Noname manuscript No.
(will be inserted by the editor)

Further on the Construction of Feedback Shift Registers with Maximum Strong Linear Complexity

Zhou Congwei · Hu Bin · Guan Jie

Received: date / Accepted: date

Abstract In this paper, we present the more accurate definition of strong linear complexity of feedback shift registers based on Boolean algebraic than before, and analyze the bound of strong linear complexity by the fixed feedback function. Furthermore, the feedback shift registers with maximum strong linear complexity are constructed, whose feedback functions require the least number of monomials. We also show that the conclusions provide particular ideas and criteria for the design of feedback shift registers.

Keywords Feedback shift register · Strong linear complexity · Non-singular · $r-$cycle

1 Introduction

The research for feedback shift registers (FSRs) can be traced back to the 1950s[1], which is still a basic theoretical problem in the field of secure communication so far. Periodic sequences generated by FSRs, such as $m-$sequences, are widely used in cryptographic algorithms and communication coding for their good statistical characteristics and implementation efficiency[2]. In recent ten years, due to the inherent linear restriction of linear feedback shift register (LFSR)[3], more and more symmetric cryptographic algorithms begin to use the non-linear feedback shift register (NLFSR) as their driving components[4]. Consequently the design standard of universality of NLFSR needs to be solved. Generally speaking, for the design of the overall architecture about cryptographic algorithms, the balance between the non-linear iterations and

PLA SSF Information Engineering University
Zhengzhou
Tel.:
Fax:
E-mail: zhoucongwei@qq.com.com

implementation efficiency may be considered. However, it is very necessary to require FSRs as the driving components to generate sequences, which have the sufficiently larger period and stronger linear complexity, or satisfy certain mutual constraints[5]. At present, because the algebraic logic of the cycle structure of NLFSR is not clear, as the only non-linear component of many lightweight cryptography algorithms, it has a great potential attack.

*The **linear complexity** is usually an important measurement of pseudo-random sequence, which is defined as the stage of the shortest LFSR that generates it.* In practice, for example, the Berlekemp-Massey's synthesis algorithm[6] can recover the whole pseudo-random sequence through the truncated sequence with twice linear complexity length. In fact, for pseudo-random sequences generated by different structures, other complexity can be defined, such as $2-$adic complexity[7]. At the same time, based on fault-tolerant mechanism, the $k - error$ linear complexity can also be defined. For the NLFSR sequences, it only can be determined that the linear complexity of de Bruijn sequence is $\left[2^{n-1}+n, 2^n-1\right]$[8]. In general, because of the selection of the initial states (the number of cycles in a state graph of FSR), we can not go through all the periodic sequences with non-equivalent translation to determine the linear complexity of them generated by the FSR with large stages.

*In the reference [9], the **strong linear complexity** of feedback register is proposed for the first time, which is defined as the stage of the shortest LFSR that generates all sequences generated by the feedback register.* Since the cycle structure of LFSR is known[10], the strong linear complexity is an important index including the period and linear complexity of sequences generated by the feedback register. At the same time, according to the reference [9], a strong linear complexity of feedback register can be derived from related monomials of its feedback function, which has a strong guiding role in the practical design of the feedback register. Based on this definition, this paper gives a more accurate definition of strong linear complexity on Boolean algebra, and complements the relevant conclusions of the upper and lower bounds of strong linear complexity of FSRs and an algorithm for finding the exact value of strong linear complexity. Furthermore, the counting of all the $r-$cycles defined in the reference [9] is given, and it is proved that when a (non-singular) FSR with the prime stage has the maximum strong linear complexity, the corresponding feedback function consists of at least $\frac{n-1}{2}\left(\frac{n+1}{2}\right)$ monomials; when a FSR has the maximum strong linear complexity, the corresponding feedback function consists of at least $\left[\left\lceil\frac{n-1}{2}\right\rceil, n-1\right]$ monomials. Meanwhile, the constructed method of the feedback function which satisfies the least number of monomials is given. To a certain extent, it provides a criterion for the design of feedback functions of FSRs.

In this paper, we first give the basic algebraic knowledge required for the definition of strong linear complexity. Secondly, studying on the strong linear complexity of FSRs as the research property, the corresponding conclusions are given respectively from the analysis and design of FSRs. Finally, the specific ideas are provided for the design of FSRs with the maximum strong linear complexity.

2 Basic Algebraic Knowledge

In this section, we refer to these references [11,13] to define a more rigorous Boolean algebra and logic algebra, and then give the set space of feedback functions of FSRs on this algebra structure. Thus, the strong linear complexity of FSRs is defined accurately. Firstly, we give the definition on Boolean algebra, function and polynomial.

Definition 1 let B be a set including at least two elements 0 and 1. If three binary operations $\vee, +, \cdot$ and one unitary operation $-$ are defined on B, then call B a Boolean algebra, denoted by $< B, \vee, +, \cdot, -, 0, 1 >$, and $\vee, +, \cdot, -$ are collectively Boolean operations. If $B = B_2 = \{0, 1\}$, then the Boolean algebra $< B_2, +, \cdot, -, >$ or $< B_2, \vee, \cdot, -, >$ is called a logical algebra.

Definition 2[11] Let B be a Boolean algebra and n be a positive integer, then the mapping $B^n \to B$ is called a $n-$variable Boolean function on B.

Definition 3[11] Every element of the Boolean algebra B is called a Boolean constant, and the variable, which be valued in B, is called a Boolean variable. The Boolean expression obtained by finite Boolean operations of some Boolean constants and variables $x_1, \cdots, x_n$ is called a $n-$variable Boolean polynomial on B, usually denoted by $f(x_1, \cdots, x_n), g(x_1, \cdots, x_n)$.

