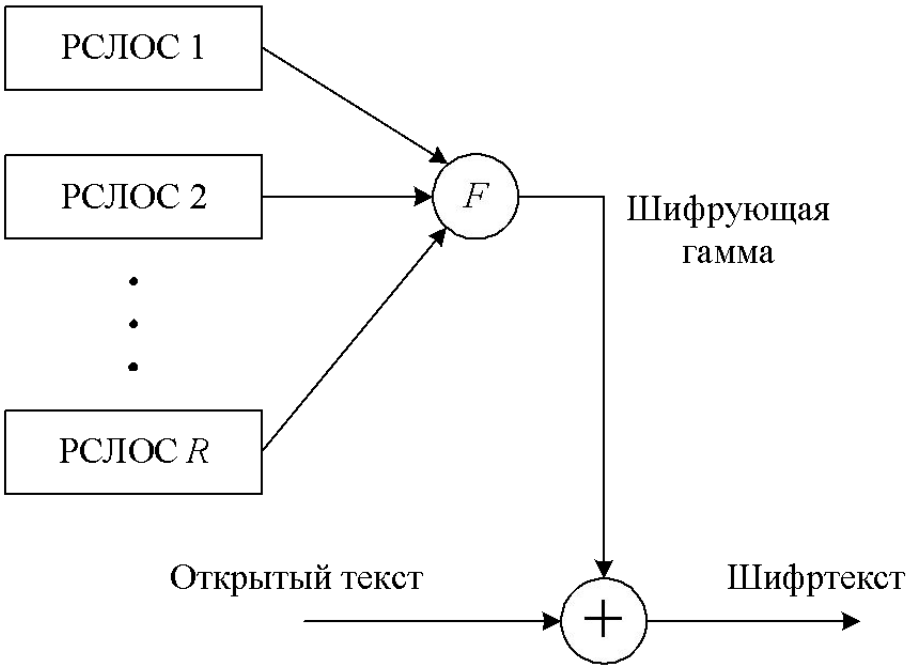
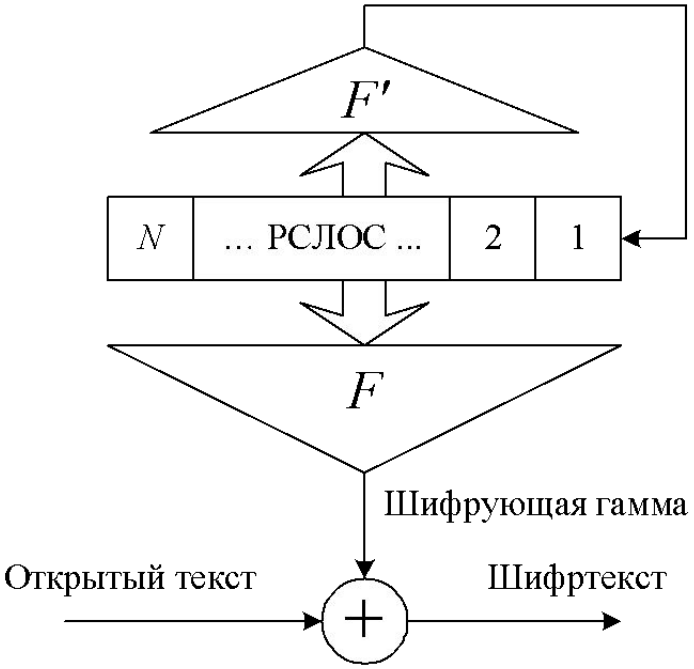
1. Основные понятия корреляционных методов криптоанализа поточных шифров

Комбинирующий генератор ШГ



Фильтрующий генератор ШГ



Нелинейные булевы функции, используемые при математическом описании комбинирующих и фильтр-генераторов, пропускают информацию о своих внутренних компонентах (входных данных) на выход (в ШГ).

