Министерство образования Республики Беларусь
Учреждение образования
«Белорусский государственный университет
информатики и радиоэлектроники»

Кафедра радиотехнических систем

С.Б. Саломатин

ПОТОЧНЫЕ КРИПТОСИСТЕМЫ

УЧЕБНОЕ ПОСОБИЕ

по курсам

**КОДИРОВАНИЕ И ЗАЩИТА ИНФОРМАЦИИ,
ОСНОВЫ КРИПТОЛОГИИ**

для студентов специальностей «Радиоэлектронные системы»,
«Радиоэлектронная защита информации» дневной формы обучения

Минск 2006

УДК 621.391.25 (075.8)
ББК 32.811.4 я 73
С 16

Р е ц е н з е н т:
профессор, д-р техн. наук В.К. Конопелько

Саломатин С.Б.

С 16 Поточные криптосистемы: Учеб. пособие по курсам «Кодирование и защита информации», «Основы криптологии» для студ. спец. «Радиоэлектронные системы», «Радиоэлектронная защита информации» дневной формы обуч. /С.Б. Саломатин. – Мн.: БГУИР, 2006. – 76 с.: ил.

ISBN 985-444-882-7

В учебном пособии рассмотрены методы построения и анализа поточных криптосистем с использованием теории дискретных функций, обладающих криптографическими свойствами. Свойства рассматриваются с использованием аппарата булевых функций, операций регистра сдвига с обратными связями, теории конечных полей. Криптографический анализ выполняется на основе спектральных и корреляционных преобразований, решения задачи линейной аппроксимации и криптографических атак.

УДК 621.391.25 (075.8)
ББК 32.811.4 я 73

ISBN 985-444-882-7

ВВЕДЕНИЕ

Криптосистемы с последовательным выполнением преобразований над элементами открытого текста называются поточными шифрами. Элементы открытого текста, как правило, имеют небольшой размер. Таким элементом может быть буква алфавита естественного языка или цифра сообщения.

Поточная криптосистема преобразует каждую цифру сообщения. Секретный ключ используется для генерации ключевой гаммы (ключевого потока), накладываемой на преобразуемое сообщение. Процедуры или устройства, генерирующие ключевой поток, называются генератором ключевого потока.

Обозначим цифры открытого текста как $m_1, m_2, m_3, \ldots$, а символы ключевого потока как $k_1, k_2, k_3, \ldots$. Аддитивный поточный шифр $\boldsymbol{c} = (c_1, c_2, c_3, \ldots)$ получается в результате последовательного суммирования по модулю 2 одного символа открытого текста с соответствующим символом ключевого потока:

$$c_i = m_i \oplus k_i,\ i = 1,2,3,\ldots,$$

где $\oplus$ – операция суммирования по модулю 2.

Поскольку суммирование и вычитание по модулю 2 совпадают, то обратная процедура расшифрования определяется как

$$m_i = c_i \oplus k_i, i = 1,2,3,\ldots.$$

Это означает, что шифрование и расшифрование могут выполняться на идентичных устройствах или с помощью одного и того же алгоритма. Если длина ключа сравнима или больше длины сообщения, то такой шифр называют *шифром Вернама.*

Если цифры ключевого потока формируются независимо и случайно, то шифр Вернама называют *одноразовым блокнотом* (*шифром одноразового применения*). Считается, что такой шифр устойчив против атаки с использованием только шифротекста. Действительно, если принять, что открытый текст, шифротекст и ключевой поток являются случайными переменными M, C и K, характеризующимися соответствующими энтропийными функциями H( ), то для рассматриваемого случая получаем H(M|C) = H(M). Это означает, что шифротекст не несет дополнительной информации об открытом тексте, взаимная информация I(M;C) = 0.

Модель Шеннона, разработанная для симметричных шифров, определяет следующее условие существования устойчивых шифров: H(K) ≥ H(M). Из этого следует, что длина секретного ключа должна быть больше длины открытого текста. Если ключ имеет длину k и символы ключа формируются случайно и независимо, тогда энтропия ключа H(K) = k ≥ H(M). Устойчивость шифра типа «одноразовый блокнот» позволяет применять его к шифрованию статистически распределенных открытых текстов, при этом считается, что такой шифр оптимален, в смысле применения секретного ключа наименьшей длины по сравнению с другими возможными способами симметричного шифрования.

Недостатком шифра одноразового применения является необходимость применения очень длинного секретного ключа, что вызывает трудности для системы управления и распределения ключей. Более удобным был бы поточный шифр, ключевой поток которого имеет псевдослучайные свойства, обеспечивает хорошее рассеивание ключа и формируется с помощью секретного ключа меньшего размера. Такие шифры не удовлетворяют требованиям абсолютной криптоустойчивости $H(K)<<H(M)$, но могут относиться к классу вычислительно стойких шифров. Их стойкость основывается на построении генератора ключевого потока, вырабатывающего псевдослучайную последовательность, которая для криптоаналитика с ограниченными вычислительными ресурсами является неотличимой от совершенно случайной последовательности. Такие шифры называются *аддитивными поточными шифрами.*

Поточная аддитивная криптосистема независимо преобразует каждую цифру сообщения. Секретный ключ управляет генератором ключевого потока и определяет очень длинную последовательность псевдослучайных цифр, вырабатываемую по детерминированному закону и имеющую очень большой период. В случае поточных шифров предположение о наличии известного исходного текста эквивалентно тому, что криптоаналитику известна часть ключевого потока. Разработка аддитивного шифра заключается в построении такого генератора ключевого потока, для которого криптоаналитик с ограниченными ресурсами по известным n битам ключевого потока не мог бы определить какую-либо неизвестную часть ключевого потока.

При использовании поточного шифра необходимо гарантировать, чтобы никакая часть ключевого потока не использовалась более одного раза. Отсюда вытекает необходимость конструирования такого конечного автомата, автономное поведение которого позволяло получить псевдослучайные последовательности необходимой длины.

1. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ДИСКРЕТНОЙ КРИПТОЛОГИИ

В криптографических исследованиях дискретные функции используются как математический объект, моделирующий преобразования информации в криптосистемах. Дискретные функции представляют собой отображения конечных алгебраических объектов – множеств, групп, колец, полей, векторных пространств.

Наиболее исследован класс булевых функций и булевых отображений. Источником исследований криптографических свойств булевых функций и отображений явились сформулированные К. Шенноном основы построения преобразований информации для криптографических систем. Важнейшими из них являются принципы:

- «перемешивания» – дает неформальную трактовку понятия для эргодических систем с конечным числом состояний;
- «рассеивания» (diffusion), посредством которого «…статистическая структура сообщений, которая приводит к избыточности в сообщениях, «распыляется» в статистику больших длин, т.е. в статистику структур, включающую длинные комбинации букв криптограмм»;
- «запутывания» (confusion), который «состоит в том, что соотношения между простыми статистиками в пространстве криптограмм и простыми подмножествами в пространстве ключей делаются весьма сложными и беспорядочными».

1.1. Булевы функции и отображения

Криптографическое преобразование информации можно формализовать в виде отображения некоторого пространства $GF(2^n)$ n-мерных векторов над полем $GF(2)$ $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ в другое пространство $GF(2^m)$ m-мерных двоичных векторов $\mathbf{y} = (y_1, y_2, \ldots, y_m)$, где для любого $i \in \{1,\ldots, n\}$ и любого $j \in \{1,\ldots, m\}$, $x_i \in GF(2)$, $y_i \in GF(2)$. Отображения такого рода удобно задавать в виде векторной булевой функции (БФ):

$$\mathbf{y} = \varphi(\mathbf{x}) : GF(2^n) \to GF(2^m),$$

которая является объединением компонентных БФ $f_i(\mathbf{x})$, выполняющих отображение $GF(2^n) \to GF(2)$, $\varphi(\mathbf{x}) = \{f_1(\mathbf{x}), f_2(\mathbf{x}),\ldots f_m(\mathbf{x})\}$.

Для описания БФ используется описание их в виде таблицы истинности и в виде полинома

$$f(\mathbf{x}) = a_0 \oplus \sum_{1\le i\le n} a_i x_i \oplus \sum_{1\le i_1<i_2\le n} a_{i_1 i_2} x_{i_1} x_{i_2} \oplus$$

$$\oplus \sum_{1\le i_1<\ldots<i_k\le n} a_{i_1\ldots i_k} x_{i_1}\ldots x_{i_k} \oplus \ldots \oplus a_{12\ldots n} x_1 x_2 \ldots x_n,$$

где $\mathbf{x} \in GF(2^n)$, а все коэффициенты $a \in GF(2)$.

Полином такого вида носит название *алгебраической нормальной формы* (АНФ).

АНФ БФ над полем $GF(2^n)$ представляет собой сумму взятых с определенными коэффициентами всевозможных произведений переменных. Количество перемножаемых переменных в крайнем правом элементе АНФ является алгебраической степенью нелинейности БФ $f(\mathbf{x})$ и обозначается как $deg(f)$.

Пример. Для функции, заданной табл. 1.1,

Таблица 1.1

№	$x_1\, x_2\, x_3$	f
0	0 0 0	1
1	0 0 1	0
2	0 1 0	1
3	0 1 1	0
4	1 0 0	0
5	1 0 1	0
6	1 1 0	1
7	1 1 1	1

БФ будет иметь вид $f(x_1, x_2, x_3) = 1 \oplus x_1 \oplus x_2 \oplus x_3 \oplus x_1 \cdot x_2 \oplus x_1 \cdot x_3$.

Функция $f(x_1, x_2, \ldots, x_n)$ называется *линейной*, если она содержит только первые степени слагаемых. Более точно функция называется линейной, если ее можно представить в виде

$$f(x_1, x_2, \ldots, x_n) = a_0 + a_1 x_1 + a_2 x_2 + \ldots + a_n x_n \ .$$

Число линейных функций n переменных равно 2^{n+1}. Если $n \geq 2$, то линейная функция в таблице истинности может содержать только четное число единиц.

Вес $wt(f)$ функции f над $GF(2^n)$ – это число наборов x из $GF(2^n)$, для которых $f(x)=1$. Функция f называется *уравновешенной*, если $wt(f) = wt(f \oplus 1) = 2^{n-1}$ (т. е. функция принимает значения 0 и 1 на одинаковом числе наборов).

Расстоянием Хэмминга $d(\mathbf{x}_1, \mathbf{x}_2)$ между двумя наборами $\mathbf{x}_1$ и $\mathbf{x}_2$ называют число компонент, в которых наборы $\mathbf{x}_1$ и $\mathbf{x}_2$ различаются. Наборы называются соседними, если $d(\mathbf{x}_1, \mathbf{x}_2) = 1$. Для двух булевых функций f_1, f_2 на $GF(2^n)$ расстояние между f_1 и f_2 определяется как $d(f_1, f_2) = \left| \left\{ \mathbf{x} \in F_2^n \mid f_1(\mathbf{x}) \neq f_2(\mathbf{x}) \right\} \right|$. Можно заметить, что $d(f_1, f_2) = wt(f_1 \oplus f_2)$. Для заданной функции f из $GF(2^n)$ минимум расстояний $d(f, \gamma)$, где γ пробегает множество всех аффинных функций из $GF(2^n)$, называется нелинейностью функции f и обозначается через N_f.

Переменные булевой функции часто называют ее входами, а принимаемые булевой функцией значения – её выходами.

1.2. Спектральное представление булевых функций

В качестве основного аппарата анализа и изучения особенностей критериев удобно выбрать преобразования Фурье и Уолша булевых функций (БФ).

Обозначим через **x**, **ω**, ***a***, **s** двоичные наборы длины n, через x_i , ω_i , a_i , s_i – координаты этих наборов. Если $f(x_1, \ldots, x_n)$ – булева функция двоичных переменных, то через $\breve{f}(x_1, \ldots, x_n) = (-1)^{f(x_1,\ldots,x_n)}$ обозначим сопряженную функцию (СФ), определенную на том же множестве. Функции f и $\breve{f}$ однозначно определяют друг друга.

Пусть $\mathbf{x} = (x_1, \ldots, x_n)$ и $\mathbf{u} = (u_1, \ldots, u_n)$ – это наборы длиной n над $GF(2)$. Скалярное произведение x и u – это целочисленная функция, которая определяется как

$$\langle \mathbf{x}, \mathbf{u} \rangle = \sum_{i=1}^{n} x_i u_i \, .$$

Преобразование Уолша булевой функции $f(\mathbf{x})$ обозначается как

$$W_f(\omega) = \sum_{\mathbf{x} \in F_2^n} f(\mathbf{x})(-1)^{\langle \mathbf{x}, \omega \rangle} \, .$$

Спектральное преобразование функции $\breve{f}(\mathbf{x})$ обозначается через $W_{\breve{f}}(\omega)$ или

$$W_{\breve{f}}(\omega) = \sum_{\mathbf{x} \in F_2^n} (-1)^{f(\mathbf{x}) + \langle \mathbf{x}, \omega \rangle}$$

и носит название преобразования Уолша–Адамара БФ.

Спектральное преобразование эквивалентно умножению матрицы Адамара вида

$$\mathbf{H}_0 = 1, \; \mathbf{H}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \; \mathbf{H}_n = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \otimes \mathbf{H}_{n-1}$$

на вектор. Здесь $\otimes$ обозначает кронекеровское произведение матриц.

Можно показать, что введенные спектры связаны соотношениями

$$W_f(\omega) = -\frac{1}{2} W_{\breve{f}}(\omega) + 2^{n-1} \delta(\omega) \, , \;\; W_{\breve{f}}(\omega) = -2 W_f(\omega) + 2^n \delta(\omega) \, ,$$

где $\delta(0\ldots0) = 1$, $\delta(\boldsymbol{\omega}) = 0$ для $\boldsymbol{\omega} \neq (0\ldots0)$.

Пример. Рассмотрим линейную булеву функцию вида

$$f_1(\mathbf{x}) = x_1 \oplus x_2 \oplus x_3 \oplus x_4 \oplus x_5 \oplus x_6 \oplus x_7 \, .$$

Последовательность **s**, порождаемая такой функцией, имеет вид

[0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 1] .

Сопряженная последовательность равна

[1., -1., -1., 1., -1., 1., 1., -1., -1., 1., 1., -1., 1., -1., -1., 1., -1., 1., 1., -1., 1., -1., -1., 1., 1., -1., -1., 1., -1., 1., 1., -1., -1., 1., 1., -1., 1., -1., -1., 1., 1., -1., -1., 1., -1., 1., 1., -1., 1., -1., -1., 1., -1., 1., 1., -1., -1., 1., 1., -1., 1., -1., -1., 1., 1., -1., 1., -1., -1., 1., -1., 1., 1., -1., 1., -1., -1., 1., 1., -1., -1., 1., -1., 1., 1., -1., 1., -1., -1., 1., 1., -1., 1., -1., -1., 1., -1., 1., 1., -1., 1., -1., -1., 1., 1., -1., -1., 1., -1., 1., 1., -1., 1., -1., -1., 1., 1., -1., 1., -1., -1., 1., -1., 1., 1., -1., 1., -1., -1., 1., 1., -1., 1., -1., -1., 1., -1., 1., 1., -1.] .

Спектр Уолша–Адамара сопряженной последовательности показан на рис. 1.1.

Пример. Рассмотрим нелинейную булеву функцию вида

$$f_2\ (\mathbf{x}) = x_1\, x_5\, x_7 \oplus x_2\, x_7 \oplus x_4\, x_6 \oplus x_4 \oplus x_3\ .$$

Последовательность **s**, порождаемая такой функцией, равна

[0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 1, 1, 1, 1, 0, 0, 1, 1, 0, 0, 0, 0, 1, 1, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 1] .

Эквивалентная сопряженная последовательность равна

[1., 1., 1., 1., -1., -1., -1., -1., -1., -1., -1., -1., 1., 1., 1., 1., 1., 1., 1., 1., -1., -1., -1., -1., -1., -1., -1., -1., 1., 1., 1., 1., 1., 1., 1., 1., -1., -1., -1., -1., 1., 1., 1., 1., -1., -1., -1., -1., 1., 1., 1., 1., -1., -1., -1., -1., 1., 1., 1., 1., -1., -1., -1., -1., 1., 1., -1., -1., -1., -1., 1., 1., -1., -1., 1., 1., 1., -1., 1., -1., -1., 1., 1., 1., -1., -1., -1., -1., 1., 1., 1., 1., -1., -1., -1., -1., 1., 1., 1., -1., -1., 1., -1., 1., 1., -1., 1., -1., -1., 1., -1., 1., 1., -1.] .

Спектр Уолша–Адамара последовательности показан на рис. 1.2.

Спектральные преобразования позволяют непосредственно оценить такие показатели качества БФ, как сбалансированность, нелинейность и корреляционную иммунность.

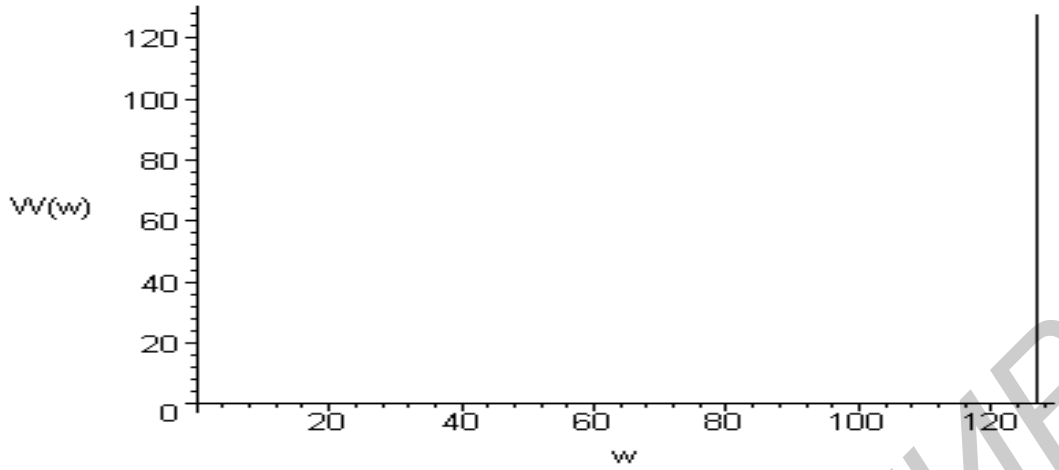

Рис. 1.1. Спектр Уолша–Адамара линейной функции вида $f_1(\mathbf{x}) = x_1 \oplus x_2 \oplus x_3 \oplus \oplus x_4 \oplus x_5 \oplus x_6 \oplus x_7$.

Рис.1.2. Спектр Уолша–Адамара нелинейной функции вида $f_2\ (\mathbf{x}) = x_1\, x_5\, x_7 \oplus \oplus x_2 x_7 \oplus x_4\, x_6 \oplus x_4 \oplus x_3$.

Рис. 1.3. Спектр Уолша нелинейной функции вида $f(\mathbf{x}) = x_1 x_2 + x_3 x_4 + x_5 x_6$.

1.3. Показатели качества булевых функций

Сбалансированность БФ. Для противодействия прямых статистических атак на криптоалгоритмы необходимо, чтобы выполнялись следующие условия:

- все компоненты БФ, реализующие преобразование, должны быть сбалансированы;

- преобразование в целом должно быть регулярным.

БФ $f(\mathbf{x})$ называется *сбалансированной*, если количество единиц в её таблице истинности равно количеству нулей, т.е. $\#\{\mathbf{x} \mid f(\mathbf{x}) = 0\} = \#\{\mathbf{x} \mid f(\mathbf{x}) = 1\} = 2^{n-1}$.

Отображение $\mathbf{y} = \varphi(\mathbf{x})$: $GF(2^n) \to GF(2^m)$ называется *регулярным* , если функции $\mathbf{y}$ ровно 2^{n-m} раза принимает все 2^m различных значений из $GF(2^m)$, в то время как $\mathbf{x}$ проходит 2^n значений из $GF(2^n)$.

Необходимым условием регулярности отображения $\mathbf{y} = \varphi(\mathbf{x})$: $GF(2^n) \to GF(2^m)$ при $n \geq m$ является сбалансированность любых линейных комбинаций компонентных БФ, реализующих векторную БФ $\varphi(\mathbf{x}) = \{f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x})\}$. Под линейной комбинацией вектора $\varphi(\mathbf{x})$ для любых $\boldsymbol{\lambda} = (\lambda_1, \ldots. \lambda_m) \in GF(2^m)$ понимается сумма $\sum_{i=1}^{m} \lambda_i f_i(x)$. Сбалансированность может быть оценена с помощью спектральных преобразований. В терминах преобразования Уолша–Адамара сбалансированность эквивалентна выполнению условий

$$W_{\tilde{f}}(\omega) = 0, \; W_f(\omega) = 2^{n-1},$$

где $\boldsymbol{\omega} = (0,\ldots,0)$. Степень отклонения БФ от сбалансированности определяет значение спектра Уолша–Адамара в точка «0».

Пример. Рассмотрим БФ $f(\mathbf{x})$, $\mathbf{x} \in GF(2^6)$, которая имеет вид $f(\mathbf{x}) = x_1 x_2 + x_3 x_4 + x_5 x_6$.

Таблица истинности данной БФ, представляемая в виде последовательности элементов, имеет вид

s = [0 0 0 1 0 0 0 1 0 0 0 1 1 1 1 0 0 0 0 1 0 0 0 1 0 0 0 1 1 1 1 0 0 0 0 1 0 0 0 1 0 0 0 1 1 1 1 0 1 1 1 0 1 1 1 0 0 0 0 1].

Спектр Уолша последовательности, соответствующей таблице истинности, показан на рис. 1.3. Нулевая компонента спектра равна 28. Нулевая компонента спектра сопряженной функции будет равна $W_{\tilde{f}}(0) = 64 - 2\,28 = 8$. Следовательно, рассматриваемая функция является несбалансированной.

Корреляционные свойства БФ. Усилением свойства сбалансированности БФ является требование сбалансированности всех частных функций, полученных из исход-

ной функции фиксированием её k или менее переменных. Указанное требование позволяет обеспечить стойкость криптографических преобразований к статистическим атакам при фиксированных значениях битов на входе преобразования. Свойство связано с показателем *корреляционной иммунности*.

БФ $f(\mathbf{x})$, $\mathbf{x} \in GF(2^n)$ называется *корреляционно-иммунной порядка* k, $1 \le k < n$, если значение компонент спектра Уолша–Адамара сопряженной функции

$$W_{\bar{f}}(\omega) = 0,$$

для всех $\boldsymbol{\omega} \in GF(2^n)$, вес Хэмминга которых удовлетворяет неравенству $1 \le wt(\boldsymbol{\omega}) \le k$.

БФ является корреляционно-иммунной порядка k, если значения $\mathbf{y} = f(\mathbf{x})$ статистически независимы от любого набора из k компонент произвольного вектора аргументов $\mathbf{x} \in GF(2^n)$. Если некоторая функция является корреляционно-иммунной порядка k, то она будет иметь корреляцию с некоторыми наборами компонент вектора $\mathbf{x}$ размером большим, чем k, т.е. существуют векторы $\mathbf{z} \in GF(2^n)$ такие, что $wt(\mathbf{z}) > k$, $W_f(\mathbf{z}) \neq 0$. Отсюда следует, что максимальный порядок корреляционной иммунности БФ не превышает значения $n - 1$.

Единственными функциями, достигающими максимального порядка $k = n - 1$, являются аффинные БФ вида

$$f(x_1, \ldots, x_n) = x_1 \oplus x_2 \oplus \ldots \oplus x_n$$

и

$$f(x_1, \ldots, x_n) = x_1 \oplus x_2 \oplus \ldots \oplus x_n \oplus 1.$$

Вид спектра одной из такой функции показан на рис.1.1.

Достижение максимуму корреляционной-иммунности является достаточно сильным требованием, и для реальных случаев можно использовать понятие корреляционно-эффективных БФ, для которых не менее чем на половине векторов веса $1 \le wt(\boldsymbol{\omega}) \le q$ значения компонент спектра Уолша–Адамара равны 0. Спектр последовательности, показанный на рис. 1.2, имеет 88 нулевых значений, что позволяет отнести БФ к классу корреляционно-эффективных.

Корреляционно-иммунная степени k булева функция называется совершенной корреляционно-иммунной степени k, или *резилентой* функцией, если она является сбалансированной функцией.

При синтезе БФ, обеспечивающих высокую стойкость к разностному, линейному и корреляционному криптоанализу, большое значение имеет автокорреляционная функция.

В случае обычного преобразования Фурье энергетический спектр и автокорреляционная функция играют важную роль и имеют широкую область приложений. Известная теорема Винера–Хинчина утверждает, что обратное преобразование энерге-

тического спектра реализуется в автокорреляционной функции. Это соотношение можно перенести и на преобразования рассматриваемых функций.

Автокорреляционная функция (АКФ) булевой функции $f(\mathbf{x})$ определяется как

$$r_f(\mathbf{a}) = \sum_{\mathbf{x}\in F_2^n} \breve{f}(\mathbf{x})\breve{f}(\mathbf{x}\oplus\mathbf{a}).$$

Преобразование Уолша–Адамара АКФ $r(\mathbf{a})$ обозначим как $R(\boldsymbol{\omega})$. Можно ввести формальное понятие энергетического спектра булевой функции $f(\mathbf{x})$ через $W_{\breve{f}}^2(\omega)$. Справедливо следующее соотношение $R(\omega) = W_{\breve{f}}^2(\omega)$ для всех $\boldsymbol{\omega}$ из $GF(2^n)$.

