Цель работы: построить РСЛОС генератор и генератор Геффа, для каждого найти длину периода генератора. Получить выборку из 38 элементов и провести для них тест последовательностей:

1) Частотный тест.
2) Тест серий.
3) Тест последовательностей для l=3.
4) Автокорреляционный тест с графиком.

1. Описание работы.

1.1. Регистры сдвига с линейной обратной связью (РСЛОС, LFSR).

Регистр сдвига с линейной обратной связью длиной L состоит из L разрядов пронумерованных $0, 1, \dots, L-1$, способных хранить один бит каждый и имеющих по одному входу и одному выходу, а также из тактового генератора, управляющего движением данных. В каждую единицу времени выполняются следующие операции:

a) содержимое разряда 0 подается на выход и формирует часть выходной последовательности;
b) содержимое разряда i сдвигается в разряд $i-1$ для каждого i, $1 \le i \le L-1$;
c) новым содержимым разряда $L-1$ становится бит обратной связи, вычисленный сложением по модулю 2 предыдущего содержимого фиксированного подмножества разрядов $0, 1, \dots, L-1$.

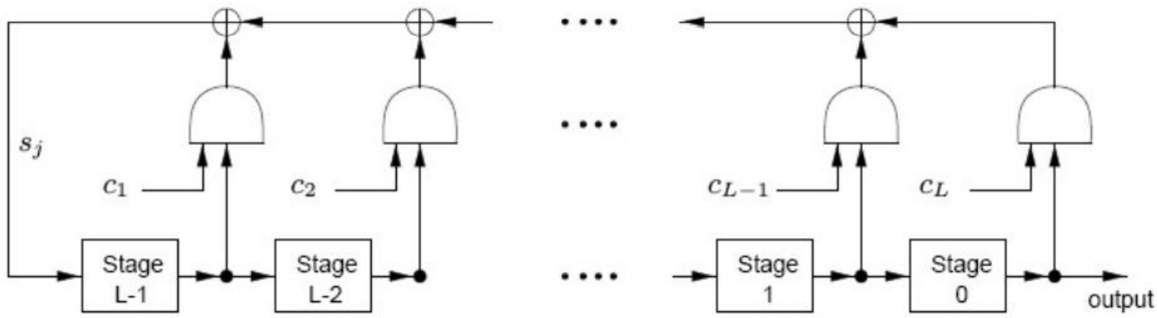

Рисунок 1 - РСЛОС длиной L

Пусть $C(D) \in Z_2[D]$ — многочлен связей степени L. Если $C(D)$ является примитивным многочленом, то каждое из $2^L - 1$ ненулевых начальных состояний невырожденного РСЛОС $\langle C(D) \rangle$ производит выходную последовательность с максимально возможным периодом $2^L - 1$.

1.2. Генератор Геффа.

Генератор Геффе определяется тремя РСЛОС с максимальной длиной, чьи длины L_1, L_2, L_3 попарно взаимно просты, с нелинейной комбинирующей функцией

$$f(x_1, x_2, x_3) = x_1 x_2 \oplus x_2 x_3 \oplus x_3$$

Генератор ключевого потока имеет период $(2^{L_1} - 1)(2^{L_2} - 1)(2^{L_3} - 1)$ и линейную сложность $L = L_1 L_2 + L_2 L_3 + L_3$.

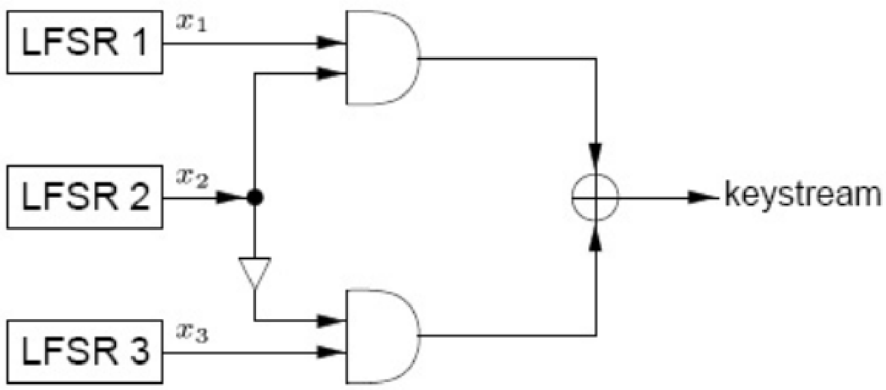

Рисунок 2 - Генератор Геффе

1.3. Базовые тесты.