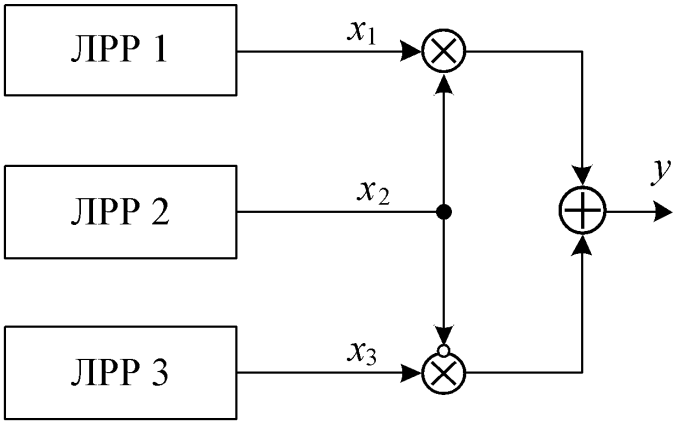
Корреляционные атаки используют корреляцию выходной последовательности схемы шифрования (ШГ) с выходной последовательностью ЛРР для восстановления их начального заполнения (вскрытия ключа).

Корреляция двух двоичных элементов x и y определяется как величина

$$R(x, y) = P(x = y) - P(x \neq y).$$

Если корреляция окажется значительно отличающейся от нуля, то вероятность успешной корреляционной атаки возрастает.

Генератор Джеффа



x_1	x_2	x_3	y
0	0	0	0
0	0	1	1
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	1

Для ЛРР 1 корреляция генератора Джеффа равна

$$R(x_{1k}, y_k) = P(x_{1k} = y_k) - P(x_{1k} \neq y_k),$$

где $P(x_{1k} = y_k) = P(x_{2k} = 1) + ½ \cdot P(x_{2k} = 0)$.

Если $P(x_2 = 1) = P(x_2 = 0) = ½$, то $P(x_{1k} = y_k) = ¾$, $P(x_{1k} \neq y_k) = ¼$.

Следовательно,

$$R(x_{1k}, y_k) = ¾ - ¼ = ½.$$

- Для выполнения корреляционного анализа генератора Джеффа поочередно перебираются ключи в ЛРР 1, и тот ключ, который дает максимальную корреляцию с выходом, принимается за истинный.
- Далее таким же образом находится ключ для ЛРР 3.
- Затем находится ключ для ЛРР 2, который даст единичную корреляцию с исходной гаммой при правильном выборе первого и второго ключей.

Таким образом, при тотальном переборе необходимо проверить

$$T_1 2^{n_1+n_2+\ldots+n_m} \text{ключей},$$

а для корреляционной атаки количество опробований будет равно

$$T_2 = 2^{n_1} + 2^{n_2} + \ldots . 2^{n_m} .$$

$T_1 >> T_2$, поэтому корреляционная атака гораздо эффективней перебора.

- ***Задание:*** *определить количество операций по подбору ключа для генератора Джеффа силовым методом и на основе корреляционной атаки.*
- *Исходные данные: ЛРР 1 – 5 разрядов;*
- *ЛРР 2 – 3 разряда;*
- *ЛРР 3 – 7 разрядов.*

Как защититься от корреляционной атаки???

Необходимо специальным образом выбрать *корреляционно нечувствительную степени l* булеву функцию $f(x_1, x_2, \ldots, x_m)$ – ненулевая корреляция существует только между гаммой и объединением не менее l выходов ЛРР.

Тогда количество опробований ключей не меньше, чем

$$T_3 = 2^{l \min n_i} >> T_2.$$

Основные классы корреляционных атак :

1) базовые корреляционные атаки:

– базовая корреляционная атака Зигенталера;

– корреляционная атака Зигенталера;

2) атаки, базирующиеся на низковесовых проверках четности:

– быстрая корреляционная атака Майера-Штаффельбаха;

– быстрая корреляционная атака Форре;

– быстрый итеративный алгоритм Михалевича-Голича;

– быстрая корреляционная атака Чепыжова-Смитса;

3) атаки, базирующиеся на использовании конволюционных кодов;