Одно из главных свойств автокорреляционной функции состоит в возможности её использования для описания критериев безопасности булевой функции через оценку вероятностных параметров:

$$\Pr(f(\mathbf{x}) \neq f(\mathbf{x}\oplus\mathbf{a})) = \Pr(\breve{f}(\mathbf{x}) \neq \breve{f}(\mathbf{x}\oplus\mathbf{a})) = \frac{1}{2} - \frac{r_f(\mathbf{a})}{2^{n+1}}$$

и

$$\sum_{\beta\in GF(2)^n,\beta\neq 0} \Pr\{\breve{f}(\mathbf{x}) \neq \breve{f}(\mathbf{x}\oplus\beta)\} = 2^{n-1} - \frac{W_{\breve{f}}^2(\omega)}{2^{n+1}}\ .$$

Связь корреляционно-иммунных булевых функций с кодами и ортогональными массивами. Напомним, что произвольное множество наборов $\mathrm{C}\in F_2^n$ называется двоичным кодом. Понятия булевой функции и кода тесно связаны. Произвольная булева функция f на F_2^n ассоциируется с её *характеристическим множеством* – кодом C : $\{x \in F_2^n \mid f(x) = 1\}$. Наоборот, произвольному коду $\mathrm{C}\in F_2^n$ может быть поставлена в соответствие его характеристическая функция $f : f(x) = \begin{cases} 1, & \text{если } x \in C, \\ 0, & \text{если } x \notin C. \end{cases}$

Код C называется линейным, если для любых $x, y, z \in F_2^n$, таких, что $z = x \oplus y$, имеем $z \in$ C, если $x, y \in$ C.

Двоичный ортогональный массив $OA(h, n, 2, m)$ – это матрица размером $h\times n$, клетки которой заполнены элементами из множества $\{0,1\}$, так что внутри любых m столбцов каждый упорядоченный поднабор двоичных символов встречается в точности в $\lambda = h/2^m$ строках. Ортогональный массив называется простым, если все строки в этом массиве попарно различны. Параметр m называется эффективностью OA. Любая двоичная матрица может рассматриваться как ортогональный массив (может быть нулевой эффективности).

Простому массиву $OA(h, n, 2, m)$ можно сопоставить код C в F_2^n, где C – это множество всех наборов из F_2^n, заданных строками OA. Если код C является линейным, то соответствующая ему максимальная эффективность OA равна (d` - 1), где d` – дуальное расстояние.

В общем случае корреляционно-иммунные функции и устойчивые отображения связаны с таким комбинаторным объектом, как ортогональная таблица. Ортогональной таблицей размером $(n \times m)$ с ограничениями уровня 2, силой t и индексом v называется таблица M над полем F_2, обладающая следующими свойствами. В любом подмножестве из t столбцов матрицы M любой из 2^t векторов пространства F_2^t встречается ровно v раз. Такая таблица обозначается как $OA_v(m, n, 2, t)$.

Для функции f определим таблицу истинности как матрицу M_f размером $wt(f) \times n$, строками которой являются наборы из F_2^n, значение функции на которых равно 1. Функция f от n переменных корреляционно-иммунная порядка t тогда и только тогда, когда её таблица истинности является ортогональной таблицей $OA_v(wt(f), n, 2, t)$.

Корреляционно-иммунная функция является частным случаем ортогонального массива, а именно, корреляционно-иммунная порядка m функция от n переменных с весом $wt(f)$ соответствует простому $OA(wt(f), n, 2, m)$ (все наборы x, для которых $f(x)=1$, задаются строками OA . Максимальный порядок корреляционной иммунности булевой функции равен максимальной эффективности соответствующего ей OA и равен дуальному расстоянию её характеристического кода, уменьшенному на 1.

Существующее соответствие между корреляционно-иммунными функциями, кодами и ортогональными массивами позволяет переносить результаты существования, справедливые для одной области в другие. Например, в теории ортогональных массивов получены классические неравенства для параметра h:

$$h \geq 2^n \left(1 - \frac{n}{2(m+1)}\right).$$

Для корреляционно-иммунных функций это означает, что

$$wt(f) \geq 2^n \left(1 - \frac{n}{2(m+1)}\right).$$

Простой ортогональный массив $OA(h, n, 2, m)$ при $h = 2^{m-1}$ не является интересным объектом для исследования в теории ортогональных массивов. Очевидно, что такой OA существует даже при $m - 1$ (соответствующая проверка на четность). Но для корреляционно-иммунных функций все обстоит как раз наоборот. Уравновешенные корреляционно-иммунные (устойчивые) функции наиболее интересны для криптографических приложений, в то время как линейная сумма по модулю 2 не является

криптографически хорошей функцией по критериям алгебраической степени и нелинейности. Поэтому важной проблемой является существование устойчивых функций, которые зависят от всех своих переменных нелинейно.

Нелинейность БФ. При анализе эффективности криптографических преобразований одной из основных характеристик является нелинейность БФ, реализующих данное преобразование, которая показывает степень удаленности БФ от множества аффинных или линейных БФ. Нелинейность БФ является важным показателем, т.к. линейные функции сравнительно легко вскрываются.

Под аффинными функциями $f_a(\mathbf{x}) \in \mathrm{A}$, где $\mathrm{A} = \{ f_a(\mathbf{x})\}$, $\mathbf{x} \in GF(2^n)$, понимаются БФ следующего вида:

$$f_a(\mathbf{x}) = \alpha_1 x_1 \oplus \alpha_2 x_2 \oplus \ldots \oplus \alpha_n x_n \oplus \mathrm{c},$$

где $\alpha, \mathrm{c} \in GF(2)$.

При $\mathrm{c} = 0$ аффинные функции образуют подмножество линейных функций

$$\Lambda = \{ f_л(\mathbf{x}) \},$$

каждая из которых является скалярным произведением вида $f_л(\mathbf{x}) = \langle \alpha, \mathbf{x} \rangle$.

Удаленность БФ от множества аффинных A или линейных Λ БФ оценивается через расстояние Хэмминга $d(f, g)$ между двумя БФ $f(\mathbf{x})$ и $g(\mathbf{x})$, которое определяет количество отличающихся значений функций:

$$d(f, g) = \#\{ f(\mathbf{x}) \neq g(\mathbf{x}) : \mathbf{x} \in GF(2)^n \} = wt(\mathbf{s}(f) \oplus \mathbf{s}(g)),$$

где $\mathbf{s}(\cdot)$ – таблица истинности БФ, представляемая в виде последовательности элементов.

Расстояние $d(f, g)$ может быть оценено через сопряженные функции $\breve{f}(\mathbf{x})$, $\breve{g}(\mathbf{x})$ и сопутствующие им последовательности $\breve{\mathbf{s}}_f$, $\breve{\mathbf{s}}_g$:

$$d(f,g) = 2^{n-1} - \frac{1}{2} \sum_{\mathbf{x} \in GF(2)^n} \breve{f}(\mathbf{x}) \breve{g}(\mathbf{x}) = 2^{n-1} - \frac{1}{2} \left\langle \breve{\mathbf{s}}_f, \breve{\mathbf{s}}_g \right\rangle,$$

а также через значение $r_\tau(f,g) = \sum_{\mathbf{x} \in GF(2)^n} \breve{f}(\mathbf{x}) \breve{g}(\mathbf{x} \oplus \tau)$ взаимно корреляционной функции при нулевом сдвиге $\tau = 0$:

$$d(f,g) = 2^{n-1} - \frac{1}{2} r_0(f,g).$$

Таким образом, под нелинейностью N_f БФ

$$f(\mathbf{x}) : GF(2^n) \rightarrow GF(2)$$

понимается значение минимального расстояния Хэмминга между функцией $f(\mathbf{x})$ и функциями $f_a(\mathbf{x}) \in \mathrm{A}$:

$$N_f = \min_{f_a \in \mathrm{A}} d(f, f_a).$$

В терминах спектрального преобразования значение нелинейности оценивается как

$$N_f = 2^{n-1} - \frac{1}{2} \max_{\omega \in GF(2^n)} \left| W_{\tilde{f}}(\omega) \right|.$$

Для оценки нелинейности БФ следует определить максимальное значение компонент спектра Уолша–Адамара.

Верхняя граница значения нелинейности.

$$N_f \le \begin{cases} 2^{n-1} - 2^{n/2-1}, \text{ для } n - \text{четного}; \\ \left\lfloor 2^{n-1} - 2^{n/2-1} \right\rfloor, \text{ для } n - \text{нечетного}, \end{cases}$$

где $\lfloor \cdot \rfloor$ – максимальное четное целое, меньшее либо равное значению аргумента.

Для сбалансированной БФ справедливы неравенства

$$N_f \le \begin{cases} 2^{n-1} - 2^{n/2-1} - 2, \text{ для } n - \text{четного}; \\ \left\lfloor 2^{n-1} - 2^{n/2-1} \right\rfloor, \text{ для } n - \text{нечетного}. \end{cases}$$

Если в преобразовании $GF(2^n) \to GF(2^m)$ используется векторная БФ $\varphi(\mathbf{x})$, то нелинейность определяется соотношением

$$N_\varphi = \min_{\mathbf{c} \in GF(2^m), \mathbf{c} \neq 0} N_{\langle \mathbf{c}, \varphi \rangle},$$

где минимум находится не только относительно всех компонентных БФ, но и любых линейных комбинаций данных БФ.

На практике используется понятие средней нелинейности векторной БФ $\varphi(\mathbf{x})$:

$$\overline{N}_f(\varphi(\mathbf{x})) = \frac{1}{2^n} \sum_{\mathbf{c} \in GF(2^n), \mathbf{c} \neq 0} N_{\langle \mathbf{c}, \varphi \rangle}.$$

Максимально нелинейные функции. Функции, для которых неравенство

$$\max_{\omega \in GF(2^n)} \left| W_{\tilde{f}}(\omega) \right| \ge 2^{n/2}$$

обращается в равенство, дальше всех других функций удалены от множества аффинных функций. В этом случае функция f называется *максимально нелинейной* (или

бент-функцией). Для неё коэффициенты спектра Уолша–Адамара равны $\pm 2^{n/2}$. Заметим, что при нечетном n максимально нелинейные функции не существуют.

В трактовке теории кодирования аффинные функции представляют собой код Рида–Маллера 1-го порядка $RM_2(n, 1)$. Понятие максимальной нелинейности функций $N_f = 2^{n-1} - 2^{(n/2)-1}$ совпадает с понятием радиуса покрытия $\rho(n, 1)$ кода $RM_2(n, 1)$.

Свойства максимально нелинейных функций:

1. Производная (разность) максимально нелинейной функции

$$\Delta f(\mathbf{x}, \mathbf{a}) = f(\mathbf{x}) \oplus f(\mathbf{x} \oplus \mathbf{a})$$

по любому направлению $\mathbf{a} \in F_2^n$ является уравновешенной функцией.

2. Функция $\tilde{f}$, определяющая знаки коэффициентов преобразования Уолша–Адамара максимально нелинейной функции f

$$W_{\tilde{f}}(\mathbf{a}) = (-1)^{\tilde{f}(\mathbf{a})} 2^{n/2},$$

сама является максимально нелинейной.

3. Степень нелинейности $\deg(f)$ или, другими словами, степень нелинейности АНФ, представляющей f, не превосходит $(n/2)$.

4. Множество максимально нелинейных функций инвариантно относительно действия элементов полной аффинной группы на пространстве $GF(2^n)$.

Несмотря на свои привлекательные свойства, максимально нелинейные функции не являются уравновешенными и поэтому не используются в чистом виде при синтезе криптосистем.

1.4. Критерии распространения изменений

Безопасность поточных криптосистем во многом зависит от характеристик булевых функций, описывающих различные преобразования информации в этих системах. Предложено несколько критериев, благодаря которым криптоанализ становится более трудным. Их основная сущность состоит в оценивании вероятности изменения значения БФ в зависимости от изменения части битов аргументов этих функций.

Критерий строгого лавинного эффекта (КСЛЭ). Пусть определена разность

$$\Delta f(\mathbf{x}, \beta) = f(\mathbf{x}) \oplus f(\mathbf{x} \oplus \boldsymbol{\beta}),$$

где $\mathbf{x}, \boldsymbol{\beta} \in GF(2^n), f(\mathbf{x}) \in GF(2)$, тогда БФ удовлетворяет:

- КСЛЭ, если разность $\Delta f(\mathbf{x}, \boldsymbol{\beta})$ является сбалансированной функцией для любых $\boldsymbol{\beta}$, вес которых $wt(\boldsymbol{\beta}) = 1$;

- КСЛЭ порядка k, если любая функция, полученная из $f(\mathbf{x})$ путем подстановки констант на места произвольных её k переменных, удовлетворяет КСЛЭ.

С точки зрения спектральных и корреляционных терминов БФ $f(\mathbf{x})$ удовлетворяет КСЛЭ, если для любых векторов $\boldsymbol{\beta} \in GF(2^n)$, вес которых $wt(\boldsymbol{\beta}) = 1$, автокорреляционная функция $r_f(\boldsymbol{\beta}) = 0$, что равносильно выполнению равенства

$$\sum_{\omega \in GF(2^n)} W_{\tilde{f}}^2(\omega)(-1)^{<\omega,\beta>} = 0 .$$

Критерий распространения (КР). Функция $f(\mathbf{x})$, $\mathbf{x} \in GF(2)$ удовлетворяет:

- КР степени $l \le n$ (КР(l)), если при замене $1 \le i \le l$ произвольных входных битов на свои дополнения функция $f(\mathbf{x})$ изменяется с вероятностью $\Pr\{f(\mathbf{x}) \ne f(\mathbf{x} \oplus \boldsymbol{\beta})\} = 0{,}5$, что эквивалентно сбалансированности разности $\Delta f(\mathbf{x}, \boldsymbol{\beta})$ для любых $\boldsymbol{\beta}$, вес которых удовлетворяет неравенству $1 \le wt(\boldsymbol{\beta}) \le l$;

- КР степени $l \le n$ и порядка $k \le n - 1$, если любая функция, полученная из БФ $f(\mathbf{x})$ фиксацией любых её k переменных, удовлетворяет КР(l).

В терминах спектральных свойств БФ удовлетворяет КР степени l, если её автокорреляционная функция $r_f(\boldsymbol{\beta}) = 0$ на всех векторах $\boldsymbol{\beta} \in GF(2^n)$, вес Хэмминга которых лежит в пределах $1 \le wt(\boldsymbol{\beta}) \le l$. Критерий распространения является обобщением КСЛЭ.

Рассмотрим БФ вида $f(\mathbf{x}) = \bigoplus_{i=1}^{n-1} \bigoplus_{j=i+1}^{n} x_i x_j$. Она удовлетворяют следующим показателям:

- КСЛЭ порядка ($n - 2$);
- КР степени k и порядка m при $k + m \le n - 1$ или ($k + m = 1$ и k – четное);
- КР степени 2 и порядка $n - 2$;
- КР степени n при четном n;
- КР степени ($n - 1$) при четном n и $f(\mathbf{x})$ – сбалансированная функция.

Пример. Зададим над векторным пространством $GF(2)^n$ $n = 8$ БФ вида

$$f(\mathbf{x}) = \bigoplus_{i=1}^{7} \bigoplus_{j=i+1}^{8} x_i x_j .$$

Таблица истинности функции имеет вид

$\mathbf{s}_f$ = [0,0,0,1,0,1,1,1,0,1,1,1,1,1,1,0,0,1,1,1,1,1,1,0,1,1,1,0,1,
0,0,0,0,1,1,1,1,1,1,0,1,1,1,0,1,0,0,0,1,1,1,0,1,0,0,0,1,0,0,0,0,0,0,1,0,1,1,1,1,
1,1,0,1,1,1,0,1,0,0,0,1,1,1,0,1,0,0,0,1,0,0,0,0,0,0,1,1,1,1,0,1,0,0,0,1,0,0,0,0,
0,0,1,1,0,0,0,0,0,0,1,0,0,0,1,0,1,1,1,0,1,1,1,1,1,1,0,1,1,1,0,1,0,0,0,1,1,1,0,1,
0,0,0,1,0,0,0,0,0,0,1,1,1,1,0,1,0,0,0,1,0,0,0,0,0,0,1,1,0,0,0,0,0,0,1,0,0,0,1,0,
1,1,1,1,1,1,0,1,0,0,0,1,0,0,0,0,0,0,1,1,0,0,0,0,0,0,1,0,0,0,1,0,1,1,1,1,0,0,0,0,
0,0,1,0,0,0,1,0,1,1,1,0,0,0,1,0,1,1,1,0,1,1,1,1,1,1,0].

Спектр Уолша–Адамара сопряженной последовательности показан на рис.1.4.

Рис. 1.4. Спектр максимально нелинейной булевой функции

Автокорреляционная функция $r_f(\boldsymbol{\beta}) = 0$ на всех векторах $\boldsymbol{\beta}$, вес которых $wt(\boldsymbol{\beta}) \geq 1$. Таким образом, БФ имеет следующие показатели: КР степени 8 и КР степени 6 и порядка 2, более того БФ является бент-функцией.

Пример. Зададим функцию g от $n = 2k + 1$ переменных:

$$g(x^{(1)},...,x^{(2k+1)}) = x^{(1)} \oplus f(x^{(1)} \oplus x^{(2)},...,x^{(2k+1)}),$$

где f – максимально нелинейная функция от n переменных. Заданная таким образом функция g удовлетворяет критерию распространения степени $2k$. Нелинейность N_g функции g удовлетворяет неравенству $N_g \geq 2^{2k} - 2^k$.

Критерий наличия линейных структур. Свидетельствует о криптографической слабости отображения или функции, имеющей в своем составе линейные структуры. Говорят, что вектор $\mathbf{u} \in GF(2^n)$ является линейной структурой отображения Φ: $GF(2^n) \to GF(2^m)$, если выполняется равенство

$$\Phi(\mathbf{x}) \oplus \Phi(\mathbf{x} \oplus \mathbf{u}) = \mathbf{c} = const, \quad \mathbf{c} \in F_2^n .$$

Все линейные структуры отображения Φ образуют подпространство линейных структур L_Φ в пространстве $GF(2^n)$. Наличие линейных структур у отображения инвариантно относительно действия на отображение полной аффинной группы.

2. КРИПТОГРАФИЧЕСКИЕ ФУНКЦИИ, ЗАДАВАЕМЫЕ ЧЕРЕЗ МОДЕЛЬ РЕГИСТРА СДВИГА С ОБРАТНОЙ СВЯЗЬЮ

2.1. Модель регистра сдвига с обратной связью

Пусть $F = GF(2) = \{0, 1\}$ и $F^{(n)} = \{(a_0, a_1, \dots, a_{n-1}) \mid a_i \in F\}$. Функция с n двоичными (бинарными) входами и одним двоичным выходом называется булевой функцией n переменных. Булева функция выполняет отображение вида $f: F^{(n)} \to F$. Для заданного значения n переменных существует 2^{2^n} различных булевых функций.

Регистр сдвига с обратной связью (РСОС) состоит из двух основных частей: тактируемого n-разрядного регистра сдвига и блока обратной связи. С каждым тактом информация, записанная в разрядах регистра сдвига, сдвигается вправо (или влево), а на освободившееся место записывается информация с выхода блока обратной связи. В фиксированный момент времени информация, записанная в разрядах регистра сдвига, определяет код состояния РСОС. Состояние из первых n символов $(a_0, a_1, \dots, a_{n-1}) \in F^{(n)}$ называется начальным. Обратная связь чаще всего формируется с помощью булевой функции n переменных:

$$f(x_0, x_1, \dots, x_{n-1}) = \sum c_{i_1 i_2 \dots i_t} x_{i_1} x_{i_2} \dots x_{i_t} .$$

Выходная последовательность символов удовлетворяет следующему рекурсивному соотношению:

$$a_{n+k} = f(a_k, a_{k+1}, \dots, a_{k+n-1}), \quad k = 0,1,\dots .$$

Если функция обратной связи f является линейной, то выходной сигнал называется последовательностью регистра сдвига с линейной обратной связью (РСЛОС). В противном случае такой сигнал носит название последовательности регистра сдвига с нелинейной обратной связью (РСНОС).

Схема РСОС позволяет формировать не только двоичный, но и многозначный или q-ичный сигнал. Для описания работы схемы используется поле $F=GF(q)$, разряды регистры сдвига оперируют с q-значными числами, а функция обратной связи выполняет отображение $f: F^{(n)} \to F$ и определяется как

$$f(x_0, x_1, \dots, x_{n-1}) = \sum c_{i_1 i_2 \dots i_{n-1}} x_{i_0}^{i_0} x_{i_1}^{i_1} \dots x_{i_{n-1}}^{i_{n-1}}, \quad c_{i_1 i_2 \dots i_{n-1}} \in F .$$

Для заданных значений n и q существует q^{q^n} различных булевых функций.

Свойство периодичности. Пусть q будет простым числом или степенью p^m простого числа, последовательность $\mathbf{a} = a_0, a_1, \dots = \{ a_i \mid a_i \in F = GF(q)\}$ – q-значная последовательность. Условие существования периода повторения определяется следующим образом. Если существуют целые числа $r > 0$ и $i_0 \geq 0$, такие, что

$$a_{i+r} = a_i \text{ для всех } i \geq i_0,$$

то последовательность имеет период повторения, а r называют длиной последовательности. Наименьшее число, удовлетворяющее условию существования, называется периодом последовательности. Если $i_0 = 0$, то последовательность называется периодической.

Пример. Последовательность 0 0 0 1 1 0 1 1 0 1 1… , формируемая 4-разрядным регистром сдвига с обратной связью $f(x_0, x_1, x_2, x_3) = x_2 \oplus x_3$ с начальным состоянием 0001, имеет период повторения равный 3, который начинается с $i_0 = 2$.

Пример. Генератор 4-разрядный с нелинейной обратной связью. Функция обратной связи задается в виде $f(x_0, x_1, x_2, x_3) = 1 + x_0 + x_1 + x_1 x_2 x_3$. Схема генератора показана на рис. 2.1.

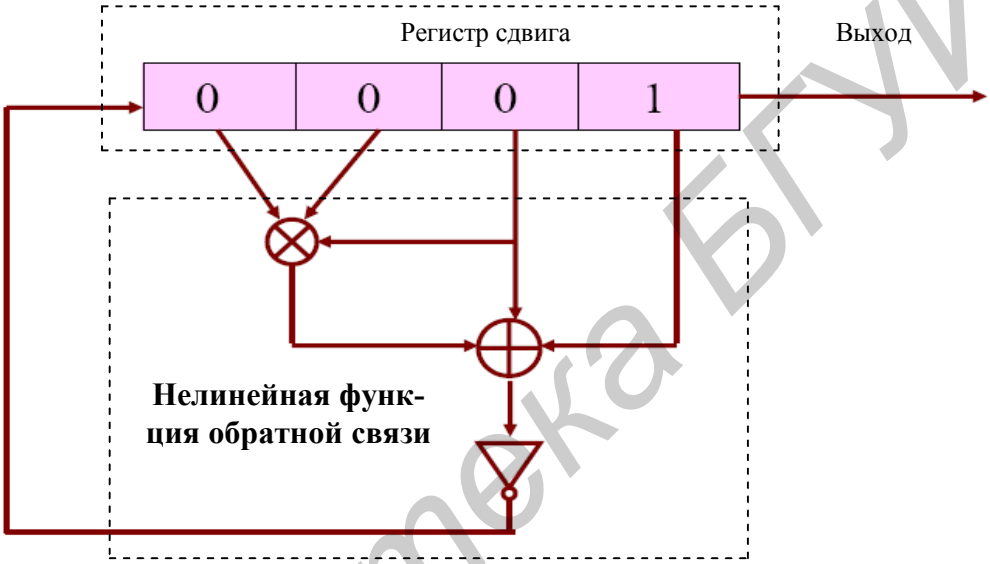

Рис. 2.1. Генератор на регистре сдвига с нелинейной обратной связью

Таблица истинности генератора (табл. 2.1) определяется как $g(x_0, x_1, x_2, x_3) = x_0 + x_1 + x_1 x_2 x_3$, $f = g + 1$.

Таблица 2.1

x_0, x_1, x_2, x_3	g	f	x_0, x_1, x_2, x_3	g	f
0000	0	1	0001	0	1
1000	1	0	1001	1	0
0100	1	0	0101	1	0
1100	0	1	1101	0	1
0010	0	1	0011	0	1
1010	1	0	1011	1	0
0110	1	0	0111	0	1
1110	0	1	1111	1	0

Диаграмма состояний генератора показана на рис. 2.2.

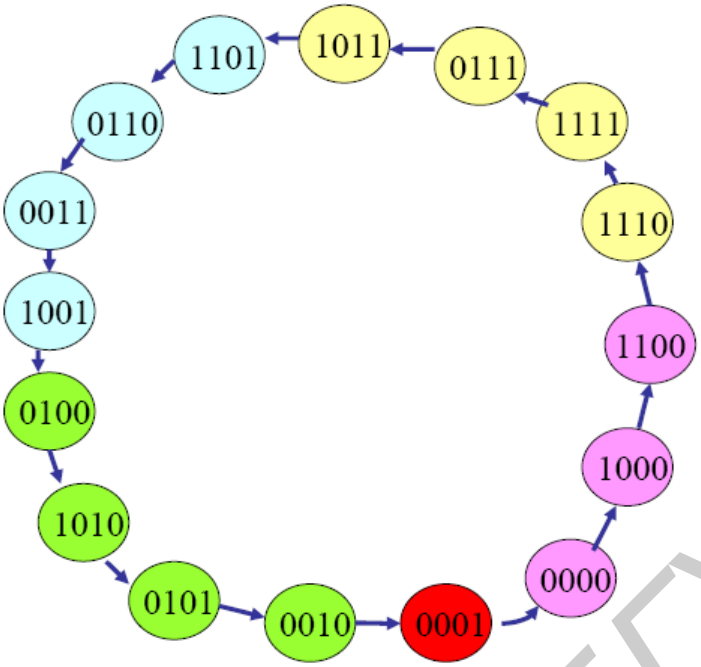

Рис. 2.2. Диаграмма состояний генератора де Брюина

Генератор формирует последовательность максимальной длины 2^n, которая носит название последовательность де Брюина.