Пусть $s = s_0, s_1, \ldots, s_{n-1}$ — двоичная последовательность длины n. Ниже будут представлены тесты, широко используемые для определения, обладает ли двоичная последовательность s некоторыми специфическими характеристиками, которые, скорее всего, демонстрировала бы истинно случайная последовательность.

1.3.1. Частотный тест.

Цель этого теста — определить, является ли примерно равным количество 0 и 1 в s, как это ожидается для случайной последовательности. Пусть n_0, n_1 обозначают количество 0 и 1 в s, соответственно. Используется статистика:

$$X_1 = \frac{(n_0 - n_1)^2}{n}$$

1.3.2. Тест последовательностей.

Цель этого теста — определить, является ли примерно равным количество вхождений 000, 001, 010, 011, 100, 101, 110, 111 в качестве подпоследовательностей в s, как это ожидается для случайной последовательности. Пусть n_0, n_1 обозначают количество 0 и 1 в s, соответственно, и пусть n_{000}, n_{001}, n_{010}, n_{011}, n_{100}, n_{101}, n_{110}, n_{111} обозначают количество вхождений 000, 001, 010, 011, 100, 101, 110, 111 в s, соответственно. Используется статистика:

$$X_3 = \frac{8}{n-2}(n_{000}^2 + n_{001}^2 + n_{010}^2 + \cdots + n_{111}^2) - \frac{2}{n}(n_0^2 + n_1^2) + 1$$

1.3.3. Тест серий.

Цель теста серий — определить, является ли количество серий различных длин в последовательности s таким, как ожидается для случайной последовательности. Ожидаемое число разрывов (или блоков) длины i в случайной последовательности длины n равно $e_i = (n - i + 3)/2^{i+2}$. Пусть k равен наибольшему целому i, для которого $e_i \geq 5$. Пусть B_i, G_i — количество разрывов и блоков длины i в $i = s$, соответственно, для каждого $1 \leq i \leq k$. Используется статистика:

$$X_4 = \sum_{i=1}^{k} \frac{(B_i - e_i)^2}{e_i} + \sum_{i=1}^{k} \frac{(G_i - e_i)^2}{e_i}$$

1.3.4. Автокорреляционный тест.

Цель этого теста — проверить корреляции между последовательностью s и ее (нециклическими) сдвигами. Пусть d — фиксированное целое число, $1 \leq d \leq \lfloor n/2 \rfloor$. Число бит в s, не равных их d-сдвигам, есть $A(d) = \sum_{i=0}^{n-d-1} s_i \oplus s_{i+d}$, где $\oplus$ обозначает операцию XOR. Используется статистика:

$$X_5 = \frac{2\left(A(d) - \frac{n-d}{2}\right)}{\sqrt{n-d}}$$

1.3.5. Тест, основанный на профиле линейной сложности.

Пусть $s = s_0, s_1, \dots, s_{n-1}$ — двоичная последовательность, и пусть L_N обозначает линейную сложность подпоследовательности $s^N = s_0, s_1, \dots, s_{N-1}$. Последовательность $L_1, L_2, \dots$ называется профилем линейной сложности для s. Аналогично, если $s = s_0, s_1, \dots, s_{n-1}$ является конечной двоичной последовательностью, то последовательность $L_1, L_2, \dots, L_n$ называется профилем линейной сложности для s^n.

Пусть s^n — конечная двоичная последовательность длины n, и пусть линейная сложность s^n равна L. Тогда единственный РСЛОС длины L, генерирующий s^n, существует тогда и только тогда, когда $L \leq n/2$.

Профиль линейной сложности последовательности может быть вычислен, используя алгоритм Берлекэмпа-Месси.