Because of the following Proposition 1, we usually do not make a distinction between the Boolean function and Boolean polynomial.

Proposition 1[11] Every $n-$variable Boolean polynomial determines a unique $n-$variable Boolean function: $(a_1, \cdots, a_2) \to f(a_1, \cdots, a_2)$, and this Boolean function is also denoted by $f(x_1, \cdots, x_n)$.

In general, the feedback function of FSR satisfies the following Proposition 2.

Proposition 2 The feedback function of $n-$stage FSR is a $n-$variable Boolean function if and only if B is a logical algebra.

Under different Boolean algebra structures, the algebraic normal form of Boolean polynomial is different. In general, we call the algebraic normal form of Boolean polynomial on the logical algebra $< B_2, \vee, \cdot, -, >$ as the complete disjunctive normal form, and on the logical algebra $< B_2, +, \cdot, -, >$ as the Zhegalkin polynomial[13]. *In the following, the Boolean algebra $< B_2, +, \cdot, -, >$ is considered, and except for in this algebra structure, $+$ means "logical addition" operation, and the rest means "real addition".*

Next by the form of Proposition 3, we give the set space of feedback functions of all FSRs on the logical algebraic $< B_2, +, \cdot, -, >$.

Proposition 3 Let Γ be a set of all $n-$variable Boolean polynomial on $< B_2, +, \cdot, -, >$, then Γ is a 2^n-dimensional vector space over F_2. One group of base vectors is the set of monomials $\{x_{i_1} x_{i_2} \cdots x_{i_k} | 1 \leq i_j \leq n, 1 \leq j \leq k\}$ and element 1 of $n-$variable Boolean polynomial, which we can call *linear variables.*

According to Proposition 3, the number of all FSRs under these definitions is 2^{2^n}. In the following, we generally make $f = f(x_1, \cdots, x_n)$ to express the feedback function of FSR, and $f(x_1, \cdots, x_n) = x_1 + g(x_2, \cdots, x_n)$ to distinguish the non-singular FSR from FSRs. At the same time, for the distinction between linear and non-linear, we require f to be a linear (an affine) Boolean polynomial if and only if the algebraic degree of f is no more than 1. In addition, they are all called nonlinear.

The reference [12] points out that $\mathit{\Gamma}$ is the $n-$variable polynomial ring over F_2, namely,

$$\mathit{\Gamma} = F_2[x_1, x_2, \cdots, x_n]/({x_i}^2 + x_i)$$

where ${x_i}^2 \equiv x_i \bmod ({x_i}^2 + x_i), 1 \le i \le n$. Thus, for the feedback function f of FSR, we can define an algebra-homomorphism δ over $\mathit{\Gamma}$, that is, Definition 4.

Definition 4[12] $\delta = \delta_f : \mathit{\Gamma} \to \mathit{\Gamma}$, such that

$$\delta(x_i) = x_{i+1}, 1 \le i < n$$

$$\delta\left(x_n\right) = f(x_1, \cdots, x_n).$$

As $\mathit{\Gamma}$ is a finite ring, there exits a minimum positive integer t such that

$$\delta^{n_0+t}(x_1) = \delta^{n_0}(x_1), n_0 \ge 0$$

and the period of δ is denoted by $p(\delta) = t + n_0$. *Thus, for any feedback function f, it can be determined the only group of ordered vectors in the vector space $\mathit{\Gamma}$, namely*

$$V^{P(\delta)} = \left\{\delta^0(x_1) = x_1, \delta^1(x_1), \cdots, \delta^{p(\delta)-1}(x_1)\right\}.$$

At the same time, it is obvious that when the feedback function f is non-singular, then $n_0 = 0$, and δ is an isomorphism[12]. In fact, we can obtain some properties of $V^{P(\delta)}$ by Proposition 4.

Proposition 4 For the ordered vector group $V^{P(\delta)}$ of feedback function f of any non-singular FSR, there exists

$$\delta^{p(\delta)-1}(x_1) = f(x_n, x_1, \cdots, x_{n-1}).$$

Proof
$$\begin{aligned} f(f(x_n, x_1, \cdots, x_{n-1}), x_1, x_2, \cdots, x_{n-1}) &= f(x_n, x_1, \cdots, x_{n-1}) + g(x_1, x_2, \cdots, x_{n-1}) \\ &= x_n + g(x_1, x_2, \cdots, x_{n-1}) + g(x_1, x_2, \cdots, x_{n-1}) = x_n \end{aligned}$$

Let the set of all periodic sequences generated by the non-singular FSR with feedback function f under 2^n different initial loadings, be $\mathit{\Omega}(f)$, and the least common multiple of the minimal periods of sequences in $\mathit{\Omega}(f)$ is denoted by $p\left(\mathit{\Omega}(f)\right)$. In the reference [12], it is proved that when the FSR is non-singular, $p\left(\mathit{\Omega}(f)\right) = p(\delta)$. *Thus, $p\left(\mathit{\Omega}(f)\right)$ can be determined by restricted to the sub-vector-space $\overset{\wedge}{\mathit{\Gamma}}$ of $\mathit{\Gamma}$ generated by $V^{P(\delta)}$ in fact.* At the same time, according to the reference [9], we can transform the strong linear complexity of FSR into the dimension of $\overset{\wedge}{\mathit{\Gamma}}$, that is Proposition 5.

Proposition 5 The strong linear complexity of FSR is equal to the dimension $|Q|$ of sub-vector-space $\overset{\wedge}{\mathit{\Gamma}}$ generated by the corresponding $V^{P(\delta)}$.