Пример. Зададим поле GF(2) и определим неприводимый в поле полином $f(x) = x^4 + x^3 + x^2 + x + 1$. Определим циклическую структуру генератора РСЛОС, использующего $f(x)$ в качестве характеристического полинома. Заметим, что $(x^5 + 1) = (x + 1)(x^4 + x^3 + x^2 + + x + 1)$. Следовательно, $f(x) \mid (x^5 + 1)$, что дает период последовательностей, равный 5. Всего РСЛОС сформирует 15/5 = 3 различных последовательностей. Диаграммы состояний показаны на рис. 2.3.

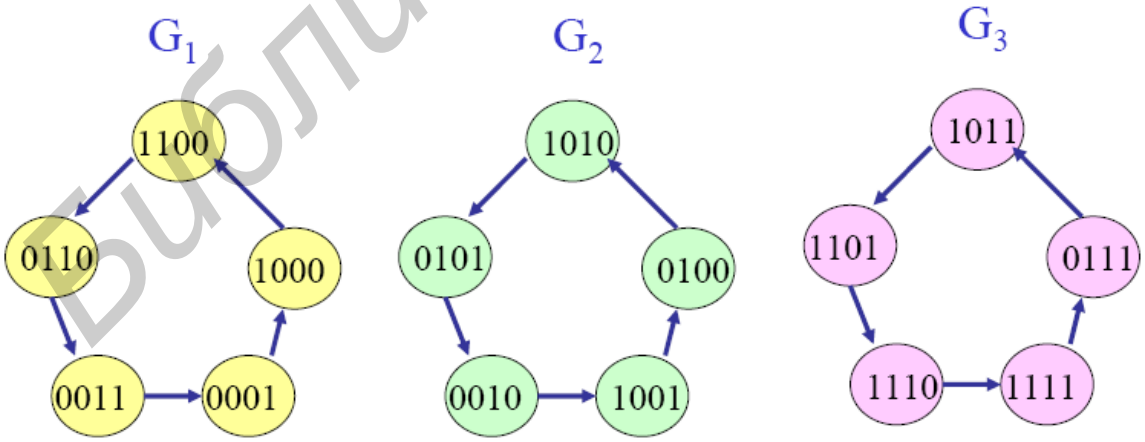

Рис. 2.3. Диаграммы состояний генератора РСЛОС

Генератор не обеспечивает максимального периода последовательности.

2.2. Регистры сдвига с линейной обратной связью

К простейшему типу регистра сдвига с обратной связью относится регистр сдвига с линейной обратной связью (РСЛОС). Обратная связь представляет собой операцию суммирования по некоему модулю (для бинарного сигнала – по модулю 2) символов, снимаемых с определенных разрядов регистра сдвига, которые определяет многочлен обратной связи.

Для обеспечения максимального периода конкретного РСЛОС соответствующий многочлен обратной связи должен быть примитивен. Степень многочлена является длиной регистра сдвига. Примитивный многочлен степени n – это неприводимый многочлен, который является делителем полинома $x^{2^{n-1}}+1$, но не является делителем x^d+1 для всех d, являющихся делителями числа 2^n-1.

Зададим конечное поле $F = GF(q)$ и примитивный полином

$$f(x) = x^n - c_{n-1}x^{n-1} - \ldots - c_1 x - c_0, \quad c_i \in F .$$

Псевдослучайная m-последовательность $\mathbf{a} = a_0, a_1, a_2, \ldots$ с периодом $q^n - 1$ над полем F задается рекурсивным выражением

$$a_{i+n} = \sum_{j=0}^{n-1} c_j a_{j+i}, \quad i = 0,1,\ldots .$$

Полином $f(x)$ определяет вид линейной обратной связи и называется характеристическим для последовательности **a**. Полином $m(x)$, имеющий наименьшую степень среди характеристических полиномов последовательности **a,** называется минимальным полиномом последовательности. *Линейная сложность* последовательности $LS(\mathbf{a})$ оценивается через степень минимального полинома $LS(\mathbf{a}) = deg(m(x))$.

Свойство s-децимации. Пусть $f(x)$ – неприводимый полином над полем $GF(q)$ степени n и s – положительное целое число и имеется последовательность $\mathrm{a} \in G(f)$. Последовательность $\mathrm{b} = \{b_i\}$, символы которой определяются как

$$b_i = a_{si}, \quad \forall i \ge 0 ,$$

называется децимированной последовательностью и обозначается как $\mathbf{b} = \mathbf{a}^{(s)}$.

Заметим, что если $f(x)$ – неприводимый полином над $GF(q)$ степени n, s – целое положительное число, α – корень $f(x)$ в расширенном поле $GF(q^n)$ и существует децимированная последовательность $\mathbf{a}^{(s)} \neq 0$, тогда минимальный полином последовательности $\mathbf{a}^{(s)}$ является минимальным полиномом и для α^s над $GF(q)$.

Рассмотрим s-децимированную m-последовательность. Последовательность будет также m-последовательностью, если выполняется условие НОД$(s, q^n - 1) = 1$.

Пример. Пусть $F = GF(2)$, $f(x) = x^4 + x + 1$, α – корень $f(x)$. Минимальным для α^s будет полиномом $f_{\alpha^s}(x)$. Для последовательности $\mathbf{a} \in G(f)$

$$\mathbf{a} = 100010011010111\ldots = (Tr(\alpha^{14}), Tr(\alpha^{14}\alpha), Tr(\alpha^{14}\alpha^{2}), \ldots)$$

получаем следующие виды децимированных последовательностей:

1) $\mathbf{a}^{(3)}$ = 1000110001…, период равен 5, минимальный полином $f_{\alpha 3}(x) = x^4 + x^3 + x^2 + x + 1$;

2) $\mathbf{a}^{(5)}$ = 101101…, период равен 3, минимальный полином $f_{\alpha 5}(x) = x^2 + x + 1$;

3) $\mathbf{a}^{(7)}$ = 111010110010001……, период равен 15, минимальный полином $f_{\alpha 7}(x) = x^4 + x^3 + 1$.

Перестановочные свойства m-последовательности. Пусть $d = (q^n - 1)/(q - 1)$, тогда:

- любая *m*-последовательность над полем *F* степени *n* может представлена в виде матрицы размером $(q - 1) \times d$, у который столбцами являются *m*-последовательности длиной $(q - 1)$ или нулевые последовательности;

- любая *m*-последовательность может быть преобразована в матрицу размером $(q^m - 1) \times v$, у которой столбцами являются *m*-последовательности степени *m* или нулевые последовательности, если *n* составное число, *m* – множитель факторизации числа *n*, и $v = (q^n - 1)/(q^m - 1)$.

Пример. Если $q = 2$, $n = 6$, $f(x) = x^6 + x + 1$. Псевдослучайная *m*-последовательность $\mathbf{a} \in GF(q)$ путем перестановки может быть преобразована в матрицы **A**(i,j) двух размеров: 3×21 или 7×9, у которых столбцами являются последовательности длинами соответственно, 3 или 7, а также последовательности, состоящие из одних нулей.

$$\mathbf{A}(3,21) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix},$$

$$\mathbf{A}(7,9) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Рассмотрим случай небинарных многозначных последовательностей. Возможно два варианта.

Вариант 1. Число $q = p$ и p – простое число.

Пример. Выберем $p = 3$, $n = 3$ и пусть $f(x) = x^3 - x^2 - 2$ – примитивный полином над полем $GF(3)$. Псевдослучайная последовательность имеет период 26 и может быть представлена в виде матрицы

$$\mathbf{A}(2,13) = \begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 0 & 2 & 1 & 1 & 2 & 1 & 0 & 1 \\ 0 & 0 & 2 & 2 & 2 & 0 & 1 & 2 & 2 & 1 & 2 & 0 & 2 \end{bmatrix}.$$

Вариант 2. Число $q = p^n$, p – простое число и $n > 1$.

Пример. Рассмотрим многозначную последовательность степени 3, периода 63, сформированную над полем $GF(2^2)$. Пусть $GF(2^2)$ определяется примитивным полиномом вида $h(x) = x^2 + x + 1$ и β – корень $h(x)$, тогда $GF(2^2) = \{0, 1, \beta, \beta^2\}$. Выберем примитивный полином степени 3 над полем $GF(2^2)$ вида $f(x) = x^3 + x^2 + \beta^2 + \beta$.

Псевдослучайная последовательность $\mathbf{a} = \{a_i\} \in GF(q)$ формируется с помощью рекурсивного соотношения

$$a_{3+k} = a_{2+k} + \beta^2 a_{1+k} + \beta\, a_k, \quad k = 0,1,2,\ldots .$$

Выберем начальное состояние (1,1,1). Псевдослучайную последовательность можно представить в виде матрицы размером 3×21:

$$\mathbf{A}(3,21) =$$

$$= \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & \beta^2 & 0 & 0 & 1 & 1 & \beta & \beta^2 & 0 & 1 & 0 & \beta^2 & 1 & \beta^2 & 1 & 1 & \beta^2 \\ \beta & \beta & \beta & 0 & \beta & 1 & 0 & 0 & \beta & \beta & \beta^2 & 1 & 0 & \beta & 0 & 1 & \beta & 1 & \beta & \beta & 1 \\ \beta^2 & \beta^2 & \beta^2 & 0 & \beta^2 & \beta & 0 & 0 & \beta^2 & \beta^2 & 1 & \beta & 0 & \beta^2 & 0 & \beta & \beta^2 & \beta & \beta^2 & \beta^2 & \beta \end{bmatrix}.$$

2.3. Представление периодических последовательностей через функцию следа

Предварительно определим понятие циклотомического смежного класса по модулю $2^n - 1$. Циклотомический смежный класс числа s представляет собой множество чисел

$$C_s = \{s, 2s, 2^2 s, \ldots, 2^{n_s - 1} s\},$$

где n_s – наименьшее положительное число, такое, что выполняется сравнение $2^{n_s} s \equiv s \bmod(2^n - 1)$. Наименьшее число в C_s называется лидером смежного класса. Заметим, что $n_s \mid n$.

Пусть $\mathbf{F}_q = GF(q)$ – конечное поле, состоящее из q элементов. Функция следа, отображающая элементы поля $\mathbf{F}_{2^n}$ в элементы поля $\mathbf{F}_2$, определяется как

$$Tr_1^n(x) = x + x^2 + \ldots + x^{2^{n-1}}.$$

Используя интерполяционную формулу Лагранжа или, что то же самое, дискретное преобразование Фурье, для любой функции $f : \mathbf{F}_{2^n} \to \mathbf{F_2}$ можно получить полиномиальное представление вида

$$f(x) = \sum c_i x^i, \quad c_i \in \mathbf{F}_{2^n}, \text{ причем } c_{i2^j} = c_i^{2^j}.$$

Иными словами, функция $f(x)$ может быть представлена как сумма функций следа $Tr_1^{n_r}(\beta_r x^r)$, где $r \in I$ – множество, состоящее из всех циклотомических смежных классов по модулю $2^n - 1$. Таким образом, имеем

$$f(x) = \sum_{r \in I} Tr_1^{n_r}(\beta_r x^r), \ \beta_r \in \mathbf{F}_{2^{n_r}}.$$

Любая бинарная последовательность $\{a_k\}$ периода $(2^n - 1)$ может быть выражена как

$$a_k = f(\alpha^k), \ k = 0,1,2,...,$$

где α – примитивный элемент в поле $\mathbf{F}_{2^n}$.

Линейная сложность бинарной последовательности $\{a_k\}$ определяется как

$$LS(\{a_k\}) = |\{0 \le i < 2^n - 1 \mid c_i \ne 0\}| = \sum_{\beta_r \ne 0} n_r,$$

где c_i – коэффициенты характеристического полинома.

Формирование многозначных последовательностей. Рассмотрим на примере q-ичной m-последовательности. Пусть $q = p^r$, где p – простое нечетное число, и пусть α – примитивный элемент поля $GF(q^n)$. Функция следа, реализующая отображение $GF(q^n) \to GF(q)$, определяется следующим образом:

$$Tr(x) = x + x^q + ... + x^{q^{n-1}}.$$

Псевдослучайная q-ичная последовательность максимальной длиной $(q^n\text{-}1)$ $S = (s_0,...,s_{q^n-2})$ вычисляется через выражение $s_i = Tr(\alpha^i)$. Один период последовательности S может быть представлен в форме $(T, \beta T, ..., \beta^{q-2} T)$, где T – q-ичный вектор длиной $v = (q^n - 1)/(q - 1)$ и $\beta = \alpha^v$ соответствует примитивному элементу поля $GF(q)$.

Пример. Для $p = 3$, $n = 2$ и примитивного полинома $x^2 + x + 2$ троичная m-последовательность имеет вид

$$\begin{array}{cccccccccc} & \alpha^0 & \alpha^1 & \alpha^2 & \alpha^3 & \alpha^4 & \alpha^5 & \alpha^6 & \alpha^7, \\ S & = & (2 & 2 & 0 & 2 & 1 & 1 & 0 & 1). \end{array}$$

Последовательность имеет линейную сложность, равную 2.

Один из приемов повышения линейной сложности это отображение многозначной последовательности в бинарную $GF(q) \rightarrow GF(2)$.

Например. Для рассмотренного выше примера введем следующее правило отображения $\rho(1) = \rho(2) = 1$ и $\rho(0) = 0$. Тогда получаем последовательность

$$S = (1\ 1\ 0\ 1\ 1\ 1\ 0\ 1),$$

имеющую период, равный 4, и линейную сложность, равную 4.

Применив правило $\rho(1) = 1$, $\rho(2) = \rho(0) = 0$, получим последовательность

$$S = (0\ 0\ 0\ 0\ 1\ 1\ 0\ 1),$$

имеющую период, равный 8, и линейную сложность, равную 8.

2.4. Моделирование генераторов на регистрах сдвига с обратной связью

Моделирование ведется в программной среде Maple.

1. *Матричная модель генератора на РСОС*

1.2. *Программа универсального генератора с РСОС*. Обозначим разряды регистра сдвига как b[i].

```
> runshift:=proc(f,m) local x, ans, b0,b;
>  if nargs>2 then b0:=args[3];   else b0:=[1,0$(m-1)]; fi;
>  ans:=[b0];  for x from 1 to 2^m do   b:=[f(b0) mod 2,op(b0[1..m-1])];
>  if member(b,ans) then break fi;   ans:=[op(ans),b]; b0:=b; od; ans:=[op(ans),b];
>  RETURN(evalm(ans)) : end:
```

Пример. > runshift(b2->((b2[1]+1)*b2[2]+(b2[2]+1)*b2[3]+b2[4]), 4);

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

2. *Модель формирования m-последовательности через функцию следа*

2.1.Формирование конечного поля GF(2^4) по полиному ($x^4 + x + 1$)

```
> alias(beta=RootOf(x^4+x+1)); (Определение примитивного элемента β );
> FieldTable := Matrix(q,2); (Задание матрицы для элементов поля).
```

2.2. Введение оператора «MakeTable» , формирующего поле $GF(2^4)$:

```
> MakeTable(16, 2, beta).
```

$$\begin{aligned}
1 &= 1 \\
\beta &= \beta \\
\beta^2 &= \beta^2 \\
\beta^3 &= \beta^3 \\
\beta^4 &= \beta + 1 \\
\beta^5 &= \beta^2 + \beta \\
\beta^6 &= \beta^3 + \beta^2 \\
\beta^7 &= \beta^3 + \beta + 1 \\
\beta^8 &= \beta^2 + 1 \\
\beta^9 &= \beta^3 + \beta \\
\beta^{10} &= \beta^2 + \beta + 1 \\
\beta^{11} &= \beta^3 + \beta^2 + \beta \\
\beta^{12} &= \beta^3 + \beta^2 + \beta + 1 \\
\beta^{13} &= \beta^3 + \beta^2 + 1 \\
\beta^{14} &= \beta^3 + 1 \\
\beta^{15} &= 1
\end{aligned}$$

2.3. Вычисление m–последовательности { g[i] } через функцию следа

```
> for i from 1 to 15 do g[i]:=Eform(Normal(beta^i+beta^(2*i)+beta^(4*i) +
+ beta^(8*i)) mod p); od;
```

$g_1 := 0, g_2 := 0, g_3 := 1, g_4 := 0, g_5 := 0, g_6 := 1, g_7 := 1, g_8 := 0, g_9 := 1,$
$g_{10} := 0, g_{11} := 1, g_{12} := 1, g_{13} := 1, g_{14} := 1, g_{15} := 0.$

3. *Моделирование генератора де Брюина*

3.1. Задание 4-разрядного генератора

```
f:=(a4,a3,a2,a1)->(( (a2*a3*a4 + a2+a1+1 mod 2),a4,a3,a2)).
```

3.2. Вычисление состояний регистра сдвига генератора

```
a:=(0,0,0,1);
for i from 0 to 15 do
        a:=f(a) od;
```

a	:=	0,0,0,1
a	:=	0,0,0,0
a	:=	1,0,0,0
a	:=	1,1,0,0
a	:=	1,1,1,0
a	:=	1,1,1,1
a	:=	0,1,1,1
a	:=	1,0,1,1
a	:=	1,1,0,1
a	:=	0,1,1,0
a	:=	0,0,1,1
a	:=	1,0,0,1
a	:=	0,1,0,0
a	:=	1,0,1,0
a	:=	0,1,0,1
a	:=	0,0,1,0

3.3. Спектр Уолша последовательности де Брюина имеет вид (рис.2.4).

8., -2., 0., -2., -4., 2., 0., -2., 0., 2., 0., 2., -4., -2., 0., 2.

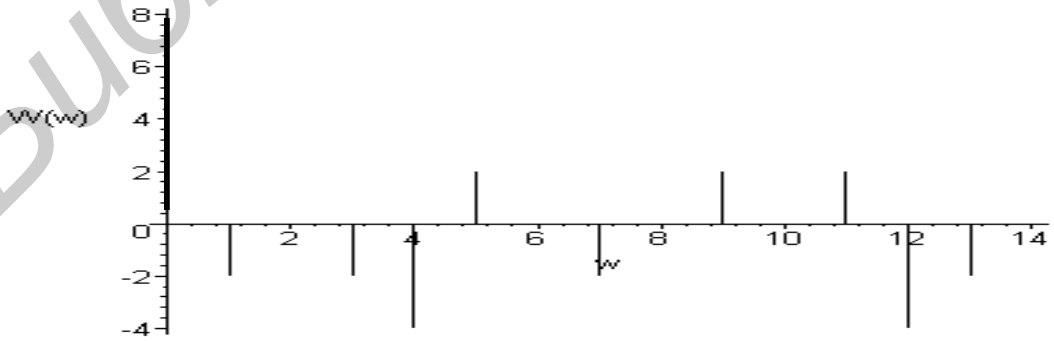

Рис. 2.4. Спектр Уолша

3. МЕТОДЫ АНАЛИЗА ПОСЛЕДОВАТЕЛЬНОСТЕЙ ПОТОЧНЫХ КРИПТОСИСТЕМ

Подходы к анализу. Существуют разнообразные оценки приемлемости шифрующей последовательности, генерируемой тем или иным поточным шифром. Все множество оценок можно разделить на две группы.

К первой группе относятся оценки статистических свойств шифропоследовательности: есть ли какой-либо дисбаланс в способе генерации этой последовательности, который позволил бы криптоаналитику предполагать следующий символ с вероятностью лучшей, чем при случайном выборе? Вторая группа оценок предоставляет аналитику возможность на основе уже имеющегося сегмента гаммы сконструировать свою собственную последовательность, которая повторяла бы шифрующую последовательность. В данном случае основной оценкой является сложность воспроизведения последовательности.

3.1. Оценка статистических свойств последовательностей

Постулаты Голомба. При многократном подбрасывании монеты можно ожидать примерно одинакового выпадения «орлов» и «решек». Аналогичным образом можно сформулировать и многие другие свойства для описания статистических характеристик последовательности, которая подразумевается сгенерированной совершенно случайным источником. Одна из первых формулировок некоторых основополагающих правил для статистических свойств периодических псевдослучайных последовательностей была представлена С. Голомбом [1, 14]. Три основных правила получили в открытой криптографии известность как *постулаты Голомба.*

G1. Количество «1» в каждом периоде должно отличаться от количества «0» не более чем на единицу.

G2. В каждом периоде половина серий (из одинаковых символов) должна иметь длину один, одна четверть должна иметь длину два, одна восьмая должна иметь длину три и т.д. Более того, для каждой из этих длин должно быть одинаковое количество серий из «1» и «0».

G3. Предположим, у нас есть две копии одной и той же последовательности периода *p,* сдвинутые относительно друг друга на некоторое значение *d.* Тогда для любой последовательности, удовлетворяющей правилу G3, функция автокорреляции должна принимать лишь два значения (быть двузначной).

Последовательность, удовлетворяющая правилам G1-G3, часто именуется *псевдослучайной последовательностью*, или ПС-последовательностью. Одних лишь этих правил недостаточно для исследования такой проблемы, как случайный вид последо-

вательности. К анализируемой последовательности применяется широкий спектр различных *статистических тестов* для исследования того, насколько хорошо она согласуется с допущением, что для генерации использовался совершенно случайный источник. Меры, обычно используемые для исследования случайности двоичной гаммы длиной *N*, проверяют нулевую гипотезу, согласно которой последовательность получена на основе *N* испытаний схемы Бернулли с вероятностью появления единицы, равной одной второй в каждом испытании.

Проблема тестирования выглядит следующим образом. Статистический тест для вырабатываемых генератором последовательностей длиной N – это функция

$$T: \mathrm{B}^N \rightarrow \{ \text{«принять», «отвергнуть»}\},$$

которая разделяет множество B^N двоичных длиной N последовательностей $s^N = s_1,\ldots,s_N$ на (обычно небольшое) множество «плохих» или «неслучайных» последовательностей

$$S_T = \{s^N : T(s^N) = << \textit{отвергнуть} >>\} \subseteq B^N$$

и остальное множество последовательностей "хороших" или "случайных. Вероятность того, что выработанная генератором последовательность будет отвергнута, выражается соотношением $\rho = \frac{|S_T|}{2^N}$. В реальных тестах ρ должно быть невелико, например $\rho \approx 0.001...0.01$.

Наиболее известны следующие виды тестов: частотный, последовательный, тест серий, автокорреляционный, универсальный, тест повторений, сравнение тестов *l*-грамм, отсечение слабых последовательностей и комплексный тест NIST.

3.2. Оценка линейной сложности

Концепция линейной сложности. Одной из первых аналитических мер качества поточных шифров стала *линейная сложность* (или *линейный размах)* шифрующей последовательности, которая определяется как длина L самого короткого регистра сдвига с линейной обратной связью, способного породить эту последовательность. Любая последовательность, сформированная конечным автоматом над конечным полем, имеет конечную линейную сложность. С помощью простого алгоритма Берлекампа–Месси (БМ) можно воссоздать РСЛОС, проверив несколько символов гаммы. Определив нужный РСЛОС, можно взломать шифр.

Эту идею можно расширить с полей на кольца и на случаи, когда выходная последовательность рассматривается как числа в поле нечетной характеристики. Обобщение приводит к понятию профиля линейной сложности, который определяет ли-

нейную сложность последовательности по мере ее удлинения. Существуют также понятия сферической и 2-адической сложности.

Линейной сложностью $LS(\mathbf{s})$ последовательности $\mathbf{s}_i(l) = (s_{0,\mathrm{i}}, s_{1,i}, \ldots, s_{l\text{-}1,\mathrm{i}})$ называется длина L самого короткого РСЛОС, который может сгенерировать **s**, когда первые L цифр последовательности **s** являются начальным заполнением регистра. Эквивалентное определение: линейная сложность $LS(\mathbf{s})$ определяется как наименьшее неотрицательное целое L, такое, что существует линейная рекуррента с фиксированными константами $c_0, c_1, \ldots, c_L$, удовлетворяющая равенству

$$s_j + c_1 s_{j-1} + \ldots + c_L s_{j-L} = 0, \quad L \le j \le 1.$$

Коэффициенты $\{c_i\}$ определяют полином обратной связи РСЛОС: $C(D) = 1 + c_1 D + c_2 D^2 + \ldots + c_L D^L$.

Рассмотрим возможность построения алгоритма БМ для бинарных последовательностей. Определим *невязку* $d_N = s_N + \sum_{i=1}^{L} c_i s_{N-i} \bmod 2$ как разность между символом s_N анализируемой последовательности и синтезированным значением на $(N + 1)$ шаге.