1.4. Пример работы программы.

Генерируемая последовательность представлена ниже.

```
LFSR:
10001000000000001011011011011011111011
Size = 38
Period = 65535
```

Рисунок 3 - Генератор РСЛОС

Частотный тест для полученной последовательности представлен ниже.

```
'0' = 20
'1' = 18
Frequency test = 0.105263
```

Рисунок 4 - Частотный тест

Тест серий для полученной последовательности представлен ниже.

```
k = 1 e_1 = 5
Number of blocks(G_1) of length 1 = 9
Number of breaks(B_1) of length 1 = 6
Batch test = 3.4
```

Рисунок 5 - Тест серий

Тест последовательностей для полученной представлен ниже.

```
n_000 = 10
n_001 = 2
n_010 = 2
n_011 = 6
n_100 = 2
n_101 = 6
n_110 = 5
n_111 = 3
Sequence_test = 11.3392
```

Рисунок 6 - Тест последовательностей

Автокорреляционный тест для полученной последовательности представлен ниже.

```
Autocorrelation test:
X_1 = -0.657596
X_2 = -1
X_3 = -3.38062
X_4 = -1.02899
X_5 = -0.696311
X_6 = -3.18198
X_7 = 0.359211
X_8 = 0
X_9 = -1.85695
X_10 = 0.755929
X_11 = 0.7698
X_12 = -1.56893
X_13 = 1.6
X_14 = 0.408248
X_15 = 0
X_16 = 1.2792
X_17 = 1.74574
X_18 = 0
X_19 = 2.29416
```