Proof According to the definition in the reference [9], the strong linear complexity of FSR is equal to the stage of the shortest LFSR that generates all sequences generated by the FSR. When the FSR is singular, its minimum period sequence contains a pre-period sequence. It follows that all minimum periodic sequences can be uniquely represented by

$$\left\{\delta^0(x_1) = x_1, \delta^1(x_1), \cdots, \delta^{p(\delta)-1}(x_1)\right\} (a_1, a_2, \cdots, a_n)$$

where $(a_1, a_2, \cdots, a_n)$ is the initial n values of the corresponding sequence. Let $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_{|Q|}$ be a base of $\overset{\wedge}{\varGamma}$, then

$$\left\{\delta^0(x_1) = x_1, \delta^1(x_1), \cdots, \delta^{p(\delta)-1}(x_1)\right\} = \left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_{|Q|}\right) \cdot A_{|Q| \times P(\delta)}.$$

Thus there exists a invertible matrix $B_{P(\delta) \times |Q|}$, which makes $\left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_{|Q|}\right) \cdot A_{|Q| \times P(\delta)} \cdot B_{P(\delta) \times |Q|}$ be exactly the generating matrix of LFSR with the $|Q|$ −stage, namely, the shortest LFSR corresponding to the generating matrix can generate all the minimum periodic sequences. It proves by the uniqueness of $V^{P(\delta)}$.

In fact, the basis of $\overset{\wedge}{\varGamma}$ *is also equal to the linear variables appearing in* $V^{P(\delta)}$*, which is illustrated by Example 1.*

Example 1 Let a feedback function of 4− stage NLFSR be

$$f = x_1 + x_{2,4} + x_{2,3,4}$$

where $x_{2,4} = x_2 x_4$ (Since $x_i \cdot x_i = x_i$, the representation method is an one-to-one corresponding), then the corresponding ordered vector group $V^{P(\delta)}$ is

$$x_1, x_2, x_3, x_4, x_1 + x_{2,4} + x_{2,3,4}, x_2 + x_{1,3} + x_{1,3,4}, x_3 + x_{1,2,4} + x_{2,3,4}, x_4 + x_{1,2,3} + x_{1,3,4},$$
$$x_1 + x_{1,2,4} + x_{2,3,4}, x_2 + x_{1,2,3} + x_{1,3,4}, x_3 + x_{2,4} + x_{1,2,4}, x_4 + x_{1,3} + x_{1,2,3}$$

It shows that these linear variables in the sub-vector-space generated by this ordered vector group $V^{P(\delta)}$ are

$$\{x_1, x_2, x_3, x_4, x_{2,4}, x_{1,3}, x_{1,3,4}, x_{2,3,4}, x_{1,2,4}, x_{1,2,3}\}$$

Therefore, the strong linear complexity of NLFSR is 10. The same result is also given in the reference [9].

If the dimension of $\overset{\wedge}{\varGamma}$ is relatively small, then the theory of LFSR can be applied to analyze NLFSRs. Therefore, it is necessary to discuss the bound of strong linear complexity of FSRs.

3 The Bound of Strong Linear Complexity for a Certain Feedback Function

For one FSR, from Definition 4 and Proposition 5, it can be seen that its feedback function uniquely determines its strong linear complexity. Let R be the set of monomials of its feedback function, and Q be the set of linear variables of the corresponding sub-vector-space $\overset{\wedge}{\varGamma}$, namely, the cardinality $|Q|$ of Q is its strong linear complexity. We identify $2^n - 1$ monomials with $2^n - 1$ index sets I, namley, monomial subscript sets (I is any non-empty subset of $\{1, \cdots, n\}$, such as $x_{2,4} \Leftrightarrow \{2, 4\}$). Then On this basis, the $r-$cycle is defined, that is, Definition 5.

Definition 5 Make the operation $+1 \bmod n$ on all elements (indicators) in any I (Here, define mod$n$ for the complete system of residues $\{1, \cdots, n\}$ of n). Through the factor of n times whole operations $+1 \bmod n$, we can get a sequence of index sets $\{I_1, \cdots, I_d(d\,|n)\}$ such that $I_1 = I_{d+1}$. If the index set $|I_1| = r(1 \leq r \leq n)$, we call the sequence of index sets, that is, the set $\{I_1, \cdots, I_d(d\,|n)\}$, as a $r-$cycle.

Let k_r be the number of all $r-$cycles, then each $r-$cycle can be expressed as ${}^rC_i(1 \leq i \leq k_r)$. To make it easier to understand, we take an example from the reference [9].

Example 2 For $n = 4$, we start with any index set I such that $r = 2$. For example, if we make the operation $+1 \bmod n$ on elements in $I_1 = \{2, 3\}$, then we can get $I_2 = \{3, 4\}, I_3 = \{1, 4\}, I_4 = \{1, 2\}$, and the next operation for I_4 will go back to I_1. So one $2-$cycle is $\{\{2, 3\}, \{3, 4\}, \{1, 4\}, \{1, 2\}\}$. In the same way, we can get another $2-$cycle $\{\{2, 4\}, \{1, 3\}\}$. Thus for $n = 4$, $k_2 = 2, {}^2C_1 = \{\{2, 3\}, \{3, 4\}, \{1, 4\}, \{1, 2\}\}, {}^2C_2 = \{\{2, 4\}, \{1, 3\}\}$.