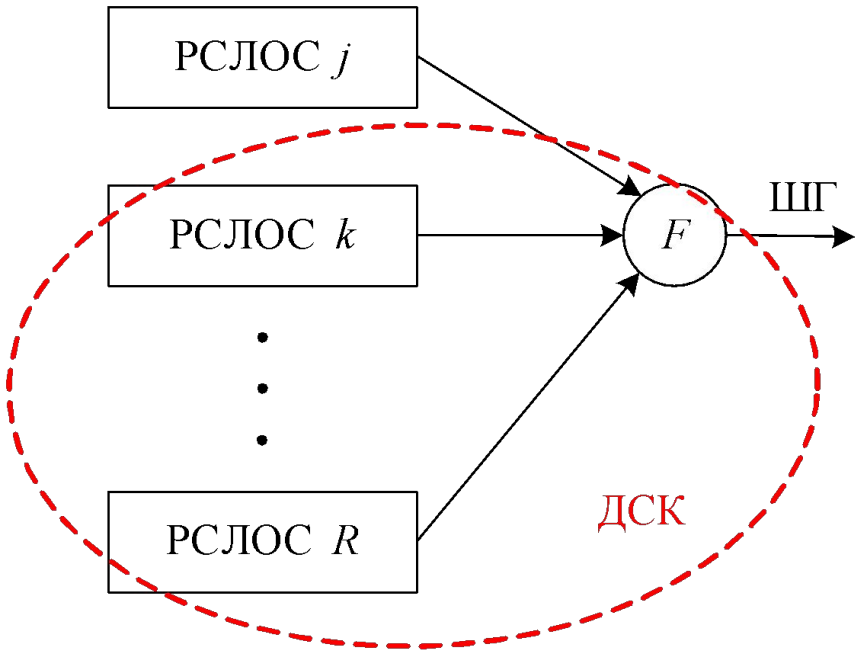
4) атаки, использующие технику турбо-кодов;

5) атаки, базирующиеся на восстановлении линейных полиномов;

6) быстрая корреляционная атака Чепыжова, Йоханссона, Смитса.

Чепыжов Владимир Викторович
1962 г.р.
д. ф.-м. н.

2. Основные корреляционные атаки



Рассмотрим ***комбинирующий генератор*** с нелинейными узлами усложнения (НУУ) F, который выдает в ШГ информацию о последовательности $a^{(j)}$, порожденной j-м регистром сдвига.

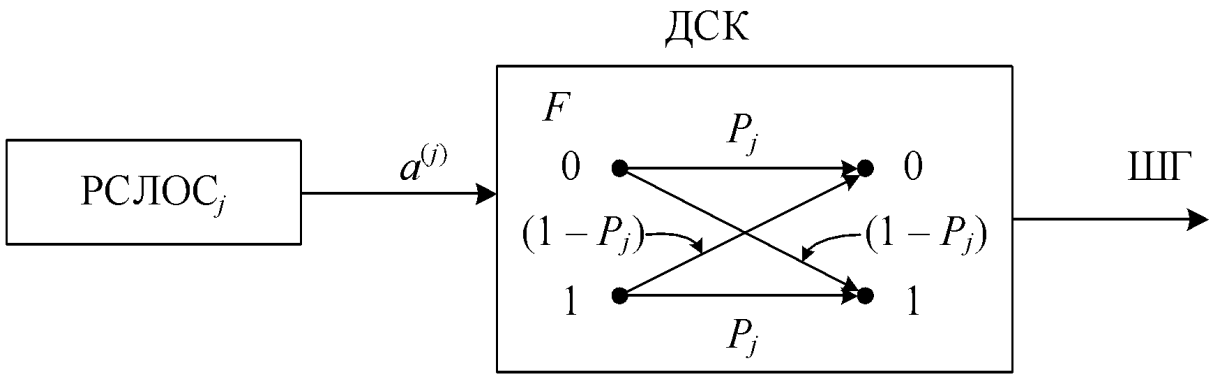
Вероятность того, что значение выходной последовательности совпадет со значением из последовательности $a^{(j)}$, порожденной j-м регистром сдвига

$P_j = P(F(A_1, A_2, \ldots, A_N) = A_j)$ – вероятность перехода (ошибка в ДСК)

Чтобы обособить эффект j-го РСЛОС на ШГ, остальная часть комбинирующего генератора моделируется как двоичный симметричный канал (ДСК) с вероятностью корреляции $P(x_i = z_i) = 1 - P_j$.