Примем следующие допущения:

- РСЛОС формирует последовательность $\mathbf{s}_i(N + 1)$ только в том случае, если невязка d_N равна нулю;

- если $d_N = 0$, то длина РСЛОС принимается равной $L(\mathbf{s}_i(N + 1)) = L$;

- предположим, что $d_N = 1$ и пусть $m < N$ наибольшее целое, такое, что длины РСЛОС, формирующие последовательности $\mathbf{s}_i(m)$ и $\mathbf{s}_i(N)$ связаны неравенством

$$L(\mathbf{s}_i(m)) < L(\mathbf{s}_i(N)).$$

Причем последовательность $\mathbf{s}_i(m)$ формируется с помощью полинома обратной связи $B(D)$. Тогда РСЛОС, имеющий наименьшую длину и генерирующий последовательность $\mathbf{s}_i(N + 1)$, имеет параметры длину L' и полином $C'(D)$, равные

$$L' = \begin{cases} L, & \text{если } L > N/2; \\ N + 1 - L, & \text{если } L \le N/2 \end{cases}$$

и $C'(D) = C(D) + B(D) D^{N-m}$.

Алгоритм БМ:

Вход: бинарная последовательность $\mathbf{s}(n) = (s_0, s_1, \ldots, s_{n-1})$ длиной n.

Выход: Линейная сложность $LS(\mathbf{s}(n))$.

1. Инициализация исходных данных $C(D) \leftarrow 1$, $L \leftarrow 0$, $m \leftarrow (-1)$, $B(D) \leftarrow 1$, $N \leftarrow 0$.
2. До тех пор пока $N < n$, выполнять следующие операции:

Вычислять невязку d, $d \leftarrow s_N + \sum_{i=1}^{L} c_i s_{N-i} \bmod 2$.

Если $d = 1$, то выполнять следующие действия:

$$T(D) \leftarrow C(D),\ C(D) \leftarrow C(D) + B(D)D^{N-m}.$$

Если $L \le N/2$, тогда $L \leftarrow N + 1 - L$, $m \leftarrow N$, $B(D) \leftarrow T(D)$.

$N \leftarrow N + 1$.

3. Выдать значение L.

Пример. Вычислим линейную сложность бинарной последовательности длиной 9 следующего вида: **s** = (0,0,1,1,0,1,1,1,0). Последовательность вычислений показана в табл. 3.1.

Таблица 3.1

s_N	d	$T(D)$	$C(D)$	L	m	$B(D)$	N
-	-	-	1	0	-1	1	0
0	0	-	1	0	-1	1	1
0	0	-	1	0	-1	1	2
1	1	1	$1 + D^3$	3	2	1	3
1	1	$1 + D^3$	$1 + D + D^3$	3	2	1	4
0	1	$1 + D + D^3$	$1 + D + D^2 + D^3$	3	2	1	5
1	1	$1 + D + D^2 + D^3$	$1 + D + D^2$	3	2	1	6
1	0	$1 + D + D^2 + D^3$	$1 + D + D^2$	3	2	1	7
1	1	$1 + D + D^2$	$1 + D + D^2 + D^5$	5	7	$1 + D + D^2$	8
0	1	$1 + D + D^2 + D^5$	$1 + D^3 + D^5$	5	7	$1 + D + D^2$	9

Последовательность может быть сформирована с помощью РСЛОС с полиномом обратной связи $1 + D^3 + D^5$. Линейная сложность равна 5.

Китайские математики Даи и Зенг показали, что алгоритм Берлекампа–Месси естественным образом следует из решения классической задачи о рациональной аппроксимации для произвольных элементов в поле

$$F_2((x)) = \{\alpha = \sum_{i=m}^{\infty} a_i x^{-i} \mid m \in Z, a_i \in F_2\}$$

формальных рядов Лорана над двоичным полем F_2. Инструментом для решения РАП служит аппарат анализа диофантовых уравнений в поле действительных чисел.

3.3. Дискретное преобразование Фурье–Галуа периодических последовательностей

Обозначим через **s** = (s_0, … , s_{N-1}) последовательность периода N над полем $GF(q)$. В расширенном поле $GF(q^n)$ определим элемент поля α порядка N. Дискретное преобразование Фурье–Галуа над конечным полем определяется через выражение

$$S(k) = \sum_{t=o}^{N-1} s_t \alpha^{tk}, k = 0,...,N-1.$$

Соответственно обратное дискретное преобразование Фурье–Галуа определяется как

$$s(t) = \frac{1}{N} \sum_{k=o}^{N-1} S(k)\alpha^{-tk}, t = 0,...,N-1.$$

Множество **S** = {S(k)} называется спектром Фурье–Галуа и имеет вид *спектральной последовательности* **S** = (S_0, … , S_{N-1}). В общем случае S(k)∈ $GF(q^n)$.

Пример. Пусть $q = 2$, $n = 3$ и $GF(2^3)$ задается примитивным полиномом $f(x) = x^3 + x + 1$, α – корень полинома f(x). Зададим бинарную последовательность периода 7 в виде **s** = (0 1 0 0 0 1 0), $N = 7$. Спектр Фурье–Галуа вычисляется следующим образом:

$$S(k) = \sum_{t=o}^{6} s_t \alpha^{tk} = \alpha^k + \alpha^{5k}, k = 0,1,...,$$

$$S(0) = \sum_{t=0}^{6} s_t = s_1 + s_5 = 0, \; S(1) = \sum_{t=0}^{6} s_t \alpha^t = \alpha + \alpha^5 = \alpha^6,$$

$$S(2) = \sum_{t=0}^{6} s_t \alpha^{2t} = \alpha^2 + \alpha^3 = \alpha^5, \ldots.$$

Спектральная последовательность равна **S** = (0,α^6, α^5, 1. α^3, 1, 1).

Обратное преобразование Фурье–Галуа $s_t = \sum_{k=0}^{6} S(k)\alpha^{-tk}, t = 0,1,...$ позволяет из спектральной восстановить исходную последовательность:

$$s_0 = \sum_{k=0}^{6} S(k) = \alpha^6 + \alpha^5 + 1 + \alpha^3 + 1 + 1 = 0,$$

$$s_1 = \sum_{k=0}^{6} S(k)\alpha^{-k} = \alpha^6\alpha^{-1} + \alpha^5\alpha^{-2} + 1\alpha^{-3} + \alpha^3\alpha^{-4} + 1\alpha^{-5} + 1\alpha^{-6} =$$

$$= \alpha^5 + \alpha^3 + \alpha^4 + \alpha^6 + \alpha^2 + \alpha = 1.$$

Для периода N, который является множителем числа (q^n - 1), можно определить циклотомические смежные классы C_t по модулю (p^r – 1): $C_t = \{t, tq, tq^2, ..., tq^{n_t - 1}\}$, где n_t – наименьшее положительное целое, такое, что $q^{n_t} t \equiv t \bmod N$. Наименьшее число в C_t называется лидером смежного класса по модулю N относительно числа q, а n_t определяет порядок смежного класса $n_t = ord(t)$.

Пример. Пусть $q = 2$, $n = 6$, $N = 21$. Циклотомические смежные класса по модулю 21 имеют вид

$C_0 = \{0\}$; $C_1 = \{1,2,4,8,16,11\}$; $C_3 = \{3,6,12\}$; $C_5 = \{5,10,20,19,17,13\}$; $C_7 = \{7,14\}$; $C_9 = \{9,18,15\}$.

Множество лидеров смежных классов $I = \{0,1,3,5,7,9\}$. Множество Z_{21} может быть выражено через объединение множеств смежных классов:

$$Z_{21} = C_0 \cup C_1 \cup C_3 \cup C_5 \cup C_7 \cup C_9 .$$

Свойство сопряженности спектральных последовательностей. Для любого k: $1 \le k \le N-1$ справедливы следующие соотношения между спектральными компонентами:

$$S(kq^j) = S(k)^{q^j}, 0 \le j \le n \text{ и } S(0) = \sum_{t=0}^{N-1} s_t .$$

Пример. Для рассмотренного выше примера, где **s** = (0 1 0 0 0 1 0) и **S** = (0, α^6, α^5, 1. α^3, 1, 1), получаем $S(1) = \alpha^6$, $S(1)^2 = (\alpha^6)^2 = \alpha^5 = S(2)$, $S(1)^4 = (\alpha^6)^4 = \alpha^3 = S(4)$, $S_3 = =1$, $S(6) = S(3)^2 = 1$, $S(5 = 12 \bmod 8) = S(3)^4 = 1$.

Таким образом, для вычисления спектра Фурье–Галуа достаточно вычислить спектральные коэффициенты, индексы которых соответствуют лидерам смежных классов, а остальные спектральные коэффициенты могут быть получены, при помощи свойства сопряженности.

Пример. Пусть $q = 2$, $n = 6$, поле $GF(2^6)$ определяется полиномом $x^6 + x + 1$, $GF(2^6) = \{\alpha^i \mid 0 \le i \le 62\}$. Бинарная последовательности периода 21 имеет вид

s = (0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0).

Зададим $\beta = \alpha^3$, для которого порядок $ord(\beta) = 21$. Дискретное преобразование Фурье–Галуа определяется соотношением

$$S(k) = \sum_{t=0}^{20} s_t \beta^{tk} = \beta^k + \beta^{8k} + \beta^{10k} + \beta^{16k} = \alpha^{3k} + \alpha^{24k} + \alpha^{30k} + \alpha^{48k}, k = 0,1,\ldots .$$

Вычисления для индексов, равных лидерам смежных классов, дают

$$S(0) =0,\ S(1) = \alpha^3 + \alpha^{24} + \alpha^{30} + \alpha^{48} = \alpha^{47},\ S(3) = \alpha^9 + \alpha^9 + \alpha^{27} + \alpha^{18} = 1,$$
$$S(5) = \alpha^{15} + \alpha^{57} + \alpha^{24} + \alpha^{51} = \alpha^{37},\ S(7) = \alpha^{21} + \alpha^{42} + \alpha^{21} + \alpha^{21} = 1,$$
$$S(9) = \alpha^{27} + \alpha^{27} + \alpha^{18} + \alpha^{54} = \alpha^9.$$

Остальные коэффициенты спектра могут быть вычислены через свойство сопряженности смежные классы. Так, для $j \in C_1 = \{1, 2, 4, 8, 16, 11\}$ получаем

$$S(1) = \alpha^{47} \to S(2) = S(1)^2 = (\alpha^{47})^2 = \alpha^{31},\ S(4) = S(1)^4 = (\alpha^{47})^4 = \alpha^{62},$$
$$S(8) = S(1)^8 = (\alpha^{47})^8 = \alpha^{61},\ S(16) = S(1)^{16} = (\alpha^{47})^{16} = \alpha^{59},$$

$$S(11) = S(1)^{32} = (\alpha^{47})^{11} = \alpha^{55}.$$

Представление последовательностей через функции следа и спектр Фурье–Галуа. Положим $\{s_i\}$ – последовательность над полем $GF(q)$ периода N, где $N \mid (q^n - 1)$. Символы последовательности могут быть выражены через обратное преобразование Фурье–Галуа следующим образом:

$$s_t = \frac{1}{N}\sum_j Tr_q^{n_j}(S(j)\alpha^{-jt}), t = 0,1,...,N-1,$$

где $n_j = ord(j)$ – порядок соответствующего смежного класса по модулю N, $Tr_q^{n_j}(x)$ – функция следа, выполняющая отображение $GF(q^{n_j}) \to GF(q)$, $S(j)$ и α^{-jt} принадлежат полю $GF(q^{n_j})$.

Пример. Для рассмотренных ранее примеров символы последовательностей могут быть выражены через следующие соотношения:

- **s** = (0, 1, 0, 0, 0, 1, 0), $GF(2^3) / x^3 + x + 1$,
- $S(1) = \alpha^6$, $S(3) = 1 \to s_t = Tr_2^3(\alpha^6\alpha^{-t}) + Tr_2^3(\alpha^{-3t}), t = 0,1,...$;
- **s** = (0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0) $\to$

$$s_t = Tr_2^6(\alpha^{47}\beta^{-t}) + Tr_2^3(\beta^{-3t}) + Tr_2^6(\alpha^{37}\beta^{-5t}) + Tr_2^3(\beta^{-7t}) + Tr_2^3(\alpha^9\beta^{-9t}), t = 0,1,...$$

3.4. Оценка линейной сложности с помощью дискретного преобразования Фурье

Если α – примитивный N-й корень из единицы в произвольном поле F, то *дискретным преобразованием Фурье* (ДПФ) последовательности из «временной области определения» $a^N = (a_0, a_1, \dots, a_{N-1})$ с компонентами из F называется последовательность из «частотной области определения» $A^N = (A_0, A_1, \dots, A_{N-1})$, где

$$A_i = \sum_{j=0}^{N-1} a_j \alpha^{ij}, \; i = 0,1,...,(N-1).$$

Соотношение для обратного ДПФ выглядит так:

$$a_j = \frac{1}{N^*}\sum_{i=0}^{N-1} A_i \alpha^{ij}, \; j = 0,1,...,(N-1),$$

где $N^* = N \bmod p$, если характеристика F равна p; и $N^* = N$, если характеристика $\mathbb{F}$ равна нулю.

Между ДПФ и линейной сложностью последовательности существует близкая взаимосвязь. Вычисление i-й компоненты S(i) в ДПФ можно рассматривать как внут-

реннее произведение между **s** и последовательностью (α^{0i}, α^{1i}, α^{2i}, … , $\alpha^{(N-1)i}$).Таким образом, можно описать ДПФ как матричное преобразование

$$\mathbf{S} = \mathbf{s}\, F,$$

где

$$F = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \alpha^{1} & \alpha^{2} & \cdots & \alpha^{(N-1)} \\ \vdots & \vdots & & \ddots & \vdots \\ 1 & \alpha^{(N-1)} & \alpha^{2(N-1)} & \cdots & \alpha^{(N-1)(N-1)} \end{bmatrix}.$$

Аналогично обратное ДПФ дается как

$$\mathbf{s} = (1/N)\, \mathbf{S}\, F^{-1},$$

где

$$F^{-1} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \alpha^{-1} & \alpha^{-2} & \cdots & \alpha^{-(N-1)} \\ \vdots & \vdots & & \ddots & \vdots \\ 1 & \alpha^{-(N-1)} & \alpha^{-2(N-1)} & \cdots & \alpha^{-(N-1)(N-1)} \end{bmatrix}.$$

Определим матрицу-циркулянт от последовательности **s**, обозначаемую $M(\mathbf{s})$, как матрицу, строки которой образованы N циклическими левыми сдвигами **s**:

$$M = \begin{bmatrix} s_0 & s_1 & \cdots & s_{N-1} \\ s_1 & s_2 & \cdots & s_0 \\ \vdots & \vdots & \ddots & \vdots \\ s_{N-1} & s_0 & \cdots & s_{N-2} \end{bmatrix}.$$

Пусть LS – линейная сложность последовательности $\hat{a} = a_0, a_1, a_2\ldots$ Тогда по определению линейной сложности LS – это наименьшее целое число, такое, что L+1 строка матрицы-циркулянт может быть записана как линейная комбинация предыдущих строк. Тогда ранг матрицы-циркулянт M равен по крайней мере L. С другой стороны, каждая строка с индексом $L < i < N - 1$ может быть записана как линейная комбинация предыдущих строк (фактически L предыдущих строк). Следовательно, ранг M равен L. Линейная сложность периодической, полубесконечной последовательности $\hat{a} = (\mathbf{s})^{\infty}$ равна рангу циркулянта M.

Матрица-циркулянт M может быть записана как

$$M = F^{-1} \cdot \mathbf{D}_s \cdot F^{-1},$$

где F^{-1} – это инвертированная матрица ДПФ, а $\mathbf{D_s}$ – диагональная матрица $N \times N$, на диагонали которой расположены элементы **s**. В условиях того, что α – примитивный N-й корень из единицы, матрицы F и F^{-1} имеют полный ранг. Следовательно, rang M = = rang ($\mathbf{D_s}$). Но ранг $\mathbf{D_s}$ равен количеству ненулевых элементов в **s** и равен *весу Хэмминга* $wt(\mathbf{s})$.

Оценка линейной сложности периодической последовательности через вес Хэмминга спектральной последовательности. Пусть $\mathbf{s} = \{s_i\}$ – последовательность периода N, $N \mid (q^n\text{-}1)$ над полем $GF(q)$ и пусть $\mathbf{S} = \{S(k)\}$ её спектральная последовательность. Тогда линейная сложность $LS(\mathbf{s})$ последовательности $\mathbf{s}$ оценивается как вес Хэмминга спектральной последовательности:

$$LS(\mathbf{s}) = wt(\mathbf{S}).$$

Если последовательность $\mathbf{s}$ имеет вес Хэмминга $wt(\mathbf{s})$, то линейная сложность спектральной последовательности равна

$$LS(\mathbf{S}) = wt(\mathbf{s}).$$

Отсюда следует, что если $\mathbf{s}$ – это бинарная последовательность периода $N = 2^n - 1$, тогда она является балансной (т.е. $wt(\mathbf{s}) = 2^n - 1$ тогда и только тогда, когда линейная сложность $LS(\mathbf{S}) = 2^n - 1$).

Пример. Зададим $q = 2$, $n = 5$, $N = 31$, поле $GF(2^5)$, которое определяется через полином $x^5 + x^3 + 1$, α – примитивный элемент $GF(2^5)$, и последовательность

$$\mathbf{s} = (1\ 1\ 0\ 0\ 0\ 1\ 0\ 1\ 0\ 1\ 1\ 1\ 1\ 1\ 1\ 0\ 1\ 0\ 1\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 1\ 0\ 0\ 1\ 1).$$

Множество лидеров смежных циклитомических классов по модулю 31 равно

$$I = \{0, 1, 3, 5, 7, 11, 15\}.$$

Представление символа последовательности через функцию следа имеет вид

$$s_t = Tr(\alpha^{25}\alpha^t) + Tr(\alpha^{5t}) + Tr(\alpha^{7t}), t = 0,1,\ldots .$$

Спектральные компоненты, имеющие индексы, равные лидерам смежных классов, равны $S(1) = \alpha^{25}$, $S(5) = 1$, $S(0) = S(3) = S(11) = S(15) = 0$. Так как порядок $n_j = 5$ для всех $0 \neq j \in I$, то вес Хэмминга спектральной последовательности равен 15 и, следовательно, линейная сложность последовательности $\mathbf{s}$ равна $LS(\mathbf{s}) = wt(\mathbf{S}) = 15$. Вес Хэмминга собственно последовательности $\mathbf{s}$ равен 16, и, следовательно, линейная сложность спектральной последовательности равна $LS(\mathbf{S}) = wt(\mathbf{s}) = 16$.

3.5. Профиль линейной сложности

Пусть LS_i – линейная сложность подпоследовательности $s^i = (s_0, s_1, \ldots, s_{i-1})$. Тогда последовательность $LS_1, LS_2, \ldots, LS_i$ называется *профилем линейной сложности* последовательности s^i. Алгоритм БМ позволяет вычислить *профиль линейной сложности.*

Пример. Рассмотрим две последовательности. Первая длиной $n = 15$ сформирована с помощью РСЛОС , полином $1 + D + D^4$:

$$\mathbf{s}_1 = (0, 0, 1, 1, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1, 0).$$

Вторая последовательность длиной $n = 21$ сформирована РСНОС с нелинейной функцией $(D^{13} + 1)\, D^{12} + (D^{12} + 1)\, D^{11} + D^{10} \mod 2$:

$$\mathbf{s}_2 = (0, 0, 1, 0, 1, 1, 1, 0, 0, 0, 1, 0, 1, 1, 1, 0, 0, 1, 0, 1).$$

Выделим в последовательностях отрезки символов $\mathbf{s}(l) = (s_0, s_1, \ldots, s_{l-1})$, l=3,4,… . К каждому из отрезков применим алгоритм БМ, вычислив тем самым линейную сложность отрезков. По результатам вычислений построим графики профилей линейной сложности (рис. 3.1).

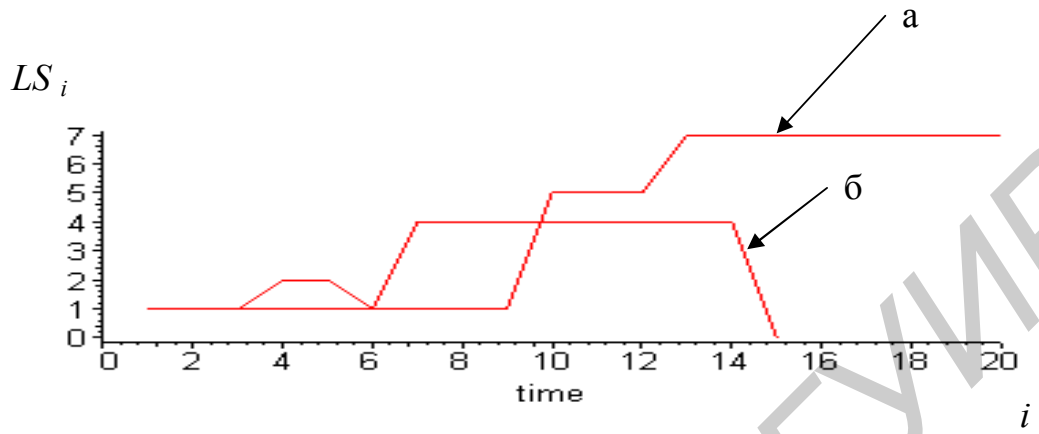

Рис. 3.1. Графики профилей линейной сложности:
а – нелинейной функции; б – линейной функции

Полученные результаты показывают, что нелинейная последовательность имеет более высокий уровень линейной сложности.

Высокая линейная сложность необязательно гарантирует стойкость генератора, но низкая линейная сложность явно указывает на недостаточную надежность.

Статистическая трактовка профиля линейной сложности. Для двоичных независимых и равномерно распределенных последовательностей L_i становится случайной величиной. Для последовательностей над произвольным конечным полем F_q дисперсия линейной сложности убывает с ростом размера алфавита последовательности; для больших значений q дисперсия примерно равна 1 / q. Таким образом, чем больше размер F_q, тем ближе профиль сложности к линии i / 2. Криптографическая последовательность должна иметь приемлемый профиль линейной сложности для каждой начальной точки отсчета. Профиль линейной сложности хорошо сконструированного генератора гаммы не отличим от профиля линейной сложности случайных последовательностей.

3.6. Корреляционные атаки

Корреляционная стойкость. Часто пытаются получить высокую линейную сложность, нелинейно объединяя результаты нескольких выходных последовательностей. В этом случае появляется опасность, что одна или несколько внутренних выходных последовательностей (выходы отдельных РСОС) могут иметь корреляцию с

объединенной гаммой и атакованы при помощи линейной алгебры (алгебраические атаки). Такие атаки называют корреляционными, или атаками «разделяй и властвуй». Можно точно оценить корреляционную стойкость и взаимосвязь выгод и потерь при выборе между корреляционной стойкостью и линейной сложностью.

Основная идея корреляционной атаки состоит в выявлении определенной корреляции между выходом генератора и выходом одной из его составных частей. Тогда, отслеживая выходную последовательность, можно получить информацию об этом промежуточном выходе. Используя эту информацию и другие корреляции, можно собрать данные о других промежуточных выходах, до тех пор пока генератор не будет взломан. Корреляционные атаки и их вариации (быстрые корреляционные атаки) предлагают компромисс между вычислительной сложностью и эффективностью.

В качестве примера рассмотрим генератор Геффе, использующий комбинацию трех РСЛОС. Первые два РСЛОС являются генерирующими, а третий РСЛОС – управляющим. Комбинирующая функция $f(x_1,x_2,x_3)$ имеет вид

$$f(x_1, x_2, x_3) = x_1 x_3 \oplus x_2 (x_3 \oplus 1).$$

Если все РСЛОС имеют максимальный период и их длины n_1 , n_2, n_3 попарно взаимно просты, то период результирующей последовательности равен произведению периодов РСЛОС, а линейная сложность гаммы генератора равна

$$LS = n_1\, n_3 + n_2\, n_1 + n_2.$$

В то же время функция f (x_1, x_2, x_3) имеет хорошие линейные приближения. Функция $f(x_1, x_2, x_3)$ совпадает с функцией x_1 (а также с функцией x_2) на ¾ всех наборов таблицы истинности. Это означает, что гамма генератора совпадает с выходом первого РСЛОС примерно для 75 % знаков. Следовательно, можно опробовать начальные состояния первого РСЛОС и статистически отсеивать «ложные» значения, генерируя знаки первого РСЛОС и наблюдая частоту совпадения этих знаков с соответствующими знаками гаммы. При ложных значениях совпадение происходит примерно для 50 % знаков. Для отсева одного ложного значения достаточно использовать порядка 15 сравнений знаков. Дальнейшее вскрытие ключа очевидно.

3.7. Алгебраические атаки

Алгебраические атаки являются важным инструментом криптоанализа поточных шифров, особенно шифров на основе РСОС и фильтрующей функции, которая формирует ключевой поток в виде булевой функции f от n переменных, связанной с состояниями регистра сдвига.

Основная задача криптоанализа криптосистем такого вида – это нахождение фазы (начального состояния) РСЛОС. Криптоанализ крайне затруднен, если:

- булева функция f описывается полиномом низкой степени или вероятностно неосуществимо найти такой полином;

- существует функция g, такая, что $f g = 0$ или $(f + 1)\, g = 0$, или $f g \neq 0$ описывается полиномом низкой степени.

Алгебраическая атака дает эффект, если булева функция f описывается полиномом с достаточно высокой степенью, и можно найти булеву функцию-множитель g, такую, что $f g = 0$ или вероятностно близко к нулю.

Критерий алгебраической безопасности. Предположим, что булева функция f описывается полиномом не низкой степени и существует булева функции g, для которой $f g \neq 0$, тогда степень полинома произведения функций $f g$, должна быть больше или равной степени функции f.

Если критерий алгебраической безопасности выполняется, говорят, что функция f инвариантна относительно алгебраической атаки.

