УДК 519.7 DOI 10.17223/20710410/58/6

ОБ ОДНОМ МЕТОДЕ ПОСТРОЕНИЯ ОДНОРОДНЫХ ПЛОСКОСТНЫХ АППРОКСИМАЦИЙ ФИЛЬТРУЮЩЕГО ГЕНЕРАТОРА

Л. А. Кущинская

Московский государственный университет им. М. В. Ломоносова, г. Москва, Россия

E-mail: lyudmila.kuschinskaja@yandex.ru

Плоскостные аппроксимации фильтрующих генераторов могут быть использованы для восстановления его начального состояния по отрезку выходной последовательности. Представлены результаты исследования одного метода построения плоскостных аппроксимаций специального вида.

Ключевые слова: *криптоанализ, восстановление ключа, фильтрующий генератор, плоскостная аппроксимация.*

THE METHOD FOR CONSTRUCTING UNIFORM PLANAR APPROXIMATIONS OF THE FILTER GENERATOR

L. A. Kuschinskaya

Lomonosov Moscow State University, Moscow, Russia

We study the possibility of constructing an approximation of filter generator to restore initial state $u^* \in V_n$ from its output sequence $z_i = f(A^i(u^*)) \in \{0,1\}$, $i = 0, \ldots, N-1$, where $A : V_n \to V_n$ is non-degenerate linear mapping, f is balanced Boolean function. The triple $(m, L_0, \mathbb{T})$ is a key element in the construction of the approximation, where $m \in \mathbb{N}$, L_0 is coset of the space V_n, $\mathbb{T} = (t_0, t_1, \ldots, t_m)$, $t_0 = 0$, $t_0 \leqslant t_1 < \ldots < t_m$. Let $(m, L_0, \mathbb{T})$ be a triple and for $b_1, \ldots, b_m \in \{0,1\}$ the probability that $f(v) = b_i$ is greater than $1/2$ for a random equiprobable choice of a vector v from $L_i = A^{t_i - t_{i-1}}(L_{i-1})$, $i = 1, \ldots, m$. Then a finite number of such triples with pairwise distinct sets L_0 makes it possible to restore the key with a complexity that is much less than the complexity of enumerating keys in some cases. In this paper, we study the possibility of constructing approximations of a special form, where all cosets L_0 have the same dimension, their union is equal to V_n, and the values of m are the same for all described triples. Expressions for the optimal values of the parameters k and δ are obtained for some enumeration method for constructing the approximations. It is shown that for $k = \left\lceil \log_2 \left((Q - \sqrt{Q^2 - \pi_0 2^{n+2}})/2\pi_0 \right) \right\rceil$ and $\delta \approx \lceil t_0 \sqrt{\Omega} \rceil$ it is possible to achieve the minimum length of the generator output sequence required to construct such approximations for a given value of the upper complexity Q and lower reliability π_0 of the initial state recovery method, where $\Omega = 2^k$, $t_0 \approx 1.19061$.

Keywords: *cryptanalysis, key recovery, filter generator, planar approximation.*

Введение

Классическими структурными элементами потоковых шифров являются фильтрующий и комбинирующий генераторы. Эти генераторы нашли широкое применение в современных потоковых шифрах, используемых на практике, например E0 [1],

A5/1 [2], SNOW 2.0 [3], Grain [4]. Анализ криптографических свойств подобных конструкций представляет существенный интерес [5 – 9].

В 2017 г. в работе [10] предложен метод восстановления ключа фильтрующего генератора, основанный на понятии плоскостной аппроксимации функции усложнения. Свойства используемой плоскостной аппроксимации напрямую влияют на численные характеристики метода и позволяют в ряде случаев восстанавливать ключ с трудоёмкостью, существенно меньшей трудоёмкости полного опробования ключей. В [10] описан только этап восстановления ключа в предположении, что необходимая для этого плоскостная аппроксимация уже построена. Вопрос построения плоскостных аппроксимаций ранее не исследовался.

В данной работе исследуется вопрос построения плоскостных аппроксимаций специального вида, которые названы *однородными*. Плоскостная аппроксимация определяет три основных численных характеристики метода восстановления ключа: трудоёмкость, надёжность и необходимое количество известных битов выходной последовательности генератора. Рассматривается задача построения однородной плоскостной аппроксимации, при которой метод восстановления ключа достигает заданных значений трудоёмкости и надёжности при минимально возможном требуемом количестве известных битов выходной последовательности. Предложена математическая модель работы переборного метода решения задачи, в рамках которой получены оптимальные значения его параметров.