From Definition 5, it can be seen that a monomial corresponds to an index set, and a $r-$cycle corresponds to d index sets with the cardinality r, that is, d monomials with the algebraic degree r. In fact, all $r-$cycles can be regarded as a partition of all $r-$order monomials. *The Corollary 2.3 in the reference [9]* points out that all elements in Q can be obtained by the shift and combination of the elements in R. To facilitate the formula derivation in the following section, we give an equivalent theorem in relation to Corollary 2.3, that is, Theorem 1.

Theorem 1 Let R be the set of index sets corresponding to all monomials in the feedback function f of FSR, and Q be the set of index sets corresponding to all linear variables in the corresponding sub-vector-space $\overset{\wedge}{\Gamma}$. Then each element in Q only satisfies the following two conditions:

1. If $I \in Q$ and $n \in I$, then for each $J \in R$, we can deduce that $J \cup \{i + 1 : i \in I, i \neq n\} \in Q$;
2. If $I \in {}^rC_i$, then $I \in {}^rC_i \Rightarrow {}^rC_i \subset Q$.

According to Definition 4, the correspongding $V^{P(\delta)}$ of any FSR contains initial linear variables $x_1, \cdots, x_n$ and these which are including in $f(x_1, \cdots, x_n)$, namely,

$$\bigcup_{1 \leq i \leq n} \{i\} \cup R \subset Q$$

Combined with Example 1, we apply Theorem 1 to calculate its strong linear complexity of FSR.

Example 3 Let a feedback function of $4-$ stage NLFSR be

$$f = x_1 + x_{2,4} + x_{2,3,4}.$$

The analysis steps are as follows.

1. $R = \{\{1\}, \{2, 4\}, \{2, 3, 4\}\} \Rightarrow \{\{1\}, \{2\}, \{3\}, \{4\}\} \cup R \subset Q$;
2. Apply the condition 2 of Theorem 1 to know that

$$\{2,4\} \in \{\{2,4\},\{1,3\}\} \subset Q, \{2,3,4\} \in$$
$$\{\{2,3,4\},\{1,3,4\},\{1,2,4\},\{1,2,3\}\} \subset Q;$$

3. These index sets I, such that $I \in Q$ and $n \in I$, are $\{4\}, \{2,4\}, \{1,2,4\}, \{1,3,4\}, \{2,3,4\}$ in the above steps. Thus apply the condition 1 of Theorem 1 to know that

$$\begin{array}{l} \{1\} \in R \Rightarrow \{\{1\},\{1,3\},\{1,2,3\},\{1,2,4\},\{1,3,4\}\} \subset Q \\ \{2,4\} \in R \Rightarrow \{\{2,4\},\{2,3,4\},\{2,3,4\},\{2,4\},\{2,3,4\}\} \subset Q \\ \{2,3,4\} \in R \Rightarrow \{\{2,3,4\},\{2,3,4\},\{2,3,4\},\{2,3,4\},\{2,3,4\}\} \subset Q \end{array} .$$

To the sum, since no new index set is added in the step 3,

$$Q = \{\{1\},\{2\},\{3\},\{4\},\{2,4\},\{1,3\},\{2,3,4\},\{1,3,4\},\{1,2,4\},\{1,2,3\}\}$$

namely, $|Q| = 10$.

In particular, if there is a constant term 1 in f, then there must be a linear variable 1 in Q. Therefore, for one FSR, if there is constant term 1 in f, its maximum strong linear complexity is 2^n; if there is no constant term 1, its maximum strong linear complexity is $2^n - 1$. Consequently, for one non-singular FSR, the lower bound of its strong linear complexity can be given immediately, that is, Theorem 2.

Theorem 2 Let R be the set of index sets corresponding to all monomials in the feedback function f of non-singular FSR. For all the $r(2 \le r \le n-1)-$cycle ${}^{r}C_i$, if there exits l $r-$cycles ${}^{r}C_j(1 \le j \le l \le k_r)$, and each of them has at least one element contained in R. It deduces that if f has a constant term 1, then its strong linear complexity is at least

$$\sum_{r=2}^{n-1}\sum_{i=1}^{l} |{}^{r}C_i| + n + 1;$$

If f has no constant term 1, then at least

$$\sum_{r=2}^{n-1}\sum_{i=1}^{l} |{}^{r}C_i| + n$$

Proof It is easy to prove by the condition 2 of Theorem 1.

For *Theorem 2.4 in the reference [9]*, the bound is widened too much because the linear term is not considered. Thus we modify it a little and get Theorem 3 in this section.

Theorem 3 Let R be the set of index sets corresponding to all monomials in the feedback function f of non-singular FSR. Let $r'(2 \le r' \le n-1)$ be the smallest integer corresponding to the lowest degree of monomials *except linear terms* in $g(x_2, \cdots, x_n)$. For all the $r'-$cycles ${}^{r'}C_j$, if there exits l $r'-$cycles ${}^{r'}C_j(1 \le j \le l \le k_{r'})$, and each of them has no element contained in R. It deduces that if f has a constant term 1, then its strong linear complexity is at most

$$2^n - 1 - \sum_{i=2}^{r'-1} \binom{n}{i} - \sum_{j=1}^{l} \left|{}^{r'}C_j\right|;$$

If f has no constant term 1, then at most

$$2^n - 2 - \sum_{i=2}^{r'-1} \binom{n}{i} - \sum_{j=1}^{l} \left| {}^{r'}C_j \right|.$$

It can be seen that for one non-singular FSR, if there is constant term 1 in f, its maximum strong linear complexity is $2^n - 1$; if there is no constant term 1 in f, its maximum strong linear complexity is $2^n - 2$. In fact, Theorem 3 also gives a criterion to judge whether one FSR is non-singular, that is, Theorem 4.

Theorem 4 If one feedback function of FSR is non-singular, then its corresponding Q does not contain the index set $\{1, \cdots, n\}$.