В основе алгебраической атаки лежат такие понятия, как дискретное преобразование Фурье булевой функции, представление булевой функции через функцию следа, полиномиальные функции и последовательности.

Дискретное преобразование Фурье булевой функции. Зададим двоичное поле $\mathbf{F_2} = GF(2)$ и расширенное поле $\mathbf{F}_2^m = \{(x_0, x_1, ..., x_{m-1}) \mid x_i \in \mathbf{F}_2\}$, где m – положительное целое число. Любую булеву функцию можно представить в виде полиномиальной функции, описывающей отображения $\mathbf{F}_{2^n} \to \mathbf{F}_2$:

$$f(x) = \sum_{i=0}^{2^n - 1} d_i x^i \qquad d_i \in \mathbf{F}_{2^n}.$$

Алгебраическая степень f определяется наибольшим значением t, для которого $d_i \neq 0$, и вес Хэмминга кода бинарного представления индекса i равен t, т.е. $wt(i) = t$. Степени булевой и полиномиальной форм эквивалентны и обозначаются как $deg(f)$.

Используя полиномиальную форму записи булевой функции, можно определить ее дискретное преобразование Фурье как

$$A_k = \sum_{x \in \mathbf{F}^*_{2^n}} [f(x) + f(0)] x^{-k}, \qquad k = 0,1,...,2^n - 1.$$

Обратное дискретное преобразование Фурье задается в виде

$$f(x) = \sum_{k=0}^{2^n - 1} A_k x^k, \qquad x \in \mathbf{F}^*_{2^n}.$$

Представление булевой функции через функцию следа. Любая ненулевая булева функция f(x), полученная через отображение $\mathbf{F}_{2^n} \to \mathbf{F}_2$, может быть записана в виде

$$f(x) = \sum_{k \in \Gamma(n)} Tr_1^{n_k}(A_k x^k) + A_{2^n - 1} x^{2^n - 1}, \quad A_k \in \mathbf{F}_{2^{n_k}}, A_{2^n - 1} \in \mathbf{F}_2, x \in \mathbf{F}_{2^n},$$

где $\Gamma(n)$ – множество, состоящее из всех лидеров смежных классов по модулю (2^n-1), n_k – размер смежного класса C_k , $n_k \mid n$, $Tr_1^{nk}(x)$ – функция следа, выполняющая отображение $\mathbf{F}_{2^n} \to \mathbf{F}_2$.

Если булева функция f имеет степень $deg(f) = r < n$, тогда функция представляется как

$$f(x) = \sum_{wt(k)<r} Tr_1^{nk}(A_k x^k),$$

где $A_k \neq 0$ как минимум для одного значения $k \subset \Gamma(n)$, такого, что $wt(k) = r$.

Рассмотрим соотношения между преобразованием Фурье и булевыми функциями, ассоциированными с бинарными последовательностями. Пусть α – примитивный элемент в $\mathbf{F}_{2^n}$. Предполагается, что $f(0) = 0$ (в противном случае функция f заменяется функцией $g = f - f(0)$). Ассоциируем функцию f с бинарной последовательностью $\{a_t\}$, элементы которой задаются соотношением $a_t = f(\alpha^t)$, $t = 0,1,...,2^n - 2$.

Период N последовательности $\{a_t\}$ определяется через множитель числа $(2^n - 1)$.

Дискретное преобразование Фурье–Галуа последовательности $\{a_t\}$ определяется как

$$A(k) = \sum_{t=0}^{N-1} a_t \beta^{-tk},\ 0 \le k < N,$$

где β – элемент поля $\mathbf{F}_{2^n}$. Обратное дискретное преобразование Фурье–Галуа задается в виде

$$a_t = \sum_{k=0}^{N-1} A(k)\beta^{kt}, \quad 0 \le t < N.$$

Выберем элемент поля $\beta = \alpha^v$, где $v = (2^n - 1) / N$. Тогда получаем, что $A(k) = A_k$, при $0 \le k < N$ и $A_{iN+j} = A_j$, $0 \le i < v$, $0 \le j < N$.

Таким образом, любой функции, полученной через отображение $\mathbf{F}_{2^n} \to \mathbf{F}_2$, или эквивалентной ей булевой функции от n переменных соответствует бинарная последовательность с периодом $N \mid (2^n - 1)$. Они имеют одинаковую последовательность спектральных коэффициентов преобразования Фурье. Это свойство позволяет определить эффективный метод линеаризации с низкой степенью аппроксимации.

Базисы линеаризации. Определим множество Ω, состоящее из всех функций, которые могут быть получены с помощью отображения $\mathbf{F}_{2^n} \to \mathbf{F}_2$, или, что эквивалентно, всех булевых функций от n переменных. Множество F можно рассматривать как линейное пространство размерностью $q = 2^n$ над полем $\mathbf{F}_2$ в том случае, если каж-

дая функция в F представляется бинарным вектором размерностью q. Для $f \in F$ возможны два пути представления f в векторной форме размерности q.

Булевый базис. Пусть f представляется в виде булевой функции. Расположим элементы в поле F_2^n в соответствии с таблицей истинности функции f. Тогда

$$f(x_0,...,x_{n-1}) = (f(\mathbf{t}_0), f(\mathbf{t}_1),..., f(\mathbf{t}_{q-1}),\ \mathbf{t}_i = (t_{i,0}, t_{i,1},..., t_{i,n-1}),\quad t_{ij} \in \mathbf{F}_2,$$

где $i = t_{i,0} + t_{i,1}2 + ... + t_{i,n-1}2^{n-1},\ \ 0 \le i < q$.

Для двух векторов $\mathbf{x} = (x_0, \ldots, x_{n-1})$ и $\mathbf{c} = (c_0, \ldots, c_{n-1})$ из F_2^n определим операцию возведения в степень как $\mathbf{x}^{\mathbf{c}} = (x_0^{c0} x_1^{c1} \cdots x_{n-1}^{cn-1})$. Булевый базис для множества Ω, определяющий линейное пространство в F_2, состоит из всех одночленов вида

$$\Delta = \{\mathbf{x}^{\mathbf{c}} \mid \mathbf{c} \in \mathbf{F}_2^n\}.$$

Полиномиальный базис. Пусть f – полиномиальная форма. Используем понятие циклической группы над F_q и представим полином $f(x)$ в виде

$$f(x) = (f(0), f(1), f(\alpha),..., f(\alpha^{q-2})).$$

Пусть

$$\Pi_k = \{Tr_1^{n_k}\left(\beta\left(\alpha^i x\right)^k\right) \mid, i = 0,1,..., n_k - 1\},\ \beta \in F_{2^{n_k}},$$

где n_k – размер циклотомического смежного класса, соответствующего k.

Заметим, что $\{\alpha^{ik} \mid i = 0, 1, \ldots, n_k - 1\}$ является базисом поля $F_{2^{n_k}}$ над F_2. Можно показать, что все мономы функции следа в Π_k, когда k пробегает все значения лидеров смежного класса по модулю $q - 1$, составляют базис Ω. Другими словами, следующее множество является базисом Ω:

$$\Pi = \cup_{k \in \Gamma(n)} \Pi_k.$$

Алгоритм вычисления полиномиального базиса. Пусть $\mathbf{a} = \{a_i\} = a_0, a_1, \ldots$ – бинарная последовательность периода N. Определим оператор $\mathbf{L}$, выполняющий левый сдвиг последовательности $\mathbf{a}$, т.е. $\mathbf{L}\{\mathbf{a}\} = a_1, a_2, \ldots, \mathbf{L}^{\mathbf{i}}\{\mathbf{a}\} = a_i, a_{i+1}, \ldots$. Определим также k-децимированную последовательность $\mathbf{a}^{(k)} = \{a_{ki}\}$, где индексы вычисляются по модулю N. Элементы $\mathbf{a}^{(k)}$ задаются через выражение $a_{ki} = Tr_1^{n_k}\left(\beta\alpha^{ki}\right)$, $i = 0, 1, \ldots$. Воздействие оператора сдвига на децимированную последовательность $\mathbf{L}^i\{\mathbf{a}^{(k)}\}$ дает бинарный вектор размерностью $q - 1$. Образуем матрицу

$$\mathbf{P}_k = \begin{bmatrix} 0, & \mathbf{a}^{(k)} \\ 0, & \mathbf{L}\{\mathbf{a}^{(k)}\} \\ \vdots & \vdots \\ 0 & \mathbf{L}^{n_k-1}\{\mathbf{a}^{(k)}\} \end{bmatrix}.$$

Строки данной матрицы будут соответствовать функциям в Π_k . Заметим, что количество лидеров смежных классов по модулю $q - 1$ равно числу неприводимых полиномов над F_2, степени которых делят n. Следовательно, для вычисления полиномиального базиса Ω достаточно вычислить |Γ(n)| децимированных последовательностей для **a** .

Определим множество $\mathbf{S}_d$, содержащее все функции полиномиального базиса с весом Хэмминга $wt(k) \le d$. Тогда $|\mathbf{S}_d| = \sum_{i=0}^{d} \binom{n}{i}$. Любая функция в Ω степени d является линейной комбинацией в $\mathbf{S}_d$ над F_2. Для функции $f \in \Omega$ можно записать

$$f(\mathbf{S}_d) = \{f(x) \cdot g(x) \,|\, g \in \mathbf{S}_d\},$$

где $f(x) \cdot g(x)$ – внутреннее произведение двух бинарных векторов f и g, имеющих размерность $q = 2^n$.

Для заданной функции $f \in \Omega$ и двух положительных чисел d и e , $1 \le d, e < d$ и в случае. Если $f(\mathbf{S}_d)$ содержит не менее $2^n + 1 - |\mathbf{S}_d|$ различных функций и все эти функции линейно независимы над F_2, тогда существует функция $g \in \Omega$ степени не больше d,такая, что степень произведения $deg(f\,g) \le e$.

С другой стороны, возможен случай, когда $|f(\mathbf{S}_d)| < |\mathbf{S}_d|$ и элементы в $f(\mathbf{S}_d)$ – линейно независимы над F_2 . В этом случае не существует функции g степени $deg(g) \le d$, такой, что $f\,g \ne 0$ и $deg(fg) \le e$.

Условие существования аппроксимации низкой степени. Пусть D – максимальное линейно независимое множество $f(\mathbf{S}_d)$, $f \in \Omega$, $1 \le d$, $e < d$. Тогда функция g ∈ Ω степени $deg(fg) \le e$ существует только в том случае, когда объединение множеств $D \cup \mathbf{S}_e$ является независимым над F_2.

Алгоритм аппроксимации:

Вход: функция $f \in \Omega$, отображающая $F_{2^n} \to F_2$; $1 \le d$, $e < d$;

$t(x) = x^n + \sum_{i=0}^{n-1} t_i x^i$, $t_i \in F_{2^n}$ – примитивный полином степени n над F_2.

Выход: функция $g \in \Omega$ степени $deg(g) \le d$ и функция $h = f\,g$, $h \ne 0$ и $deg(h) \le e$, если существуют такие g и h, в противном случае $g = 0$ и $h = 0$.

Процедура вычислений:

1. Выбираем случайным образом вектор начального состояния (a_0, a_1,…, a_{n-1}), $a_i \in F_2$, вычисляются элементы последовательности по формуле

$$a_{n+i} = \sum_{j=0}^{n-1} t_j a_{j+i}, i = 0,1,...,2^n - 1 - n.$$

2. Вычисляется k – лидер циклотомического смежного класса по модулю $2^n - 1$ и n_k – размер класса. Множество $I = \{(k, n_k) \mid \Gamma(n)\}$ содержит все лидеры смежных классов.

3. Задается значение $m = max(d, e)$ и формируется $\mathbf{S}_m$ следующим образом:

$P_0 = (1, 1,…, 1)$; для $0 \neq k$ в области $\Gamma(n)$ с $wt(k) \leq d$ выполняется вычисление децимированной последовательности $\mathbf{a}^{(k)} = (a_0, a_k, ..., a_{k(2^n-2)})$ и ее сдвига влево с помощью оператора сдвига. Определяется матрица P_k как подматрица размером $n_k \times 2^n$ от $\mathbf{S}_m$ для всех k, имеющих вес $0 \leq wt(k) \leq m$.

4. Методом исключения Гаусса определяется ранг $f(\mathbf{S}_d)$ и максимально независимое множество $f(\mathbf{S}_d)$, которое обозначают как UD.

5. Метод исключения Гаусса применяется к матрице

$$\begin{bmatrix} UD \\ \mathbf{S}_e \end{bmatrix},$$

если ранг матрицы равен $(t + l)$, где $t = |UD|$, $l = |\boldsymbol{S}_e|$, тогда $g = 0$, $h = 0$ и осуществляется переход к п.6. В противном случае вычисляются коэффициенты c_i, $i = 1,…, t$, удовлетворяющие равенству

$$\sum_{i=1}^{t} c_i f g_i + \sum_{i=1}^{l} c_{t+i} h_i = 0.$$

Определяются полиномы $g = \sum_{i=1}^{t} c_i g_i$, $h = \sum_{i=1}^{i} c_{t+i} h_i$.

6. Вывод g и h.

Инвариантность относительно алгебраической атаки. Говорят, что функция f инвариантна относительно алгебраической атаки, если $\forall g \in \Omega$ и выполняются соотношения

$deg(f g) \geq deg(f)$, причем $f g \neq 0$ или $f g = 0 / (f + 1)g = 0$ причем $deg(g) \geq deg(f)$.

Если эти соотношения верны только для функций g, у которых $deg(g) \leq k$, то говорят, что $f(x)$ обладает инвариантностью k-порядка относительно алгебраической атаки.

Пример. Пусть F_{2^n} определяется примитивным полиномом $t(x) = x^4 + x + 1$ и α – примитивный корень полинома $t(x)$. Функция следа, отображающая элементы поля F_{2^n} в элементы поля F_2, выглядит как $Tr(x) = x + x^2 + x^4 + x^8$. Зададим функцию f в виде $f(x) = Tr(\alpha x^3)$. В булевой форме функция задается как

$$f(x) = x_2 \oplus x_0 x_1 \oplus x_0 x_2 \oplus x_0 x_3 \oplus x_1 x_2 \oplus x_1 x_3 \oplus x_2 x_3$$

и относится к классу бент-функций.

В соответствии с вышеизложенной теорией любая функция $g \in \Omega$ и $g(0)=0$ может быть представлена в виде

$$g(x) = Tr(bx) + Tr(cx^3) + Tr_1^2(dx^5) + Tr(ex^7) + \omega x^{15},\ b, c, e \in F_{2^4},\ d \in F_{2^2},\ \omega \in F_2,$$

где $Tr_1^2(x) = x + x^2$ – функция следа отображающая $F_{2^2} \to F_2$.

Умножив f на каждый моном функции следа для g, получим

$$Tr(bx)Tr(\alpha x^3) = Tr(b^4\alpha^4 x) + Tr_1^2((b^2\alpha + b^8\alpha^4)x^5) + Tr((b\alpha^2 + b^4\alpha)x^7),$$

$$Tr(cx^3)Tr(\alpha x^3) = Tr((c^2\alpha^4 + c^4\alpha^2 + c^8\alpha^8)x^3) + Tr(c\alpha^4),$$

$$Tr_1^2(dx^5)Tr(\alpha x^3) = Tr(d^2\alpha^2 x + d^2\alpha^4 x^7),$$

$$Tr(ex^7)Tr(\alpha x^3) = Tr((e^4\alpha^4 + e\alpha^8)x) + Tr_1^2((e^2\alpha^2 + e^8\alpha^8)x^5) + Tr(e^4\alpha^8 x^7).$$

Положим для определенности $\omega = 0$. Произведение функций $f(x)$ $g(x)$ может быть представлено как

$$f(x)g(x) = Tr(Ax + Bx^3 + Dx^7) + Tr_1^2(Cx^5) + E,$$

где

$A = b^4\alpha^4 + d^2\alpha^2 + e^4\alpha + e\alpha^8$; $B = c^2\alpha^4 + c^4\alpha^2 + c^8\alpha^8$; $C = b^2\alpha + b^8\alpha^4 + e^2\alpha^2 + e^8\alpha^8$;
$E = Tr(c\alpha^4)$.

Если предположить, что $fg \neq 0$, тогда $deg(fg) = 1$ только в том случае, если

$$B = C = D = 0.$$

Можно показать, что система таких уравнений не имеет решений для любого выбора b, c, d. Следовательно, $deg(fg) \geq 2$. Таким образом, можно сделать вывод, что f инвариантна относительно алгебраической атаки.

Отметим, что в качестве инвариантов относительно алгебраической атаки представляет интерес класс так называемых гипербент-функций.

4. ПОТОЧНЫЕ КРИПТОСИСТЕМЫ

Поточные шифры применяют изменяющиеся во времени преобразования к отдельным цифрам открытого текста. Теория автоматов [1] хорошо подходит для описания систем поточного шифрования. Обозначим алфавиты открытого и закрытого текстов соответственно как X и Y, g – алфавит шифрующей последовательности, Z – пространство состояний поточного шифра, K – пространство ключей. Пусть x_i, y_i, g_i, z_i означают соответственно цифру открытого текста, цифру шифротекста, цифру гаммы и внутреннее состояние в момент времени i. Ключ $k \in K$ выбирается в соответствии с вероятностным распределением P_K. Обычно ключ выбирают в соответствии с равномерным распределением, но в некоторых случаях может быть недостижимо выбрать ключ совершенно случайно.

В общем виде поточный шифр может быть описан уравнениями

$$z_{i+1} = F(k_i, z_i, x_i),$$
$$y_i = f(k_i, z_i, x_i),$$

где F – функция состояния,

f – функция выхода.

Последовательность $\{g_i = f(k_i, z_i), i \geq 1\}$ называется цифровой гаммой. Если выполняется условие $y_i = x_i + F(k_i, z_i)$, то поточный дешифратор работает без задержки.

Поточные шифры подразделяются на синхронные и самосинхронизирующиеся.

4.1. Синхронные поточные шифры

В *синхронном поточном шифре* шифрующая функция генерируется независимо от потоков открытого текста и шифротекста. Процесс шифрования в синхроннной криптосистеме можно описать с помощью уравнений

$$\sigma_{i+1} = f(\sigma_i, k); \quad z_i = g(\sigma_i, k); \quad c_i = h(z_i, m_i),$$

где σ_0 – начальное состояние, зависящее от функции ключа k; f – функция следующего состояния; g – функция ключевого потока; z_i, h – функции, осуществляющие преобразование сообщения открытого текста m_i в символ шифра c_i (рис. 4.1).

Процесс расшифрования выполняется в обратном порядке с использованием обратной функции h^{-1} (рис. 4.2).

Функционирование генератора гаммы описывается двумя уравнениями:

$$z_{i+1} = F(k_i, z_i),$$
$$g_i = f(k_i, z_i).$$

Начальное состояние z_0 может быть функцией от ключа k или от некоторой рандомизированной переменной. Цель генератора гаммы – развернуть короткий слу -

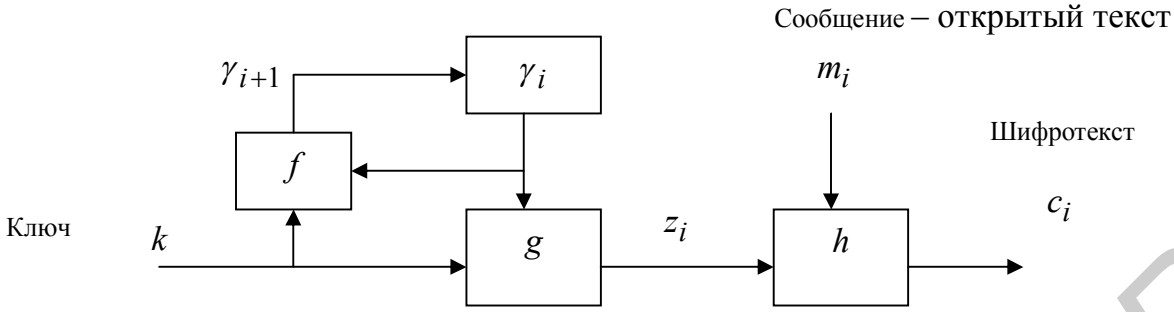

Рис. 4.1. Схема шифрования синхронным поточным шифром

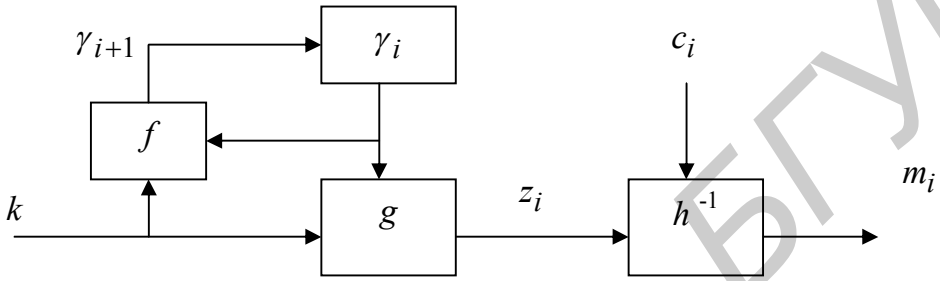

Рис. 4.2 Схема расшифрования синхронным поточным шифром

чайный ключ k в длинную псевдослучайную последовательность $g^l = g_1, \ldots, g_l$. Для двоичного ключа k^n длиной n битов гамма длиной l генерирует $(n\ l)$ последовательностей, каждая из которых это функция отображения множества $\{0, 1\}^n$ в множество $\{0, 1\}^l$. В теории сложности генератор задается асимптотически как бесконечный класс генераторов $(n, l(n))$-последовательностей, где l – это полиномиальная функция с индексом n, а время вычисления каждого генератора ограничено сверху полиномиальной функцией n.

Возможны следующие режимы работы поточных синхронных шифраторов.

Режим счетчика. Функция следующего состояния не зависит от битов ключа, но гарантированно проходит через все пространство состояний (или его большую часть):

$$z_{i+1} = F(z_i),\ g_i = f(k_i, z_i).$$

Примеры такой функции F – обычные счетчики и генераторы последовательностей максимальной длины на регистрах сдвига с линейной обратной связью.

Режим обратной связи от выхода (внутренней обратной связи). Выходная функция f не зависит от ключа

$$z_{i+1} = F(k_i, z_i),\ g_i = f(z_i).$$

Часто f состоит из 1-битного предиката текущего состояния (например самый младший бит или бит четности). Иногда рассматривается вариант такого режима, когда ключ задает начальное состояние генератора.

Синхронный поточный шифр имеет свойство не распространять ошибки. Расшифрование искаженной цифры шифротекста влияет только на соответствующую цифру открытого текста. С точки зрения синхронизации это положительное свойство. Но с другой стороны, ограничивается возможность обнаружения ошибки при расшифровании, но что еще более важно, противник имеет возможность произвести управляемые изменения в части шифротекста, совершенно точно зная, какие это вызовет изменения в соответствующем открытом тексте.

Процесс расшифрования выполняется в обратном порядке с использованием обратной функции h^{-1} .

Свойства синхроннных криптосистем:

1. *Требование синхронизации.* Шифрующий и расшифровывающий блоки должны быть засинхронизированы, т.е. в они в одно и то же время должны быть установлены в одно и то же начальное состояние. В противном случае криптосистема переходит в режим ложного шифрования и требует выполнения процесса ресинхронизации (повтороной синхронизации).

2. *Распространение ошибок.* Символы шифротекста в процессе передачи по каналам связи не влияют на другие символы шифротекста. Криптосистема не обладает свойством распространения ошибок.

3. *Устойчивость к активным атакам.* В соответствии с первым свойством криптосистемы атаки типа вставки, исключения, повторения символа шифротекста приводят к потери синхронизации, следовательно, могут быть обнаружены на приемной стороне. В соответствии со вторым свойством криптосистемы, активный злоумышленник может выборочно изменить символы шифротекста, точно зная, как они могут изменить открытый текст. Для противодействия атаки такого типа нужно дополнительно использовать механизмы аутентификации и контроля целостности сообщений.

Наиболее известным примером синхронных криптосистем является аддитивный бинарный поточный шифр, в котором ключевой поток, открытый текст и шифротекст представляют собой числа, представленные в двоичных кодах, а выходная функция h выполняет операцию суммирования по модулю 2. Функции f и g совмещены в генераторе ключевого потока.

Генератор гаммы синхронной криптосистемы должен выдавать одинаковую гамму как на шифрующей, так и на расшифровывающей стороне, и, следовательно, выход генератора гаммы должен быть детерминирован. Реальный генератор формирует повторяющуюся гамму. Такие генераторы называются периодическими и к ним относятся все генераторы гаммы. Период гаммы должен намного превышать число символов, формируемых между заменами ключей. Если это условие не выполняется, то различные части открытого текста будут зашифрованы одинаково, что значительно

снижает надежность системы. Если криптоаналитику известна часть открытого текста, он может восстановить часть гаммы и использовать ее для дальнейшей реконструкции открытого текста. Даже если криптоаналитик располагает только шифротекстом, он может выполнить операцию h^{-1} (суммирования по модулю 2) над разделами, шифрованными одинаковыми гаммами, и получить значение соответствующих участков открытого текста.

Атака вставкой. Допустим, криптоаналитик записал поток шифротекста, но не знает ни открытого текста, ни гаммы, использованной для шифрования открытого текста.

Оригинальный открытый текст: $m_1, m_2, m_3, m_4 \ldots$.

Оригинальная гамма : $k_1, k_2, k_3, k_4 \ldots$.