1. Основные определения и обозначения

Пусть $\mathbb{F}_2$ — поле из 2 элементов, $V_n = \mathbb{F}_2^n$ — линейное пространство размерности n над полем $\mathbb{F}_2$. *Носителем вектора* $x = (x_0, \ldots, x_{n-1}) \in V_n$ называется множество $\operatorname{supp}(x) = \{i \in \{0, \ldots, n-1\} : x_i = 1\}$. Тот факт, что $L \subseteq V_n$ является подпространством пространства V_n, будем обозначать так: $L < V_n$; линейную оболочку векторов $v^{(1)}, \ldots, v^{(k)}$ из V_n обозначим $L(v^{(1)}, \ldots, v^{(k)})$ [11]. *Плоскостью* в пространстве V_n будем называть смежный класс по подпространству этого пространства, а её размерностью — размерность этого подпространства.

Булевой функцией f от n переменных называется отображение $f : V_n \to \mathbb{F}_2$. Множество всех булевых функций от n переменных будем обозначать через $\mathcal{F}_n$. Носителем функции $f \in \mathcal{F}_n$ называется множество $1_f = \{x \in V_n : f(x) = 1\}$. *Весом* $\operatorname{wt}(f)$ булевой функции $f \in \mathcal{F}_n$ называется мощность её носителя.

Через $\mathbb{N}_0$ обозначим множество $\mathbb{N} \cup \{0\}$. Под *фильтрующим генератором* будем понимать отображение из $\mathbb{N}_0 \times V_n$ в $\mathbb{F}_2$, определяющееся невырожденным линейным отображением $A : V_n \to V_n$ и уравновешенной булевой функцией $f \in \mathcal{F}_n$, которое ставит в соответствие числу i и вектору $u^* \in V_n$ бит $z_i = f(A^i(u^*))$. Вектор u^* будем называть *ключом* или *начальным заполнением* фильтрующего генератора, а последовательность битов $z_0, z_1, \ldots$ — его *выходной последовательностью*. Результатом зашифрования открытого текста $x \in V_N$ на ключе $u^* \in V_n$ с помощью потокового шифра, построенного на основе фильтрующего генератора, является вектор $y \in V_N$, такой, что $y_i = x_i \oplus z_i$ для любого $i \in \{0, \ldots, N-1\}$. Другими словами, $y = x \oplus z$, где $z = (z_0, z_1, \ldots, z_{N-1}) \in V_N$ — начальный отрезок длины N выходной последовательности фильтрующего генератора.

Определение 1 [12]. Пусть имеется случайная выборка из n элементов без возвращения из конечной совокупности мощности N, при этом ровно D из них обладают заданным свойством. Пусть случайная величина x — число элементов из выборки, обладающих заданным свойством. Говорят, что случайная величина x имеет *гипергео-*

метрическое распределение с параметрами N, D, n ($x \sim HG(D, N, n)$), если справедливо соотношение

$$\mathsf{P}[x = k] = \binom{D}{k}\binom{N-D}{n-k}\Big/\binom{N}{n}.$$

2. Предварительные сведения

Метод из работы [10], как было отмечено, позволяет восстанавливать ключ фильтрующего генератора тем эффективнее, чем более «точная» плоскостная аппроксимация им используется. В настоящем разделе вводится понятие однородной плоскостной аппроксимации (п. 2.1), кратко описывается метод восстановления ключа на её основе (п. 2.2), приводятся соотношения между параметрами плоскостной аппроксимации и основными численными характеристиками метода (п. 2.3).

2.1. Плоскостные аппроксимации

Для дальнейшего изложения приведём необходимые определения основных конструкций, введённых в [10].

Всюду далее A — линейное отображение из V_n в V_n, а f — функция из $\mathcal{F}_n$.

Пусть $m \in \mathbb{N}$, $\mathbb{L} = (L_0, \ldots, L_m)$, где все L_i являются плоскостями в V_n, и $\mathbb{T} = (t_0, t_1, \ldots, t_m) \in \mathbb{Z}^{m+1}$, где $t_0 = 0$, $t_0 \leqslant t_1 < \ldots < t_m$. Тройка $\mathrm{Traj} = (m, \mathbb{L}, \mathbb{T})$ называется *траекторией* для A, если выполнены соотношения $L_i = A^{t_i - t_{i-1}}(L_{i-1})$, $i = 1, \ldots, m$. При этом m называется *длиной траектории*, а L_0 — *начальной плоскостью*.