Proof It supposes that Q contain the index set $\{1, \cdots, n\}$. According to the condition 1 of Theorem 1, because the indicator $1 \notin \{i+1 : i \in I, i \neq n\}$, $1 \in J$. It follows that the feedback function f is non-singular, hence when merging to get index set $\{1, \cdots, n\}$, $J = \{1\}$. Therefore, $\{i+1 : i \in I, i \neq n\} = \{2, \cdots, n\}$, that is, $I = \{1, \cdots, n\}$. So we can see that $\{1, \cdots, n\} \in R$. It is contradictory, and the theorem is proved.

From Theorem 4, we can also deduce the upper and lower bounds of strong linear complexity of FSR, that is, if there is a constant term 1 in f, then the value range of $|Q|$ is

$$\left[\sum_{r=2}^{n-1} \sum_{i=1}^{l} |{}^{r}C_i| + n + 1, 2^n - \sum_{i=2}^{r'-1} \binom{n}{i} - \sum_{j=1}^{l} \left| {}^{r'}C_j \right| \right];$$

if there is no constant term 1 in f, then the value range of $|Q|$ is

$$\left[\sum_{r=2}^{n-1} \sum_{i=1}^{l} |{}^{r}C_i| + n, 2^n - 1 - \sum_{i=2}^{r'-1} \binom{n}{i} - \sum_{j=1}^{l} \left| {}^{r'}C_j \right| \right].$$

In fact, from Theorem 2 to Theorem 3, in order to give the exact value of $|Q|$, the key is to study on those $r-$cycles ${}^{r}C_i$, where $n \in I \in {}^{r}C_i$. Therefore, the Algorithm 1 can be given to calculate the value of $|Q|$ accurately, based on the steps of *Example 3*.

According to actual application requirements, a FSR is designed to satisfy the maximum strong linear complexity, and it is necessary to minimize the number of monomials in its feedback function f as much as possible. According to the existence, let f require at least $|R|_{min}$ monomials to make its FSR have the maximum strong linear complexity. In the next section, we will give the corresponding conclusions and construction methods.

4 Construction of FSRs with the Maximum Strong Linear Complexity

In this section, we mainly discuss at least which combinations of monomials are required for feedback functions of FSRs with the maximum strong linear complexity. According to the condition 2 of Theorem 1, $|R|_{min}$ has its upper bound, that is, $|R|_{min}$ is not more than the sum of number k_r of all $r(1 \leq$

Algorithm 1 Calculate the strong linear complexity of $n(n \geq 3)-$stage FSR.

```
Require:
    The set $R$ of index sets corresponding to all monomials in the feedback function $f$ of FSR;
Ensure:
    The strong linear complexity of FSR, $|Q|$;
1: Make the operation +1 mod $n$ on all elements in $I_0^i(1 \leq i \leq |R|) \in R$, sieve out duplicate index sets, and set $r(2 \leq r \leq n)-$cycles ${}^rC_{1_0}, {}^rC_{2_0} \cdots, {}^rC_{l_0}(0 \leq l_0 \leq k_r)$;
2: For $I_1 \in \bigcup_{\substack{1_0 \leq i_0 \leq l_0 \\ 2 \leq r \leq n}} {}^rC_{i_0}$ such that $n \in I_1$, set $I_2 = I_0^j \cup \{i+1 : i \in I_1, i \neq n\}(1 \leq j \leq |R|)$, which is different from the elements in $\bigcup_{\substack{1_0 \leq i_0 \leq l_0 \\ 2 \leq r \leq n}} {}^rC_{i_0}$. Make the operation +1 mod $n$ on all elements in $I_2$ again, sieve out duplicate index sets, and set $r(3 \leq r \leq n)-$cycles ${}^rC_{1_1}, {}^rC_{2_1} \cdots, {}^rC_{l_1}(0 \leq l_1 \leq k_r)$.;
3: Repeat STEP2 until a certain $p(n-3 \geq p \geq 0)$ such that
       $$\bigcup_{\substack{1_{p+1} \leq i_{p+1} \leq l_{p+1} \\ 3+p \leq r \leq n}} {}^rC_{i_{p+1}} \subseteq \bigcup_{\substack{1_p \leq i_p \leq l_p \\ 2+p \leq r \leq n}} {}^rC_{i_p};$$
4: return $|Q| = \sum_{\substack{0 \leq j \leq p \\ 1_j \leq i_j \leq k_j \\ 2 \leq r \leq n}} \left|{}^rC_{i_j}\right| + n$;
```

$r \leq n)-$cycles. So we first give that k_r should satisfy the law by the form of Theorem 5.

Theorem 5 Let $p_1, \ldots, p_m$ be all nontrivial positive factors of n such that $1 < p_i \leq r(1 \leq i \leq m)$. It follows that when $1 < r \leq \lceil \frac{n-1}{2} \rceil$, the cardinality of each $r-$cycle can only be $n, \frac{n}{p_1}, \ldots, \frac{n}{p_m}$. If let the number of $r-$cycles corresponding to its cardinality be $y_0, y_1, \ldots, y_m$ respectively, then

$$\binom{n-1}{r} = (n-r) \cdot (y_0 + \sum_{1 \leq i \leq m} \frac{y_i}{p_i})$$

and when $\lceil \frac{n+1}{2} \rceil < r \leq n-1$, the result of similar equation is symmetrical. In particular, when n is a prime, $k_r = \binom{n}{r}/n$.