Оригинальный шифротекст : $c_1, c_2, c_3, c_4 \ldots$.

Предположим, что криптоаналитик вставляет один известный ему символ $m^{\#}$ в открытый текст после m_1, а затем ухитряется получить модифицированный открытый текст, зашифрованный той же гаммой. Новый шифротекст имеет следующий вид.

Новый открытый текст: $m_1, m^{\#}, m_2, m_3, m_4 \ldots$.

Оригинальная гамма : $k_1, k_2, k_3, k_4, k_5 \ldots$.

Оригинальный шифротекст : $c_1, c_{2\#}, c_{3\#}, c_{4\#}, c_{5\#} \ldots$.

Так как криптоаналитику известно значение $m^{\#}$, он может определить весь открытый текст, следующий за этим битом, по оригинальному и новому шифротекстам:

$$k_2 = c_{2\#} \oplus m^{\#}, \text{ тогда } m_2 = c_2 \oplus k_2 ;$$

$$k_3 = c_{3\#} \oplus m_2, \text{ тогда } m_3 = c_3 \oplus k_3 ;$$

$$k_4 = c_{3\#} \oplus m_3, \text{ тогда } m_4 = c_4 \oplus k_4 .$$

Для предотвращения такого вскрытия нельзя использовать одну и ту же гамму для шифрования двух разных сообщений.

4.2. Самосинхронизирующиеся поточные криптосистемы

В самосинхронизирующихся поточных криптосистемах каждый символ гаммы представляет собой функцию фиксированного числа предыдущих символов шифротекста. Иногда такая система называется автоключом шифра. Внутреннее состояние зависит от нескольких n предыдущих символов шифротекста. Выходная функция (g, h) должна иметь определенную криптографическую сложность. Так как внутреннее состояние полностью зависит от предыдущих символов шифротекста (система с памятью), то расшифровывающий генератор гаммы, приняв символы шифротекста, синхронизируется с шифрующим генератором. В интеллектуальных реализациях этой системы каждое сообщение начинается случайным (синхронизирующим) заголовком длиной n. Этот заголовок шифруется, передается, а затем расшифровывается. Рас-

шифрование сначала будет некорректным, но после приема *n* символов оба генератора гаммы синхронизируются.

Недостаток самосинхронизирующегося шифра – *распространение ошибок.* Для каждого символа шифротекста, искаженного при передаче, расшифровывающий генератор выдает *n* некорректных символов гаммы. Следовательно, пока испорченный символ влияет на внутреннее состояние, каждой ошибке шифротекста будут соответствовать *n* ошибок открытого текста.

Свойства самосинхронизирующихся поточных криптосистем:

1. *Самосинхронизация.* Режим самосинхронизации возможен, даже если произошло выпадение или вставка символов в шифротекст, поскольку процесс расшифрования зависит только от фиксированного числа *n* предыдущих символов шифротекста. Синхронизация восстанавливается автоматически после ее потери.

2. *Распространение ошибок.* После потери режима синхронного приема будут некорректно расшифрованы *n* символов открытого текста.

4. *Активные атаки.* В соответствии с первым свойством на приемной стороне труднее (по сравнению с синхронными системами) обнаружить вставку, устранение или повтор символов шифротекста активным злоумышленником. Например, злоумышленник записывает несколько битов шифротекста и спустя некоторое время вставляет эту запись в передаваемый сигнал. На приемной стороне после самосинхронизации старый шифротекст будет расшифрован как нормальный. Механизмом защиты от такой атаки служат метки времени.

5. *Диффузия статистики открытого текста.* Поскольку каждый символ открытого текста влияет на определенное число символов шифротекста, то статистические свойства открытого текста менее заметны в шифротексте (эффект рассеивания). Следовательно, самосинхронизирующиеся криптосиситемы могут быть более устойчивы, чем синхронные системы, к атакам на основе избыточности открытого текста.

Рассмотрим пример самосинхронизирующей криптосистемы. Пусть криптографическое преобразование T_k воздействует на исходное сообщение X:

$$Y = T_k \{X\}.$$

Очевидно, что в этом случае должно быть справедливо следующее соотношение:

$$X = T_k^{-1} \{Y\},$$

где T^{-1} – обратное по отношению к T криптографическое преобразование.

Схемы шифратора и расшифровывающего устройства приведены на рис. 4.3, 4.4.

Прямое и обратное преобразования должны удовлетворять определенным требованиям. Если в составе криптопреобразований использовать операцию сложения по модулю 2, прямые и обратные преобразования должны быть такими, чтобы m-й знак исходного сообщения X формировался бы из последовательности, состоящей из μ предыдущих знаков преобразованного сообщения Y и m-го знака этого же сообщения

Y, а определение m-го знака преобразованного сообщения Y осуществлялось бы по m-му знаку исходного сообщения X и последовательности, состоящей из μ предыдущих знаков Y.

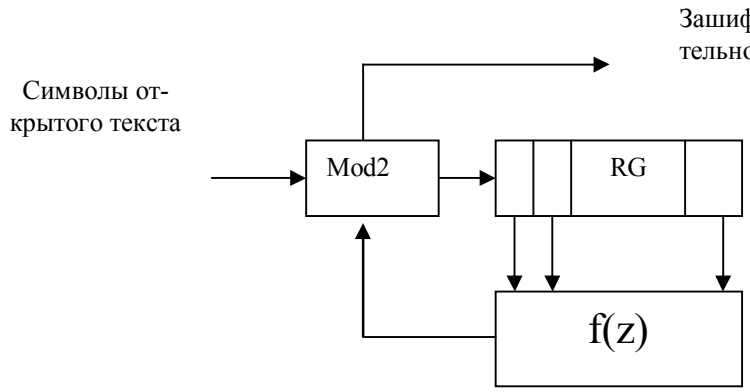

Рис. 4.3. Схема шифратора

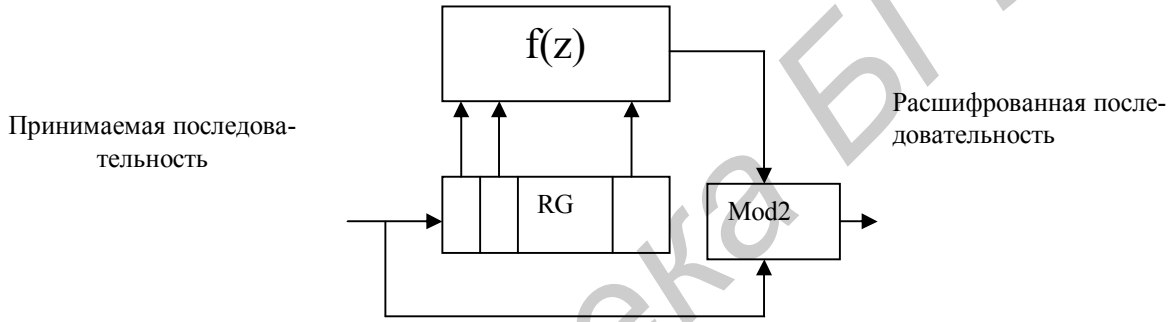

Рис. 4.4. Схема расшифровывающего устройства

Действительно, если m-й знак преобразованного сообщения Y определяется по формуле

$$y_m = [x_m + f(y_{m-\mu}, y_{m-\mu+1}, \ldots, y_{m-1})] \bmod 2,$$

где $f(y_{m-\mu}, y_{m-\mu+1}, \ldots, y_{m-1})$ – некоторая функция, принимающая значение 0 или 1 в зависимости от содержания последовательности $(y_{m-\mu}, y_{m-\mu+1}, \ldots, y_{m-1})$ шифротекста Y, то m-й знак исходного сообщения X может быть определен с помощью следующей формулы:

$$x_m = [y_m + f(y_{m-\mu}, y_{m-\mu+1}, \ldots, y_{m-1})] \bmod 2 =$$

$$= \{[x_m + f(y_{m-\mu}, y_{m-\mu+1}, \ldots, y_{m-1})] + f(y_{m-\mu}, y_{m-\mu+1}, \ldots, y_{m-1})\} \bmod 2 = x_m.$$

Приведенные сооотношения справедливы для $m \geq \mu+1$. Для того чтобы соотношения выполнялись и для $m < \mu + 1$, необходимо использовать для формирования знаков x_i, y_i некоторую кодовую последовательность R, состоящую из μ знаков. Тогда значение первого знака шифротекста Y будет определяться по формуле

$$y_1 = [x_1 + f(r_1, r_2, \ldots, r_\mu)] \bmod 2,$$

значение второго знака – по формуле $y_2 = [x_2 + f(r_2, r_3, \ldots, y_1)] \bmod 2$, а значение μ знака – по формуле $y_\mu = [x_\mu + f(r_\mu, y_1, \ldots, y_{\mu-1})] \bmod 2$.

Аналогичные значения функций $f(r_1, r_2, \ldots, r_\mu)$, $f(r_2, r_3, \ldots, y_1)$ и $f(r_\mu, y_1, \ldots, y_{\mu-1})$ должны использоваться при расшифровании текста.

При искажении одного знака шифротекста Y вследствие действия помех на приемной стороне будет наблюдаться искаженное «пачкой» ошибок исходное сообщение X, число ошибок в котором не будет превышать величины $\mu+1$. При этом действительное число ошибок в пачке и их расположение будут определяться значениями функции f(z) и видом Y. Значение величины μ определяет, с одной стороны, длину пачки ошибок и, следовательно, время синхронизации криптосистемы в условиях воздействия на канал передачи информации помех, а с другой стороны, – криптографическую стойкость системы. Если выполняется условие, что функция f(z) с равной вероятностью принимает значения 0 или 1 и такое условие выполняется для знаков шифротекста, то вероятность P_1 угадывания криптоаналитиком значений функции f(z) (являющихся по сути ключом криптографического преобразования) с первой попытки можно оценить по формуле

$$P_1 = 2^{-q}, q = 2^{\mu}.$$

Так, при $\mu = 4$ вероятность угадывания равна $P_1 = 2^{-16}$, а при $\mu = 10$ – $P_1 = 2^{-1024}$. Таким образом, для обеспечения нераскрываемости криптографической системы необходимо, чтобы выполнялось условие $2^q = 1{,}3 \times 10^{51}$. Следовательно, q = 169,8, а $\mu = 7{,}41$. Поскольку μ должно быть целым числом, то искомое значение $\mu = 8$.

4.3. Криптосистема RC4

Система RC4 реализует поточный шифр с переменной длиной ключа, разработана в 1987 г. Рональдом Ривистом (Ronald Rivest) для компании RSA Data Security. Пригодна для быстрого магистрального шифрования. Очень компактна в терминах размера кода и особо удобна для процессоров с побайтно-ориентированной обработкой. Криптогенератор функционирует независимо от открытого текста. Генератор имеет подстановочную таблицу (*S*-блок 8 x 8): S_0, S_1, . . ., S_{255}. Входами генератора являются замененные по подстановке числа от 0 до 255, и эта подстановка является функцией от ключа изменяемой длины. Генератор имеет два счетчика *i* и *j*, инициализируемых нулевым значением.

RC4 представляет собой семейство алгоритмов, задаваемых параметром *n*, который является положительным целым с рекомендованным типичным значением n=8. Внутреннее состояние генератора RC4 в момент времени *t* состоит из таблицы S_t=(S_t(l); $l = 1 \ldots 2^n - 1$), содержащей 2^n *n*-битных слов и из двух *n*-битных слов-указателей i_t и j_t.

Таким образом, размер внутренней памяти составляет M = $n\,2^n + 2n$ битов. Пусть выходное n-битное слово генератора в момент t обозначается как Z_t и начальные значения $i_0 = j_0 = 0$. Тогда функция следующего состояния и функция выхода RC4 для каждого $t \geq 1$ задается следующими соотношениями:

$i_t = i_{t-1} + 1$; $j_t = j_{t-1} + S_{t-1}(i_t)$; $S_t(i_t) = S_{t-1}(j_t)$, $S_t(j_t) = S_{t-1}(i_t)$; $Z_t = S_t(S_t(i_t) + S_t(i_t))$, где все сложения выполняются по модулю 2^n.

Подразумевается, что все слова, кроме подвергнутых своппингу, остаются теми же самыми. Выходная последовательность n-битных слов обозначается как $Z = (Z_t ; t = 1 \ldots \infty)$. Начальная таблица S_0 задается в терминах ключевой последовательности $K = (K_i; i = 0 \ldots 2^n - 1)$ с использованием той же самой функции следующего состояния, начиная от таблицы единичной подстановки $(l; l = 0 \ldots 2^n - 1)$.

Более строго, пусть $j_0 = 0$ и для каждого $1 \leq t \leq 2$ вычисляется $j_t = (j_t - 1 + S_t - 1(t-1) + K_t - 1) \bmod 2^n$, а затем переставляются местами $S_t - 1(t - 1)$ и $S_t - 1(j_t)$. На последнем шаге порождается таблица, представляющая S_0. Ключевая последовательность K составляется из секретного ключа, возможно повторяющегося, и рандомизированного ключа, передаваемого в открытом виде в целях ресинхронизации.

Для последовательностей, генерируемых RC4, не подходят методы статистического анализа, применяющиеся к генераторам на базе РСЛОС. Но, с другой стороны, для блоков, размер которых превышает M (размер внутренней памяти генератора), всегда существует линейная статистическая слабость или так называемая «линейная модель». Такую модель можно эффективно определять с помощью метода аппроксимации линейной последовательной схемой – АЛПС . *Линейная статистическая слабость* – это линейное соотношение между битами гаммы, которое выполняется с вероятностью, отличающейся от 1/2. С помощью метода АЛПС можно вывести линейные модели для RC4. Метод АЛПС заключается в нахождении и решении последовательной линейной схемы, которая аппроксимирует генератор гаммы и приводит к линейным моделям с относительно большим корреляционным коэффициентом c, где вероятность соответствующего линейного соотношения между битами гаммы составляет $(1 + c)/2$.

При анализе используется техника *двоичных производных*. Пусть $\{z_t; t = 1 \ldots \infty\}$ обозначает последовательность самых младших битов слов выхода RC4 и пусть

$$\dot{z} = (\dot{z}_t = z_t + z_{t+1}; t = 1 \ldots \infty); \qquad \ddot{z} = (\ddot{z}_t = z_t + z_{t+2}; t = 1 \ldots \infty)$$

обозначают ее первую и вторую двоичные производные соответственно. Показано, что $\dot{z}$ не коррелирует ни с 1, ни с 0, но $\ddot{z}$ коррелирует с 1 с корреляционным коэффициентом, близким к $(15\ 2^{-3n})$ при больших 2^n. Поскольку длина выходной последовательности, требуемая для выявления статистической слабости с корреляционным коэффициентом c, составляет $O(c^{-2})$, то эта длина равна примерно $64^n/225$. Например,

если $n = 8$, как рекомендуется в большинстве приложений, то требуемая длина близка к $2^{40} \approx 10^{12}$.

Поскольку результирующий коэффициент корреляции существенно превышает величину $2^M/2$, то установленную линейную модель следует рассматривать как статистическую слабость генератора, по крайней мере в теоретическом аспекте.

С практической точки зрения данная линейная модель может быть использована для выделения по шифротексту генератора RC4 среди других криптосистем, а также для восстановления параметра n.

Пример формирования шифра RC 4 показан на рис. 4.5.

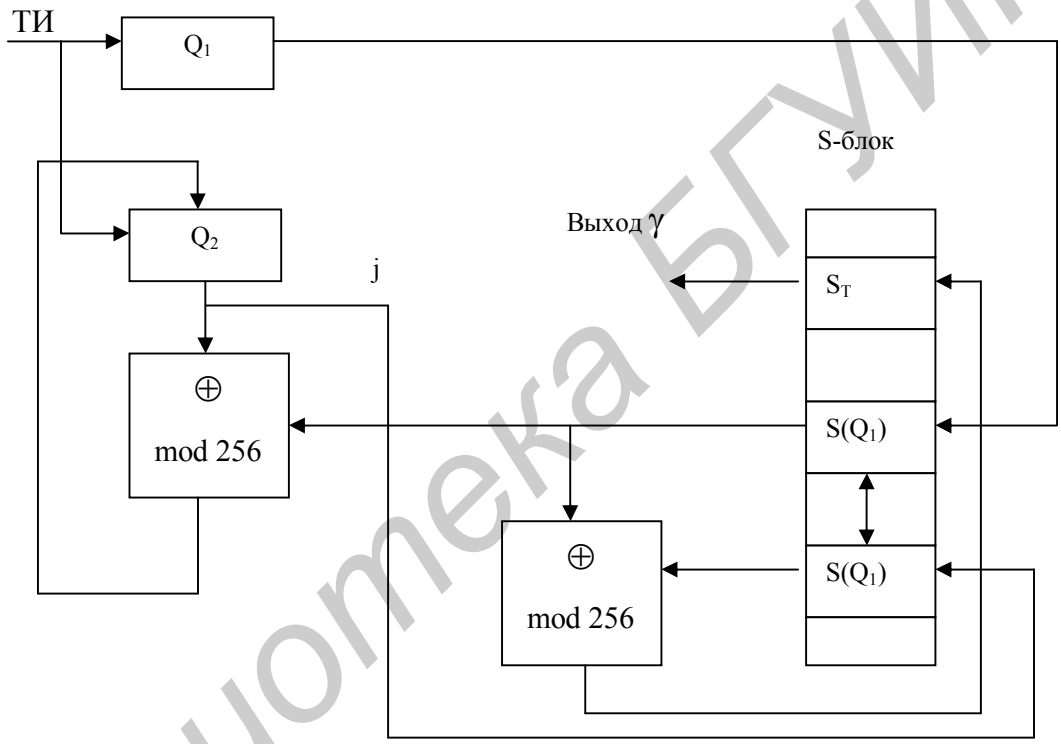

Рис.4.5. Схема формирования RC 4

Алгоритм работы:

1. Такт работы первого счетчика:

$$Q_1 = (Q_1 + 1) \bmod 2^8.$$

2. Такт работы второго счетчика

$$Q_2 = (Q_2 + S(Q_1)) \bmod 2^8.$$

3. Ячейки таблицы замен S-блока с адресами Q_1 и Q_2, обмениваются (свопируют) своим содержимым (операция swap):

$$S(Q_1) \leftrightarrow S(Q_2).$$

4. Вычисление суммы содержимого ячеек таблицы замен S-блока с адресами Q_1 и Q_2:

$$T = (S(Q_1) + S(Q_2).\bmod 2^8.$$

5. Считывание содержимого ячейки таблицы замен S-блока с адресом T: $\gamma = S_T$.

Алгоритм *инициализации* таблицы замен S-блока включает в себя запись в каждую ячейку таблицы замен S-блока ее собственного адреса, заполнение байтами ключа другой 256-байтной таблицы, инициализацию индекса j и перемешивание таблицы замен S-блока. Число состояний RC 4 приблизительно равно $2^{1700}(256! \times \times 256^2)$.

Пример. Зададим начальную фазу (зерно) ключевой последовательности в виде числа 427419669081. Используя процедуру случайного генерирования байтов, сформируем ключевую последовательность K из 32 цифр вида

K = 50, 33, 191, 16, 214, 33, 21, 158, 243, 219, 154, 0, 211, 175, 112, 31, 203, 146, 84, 110, 160, 148, 12, 207, 42, 90, 118, 207, 171, 147, 40, 186.

Выходная последовательность шифра RC4 длиной 256 бит имеет вид

[0, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0, 1, 0, 0, 0, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 1, 1, 0, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 1, 1, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 1, 1, 0, 1, 0].

На рис. 4.6 приведен спектр Уолша–Адамара выходной последовательности.

Рис. 4.6. Спектр Уолша–Адамара

На рис. 4.7 показан профиль линейной сложности LS сформированного сигнала.

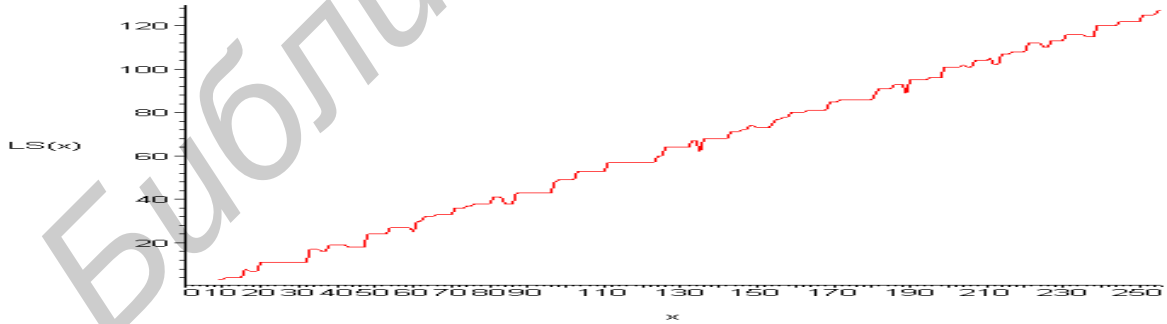

Рис.4.7. Профиль линейной сложности

Полученная последовательность может быть оценена как нелинейная булева функция и как псевдослучайная последовательность с линейной зависимостью профиля линейной сложности

5. ПОТОЧНЫЕ ШИФРЫ НА ОСНОВЕ РСОС

В основе проектирования генераторов гаммы лежит следующий подход. Выбирается один или несколько РСЛОС, обычно с различными длинами и различными многочленами обратной связи. Если длины взаимно просты, а все многочлены обратной связи примитивны, то у сконструированного генератора будет максимальная длина. Начальные состояния генераторов образуют ключ. Каждый раз для получения нового бита следует однократно сдвинуть РСЛОС (провести тактирование генераторов). Выходной бит представляет собой функцию, желательно нелинейную, некоторых битов РСЛОС. Эта функция называется комбинирующей функцией, а генератор в целом – *комбинирующим генератором*. Если выходной бит является функцией единственного РСЛОС, то генератор называется *фильтрующим* генератором. Иногда вводят усложнение структуры генератора. Для различных РСЛОС используется различная тактовая частота. Иногда тактирование одного генератора зависит от выхода другого. Такого типа генераторы называются генераторами с управляемым тактированием (генераторами с неравномерным движением). Управление тактированием может быть с прямой или обратной связью.

Наболее известны генераторы Геффе, Дженнингса, пороговые, различные генераторы «старт-стоп», самопрореживающие, многоскоростные скалярного произведения, суммирующие, сжимающие и самосжимающие генераторы.

В зависимости от структурных и алгебраических свойств некоторые из комбинирующих генераторов могут быть уязвимы для атак *встраиванием* и вероятностной корреляции.

5.1. Классические генераторы

Генераторы с управляемым тактированием

Основная идея состоит в том, что для обеспечения максимальной длины ключевого потока и других свойств необходимы определенные математически строгие структуры типа РСЛОС, но чтобы предотвратить раскрытие содержимого регистра и вскрытие алгоритма, следует внести некоторый сложный нелинейный беспорядок.

Один из возможных методов состоит в изменении тактирования РСЛОС. Например, когда тактирование одного РСЛОС 1 зависит от состояния другого РСЛОС 2 (рис. 5.1).

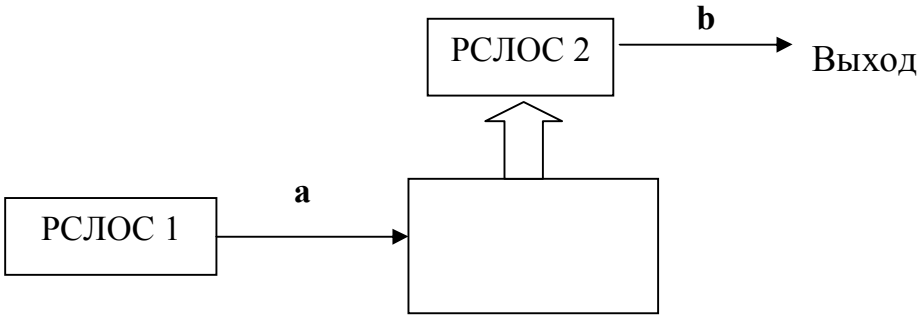

Рис. 5.1. Генератор с управляемым тактированием

Генератор «стоп – старт»

Простая модель генератора с управляемым тактированием – это генератор «стоп–старт». Положим, что два РСЛОС формируют последовательности **a** = $\{a_i\}$ и **b** = $\{b(t)\}$. Выходной является последовательность **u** = $\{u(t)\}$. Управление тактированием осуществляется следующим образом. Если в момент времени t символ $a_i = 1$, тогда на выходе генератора формируется символ $u(t) = b(i_t)$. В остальных случаях формируется символ $u(t) = b(i_t - 1)$, соответствующий состоянию в предыдущий момент времени. В общем случае формирование сигнала происходит в соответствии с формулой

$$u(t) = b(\sum_{i=0}^{t} a_i), \quad t = 0,1,\ldots .$$

Пример. Предположим, что последовательность **a** = 1 0 0 1 0 1 1, а последовательность **b** = ($b(0)$, $b(1)$, …, $b(6)$). Выходная последовательность имеет вид

$$u(0) = b(a_0) = b(1);$$
$$u(1) = b(a_0 + a_1) = b(1);$$
$$u(2) = b(a_0 + a_1 + a_2) = b(1);$$
$$u(3) = b(a_0 + a_1 + a_2 + a_3) = b(2) \ldots .$$

Сжимающий генератор

В сжимающих генераторах используется следующий способ управления тактированием. Выбираются два РСЛОС: РСЛОС 1 и РСЛОС 2. Тактовые импульсы подаются на оба генератора. Если выходом РСЛОС-1 является 1, то выходом генератора является выход РСЛОС-2. Если выход РСЛОС-1 равен 0, оба бита сбрасываются, оба РСЛОС тактируются заново и все повторяется.