Проблема криптоанализа – проблема декодирования некоторого кода с присутствующим в ДСК сильным шумом.

Комбинирующий генератор представляется в виде модели «РСЛОС + ДСК»



ШГ рассматривается как искаженная версия последовательности регистра сдвига $a^{(j)}$.

Задача криптоанализа сокращается до нахождения верной фазы $a_0^{(j)}$ (начального заполнения регистра), исходя из фрагмента гаммы Γ_n конечной длины и избыточности, содержащейся в $a^{(j)}$ (т. е. линейных соотношений, управляющих поведением $a^{(j)}$).

Базовая корреляционная атака Зигенталера («разделяй и вскрывай»)

Криптоаналитику известно полное описание комбинирующего генератора, за исключением ключа (начального заполнения ЛРР).

Алгоритм предполагает тотальный перебор начальных состояний каждого отдельного ЛРР и оценку расстояния Хэмминга между двумя двоичными последовательностями одинаковой длины.

Вычислительная сложность полного перебора ЛРР длины n:

$$O = \prod_{i=1}^{k} 2^{n_i}$$

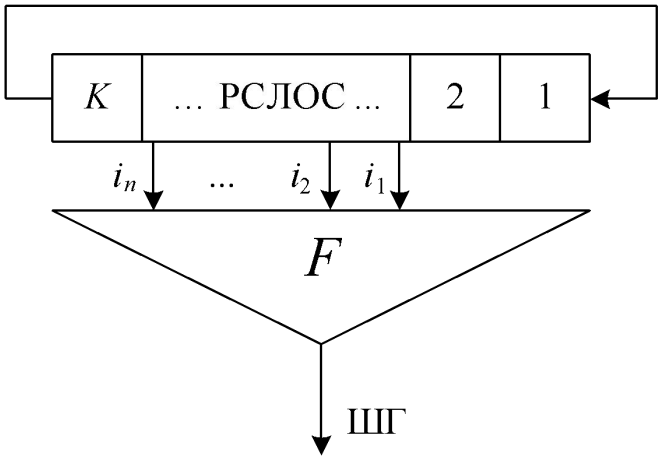
Томас Зигенталер

Вычислительная сложность атаки Зигенталера:

$$O = \sum_{i=1}^{k} 2^{n_i}$$

Такая атака возможна только для значений длин регистров $n \leq 50$.

Корреляционная атака Зигенталера



Заключается в анализе ***фильтр-генератора***, формирующего $ШГ_k$, и нахождении эквивалентной схемы, которая бы генерировала аналогичную выходную последовательность при условии, что криптоаналитику известен примитивный образующий полином.

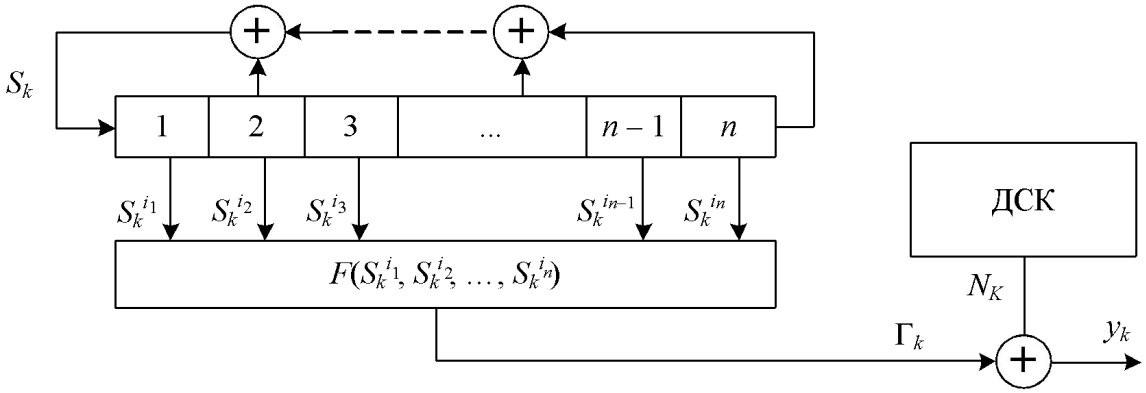
Количество точек съема n, позиций ячеек точек съема i_1, i_2, …, i_n, вид нелинейной функции F и начальное заполнение считаются неизвестными.