Proof The cardinality of each $r-$cycle can only be a positive factor of n , thus when $1 < p_i \leq r(1 \leq i \leq m)$ and $1 < p_i \leq r(1 \leq i \leq m)$, the cardinality of each $r-$cycle can only be $n, \frac{n}{p_1}, \ldots, \frac{n}{p_m}$. According to the definition of combination number, it follows that

$$\binom{n}{r} = y_0 \cdot n + \sum_{1 \leq i \leq m} y_i \cdot \frac{n}{p_i}.$$

At the same time, the number of index $j(1 \leq j \leq n)$ in $r-$cycles, whose cardinality is the above value, is exactly $r, \frac{r}{p_1}, \ldots, \frac{r}{p_m}$ respectively, so it can be obtained that

$$\binom{n-1}{r-1} = y_0 \cdot r + \sum_{1 \leq i \leq m} y_i \cdot \frac{r}{p_i}.$$

By associating with the above two equations, we can obtain that

$$\binom{n-1}{r} = (n-r)\cdot(y_0 + \sum_{1\le i\le m} \tfrac{y_i}{p_i}).$$

For $\lceil \frac{n+1}{2} \rceil < r \le n-1$, we only need to verify the complement set of each index set on the $\{1,\cdots,n\}$ in arbitrary $(n-r)-$cycles, then it follows that the result is symmetrical. In particular, when n is a prime, because $y_1 = \cdots = y_m = 0$, $k_r = \sum\limits_{0\le i\le m} y_i = y_0 = \binom{n}{r}/n$.

In fact, Theorem 5 gives a simple method to solve the value of k_r for a certain n, r. Thus we use Theorem 5 to give the value of k_r for $r = 2$ in the form of Corollary 1, which also needs to be explicit in the following conclusions.

Corollary 1 $k_2 = \lceil \frac{n-1}{2} \rceil$.

Proof When n is odd, $k_2 = y_0 = \binom{n}{2}/n = \frac{n-1}{2}$; When $n(n \ge 3)$ is even, from Theorem 5 it can be seen that

$$2y_0 + y_1 = n - 1.$$

And because the number of index $j(1 \le j \le n)$ in $r-$cycles, whose cardinality is $\frac{n}{2}$, is exactly 1, if and only if there is only one $r-$cycle $\{\{1, \frac{n}{2}+1\}, \cdots, \{\frac{n}{2}, n\}\}$, namely, $y_0 = \frac{n-2}{2}, y_1 = 1$. To the sum, the corollary is proved.

At the same time, if the n numbers $\{1,\cdots,n\}$ are linked end-to-end, then any index set corresponging to the two numbers with the same distance is in the same $2-$cycle. Thus k_2 is actually equal to the number of possible value of distance between two numbers. Because the distance between two numbers is at most $\lceil \frac{n-1}{2} \rceil$, $k_2 = \lceil \frac{n-1}{2} \rceil$. Through this method of distance selection, we can quickly get ***representative elements of index sets*** for all $2-$cycles. (*The typically of representative elements of index sets in $r-$cycles means that when the selected index set from a certain $r-$cycle appears in Q, according to the condition 2 of Theorem 1, all index sets in this $r-$cycle are included in Q.*)

In fact, we can also get the sum $\sum\limits_{r=1}^{n} k_r$ of numbers k_r of all $r(1 \le r \le n)-$cycles, that is, Theorem 6.

Theorem 6 The sum of numbers k_r of all defined $r(1 \le r \le n)-$cycles for one FSR is

$$\sum_{r=1}^{n} k_r = Z(n) - 1 = \frac{1}{n}\sum_{d|n} \phi(d) 2^{\frac{n}{d}} - 1$$

where ϕ is the Euler function.

Proof In the cycle structure of $n-$tage pure circulating shift register, the state with different weights is not on the same cycle. Its each state consists n components. The position of the rightmost component of states is considered as

1, and the position of the leftmost component is n. Note that the set of positions in any state with the weight $r(r \geq 1)$, where the component value is 1, is $J = \{j_1, \cdots, j_r\}$. Then its states on the same cycle can be represented by $J = J_1, \cdots, J_d(d\,|n)$. Let the set of states with weight r on any cycle be ${}^rD_i$. Since there is an one-to-one correspondence between the elements contained in ${}^rD_i$ and ${}^rC_i(1 \leq i \leq k_r)$, and the corresponding form means that each element in the ${}^rD_i = \{J_1, \cdots, J_k(d\,|n)\}$ and $r-$cycle ${}^rC_i = \{I_1, \cdots, I_d(d\,|n)\}$ completely coincides under the corresponding situation of indexes, the sum of the number of all $r-$cycles is equal to the number of cycles in the cycle structure of $n-$tage pure circulating shift register. It follows that $\emptyset$, namely, the all-zero state cycle, is not included in ${}^rD_i$, and it can be seen from the reference [10] that the number of cycles in the cycle structure of $n-$tage pure circulating shift register is $Z(n)$. Thus the equation in the theorem is true.

It can be seen from Theorem 1 to Theorem 2 that these elements of Q, which are not in ${}^rC_j(1 \leq j \leq l \leq k_r)$, must be merged with the index sets in R to upgrade a certain index set in $r+i(i \geq 1)-$cycles through the condition 1 of Theorem 1, that is, those *representative elements of index sets* in ${}^rC_j$ such that $n \in I$. In fact, when n is a prime, we can get the following Theorem 7 and 8 by using the properties of *representative elements of index sets.*

Theorem 7 When $n(n \geq 3)$ is a prime, the feedback function f of nonsingular FSR requires $|R|_{min} = \frac{n+1}{2}$ monomials to maximize its strong linear complexity.

