015, pp. 1-4.

[30] T. Rachwalik, J. Szmidt, R. Wicik and J. Zabłocki, "Generation of nonlinear feedback shift registers with special-purpose hardware," *Proceedings of the 2012 Military Communications and Information Systems Conference (MCC)*, Gdansk, 2012, pp. 1-4.

[31] S. Moriyama and A. Tsuneda, "A study on construction of low-density parity-check codes using nonlinear feedback shift registers," *Proceedings of the 2016 International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju, 2016, pp. 697-699.

[32] A. Kuznetsov, Y. Gorbenko, A. Andrushkevych and I. Belozersev, "Analysis of block symmetric algorithms from international standard of lightweight cryptography ISO/IEC 29192-2," *Proceedings of the 2017 4th International Conference Problems of Infocommunications. Science and Technology (PIC S&T)*, Kharkov, 2017, pp. 203-206.

[33] J. Zhong and D. Lin, "On minimum period of nonlinear feedback shift registers in grain-like structure," *IEEE Transactions on Information Theory*, vol. 64, no. 9, pp. 6429-6442, 2018.

[34] A. M. Arshad, H. Ino, C. Ogawa and Y. Nogami, "Linear complexity of signed binary sequence over odd characteristic field," *Proceedings of the 2016 19th International Conference on Computer and Information Technology (ICCIT)*, Dhaka, 2016, pp. 266-269. DOI: 10.1109/ICCITECHN.2016.7860207

[35] G.L. Mayhew, S.W. Golomb, "Linear spans of modified de Bruijn sequences," *IEEE Trans.*

*Inform. Theory*., vol. 36, no. 5, pp. 1166–1167, 1990.

[36] P. Rizomiliotis, "Constructing periodic binary sequences with maximum nonlinear span," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 4257-4261, Sept. 2006.

[37] V.I. Dolgov, I.V.Lisitska, K.Ye. Lisitskyi, "The new concept of block symmetric ciphers design," *Telecommunications and Radio Engineering*, vol. 76, issue 2, pp. 157-184, 2017.

[38] K. Fukuda and A. Tsuneda, "Key-sensitivity improvement of block cipher systems based on nonlinear feedback shift registers," *Proceedings of the 2012 IEEE Asia Pacific Conference on Circuits and Systems*, Kaohsiung, 2012, pp. 100-103.

[39] I. Gorbenko, O. Nariezhnii and I. Kudryashov, "Construction method and features of one class of cryptographic discrete signals," *Proceedings of the 2017 4th International Conference Problems of Infocommunications. Science and Technology (PIC S&T)*, Kharkov, 2017, pp. 156-160.

[40] A. A. Zadeh and H. M. Heys, "Simple power analysis applied to nonlinear feedback shift registers," *IET Information Security*, vol. 8, no. 3, pp. 188-198, May 2014.

[41] V. A. Krasnobayev, "Method for realization of transformations in public-key cryptography," *Telecommunications and Radio Engineering*, vol. 66, issue 17, pp. 1559-1572, 2007. DOI: 10.1615/TelecomRadEng.v66.i17.60

[42] A. A. Kuznetsov, A. V. Potii, N. A. Poluyanenko, S. G. Vdovenko, "Combining and filtering functions based on the nonlinear feedback shift registers," *Telecommunications and Radio Engineering*, vol. 78, issue 10, pp. 853-868, 2019. DOI: 10.1615/TelecomRadEng.v78.i10.20

[43] A. A. Kuznetsov, A. V. Potii, N. A. Poluyanenko, I. V. Stelnik, "Nonlinear functions of complication for symmetric stream ciphers," *Telecommunications and Radio Engineering*, vol. 78, issue 9, pp. 743-458, 2019. DOI: 10.1615/TelecomRadEng.v78.i9.10

***Alexandr A. Kuznetsov,*** *Doctor of Sciences (Engineering), Full Professor, Professor of the Department security information systems and technologies of the V. N. Karazin Kharkiv National University.*

*Areas of scientific interests: cryptography and authentication, steganography, cybersecurity.*

***Oleksandr V. Potii****, Doctor of Sciences (Engineering), Full Professor, Professor of the Department security information systems and technologies of the V. N. Karazin Kharkiv National University. Areas of scientific interests: cryptography and authentication, cybersecurity.*

***Nikolay O. Poluyanenko****, Candidate of Technical Sciences (Engineering), Researcher of the Department security information systems and technologies of the V. N. Karazin Kharkiv National University. Areas of scientific interests: cryptography and authentication, cybersecurity.*

***Serhii M. Ihnatenko,*** *Employee of Security Service of Ukraine. Areas of scientific interests: applied cryptography and coding, post-quantum cryptography, security information systems and technologies.*

***Igor V. Stelnyk,*** *Deputy Director of the Department of Information Protection Administration of the State Service of Special Communication and Information Protection of Ukraine. Areas of scientific interests: information and cybersecurity.*

***Danylo V. Mialkovsky,*** *Deputy Director of the Department of Information Protection Administration of the State Service of Special Communication and Information Protection of Ukraine. Areas of scientific interests: information and cybersecurity.*