Пусть выход генератора в предыдущий момент времени равен $u(i\text{-}1)$,

где $i = \sum_{j=0}^{t-1} a_j$. Начальные состояния генератора $u(0) = b(s)$, $a_0 = a_1 = \ldots a_{s\text{-}1} = 0$, $a_s = 1$.

Если $a_t = 1$, то выход генератора равен $b(t)$: $u(i) = b(t)$, $i > 0$. В остальных случаях генератор пропускает значение $b(t)$.

Пример. Если **b** = ($b(0)$, $b(1)$, …, $b(6)$), **a** = 1 0 0 1 0 1 1, то выходная последовательность имеет вид **u** = (($b(0)$, $b(3)$, $b(5)$, $b(6)$).

Одна из проблем реализации состоит в том, что скорость выдачи результирующего сигнала непостоянна. Если РСЛОС-1 формирует последовательность нулей, то на выходе генератора ничего нет.

Пример. Управляющий генератор представляет собой РСЛОС, построенный на основе неприводимого полинома $f_1(x) = x^3 + x + 1$. Управляемый генератор – РСЛОС на основе полинома $f_2(x) = x^5 + x^3 + 1$. Начальные состояния соответственно раны (1,0,0) и (0,0,1,0,1). Первый генератор формирует последовательность периода 7 следующего вида: **a** = (0,0,1,1,1,0,1). Второй генератор формирует m-последовательность периода 31 вида

b = (1,0,1,0,0,0,0,1,0,0,1,0,1,1,0,0,1,1,1,1,1,0,0,0,1,1,0,1,1,1,0).

Выходная последовательность ключевого потока равна

s = 1,0,0,0,0,1,0,1,1,1,1,0,1,1,1,0,… .

Основные свойства генератора:

1. Если РСЛОС имеют длины соответственно L_1 и L_2 и НОД (L_1, L_2) = 1, то ключевая последовательность **s** имеет период $(2^{L_2} - 1)2^{L_1 - 1}$.

2. Линейная сложность LS(**s**) удовлетворяет неравенству

$$L_2 2^{L_1 - 2} < LS(s) \le L_2 2^{L_1 - 1}.$$

3. Предположим, что полиномы обратных связей для РСЛОС выбраны случайно из всего множества примитивных полиномов степени L_1 и L_2 над Z_2. Тогда распределение шаблонов битов в выходной последовательности подчиняется почти равномерному закону. Иными словами, если **p** – любой бинарный вектор длиной t битов и $s(t)$ обозначает любой набор упорядоченных t битов в **s** , то вероятность того, что $s(t)$ = **p,** равна $(1/2)^t + O(t/2^{L_2})$.

Вычислительная сложность взлома генератора оценивается следующим образом. В том случае, если полиномы РСЛОС известны, но неизвестны начальные состояния, то наилучшая атака по восстановлению секретного ключа потребует $O(2^{L_1} L_2^3)$ шагов. С другой стороны, если используется неизвестные или изменяемые полиномы в РСЛОС, то наилучшая атака требует $O(2^{2L_1} L_1 L_2)$ шагов. Атака через оценку линейной сложности потребует $O(2^{L_1} L_2^2)$ шагов, но для её организации требуется выборка из $2^{L_1} L_2$ последовательных битов. Максимальная безопасность обеспечивается, когда используются последовательности максимальной длины. Если предположить, что $L_1 = L_2 = l$, то безопасность шифра приблизительно равна 2^{2l} и при $l \approx 64$ будет достаточно высокой.

Фильтрующий генератор гаммы

Конструкция генератора

Параметры:

1. Выборочный вектор **a** $=\{a_i\}$ в виде m-последовательности степени n над полем $GF(2)$.

2. Положительное число m с ограничением $m \le n$.

3. Ключевая последовательность (начальные состояния регистра сдвига) в виде m положительных целых чисел $0 \le d_1 < d_2 < \ldots < d_m$.

4. Булева функция $f(x_1, x_2, \ldots, x_m)$ от m переменных $x_1, x_2, \ldots x_m$.

Выходная последовательность **s** $= \{s_i\}$ фильтрующего генератора вычисляется по формуле

$$s_i = f(a_{d_i+i}, a_{d_2+i}, \ldots, a_{d_m+i}),\ i = 0,1,\ldots .$$

Схема генератора приведена на рис. 5.2

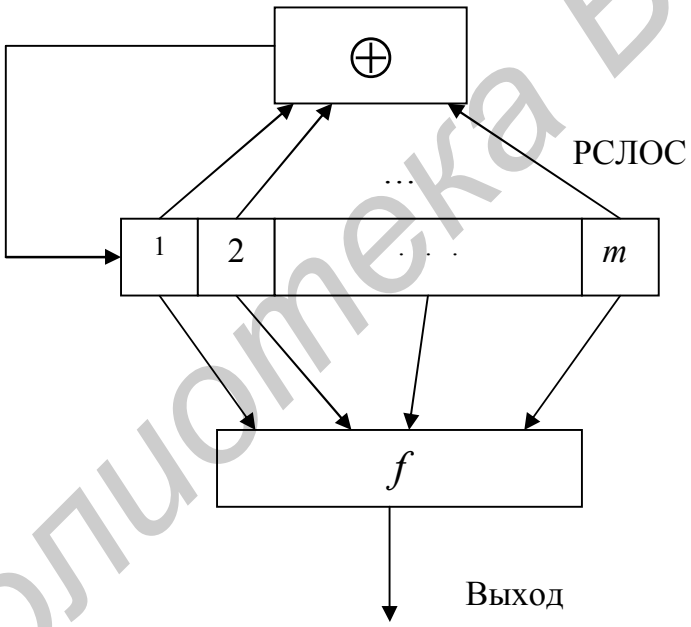

Рис. 5.2. Фильтрующий генератор гаммы

Профиль параметров последовательности фильтрующего генератора:

1. Период: $2^n - 1$.

2. Оценка линейной сложности: *LS*: удовлетворяет верхней границе

$$LS \le \sum_{k=1}^{m} \binom{n}{k},$$

в частности, для $m = 2$ оценка линейной сложности равна $LS = n^2$.

3. Уровень рандомизированности: в настоящее время оценка недостаточно ясная.

5.2. Генераторы на основе функций следа

Генератор на основе суммы функции следа. Для нечетного n ≥ 5 и $n = 2m + 1$ с периодом $p = 2^n - 1$ бинарная последовательность определяется как

$$s_i = Tr(\alpha^i) + Tr(\alpha^{q1i}) + Tr(\alpha^{q2i}), i = 0,1,\ldots ,$$

где α – примитивный элемент поля $GF(2^n)$, $q_1 = 2^m + 1$, $q_2 = 2^m + 2^{m-1}$ 1.

Для получения подобной последовательности $\{s_i\}$ можно использовать соотношение

$$q_2 \equiv q_1^2 \bmod p .$$

Комбинирующая последовательность представляет собой комбинацию трех последовательностей формируемых РСЛОС (рис. 5.3) и имеет двухуровневую автокорреляционную функцию.

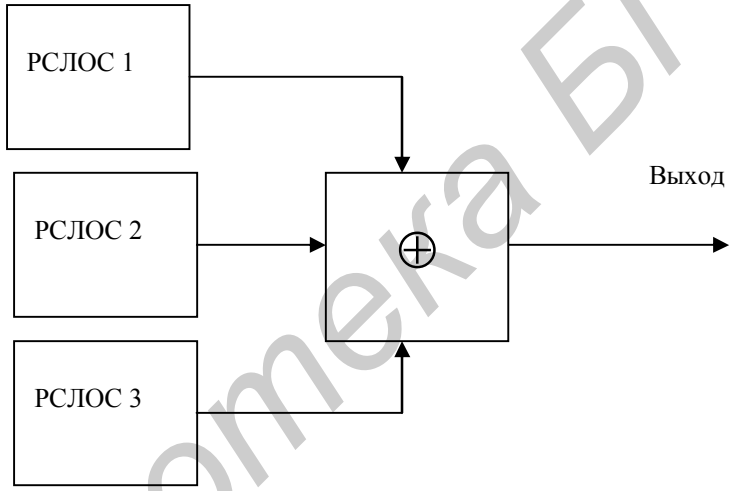

Рис. 5.3. Генератор на основе функции следа

Предварительные вычисления:

1. Выбор нечетного числа $n = 2m + 1$.
2. Выбор функции обратной связи в виде примитивного над полем $GF(2)$ полинома $f_0(x) = x^n + \sum_{i=0}^{n-1} c_i x^i$ и определения его корня α.
3. Вычисление децимирующих чисел $q_1 = 2^m + 1$, $q_2 = 2^m + 2^{m-1}$ 1.
4. Вычисление минимальных полиномов для α^{q1} и α^{q2}:

$$f_1(x) = x^n + \sum_{i=0}^{n-1} d_i x^i \text{ и } f_2(x) = x^n + \sum_{i=0}^{n-1} e_i x^i .$$

5. Вычисление трех начальных состояний:
   - $a_i = Tr(\alpha^i)$, $0 \le i \le n-1$, $\mathbf{a} = (a_0, \ldots , a_{n-1})$;

- $t_i = Tr(\alpha^{q_1 i})$, $0 \le i \le n-1$, $\mathbf{t} = (t_0, \ldots, t_{n-1})$;

- $v_i = Tr(\alpha^{q_2 i})$, $0 \le i \le n-1$, $\mathbf{v} = (v_0, \ldots, v_{n-1})$.

Формирование выходной последовательности осуществляется в соответствии с выражением

$$s_i = a_i + t_i + v_i, \; i = 0,1,\ldots,$$

где $a_i = \sum_{j=0}^{n-1} c_j a_{i+j}$, $t_i = \sum_{j=0}^{n-1} d_j t_{i+j}$, $v_i = \sum_{j=0}^{n-1} e_j v_{i+j}$.

Профиль параметров комбинирующей последовательности:

1. Период равен $2^n - 1$.
2. Сбалансированность: количество единиц равно 2^{n-1}, а нулей – $(2^{n-1} - 1)$.
3. Функция автокорреляции имеет двухуровневую форму.
4. Линейная сложность равна $3n$.
5. Количество возможных различных сдвигов последовательности определяется через функцию Эйлера $\varphi\,(2^n - 1) / n$.

GMW-последовательности. GMW-последовательности формируются с помощью функции следа $Tr_m^n(x)$, выполняющей отображение вида $GF(2^n) \rightarrow GF(2^m)$, m и n – целые положительные числа, причем число m делит n.

Определим два примитивных полинома. Первый полином вида

$$h(x) = x^m + h_{m-1}x^{m-1} + \ldots + h_1 x + h_0$$

над полем $GF(2^m)$, который будем использовать для формирования расширенного поля $GF(2^m)$. Второй примитивный полином определим над полем $GF(2^m)$:

$$g(x) = x^l + c_{l-1}x^{l-1} + \ldots + c_1 x + c_0,$$

где степень $l = n / m$.

Зададим число s, принимающее значения из интервала $1 < s < 2^m - 1$, взаимно простое с $(2^m - 1)$.

Алгоритм формирования. Бинарная GMW последовательность $\mathbf{a} = a_0, a_1, \ldots$, периода $N = 2^n - 1$ может быть сформирована следующим образом.

1. Определяется конечное расширенное поле $GF(2^m)$ полинома $h(x)$.
2. Проводятся в поле $GF(2^m)$ следующие вычисления:

$$b_{i+l} = \sum_{j=0}^{l-1} c_j b_{j+i},$$

вектор **b** представляет собой m-последовательность степени l над полем $GF(2^m)$;

$$d_i = b_i^s;$$

$$a_i = Tr_1^m(d_i), \quad i = 0,1,....$$

3. Формируется вектор $\mathbf{a} = (a_0, a_1, \ldots, a_{t-1})$, $t = 2^n - 1$.

Профиль рандомизированности GMW-последовательности:

1. Период последовательности равен $(q^n - 1)$.
2. Идеальные балансные свойства распределения m-грамм.
3. Двухуровневая автокорреляционная функция.
4. Линейная сложность оценивается как $LS(\mathbf{a}) = ml^{wt(s)}$, где $wt(s)$ – вес Хэмминга числа s.
5. Количество сдвигов различных последовательностей равно

$$N(GMW, n) = \frac{\phi(2^n - 1)}{n} \prod_{m|n} \left(\frac{\phi(2^n - 1)}{m} - 1 \right),$$

где $\phi(k)$ – функция Эйлера.

Пример. Зададим $n = 6$ и $m = 3$. Примитивный полином над полем $GF(2)$ равен $h(x) = x^3 + x^2 + 1$. Зададим $s = 3$. Выберем примитивный полином над полем $GF(2^3)$ в виде $g(x) = x^2 + \alpha^3 x + \alpha$, причем $h(\alpha) = 0$. Начальное состояние генератора m-последовательности $(b_0, b_1) = (0, \alpha^3)$.

Следующие символы m-последовательности вычисляются по формуле

$$b_{i+2} = \alpha^3 b_{i+1} + \alpha b_i, \quad i = 0,1,\ldots.$$

Проводя соответствующие вычисления в поле $GF(2^3)$, получим

$\mathbf{b} = \alpha^0, \alpha^3, \alpha^6, \alpha^5, \alpha^5, \alpha^2, \alpha^3, \alpha^5, \alpha^3, \alpha^0, \alpha^4, \alpha^0, \alpha^6, \alpha^6, \alpha^3, \alpha^4, \alpha^6, \alpha^4, \alpha^0, \alpha^5, \alpha^1, \alpha^0, \alpha^0, \alpha^4, \alpha^5, \alpha^0, \alpha^5, \alpha^0, \alpha^6, \alpha^2, \alpha^1, \alpha^1, \alpha^5, \alpha^6, \alpha^1, \alpha^6, \alpha^0, \alpha^0, \alpha^3, \alpha^2, \alpha^2, \alpha^6, \alpha^0, \alpha^2, \alpha^0, \alpha^0, \alpha^1, \alpha^4, \alpha^3, \alpha^3, \alpha^0, \alpha^1, \alpha^3, \alpha^1, \alpha^0, \alpha^2, \alpha^5, \alpha^4, \alpha^4, \alpha^1, \alpha^2, \alpha^4, \alpha^2.$

Отображение вида $d_i = b_i^3$ дает последовательность элементов расширенного поля

$\mathbf{d} = \alpha^0, \alpha^2, \alpha^4, \alpha^1, \alpha^1, \alpha^6, \alpha^2, \alpha^1, \alpha^2, \alpha^0, \alpha^5, \alpha^0, \alpha^4, \alpha^4, \alpha^2, \alpha^5, \alpha^4, \alpha^5, \alpha^0, \alpha^1, \alpha^3, \alpha^0, \alpha^0, \alpha^5, \alpha^1, \alpha^0, \alpha^1, \alpha^0, \alpha^4, \alpha^6, \alpha^3, \alpha^3, \alpha^1, \alpha^4, \alpha^3, \alpha^4, \alpha^0, \alpha^0, \alpha^2, \alpha^6, \alpha^6, \alpha^4, \alpha^0, \alpha^6, \alpha^0, \alpha^0, \alpha^3, \alpha^5, \alpha^2, \alpha^2, \alpha^0, \alpha^3, \alpha^2, \alpha^3, \alpha^0, \alpha^6, \alpha^1, \alpha^5, \alpha^5, \alpha^3, \alpha^6, \alpha^5, \alpha^6.$

Отображение вида $a_i = Tr_1^3(d_i)$ дает бинарную последовательность

$\mathbf{a} = 0\ 1\ 1\ 1\ 1\ 0\ 1\ 1\ 1\ 0\ 0\ 1\ 1\ 1\ 1\ 0\ 1\ 0\ 0\ 1\ 0\ 1\ 1\ 0\ 1\ 1\ 1\ 0\ 1\ 0\ 0\ 0\ 1\ 1\ 0\ 1\ 0\ 1\ 1\ 0\ 0\ 1\ 1\ 0\ 1\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0.$

Профиль параметров GMW–последовательности (n = 6):

1. Период равен 63.
2. Сбалансированность, идеальное распределение 3-грамм.

3. Двухуровневая функция периодической автокорреляции.
4. Линейная сложность оценивается как s =3, LS(**a**) = 3 2^2 = 12.
5. Количество возможных сдвигов N(GMW, 6) = 6.

3-термовые последовательности. Для формирования используют три разных генератора m-последовательности, выходы которых объединяются сумматором по модулю два. Бинарные последовательности **a** = $\{a_i\}$ для нечетного числа $n \geq 5$ и $n = 2m + 1$ имеют период $p = 2^n - 1$. Элемент последовательности задается выражением из трех терм:

$$a_i = Tr(\alpha^i) + Tr(\alpha^{q1 i}) + Tr(\alpha^{q2 i}),\ i=0,1,\ldots,$$

где α – примитивный элемент поля $GF(2^n)$ и $q_1 = 2^m + 1$, $q_2 = 2^m + 2^{m-1} + 1$ или $q_2 \equiv q_1^2 \bmod p$.

Последовательность имеет двухуровневую периодическую автокорреляционную функцию.

Синтез генератора последовательности можно вести в следующей последовательности.

1. Выбирается нечетное число $n = 2m + 1$.
2. Выбирается полином $f_0(x) = x^n + \sum_{i=0}^{n-1} c_i x^i$, который примитивен над полем $GF(2)$, элемент α расширенного поля является корнем выбранного полинома $f_0(\alpha)=0$.
3. Вычисляются децимирующие числа $q_1 = 2^m + 1$, $q_2 = 2^m + 2^{m-1} + 1$.
4. Вычисляются минимальные полиномы для α^{q1} и α^{q2}:

$$f_1(x) = x^n + \sum_{i=0}^{n-1} d_i x^i,\ f_2(x) = x^n + \sum_{i=0}^{n-1} e_i x^i.$$

5. Вычисляются три начальных состояния:

$s_i = Tr(\alpha^i), 0 \leq i \leq n-1, \mathbf{s}_0 = (s_0, s_1, \ldots, s_{n-1})$ – для первого генератора,

$t_i = Tr(\alpha^{q_1 i}), 0 \leq i \leq n-1, \mathbf{t}_0 = (t_0, t_1, \ldots, t_{n-1})$ – для второго генератора,

$u_i = Tr(\alpha^{q_2 i}), 0 \leq i \leq n-1, \mathbf{u}_0 = (u_0, u_1, \ldots, u_{n-1})$ – для третьего генератора.

6. Вычисляются элементы составляющих последовательностей

$$s_{i+n} = \sum_{j=0}^{n-1} c_j s_{i+j}, t_{i+n} = \sum_{j=0}^{n-1} d_j t_{i+j}, u_{i+n} = \sum_{j=0}^{n-1} e_j u_{i+j},\ i = 0,1,\ldots.$$

7. Вычисляются элементы результирующей последовательности

$$a_i = s_i + t_i + u_i,\ i = 0,1,\ldots.$$

Одна из возможных схем генератора последовательности приведена на рис.5.3.

Профиль параметров 3-термовой последовательности:

1. Период равен $2^n - 1$.
2. Сбалансированность нулей и единиц. Последовательность состоит из 2^{n-1} единиц и $(2^{n-1} - 1)$ нулей.
3. Двухуровневая периодическая автокорреляционная функция.
4. Линейная сложность оценивается как $3n$.
5. Количество сдвигов различных 3-термовых последовательностей

$$N(T3, n) = \phi(2^n - 1)/n .$$

5.3. Схемы на регистрах сдвига с операцией переноса

В течение 90-х гг. два американских алгебраиста-криптографа Э. Клаппер и М. Горецки разрабатывали теорию новой и очень простой архитектуры для генерации гаммы на основе регистров сдвига, которую они назвали *регистры сдвига с* «*обратной связью с переносом*» (РСОСП) .

РСОСП – это регистр сдвига, имеющий небольшое количество добавочной памяти. В своей простейшей форме ячейки регистра заполняют биты 0 и 1, а память m_{n-1} содержит неотрицательное целое число (рис.5.4).

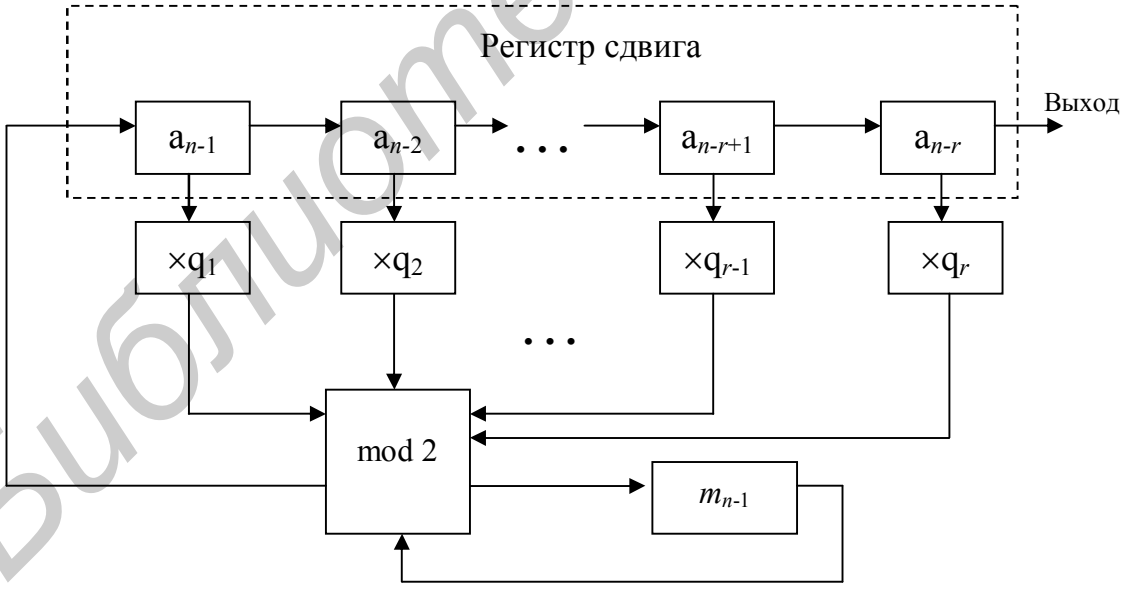

Рис. 5.4. Регистр сдвига с обратной связью с переносом

Содержимое ячеек с точками съема прибавляется *как целые числа* к содержимому памяти для формирования суммы **σ**. Бит четности (**σ** mod 2) суммы **σ** в качестве обрат-

ной связи поступает на первую ячейку регистра, а старшие биты ($\lfloor \sigma / 2 \rfloor$) используются для образования нового значения памяти.

Как оказалось, для генерируемых РСОСП-последовательностей характерны многие важные свойства, присущие линейным рекуррентным (РСЛОС) последовательностям. Но для анализа РСОСП-последовательностей используется арифметика 2-адических чисел.

Алгебраические свойства РСОСП-последовательностей:

1. В РСОСП длиной r отводы ячеек $q_1, q_2, \ldots, q_r$ задают *целое число обратной связи*:

$$q = q_r 2^r + q_{r-1} 2^{r-1} + \ldots + q_1 2 - 1.$$

Период и многие другие свойства РСОСП-последовательности могут быть выражены в терминах теоретико-числовых свойств этого целого числа.

2. Если периодическая последовательность $\boldsymbol{a} = (a_0, a_1, a_2 \ldots)$ получена от РСОСП с целым обратной связи q и если $\gamma = 2^{-1} \in Z/(q)$ – мультипликативная инверсия элемента 2 в кольце целых по модулю q, то существует $A \in Z/(q)$, такое, что для всех $i = 0, 1, 2, \ldots$ имеем

$$a_i = (A\, \gamma^{\,i} (\mathrm{mod}\ q)) (\mathrm{mod}\ 2).$$

3. Всякая бесконечная последовательность $\boldsymbol{a} = (a_0, a_1, a_2 \ldots)$ может быть выражена с помощью формального степенного ряда $\alpha = \sum_{i=0}^{\infty} a_i 2^i$, являющегося элементом кольца 2-адических чисел Z_2. Последовательность $\boldsymbol{a}$, в конечном счете, периодична тогда и только тогда, когда 2-адическое число α является рациональным, т. е. если существуют такие целые r и q, что

$$\alpha = r / q \in Z_2.$$

В этом случае знаменатель q – это целое обратной связи для РСОСП, порождающей периодическую часть последовательности $\boldsymbol{a}$. Последовательность $\boldsymbol{a}$ строго периодична, если выполняются неравенства $\alpha < 0$ и $|r| < |q|$.

4. Размер наименьшего РСОСП, порождающего заданную периодическую последовательность $\boldsymbol{a}$ назван *2-адическим размахом* $\boldsymbol{a}$. Алгоритм формирования, использующий регистр сдвига минимального размера, может быть найден эффективным способом с использованием теории 2-адической аппроксимации. Требуется знание всего $2M + 2\log(M)$ битов последовательности (где M обозначает 2-адический размах $\boldsymbol{a}$). Алгоритм является *адаптивным*: каждый раз, когда определен новый бит (например при атаке с открытым текстом), он используется для быстрого обновления определенного на предыдущем шаге РСОСП. Таким образом, количество битов не требуется знать заранее.