Proof When n is prime, according to Corollary 1, any FSR requires $\frac{n-1}{2}$ *representative elements of index sets* to ensure the existence of corresponding monomials in all 2$-$cycles. (Otherwise, according to the condition 1 of Theorem 1, it is necessary to add a corresponding linear term so that the combined index set is just the missing *representative element of index sets* in 2$-$cycles.) For $n-2 \geq r \geq 2$ and $n \geq 5$, according to Theorem 5, since n is a prime, it can be seen that the number of index n in each $r+1-$cycle is exactly $r+1$, and the cardinality of each $r+1-$cycle is exactly n. Then there exists at least one *representative element of index sets* in each $r+1-$cycle is

$$C = \left\{ a_1, a_2, \cdots, a_r, n \,\middle|\, \begin{matrix} \textit{when } 1 \leq i < j \leq r, 2 \leq a_i < a_j \leq n-1 \textit{ and} \\ \textit{there is at least one } a_k : \frac{n+1}{2} \leq a_k \leq n-1 \end{matrix} \right\}.$$

Otherwise, there are only three kinds of forms C' such that $n \in C'$, namely,

$$C' = \begin{cases} \left\{a_1, \cdots, a_r, n \,\middle|\, \textit{when } 1 \leq i < j \leq r \leq \frac{n-3}{2}, 2 \leq a_i < a_j \leq \frac{n-1}{2} \right\} \\ \left\{1, a_2, \cdots, a_r, n \,\middle|\, \textit{when } 2 \leq i < j \leq r, 3 \leq a_i < a_j \leq n-1 \right\} \\ \left\{1, 2, a_3, \cdots, a_r, n \,\middle|\, \textit{when } 3 \leq i < j \leq r, 3 \leq a_i < a_j \leq n-1 \right\} \end{cases}$$

For the first form C', we can make the operation $(-1 \bmod n)$ a_r times on all elements in it to obtain the form C, where $a_k = n - a_r$; For the second form C', we only make the operation $(-1 \bmod n)$ one time on all elements in it to obtain the form C, where $a_k = n-1$; For the third form C', if $a_3 \neq 3$, then we can make the operation $(-1 \bmod n)$ one time on all elements in it to obtain the second form C', and otherwise, because of $r \leq n-2$, the second form C'

can always be obtained by making the operation $(-1 \bmod n)$ less than n times on all elements in it.

It follows that we can get the value range of a_k is $\frac{n+1}{2} \leq a_k \leq n-1$. Thus we can select the corresponding index set

$$\{a_1 - 1, \cdots, a_{k-1} - 1, a_{k+1} - 1, \cdots, a_r - 1, n\}$$

in any $r-$cycle to merge with the index set $\left\{a_k, n \left| a_k : \frac{n+1}{2} \leq a_k \leq n-1 \right.\right\}$ in $2-$cycles to obtain the form C through the condition 1 of Theorem 1. At the same time, these index sets $\left\{a_k, n \left| a_k : \frac{n+1}{2} \leq a_k \leq n-1 \right.\right\}$ can be considered as *representative elements of index sets* for all $2-$cycles.

Therefore, according to the mathematical induction, when we select the $\frac{n-1}{2}$ *representative elements of index sets*

$$\left\{\left\{\frac{n+1}{2}, n\right\}, \left\{\frac{n+3}{2}, n\right\}, \cdots, \{n-1, n\}\right\}$$

in $2-$cycles, then we can get *representative elements of index sets* in each $r+1(n-2 \geq r \geq 2)-$cycle. According to the condition 2 of Theorem 1, all elements in each $r(n-1 \geq r \geq 1)-$cycle can be included in Q. Therefore, if f includes at least $\frac{n-1}{2}$ monomials corresponding to the $\frac{n-1}{2}$ *representative elements of index sets*

$$\left\{\left\{\frac{n+1}{2}, n\right\}, \left\{\frac{n+3}{2}, n\right\}, \cdots, \{n-1, n\}\right\}$$

in $2-$cycles and x_1, then we can maximize its strong linear complexity. And for $n = 3$, $f = x_1 + x_2 x_3$. So the theorem is proved.

Similarly, since all linear terms are directly contained in Q, Corollary 2 can be obtained.

Corollary 2 When $n(n \geq 3)$ is a prime, the feedback function f of FSR requires $|R|_{min} = \frac{n-1}{2}$ monomials to maximize its strong linear complexity.

For Corollary 2, f can be selected these monomials corresponding to the $\frac{n-1}{2}$ *representative elements of index sets*

$$\left\{\{1, n\}, \{2, n\}, \cdots, \left\{\frac{n-1}{2}, n\right\}\right\}$$

in $2-$cycles. For one FSR with an arbitrary stage, we have the following Theorem 8 to describe.

Theorem 8 For $n \geq 3$, the feedback function f of FSR requires $|R|_{min}$ monomials to maximize its strong linear complexity, where

$$\left\lceil \frac{n-1}{2} \right\rceil \leq |R|_{min} \leq n-1$$

Proof When $n-2 \geq r \geq 1$, let the selected *representative element of index sets* in $r+1-$cycles be $\left\{a_1, \cdots, a_r, n \left| when\ 1 \leq i < j \leq r, 1 \leq a_i < a_j \leq n-2 \right.\right\}$. Then by merging with the index set $\{1, n\}$ through the condition 1 of Theorem 1, there must exist one *representative element of index sets*

$$C = \left\{1, a_1 + 1, \cdots, a_r + 1, n \left| when\ 1 \leq i < j \leq r, 1 \leq a_i < a_j \leq n-2 \right.\right\}$$

in $r+2-$cycles. Note that those *representative elements of index sets* in $r+2-$cycles, which do not satisfy the form C and contain the index n, can be considered as

$$D = \left\{b_1, \cdots, b_{r+1}, n \left| when\ 1 \le i < j \le r, 2 \le b_i < b_j \le n-1 \right.\right\}.$$