Рисунок 7 - Автокорреляционный тест

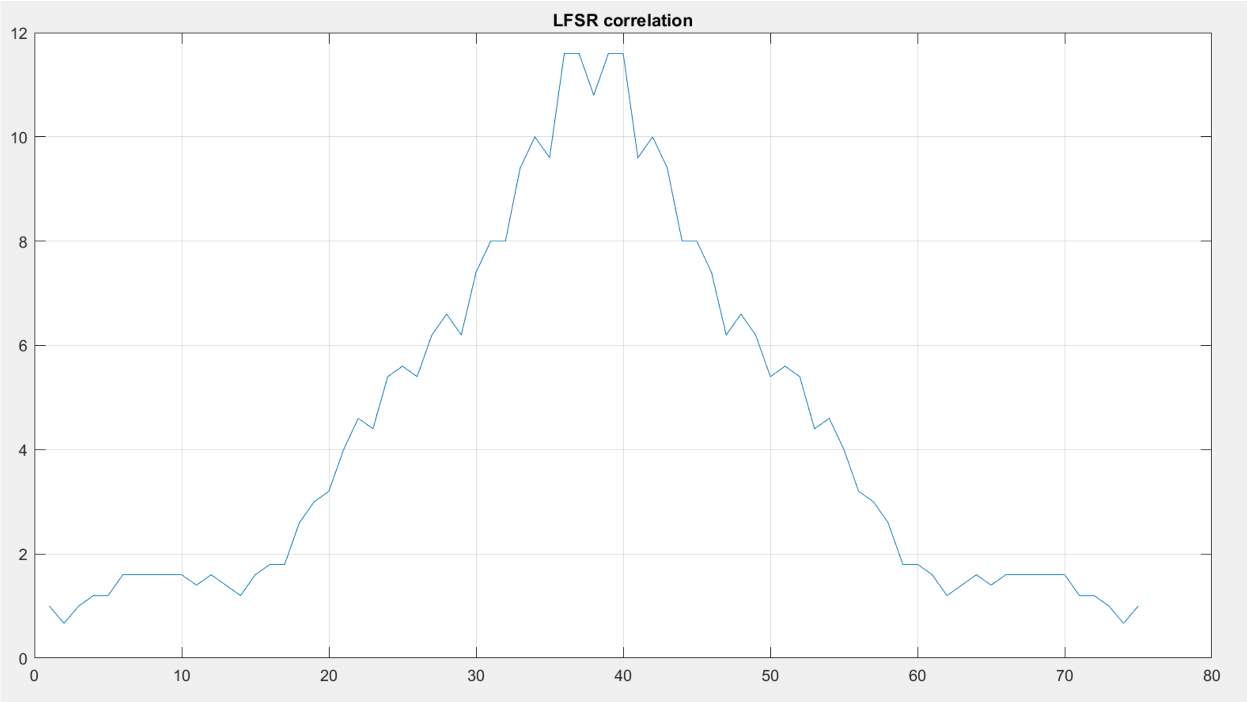

График 1 - График корреляции

Для уровня значимости $\alpha = 0.05$, сравним пороговые значения каждого теста и полученные:

Таблица 1 - Сравнение значений

Название теста	Пороговое значение	Полученное значение
Частотный тест	3,8415	0,1052
Тест серий	9,4877	3,4
Тест последовательностей	5,9915	11,3392
Автокорреляционный тест	11,0705	Максимальное значение = 2,294

Таким образом, можно сделать вывод, что пройдены все тесты кроме теста последовательностей.

Последовательность, полученная генератором Геффе с использованием трех РСЛОС представлена ниже.

```
GEFFE:
polynom: x^18 + x^5 + x^2 + x + 1  Lfsr3 = 11000000000000000010110001011000101111
Size = 38
Period = 262143

polynom: x^19 + x^5 + x^2 + x + 1 Lfsr2 = 01000000000000000000110111100100001101
Size = 38
Period = 524287

polynom: x^17 + x^3 + 1 Lfsr1 = 10000000000000000100100100100100101100
Size = 38
Period = 131071

Geffe = 01000000000000000001011010111100001111
```

Рисунок 8 - Генератор Геффе

Период такого генератора будет равен $262\,143 * 524\,287 * 131\,071$.

Частотный тест для полученной последовательности представлен ниже.

```
'0' = 24
'1' = 14
Frequency test = 2.63158
```

Рисунок 9 - Частотный тест

Тест серий для полученной последовательности приведен ниже.

```
k = 1 e_1 = 5
Number of blocks(G_1) of length 1 = 6
Number of breaks(B_1) of length 1 = 3
Batch test = 1
```

Рисунок 10 - Тест серий

Тест последовательностей для полученной приведен ниже.

```
n_000 = 16
n_001 = 2
n_010 = 3
n_011 = 3
n_100 = 2
n_101 = 3
n_110 = 2
n_111 = 5
Sequence_test = 31.4795
```

Рисунок 11 - Тест последовательностей

Автокорреляционный тест для полученной последовательности представлен ниже.

```
Autocorrelation test:
X_1 = -2.30159
X_2 = -3
X_3 = -1.69031
X_4 = 0
X_5 = -1.04447
X_6 = -1.06066
X_7 = -1.43684
X_8 = -2.19089
X_9 = -1.48556
X_10 = -0.377964
X_11 = -0.3849
X_12 = 0.784465
X_13 = 0.8
X_14 = 0
X_15 = 0.834058
X_16 = 0
X_17 = 0
X_18 = 1.34164
X_19 = 0.458831
```

Рисунок 12 - Автокорреляционный тест

График автокорреляции для генератора Геффе представлен ниже.