5. Рассмотрим схему сумматора, в котором две бесконечные периодические последовательности $\boldsymbol{a} = (a_0, a_1, a_2 \ldots)$ и $\boldsymbol{b} = (b_0, b_1, b_2 \ldots)$ складываются с помощью опе-

рации *переноса*. Результирующая последовательность $\boldsymbol{c} = (c_0, c_1, c_2 \ldots)$ задается сложением $\gamma = \alpha + \beta \in Z_2$ в кольце 2-адических целых (где $\gamma = \sum_{i=0}^{\infty} c_i 2^i$). Доказано, что 2-адический размах последовательности ***c*** аппроксимировано ограничен суммой 2-адических размахов последовательностей ***a*** и ***b***.

6. По определению, *l-последовательность* – это РСОСП-последовательность максимально возможного периода $T = q - 1$ (где q – целое обратной связи РСОСП). Такие *l*-последовательности генерируются РСОСП с целыми обратной связи q, для которых 2 является примитивным корнем. Отдельный период *l*-последовательности – это циклический сдвиг последовательности, образованной *реверсированием* (обращением) отдельного периода в двоичном разложении дроби $1/q$. Последовательности, основанные на разложении дробей, изучаются еще со времен Гаусса и имеют примечательные вероятностные и корреляционные свойства, параллельные свойствам *m*-последовательностей.

Важным результатом с криптоаналитической точки зрения стал разработанный на данном базисе метод атаки криптосхемы сумматора. Известно, что линейный размах результирующей последовательности ***c*** равен произведению линейных размахов ***a*** и ***b***, но, с другой стороны, из свойства (5) следует, что 2-адический размах ***c*** равен лишь сумме 2-адических размахов ***a*** и ***b***. Более того, алгоритм рациональной аппроксимации (4) отыскивает эквивалентный РСОСП, который генерирует последовательность ***c*** по малому количеству битов.

Самый же интересный результат – это то, что последовательности, генерируемые на основе РСОСП-архитектуры, в принципе поддаются строгому анализу.

P-адические псевдослучайные последовательности. Один из методов построения криптогенераторов связан с теорией p-адических чисел и функций. Определим R – коммутативное кольцо, не имеющее нулевых делителей; поле F и простой элемент $\pi \in R$. Идеал, сформированный π, запишется как $I = (\pi)$.

Например, пусть $\pi^2 + 2\pi = 2$; $R = \mathbf{Z}[\pi] = \mathbf{Z} + \pi\mathbf{Z}$; $I = \pi\mathbf{Z} + 2\mathbf{Z}$; тогда $F = \mathbf{Q}[\pi] = \mathbf{Q}[\sqrt{3}]$ – поле квадратических чисел.

Представляет интерес случай, когда целая часть $K = R/(f)$ конечна. В этом случае K представляет собой поле вычетов (R, π). Множество $\{(\pi^i)\}$ определяет базис, и такая топология известна как π-адическая топология на R.

Класс псевдослучайных последовательностей (ПСП) задается кортежом (R, f, A, T), где $f \in R$; A, T являются подмножествами R, каждое из которых в свою очередь является множеством представителей для $R/(f)$. Определим полином $q(x) = \sum_{i=1}^{r} q_i f^i - q_0$

для некоторого значения r и $q_i \in A$; q_0 – инвертирумое относительно модуля f. С точки зрения теории чисел коэффициенты $q_0, q_1, ..., q_r$ являются коэффициентами последовательности f-адического расширения рациональной функции u/q с $u \in R$. Известно [6], что рациональная функция со знаменателем q может быть записана как f-адическое число

$$\frac{u}{q} = \sum_{i=0}^{\infty} a_i f_i ,$$

причем $a_i(x)$ является элементом A.

ПСП S может интерпретироваться как последовательность коэффициентов f-адического числа $s_i = a_i$ и формируется с помощью регистра сдвига со специальной, алгебраической обратной связью, которая задается элементами $q_0, q_1, ..., q_r$ и элементом памяти m. Состояния регистра сдвига определяется как $(a_0, a_1, ..., a_{r-1}; m)$, $a_i \in A$ и $m \in R$. Изменения состояний регистра сдвига происходит следующим образом. Предполагается, что существуют уникальные элементы $a_r \in A$ и $m' \in R$ такие, что выполняется равенство

$$q_0 a_r + fm' = m + \sum_{i=1}^{r} q_i a_{r-i} .$$

Новое состояние регистра сдвига определяется как $(a_1, a_2, ..., a_r; m')$.

Заметим, что ПСП, формируемые с помощью регистра сдвига с линейной обратной связью можно рассматривать как частный случай $R = F[x]$ для некоторого поля $F, f = x, A = T = F$ и $q_0 = 1$.

ПСП, формируемые с помощью регистров сдвига со специальной обратной связью [3], можно рассматривать как вариант:

$$R = \mathbf{Z}, f = N \in \mathbf{Z}\ A = T = \{0,1,...,N-1\} .$$

В качестве примера рассмотрим формирование ПСП по следующему алгоритму.

1. Вычисляется

$$\tau = \sum_{i=1}^{r} q_i a_{r-i} + m .$$

2. Находится $a_r \in A$, такое, что выполняется сравнение $q_0 a_r \equiv \tau \mod \pi$.

3. Состояния регистра сдвига $(a_0, ..., a_{r-1})$ меняются на состояния $(a_1, ..., a_r)$, а в память m заносится значение $quo(\tau - q_0 a_r, \pi)$,

где

$$quo(\alpha,\pi)=\sum_{i=0}^{\infty}a_{i+1}\pi^{i}\ .$$

Положим $R=\mathbf{Z}$, $F=\mathbf{Q}$, $\pi=p=2$ и, следовательно, $K=GF(2)$, A=T={0,1}. Зададим r = 3 и $q=\pi^2+\pi+1$. Структурная схема соответствует схеме, показанной на рис. 5.4, а пространство состояний генератора ПСП приведено в табл.5.1.

Таблица 5. 1

омер шага	остояния регистра	амять
0	111	0
1	110	1
2	100	1
3	000	1
4	001	0
5	011	0

Псевдослучайные последовательности, сформированные с помощью алгебраической обратной связи, представляются как след степени примитивного элемента $a_i=(\delta\gamma^{i} \bmod q)\bmod\pi$ для некоторого δ, где γ – инверсия π по модулю q.

В важном для практики случае p = 2 (бинарные последовательности) можно ввести понятие 2-адической линейной сложности $\lambda_2(S)$ и 2-адической вычислительной сложности $\Phi_2(S)$ [5].

Обозначим через $\sigma(S)$ количество ненулевых коэффициентов преобразования Фурье последовательности S. Тогда 2-адическая вычислительная сложность удовлетворяет границе

$$\Phi_2(S)<\sigma(S)+2^{\omega(N)-1},$$

где $\omega(N)$ – число положительных простых делителей N.

6. ПРОЕКТИРОВАНИЕ ПОТОЧНЫХ СТРУКТУР НА ОСНОВЕ ТЕОРИИ СЛОЖНОСТИ

В основе теоретико-сложностного подхода лежит идея о вычислительно достижимой информации. Если две случайных величины статистически независимы, то не существует способа вычислить информацию о второй величине, наблюдая первую. Но даже если существует полная статистическая зависимость, она может не быть вычислительно достижима.

6.1. Базовые идеи и концепции

Битовый генератор G – это последовательность множества $\{G_n: n \geq 1\}$ алгоритмов полиномиального времени G_n. Каждый оператор G_n: $\{0, 1\}^n \rightarrow \{0, 1\}^l$ разворачивает некоторый случайный ключ x^n длиной n в псевдослучайную последовательность z^l длиной $l(n)$, где l – полиномиальная функция от n. Пусть $z^l = =G_n(x^n)$ обозначает последовательность, выработанную алгоритмом G_n от входа x^n.

Пусть $\mu_{R,l}$ обозначает равномерное вероятностное распределение на множестве l-битных последовательностей, то есть $\mu_{R,l}(s^l) = 2^{-l}$. Соответственно, пусть $\mu_{G,l(n)}$ означает вероятностное распределение последовательностей гаммы z^l, генерируемых G_n при случайно выбираемых ключах (то есть при ключах, выбираемых в соответствии с $\mu_{R,n}$). Тогда вероятность отдельной псевдослучайной последовательности z^l

$$\mu_{G,l}(z^l) = 2^{-n} \#\{x^n: G_n(x^n) = z^l\}.$$

где # определяет количество элементов множества.

Пусть μ_G – последовательность вероятностных распределений $\{\mu_{G,l(n)}: n \geq 1\}$, генерируемых G. Практическое требование к стойкости генератора гаммы – это его непредсказуемость. При наличии фрагмента z^i из i битов невозможно развить эту последовательность за пределы i. В противном случае добытый фрагмент гаммы позволил бы успешно дешифровать некоторую часть шифротекста без знания ключа. Концепция непредсказуемости может быть формализована с помощью идеи «теста следующего бита» (или «предсказателя»).

Предсказатель (*тест следующего бита*) $C = \{C_n: n \geq 1\}$ – это последовательность вероятностных полиномиального размера схем C_n с $i_n < l(n)$ входами и одним двоичным выходом. Предсказатель C способен предсказать псевдослучайный генератор G, если доля тех случаев, когда выходной бит C_n согласуется с z_{i+1}, «существенно» отличается от ½. Более строго, говорят, что C предсказывает псевдослучайный генератор G (или «G не проходит тест следующего бита C»), если существует многочлен $P(n)$, такой, что для бесконечно большого числа n вероятность предсказания

$$\Pr(C_n(z^i) = z_{i+1}) \geq \frac{1}{2} + \frac{1}{P(n)},$$

где z^l появляются согласно $\mu_{G,l(n)}$.

Напротив, G проходит C, если для всех, кроме конечного числа n, и для всех многочленов $P(n)$ выполняется неравенство

$$\Pr(C_n(z^i) = z_{i+1}) < \frac{1}{2} + \frac{1}{P(n)}.$$

Аналогично G непредсказуем, если для всех тестов следующего бита C, для всех, кроме конечного числа n, для всех многочленов $P(n)$ и для всех $i < l(n)$

$$\Pr(C_n(z^i) = z_{i+1}) < \frac{1}{2} + \frac{1}{P(n)}.$$

Генератор гаммы пытается эффективно имитировать случайность. Если бы псевдослучайные последовательности, которые он генерирует, были эффективно отличимы от чисто случайных последовательностей, то нельзя было бы утверждать, что данный генератор – симулятор случайности. С другой стороны, выход генератора всегда может быть отличён от чисто случайных последовательностей простым перебором всех ключей и сравнением получающихся последовательностей с той последовательностью, что попала в руки аналитика. Но такой метод требует экспоненциального времени и считается реально недостижимым.

Способность отличить псевдослучайную последовательность от случайной последовательности представляется наиболее фундаментальным шагом, предшествующим любому другому шагу при анализе псевдослучайных последовательностей. Концепция неотличимости может быть формализована с помощью идеи статистического теста .

Статистический тест $T = \{T_n: n \geq 1\}$ для битового генератора – это последовательность вероятностных полиномиального размера схем T_n с $l(n)$ входами и одним двоичным выходом. Тест T способен отличить псевдослучайный генератор G, если доля тех случаев, когда он дает на выходе 1 при несовпадении с последовательностями r^l, берущимися равновероятно из $\mu_{R,l}$, «существенно» отличается от доли тех случаев, когда он дает на выходе 1 при несовпадении с последовательностями z^l, берущимися в соответствии с распределением выхода генератора $\mu_{G,l(n)}$.

Считается, что тест T отличает G, если существует многочлен $P(n)$, такой, что для бесконечно большого числа n

$$| p_n^{T,G} - p_n^{T,R} | \geq \frac{1}{P(n)},$$

где $p_n^{T,G}$, $p_n^{T,R}$ обозначают вероятности того, что T выдает 1 при подаче на вход последовательностей, берущихся согласно $\mu_{G,l(n)}$ и $\mu_{R,l}$.

Напротив, генератор G проходит статистический тест T, если для всех многочленов $P(n)$ и для всех, кроме конечного числа n, выполняется неравенство

$$| p_n^{T,G} - p_n^{T,R} | < \frac{1}{P(n)}.$$

По сути идеи предсказателя и статистического теста эквивалентны. Битовый генератор G проходит все тесты следующего бита C тогда и только *тогда,* когда он проходит все статистические тесты T. Тогда можно определить как *совершенный* битовый генератор G, если он проходит все статистические тесты T полиномиального размера.

Общая схема для конструирования битовых генераторов: пусть $f = \{f_n: X_n \to \to X_n\}$ – однонаправленная подстановка, используемая в качестве функции следующего состояния генератора. Пусть $B = \{B_n: X_n \to \{0, 1\}\}$ – двоичный предикат с областью определения X_n, используемый в качестве выходной функции. Случайно выбираем элемент $x \in X_n$ в качестве зерна (начального значения), итерируем f_n от x, и берем на выходе $z_i = B_n(f_n^i(x))$ для $1 \le i \le l(n)$. Если B – непредсказуемый предикат для f, то последовательность z^l непредсказуема (тестами следующего бита) влево и неотличима любым статистическим тестом T (в частности она также непредсказуема вправо).

6.2. Генераторы поточных структур

Генератор Блюма–Микали. **Безопасность генератора определяется трудностью вычисления дискретного логарифма. Пусть g – простое число, а p – нечетное простое число. Ключ x_0 начинает процесс:**

$$x_{i+1} = g^{x_i} \bmod p.$$

Если $x_i < (p-1)/2$, выходом генератор будет 1, и 0 – в противном случае. Если p достаточно велико, чтобы вычисление дискретных логарифмов по модулю p стало физически невозможным, этот генератор можно считать надежным.

Генератор RSA. Генератор RSA представляет собой модификацию генератор Блюма–Микали. Начальным параметром служат модули n, равные произведению двух больших простых чисел p и q, и целое число e взаимно простое с (p –1)(q –1), а также случайное начальное число x_0 меньшее n:

$$x_{i+1} = x_i^e \bmod n .$$

Выход генератора представляет собой младший значащий бит x. Безопасность этого генератора опирается на сложность вскрытия криптосистемы RSA и следующих предположений и гипотез.

1. Всякий вероятностный алгоритм, который, имея на входе $x^e \bmod N$, e и N, способен угадать самый младший бит числа x с вероятностью, по крайней мере ½ + ε(n), эквивалентен статистическому тесту T, который способен различать такие распределения, как равновероятное распределение на [1, N] и распределение $x^e \bmod N$ для случайного, четного $x \in [1, N]$ с преимуществом ε(n).

2. Два приведенных выше распределения при случайно выбираемом RSA-модуле размером n неразличимы статистическими тестами полиномиального времени. Итерационное применение $x^e \bmod N$ в генераторе RSA приводит к последовательности z^l, которая также неотличима от равновероятной статистическими тестами полиномиального времени, и, следовательно, RSA-генератор является «совершенным». Эти же аргументы можно распространить на log n самых младших бит числа x, которые неотличимы от равновероятных, если RSA полагать стойким. Таким образом, без потери «совершенства» можно выделять log n битов при каждом модульном экспоненциировании RSA-генератора.

Модифицированный RSA-генератор. Генерация всего лишь одного бита (или log n битов) гаммы на одно модульное экспоненциирование – это слишком медленно для большинства приложений. Чтобы преодолеть этот недостаток, Микали и Шнорр ввели следующую, намного более сильную гипотезу.

RSA-гипотеза. Пусть $e \geq 3$ – нечетное целое число. Для случайных модулей N размером n (которые являются произведениями двух простых чисел, размер каждого из которых n/2), таких, что НОД(e, φ (N)) = 1, и всех M, пропорциональных $N^{2/e}$, следующие распределения на [1, N] неотличимы статистическими тестами полиномиального времени:

а) равновероятное распределение на [1, N];

б) распределение $x^e \bmod N$ для случайного $x \in [1, M]$.

В *модифицированном* RSA-генераторе при каждом экспоненциировании берется доля (e – 2)/e битов, и только доля 2/e нужна для следующей итерации. Предполагается, что итерационное применение $x^e \bmod N$ в генераторе RSA приводит к последова-

тельности z^l, которая также неотличима от равновероятной статистическими тестами полиномиального времени.

Генератор квадратичных вычетов. Пусть N будет произведением двух различных нечетных простых чисел p и q. Элемент $y \in Z_N^*$ называется квадратичным вычетом по модулю N, если $y = x^2 \mod N$ для некоторого $x \in Z_N^*$. Обозначим множество квадратичных вычетов по модулю N как QR_N. Каждый элемент $y \in QR_N$ имеет ровно четыре квадратных корня. Если $p = q = 3 \mod N$, то в точности один из этих четырех квадратных корней принадлежит QR_N. Как следствие, отображение

$$QR_N \ni x \mapsto x^2 \mod N \in QR_N$$

является отображением «на» и «один в один» и имеет соответствующее обратное отображение

$$QR_N \ni y \mapsto \sqrt{y} \mod N \in QR_N .$$

Рабин М. показал, что факторизация N и вычисление квадратных корней являются эквивалентными проблемами .

Алгоритм генератора квадратичных вычетов:

Вход: параметры: модуль N размером n.

Ключ: случайно выбираемое $x_1 \in QR_N$.

Для $i = 1, 2, \ldots, l$ выполнить

а) $z_i = \mathrm{lsb}(x_i)$;

б) $x_{i+1} = x_i^2 \mod N$.

Выход: последовательность из z_i, $i = 1, 2, \ldots, l$.

Если $p = q = 3 \mod 4$, то в точности половина элементов Z_N^* имеет символ Якоби (+1), а другая половина имеет символ Якоби (–1).

Проблема квадратичный вычетности при входе N и $x \in Z_N^*(+1)$ заключается в решении вопроса о том, является ли x квадратичным вычетом по модулю N. Стойкость генератора квадратичных вычетов основана на сложности решения проблемы квадратичной вычетности. Заметим, что нахождение эффективной процедуры, решающей проблему квадратичной вычетности, – это открытый вопрос в математике.

Предположение о квадратичной вычетности. Для любых многочленов P и P', для любой схемы C и всех (кроме конечного числа) $n \geq 1$ доля модулей N размером n, для которых C способна определить квадратичную вычетность с вероятностью, большей чем $(1 - 1/P(n))$, ограничена сверху величиной $1/P'(n)$. Данная гипотеза истинна тогда и только тогда, когда генератор квадратичных вычетов непредсказуем (влево).

Генератор BBS. Простейший и наиболее эффективный генератор, использующий сложностно-теоретический подход, называется генератором Блюм–Блюма–Шуба (BBS). Иногда его называют генератором квадратичных вычетов по модулю n.

Определим два простых числа p и q, которые сравнимы с 3 по модулю 4. Произведение этих чисел n будет числом Блюма. Выберем другое случайное целое число x, взаимно простое с n. Вычислим начальное значение генератора

$$x_0 = x^2 \bmod n .$$

Псевдослучайным битом с номером i будет младший значащий бит x_i, где

$$x_i = x_{i-1}^2 \bmod n .$$

Заметим, что для получения i-го бита не нужно вычислять $(i - 1)$ предыдущих битов. Если известны значения p и q, можно вычислить i-й бит b_i напрямую

$$b_i - \text{младший значащий бит } x_i = x_0^{2^i \bmod((p-1)(q-1))} .$$

Пример. Выберем p = 4139, q = 1279, n = 5293781. Случайным образом выбираем начальное значение x_0 = 3161425, гамма генератора имеет вид

s = (0, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0,…).

Профиль линейной сложности последовательности имеет вид (рис. 6.1)

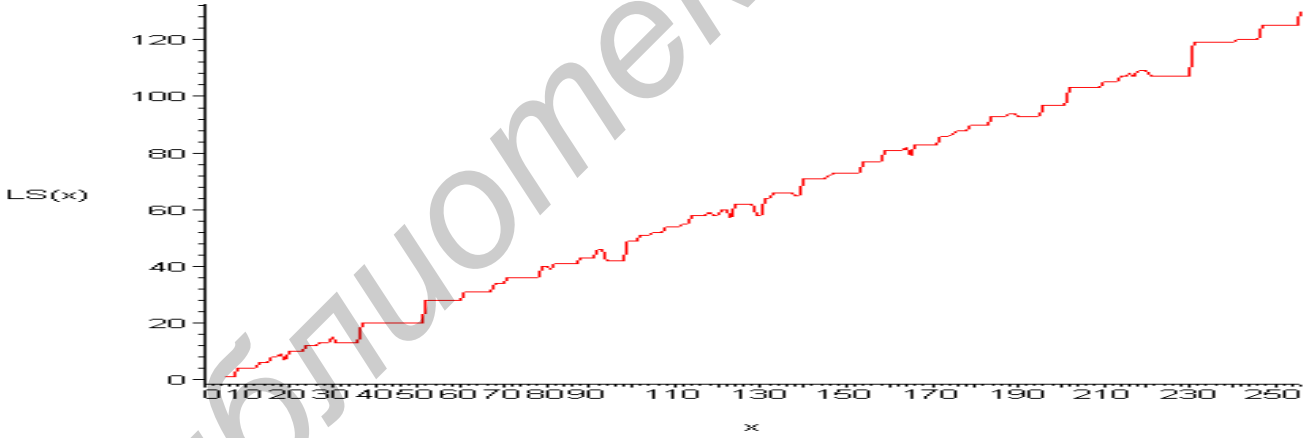

Рис. 6.1. Профиль линейной сложности генератора BBS

Рассмотренный генератор псевдослучайных чисел можно использовать в качестве элемента поточной криптосистемы с произвольным доступом. Устойчивость схемы BBS к вскрытию основана на сложности разложения n на множители и на непредсказуемости влево и вправо. Это означает, что, получив последовательность, выданную генератором, криптоаналитик не может предсказать ни следующий, ни предыдущий бит последовательности.

ЛИТЕРАТУРА

1. Фомичев В.М. Дискретная математика и криптология: Курс лекций /Под общ. ред. Н.Д. Падуфалова. – М.: ДИАЛОГ-МИФИ, 2003. – 400 с.

2. Шеннон К. Теория связи в секретных системах. Работы по теории информации и кибернетики. – М.: Иностранная лит., 1963.

3. Харин Ю.С., Берник В.И., Матвеев Г.В. Математические основы криптологии. – Мн.: БГУ, 1999.

4. Лил Р., Нидеррайтер Г. Конечные поля: В 2 т. – М.:Мир, 1988.

5. Stinson D. Cryptography: theory and practice. CRC Press, 1995.

6. Боревич З.И., Шафаревич И.Р. Теория чисел. – М.: Наука, 1985.

7. Massey J. Cryptography and system theory, Proc 24th Allerton Conf. Commun., Control, Comp., Oct. 1–3, 1986.

8. Goresky M., Klapper A. A New Class of Pseudonoise Sequences. – ISIT 2003., Yokohama, Japan, 29June-4July, 2003.

9. Klapper A., Goresky M. Algebraic feedback shift registers, Theoretical Comp. Sci. 226 (1999) P. 61 – 93.

10. Klapper A., Goresky M. Feedback shift registers, 2-adic span, and combiners with memory. J.Crypt. 10 (1997). P. 111 – 147.

11. Menezes A., Van Oorschot P., Vanstone S. Handbook of Applied Cryptography. CRC Press, 1996.

12. Kumar P.V.,Scholtz R.A. Bounds of the linear span of bent sequences. IEEE Trans. Inform. Theory, vol. IT-29, Nov, 1983.

13. Olsen J.D., Scholtz R.A.,Welch L.R. Bent-Function Sequences. IEEE Trans. Inform. Theory, vol. IT-28 № 6, Nov., 1982.

14. Gong G. Sequences Analysis. Lecture Notes for CO739x, Winter, 1999,Waterloo.

15. Иванов М.А. Криптографические методы защиты информации в компьютерных системах и сетях. – М.:КУДИЦ-ОБРАЗ, 2001. – 368 с.

16. Шнайер Б. Прикладная криптография. Протоколы, алгоритмы, исходные тексты на языке Си. – М.: ТРИУМФ, 2002. – 816 с.

17. Молдовян Н.А., Молдовян А.А., Еремеев М.А. Криптография. От примитивов к синтезу алгоритмов. – СПб.: БХВ-Петербург, 2004. – 448 с.

18. Бабаш А.В., Шанкин Г.П. Криптография / Под ред. В.П.Шерстюка и др. – М.: СОЛОН-Р, 2002. – 512 с.

19. Armknecht F., Krause M. Algebraic attacks on combiner with memory. Advances in Cryptology-Crypto'2003, Lecture Notes in Computer Science. № 2729 P. 162 – 175, Springer – Verlag, 2003.

СОДЕРЖАНИЕ

Св. план 2005, поз. 75

Учебное издание

Саломатин Сергей Борисович

ПОТОЧНЫЕ КРИПТОСИСТЕМЫ

УЧЕБНОЕ ПОСОБИЕ

по курсам

КОДИРОВАНИЕ И ЗАЩИТА ИНФОРМАЦИИ, ОСНОВЫ КРИПТОЛОГИИ

для студентов специальностей

«Радиоэлектронные системы», «Радиоэлектронная защита информации»

дневной формы обучения

Редактор Е.Н. Батурчик
Корректор Н.В. Гриневич

Подписано в печать 03.01.2006. Формат 60x84 1/16. Бумага офсетная.
Гарнитура «Таймс». Печать ризографическая. Усл. печ. л. 4,65.
Уч.-изд. л. 4,0. Тираж 150 экз. Заказ 239.

Издатель и полиграфическое исполнение: Учреждение образования
«Белорусский государственный университет информатики и радиоэлектроники»
Лицензия на осуществление издательской деятельности №02330/0056964 от 01.04.2004.
Лицензия на осуществление полиграфической деятельности №02330/0131518 от 30.04.2004.
220013, Минск, П. Бровки, 6