Данная атака позволяет представить любую криптосхему в виде ее *эквивалента* и является универсальной атакой на различные криптосистемы.

Эквивалентные схемы существуют всегда.

Построение эквивалентных схем возможно только при известном полиноме.

Схема фильтр-генератора



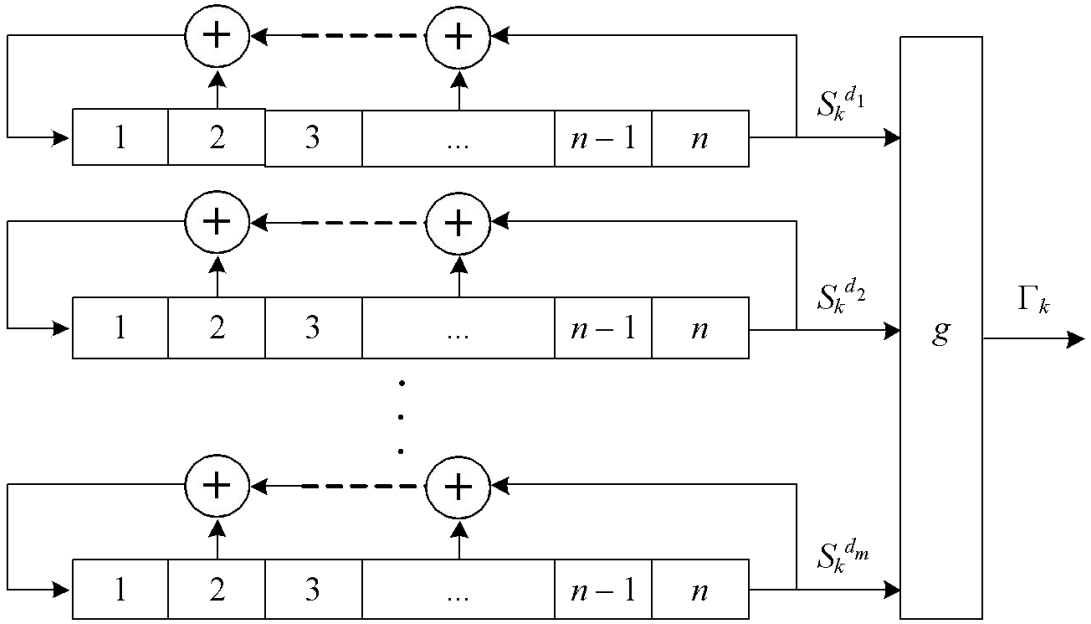
Эквивалентная схема будет содержать m ЛРР, построенных согласно известному полиному, с различными начальными состояниями ($1 \le m \le n$).

n фаз, снимаемых функцией F, загружаются в отдельные регистры.

Полагается, что $g = F$.

При анализе схемы рассматривается функция кросс-корреляции между последовательностями S_k и Γ_k.

Эквивалент схемы фильтр-генератора



Атака применима для значений длин ЛРР не более 50.

Быстрая корреляционная атака

Быстрые корреляционные атаки – атаки, вычислительная сложность которых значительно меньше сложности силовых атак.

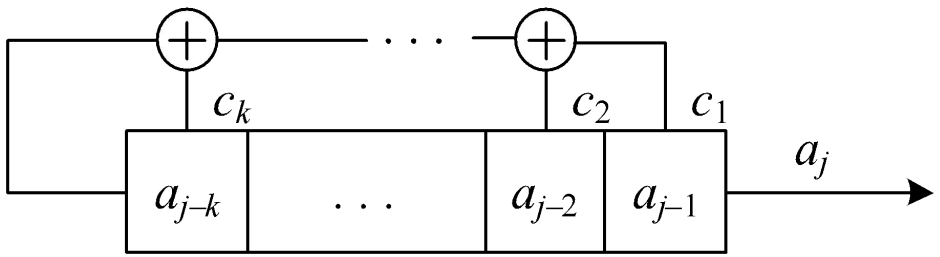
- *Условие применимости*: количество точек съема ЛРР невелико ($t \leq 10$).
- *Применимость*: комбинирующие и фильтр-генераторы.
- *Основа атаки*: использование линейных соотношений – *уравнений проверки четности* для полинома обратных связей.