Then we only merge the *representative element of index sets*

$$\left\{b_1 - 1, \cdots, b_{k-1} - 1, b_{k+1} - 1, \cdots, b_{r+1} - 1, n \left| when\ 1 \le i < j \le r, 2 \le b_i < b_j \le n-1 \right.\right\}$$

in $r+1-$cycles and $\{b_k, n\,|b_k : 2 \le b_k \le n-1\}$ to obtain the form D through the condition 1 of Theorem 1. Therefore, according to the mathematical induction, for f, we can select at most $n-1$ monomials corresponding to these $n-1$ *representative elements of index sets*

$$\{b_i, n\,|b_i : 1 \le b_i \le n-1\}$$

in $2-$cycles to maximize its strong linear complexity. Combined with Corollary 2, Theorem 8 is proved.

In the process of practical application, it can be seen from the proof of Theorem 8 that we can find all *representative elements of index sets* in $r(\left\lceil\frac{n-1}{2}\right\rceil \ge r \ge 3)-$cycles, which do not satisfy the form C and satisfy the form D (For $n-1 \ge r \ge \left[\frac{n-1}{2}\right]$, use the symmetry to know the existence of the form C inevitably). Through these *representative elements of index sets*, we can find the maximum union from

$$\{b_k, n\,|b_k : 2 \le b_k \le n-1\}.$$

Finally, the feedback function with maximum strong linear complexity can be obtained by adding the above union, *representative elements of index sets* in the missing $2-$cycles and $\{1, n\}$. This construction method is illustrated in the following Example 4.

Example 4 For $n = 8$ and $r = 3$, there are only two $3-$cycles, where index sets do not satisfy the form C and satisfy the form D, namely,

$$\begin{array}{c}\{\{1,3,5\},\{2,4,6\},\{3,5,7\},\{4,6,8\},\{1,5,7\},\{2,6,8\},\{1,3,7\},\{2,4,8\}\},\\ \{\{1,3,6\},\{2,4,7\},\{3,5,8\},\{1,4,6\},\{2,5,7\},\{3,6,8\},\{1,4,7\},\{2,5,8\}\}\end{array}.$$

For $r = 4$, there is only one $4-$cycle, where index sets do not satisfy the form C and satisfy the form D, namely,

$$\{\{1,3,5,7\},\{2,4,6,8\}\}.$$

Through the analysis, the *representative elements of index sets*

$$\{2,4,8\},\{2,5,8\},\{2,4,6,8\}$$

and the corresponding maximum union $\{2, 8\}$ are selected. Then for $n = 8$, one set of monomials of feedback function with maximum strong linear complexity is

$$\{\{1,8\},\{2,8\},\{3,8\},\{4,8\}\}.$$

5 Conclusion

In fact, this paper extends the conclusion of *Corollary 2.5 in the reference [9]*, reveals the monomial structure of feedback function of FSR with the maximum strong linear complexity, and further guarantees that when its corresponding sub-vector-space generated by $V^{P(\delta)}$ is itself of Γ, it provides the structure of feedback function with the least number of monomials. Thus in order to design a NLFSR in cryptographic algorithms, some design criteria are established. At the same time, the next step of research will continue to study on the maximum strong linear complexity of non-singular FSR with any stage, whose the feedback function structure requires the least number of monomials, and provide theoretical support for the operational non-singular FSR.

6 Acknowledgements

The authors would like to thank all who presented the related definitions in the reference [9]. Meanwhile this research was supported by the National Natural Science Foundation of China (Grant No. 61572516, 61802437).

References

1. Zierler N. Linear recurring sequences[J]. Journal of the Society for Industrial and Applied Mathematics, 1959, 7(1): 31-48.
2. Xianyong W, Zhou X. A kind of generating method of m-sequence pseudo-code generator[J]. Meas. Control. Technol, 2003, 22(9): 56-58.
3. Rueppel R A. New approaches to stream ciphers[D]. ETH Zurich, 1984.
4. ECRYPT. eSTREAM: The ECRYPT stream cipher project(IST-2002-507932)[EB/OL]. http://www. ecrypt.eu.org/stream[2014-01-01].
5. Ding C, Xiao G. Stream cipher cryptography and its application[M]. National Defense Industry Press, 1994.
6. Berlekamp E. Algebraic coding theory[M]. New York: McGraw-Hill, 1968.
7. Klapper A, Goresky M. Feedback shift registers, 2-adic span, and combiners with memory[J]. Journal of Cryptology, 1997, 10(2): 111-147.
8. Chan A H, Games R A, Key E L. On the complexities of de Bruijn sequences[J]. Journal of Combinatorial Theory, 1982, 33(3):233-246.
9. Chan A H, Goresky M, Klapper A. On the Linear Complexity of Feedback Registers[C]// Workshop on the Theory and Application of Cryptographic Techniques. Springer, Berlin, Heidelberg, 1989.
10. Golomb S W. Shift register sequences[M]. Aegean Park Press, 1967.
11. Whitesitt J E. Boolean algebra and its applications[M]. Courier Corporation, 2012.
12. Kjeldsen K. On the cycle structure of a set of nonlinear shift registers with symmetric feedback functions[J]. Journal of Combinatorial Theory, Series A, 1976, 20(2): 154-169.
13. Pospelov, D A. Logical Methods of Scheme Analysis and Synthesis[M]. Moscow: Energiya, 1974.