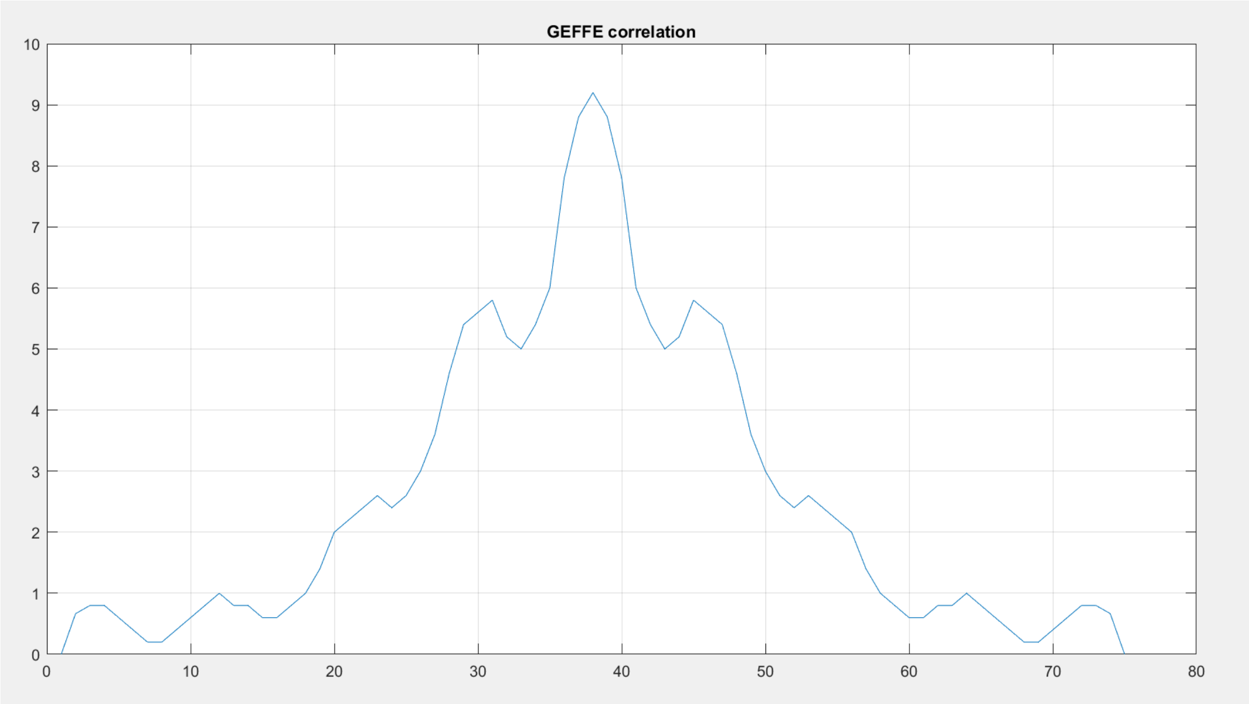

График 2 - График корреляции

Для уровня значимости $\alpha = 0.05$, сравним пороговые значения каждого теста и полученные:

Таблица 2 - Сравнение значений

Название теста	Пороговое значение	Полученное значение
Частотный тест	3,8415	2,6315
Тест серий	9,4877	1
Тест последовательностей	5,9915	31,4795
Автокорреляционный тест	11,0705	Максимальное значение = 1,34164

Таким образом, можно сделать вывод, что пройдены все тесты кроме теста последовательностей.

Выводы: криптографических алгоритмы, созданные на основе РСЛОС, смогут обеспечить высокое быстродействие. Однако, одна из главных проблем РСЛОС в том, что их программная реализация крайне неэффективна: приходится избегать разреженных многочленов обратной связи, так как они приводят к облегчению взлома корреляционным вскрытием, а плотные многочлены очень медленно просчитываются.

Также, линейность последовательности на выходе регистра позволяет однозначно определить многочлен обратной связи по последовательным битам с помощью алгоритма Берлекэмпа — Мэсси. Наконец, относительная лёгкость анализа алгебраическими методами не только облегчает разработку, но и увеличивает шансы на взлом генератора на базе РСЛОС.

Что же касается генератора Геффа, генератор криптографически слаб, потому что информация о состояниях генераторов РСЛОС содержится в его выходной последовательности. По этой причине, несмотря на длинный период и достаточно высокую линейную сложность, генератор Геффа поддаётся атакам.

Список используемой литературы:

1. Овчинников, А. А. Криптографические методы защиты информации: учеб. пособие / А. А. Овчинников. – СПб.: ГУАП, 2021. – 133 с.

2. Беззатеев С.В., Крук Е.А., Овчинников А.А. Блоковые шифры: Учеб.пособие/СПб.:Изд-во Нестор, 2003, 64 с.

3. Alfred J. Menezes, Paul C. van Oorschot and Scott A. Vanstone. Handbook of Applied Cryptography. – CRC Press, 1996. – ISBN 0-8493-8523-7.