Атака Майера-Штаффельбаха – базовая для всех быстрых корреляционных атак.

Пусть последовательность a_n порождается РСЛОС, имеющим t точек ***обратной связи***, и примитивным многочленом $p(x)$ степени k:

$$p(x) = c_0 + c_1 \cdot x + c_2 \cdot x^2 + \ldots + c_k \cdot x^k,$$

где $c_0 = 1$ и $c_1, c_2, \ldots, c_k \in \{0,1\}$.



Линейное соотношение можно переписать как *уравнение проверки четности*, состоящее из $(k + 1)$ членов РСЛОС-последовательности a_j:

$$L = a_0 + a_1 + a_2 + \ldots + a_k = 0$$

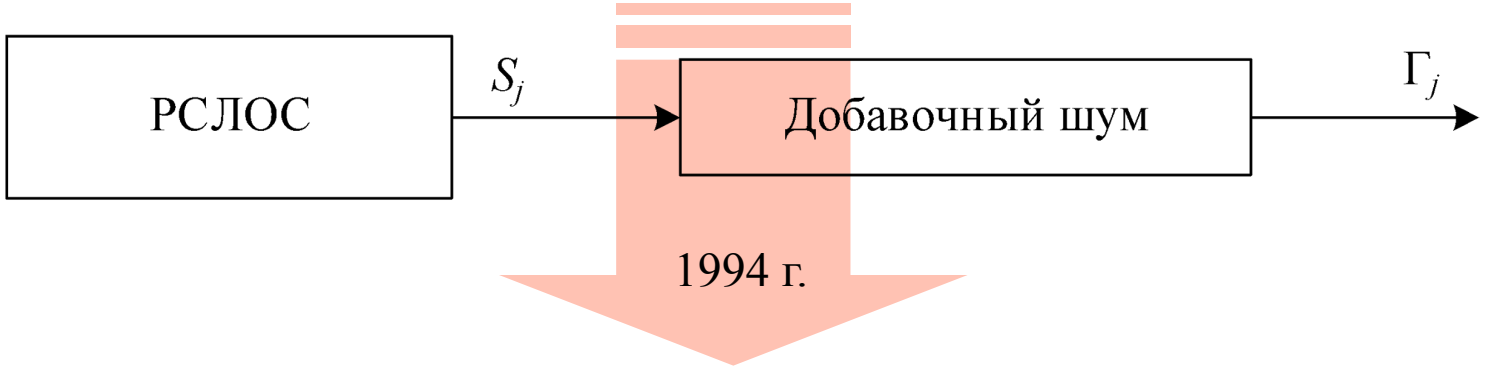
где члены a_i – значения в ячейке с отводом обратной связи.

Сущность быстрой корреляционной атаки:

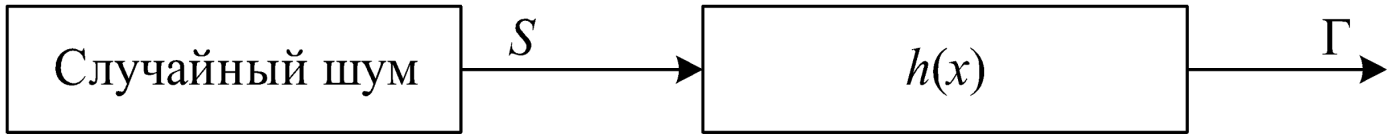
Поиск начального состояния ЛРР осуществляется методом перебора, но не из всех возможных вариантов. Для анализа будут использоваться фазы, значения уравнения проверки на четность которых совпадают со значениями уравнения проверки на четность шифргаммы.

3. Оптимальная корреляционная атака Андерсона

Стандартная модель для корреляционной атаки Зигенталера



Модель Андерсона



Росс Андерсон
1956 г. р.

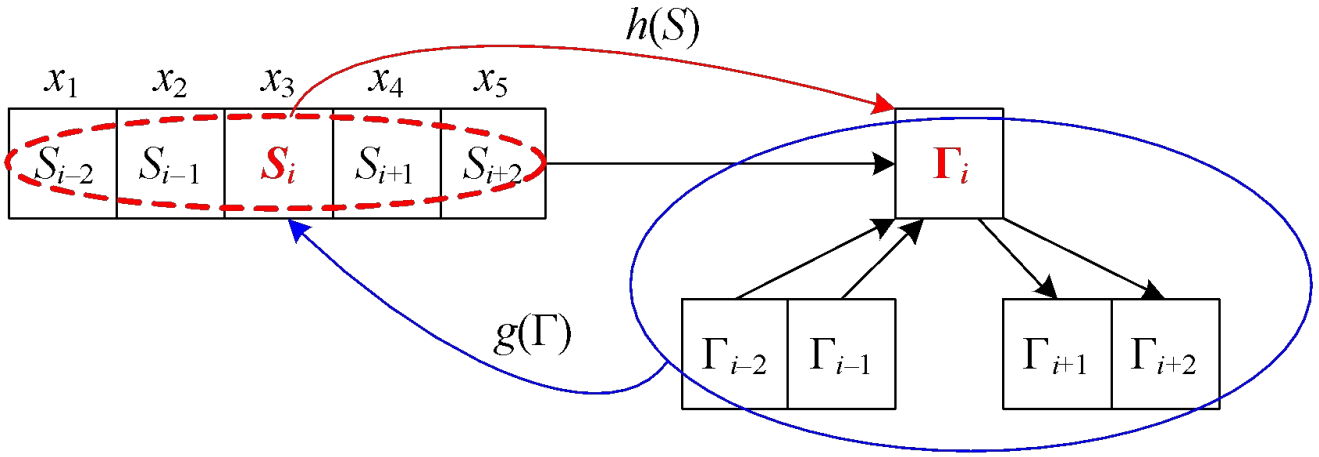
Основная цель модели Андерсона – определить, сколько информации о некотором произвольном сигнале S просачивается через нелинейную комбинирующую функцию $h(x)$ в гамму Г.

Если $\Gamma_i = S_i + S_{i+1}$, то знание Γ_i ничего не говорит о S_i.
Но если $\Gamma_i = S_i \cdot S_{i+1}$ всякий раз, когда $\Gamma_i = 1$, и $S_i = 1$.
При атаке на фильтр-генератор всегда можно снять влияние линейной функции путем перехода к другой фазе исходного ЛРРС.

Пример нелинейной комбинирующей функции h:

$$h(x_1, x_2, x_3, x_4, x_5) = x_1 + x_2 + (x_1 + x_3)\cdot(x_2 + x_4 + x_5) + (x_1 + x_4)\cdot(x_2 + x_3)\cdot x_5.$$

Данная функция сбалансированная и корреляционно иммунная 2 порядка.

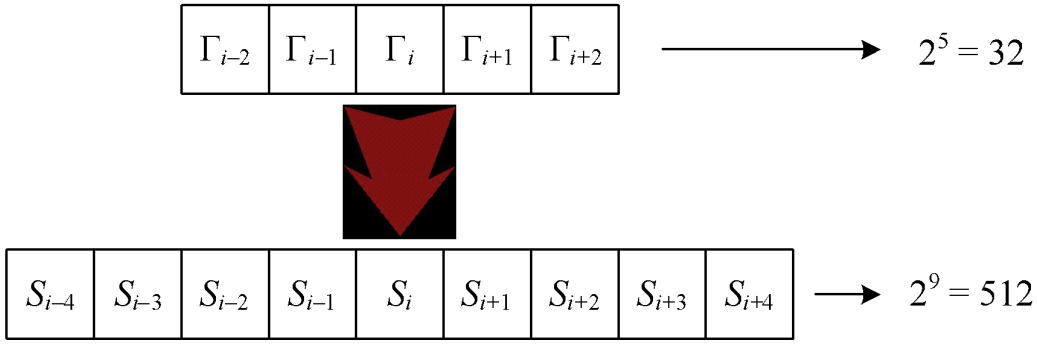


Если $\Gamma_i = h(S_{i-2}, S_{i-1}, S_i, S_{i+1}, S_{i+2})$, то биты гаммы $\Gamma_{i-2}\ldots\Gamma_{i+2}$ зависят от Γ_i. При аппроксимации $S_i = g(\Gamma_{i-2}, \Gamma_{i-1}, \Gamma_i, \Gamma_{i+1}, \Gamma_{i+2})$.

Однако биты $\Gamma_{i-2}\ldots\Gamma_{i+2}$ зависят от девяти бит $S_{i-4}\ldots S_{i+4}$.

Поэтому анализируется влияние входных 9-грамм на любой 5-битовый фрагмент выходной гаммы Г.

Функция, переводящая 9 бит в 5, называется пополненной функцией $\overline{h}$

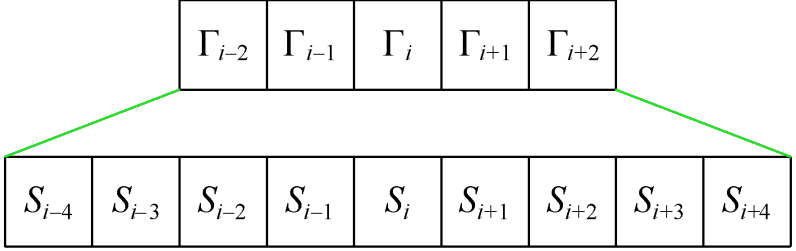


Для Г = 26:

0	0	1	0	1	0	**1**	0	1
0	0	1	1	1	0	0	0	1
0	0	1	1	1	0	0	1	0
1	0	0	1	1	0	0	0	1
1	0	0	1	1	0	0	1	0
1	0	1	0	**0**	**1**	0	1	1
1	0	1	1	1	0	0	0	1
1	0	1	1	1	0	0	1	0
1	1	0	1	1	0	0	0	1
1	1	0	1	1	0	0	1	0

Если 5 бит гаммы равны «1 1 0 1 0» (26), то $S_i = 1$, $S_{i+1} = 0$, $S_{i+2} = 0$ с вероятностью 0,9.

Значения гаммы	Количество исходных состояний	Значения гаммы	Количество исходных состояний
0	18	16	16
1	16	17	18
2	14	18	16
3	20	19	18
4	16	20	12
5	14	21	**10**
6	21	22	15
7	17	23	15
8	11	24	23
9	17	25	17
10	12	26	**10**
11	12	27	18
12	23	28	17
13	13	29	15
14	13	30	19
15	19	31	17



Для Γ = 26 = (1 1 0 1 0):

0	0	1	0	1	0	1	0	1
0	0	1	1	1	0	0	0	1
0	0	1	1	1	0	0	1	0
1	0	0	1	1	0	0	0	1
1	0	0	1	1	0	0	1	0
1	0	1	0	0	1	0	1	1
1	0	1	1	1	0	0	0	1
1	0	1	1	1	0	0	1	0
1	1	0	1	1	0	0	0	1
1	1	0	1	1	0	0	1	0

Столбцы 5 и 6 являются инвертированными копиями друг друга. Значит

$$S_i = 1 + S_{i+1}$$

Если 5 бит гаммы равны «0 1 0 0 1» (9), то

$S_{i-1} = 0$ с вероятностью 1

(т.к. все 17 состояний имеют «0» в 4-м бите).

Значения гаммы	Количество исходных состояний
0	18
1	16
2	14
3	20
4	16
5	14
6	21
7	17
8	11
9	**17**
10	12
11	12
12	23
13	13
14	13
15	19

Значения гаммы	Количество исходных состояний
16	16
17	18
18	16
19	18
20	12
21	**10**
22	15
23	15
24	23
25	17
26	**10**
27	18
28	17
29	15
30	19
31	17