Пусть $\mathrm{Traj} = (m, \mathbb{L}, \mathbb{T})$ — траектория для некоторого линейного преобразования. Пусть также $\mathbb{B} = (b_1, \ldots, b_m)$, где все $b_i \in \mathbb{F}_2$, и $\mathbb{P} = (p_1, \ldots, p_m)$, где все $p_i \in [0; 1]$. Пару $(\mathbb{B}, \mathbb{P})$ будем называть *характеристикой* траектории Traj относительно функции $f \in \mathcal{F}_n$, если p_i — вероятность того, что при случайном равновероятном выборе вектора v из L_i значение $f(v)$ совпадает с константой b_i.

Отметим, что каждой траектории длины m соответствует 2^m характеристик, среди которых существует не более одной характеристики, у которой $p_i > 1/2$ для всех i. Характеристику, обладающую таким свойством, будем называть *положительной* относительно f. Траектория, для которой существует положительная характеристика, будем называть *подходящей траекторией*.

В целях визуализации понятия подходящей траектории на рис. 1 для $n = 6$ схематично изображена траектория длины 3. Включение векторов в плоскость носит условный характер, так как на практике плоскость может состоять из нескольких несвязанных областей из векторов. Плоскости пространства V_6 условно заключены в рамку; плоскости, вошедшие в траекторию, выделены отдельно.

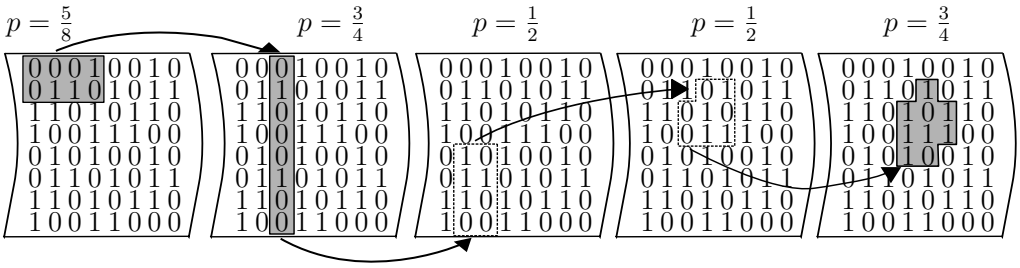

Рис. 1. Траектория $\mathrm{Traj} = (3, \mathbb{L}, \mathbb{T})$, где $\mathbb{T} = (0, 0, 1, 4)$, характеристика ($\mathbb{B} = (0, 0, 1)$, $\mathbb{P} = (5/8, 3/4, 3/4)$)

Конечный набор траекторий $\mathrm{Traj}^{(1)}, \ldots, \mathrm{Traj}^{(s)}$ для A, являющихся подходящими относительно функции f, с попарно различными начальными плоскостями будем называть *плоскостной аппроксимацией* функции f относительно отображения A.

Для плоскостной аппроксимации через $\mathbb{L}_{\mathrm{start}}$ будем обозначать множество, состоящее из начальных плоскостей входящих в аппроксимацию траекторий, т. е. $\mathbb{L}_{\mathrm{start}} = \{L_0^{(1)}, \ldots, L_0^{(s)}\}$.

Определение 2. *Однородной плоскостной аппроксимацией* будем называть плоскостную аппроксимацию, удовлетворяющую следующим условиям:

1) начальные плоскости аппроксимации являются смежными классами по одному и тому же подпространству и их объединение совпадает с ключевым пространством V_n, т. е. $\bigcup\limits_{L \in \mathbb{L}_{\mathrm{start}}} L = V_n$, и $\dim(L) = k$ для любого $L \in \mathbb{L}_{\mathrm{start}}$;
2) длины m_i всех траекторий равны между собой.

Всюду далее речь идёт именно об этом типе плоскостных аппроксимаций, которую для краткости будем называть однородной аппроксимацией.

2.2. Восстановление ключа с помощью однородной плоскостной аппроксимации

Под задачей восстановления ключа фильтрующего генератора по известному отрезку его выходной последовательности длины N будем понимать задачу поиска по известным $z_0, \ldots, z_{N-1}$, где $z_i \in \{0, 1\}$, такого значения $u^* \in V_n$, что

$$f(A^i(u^*)) = z_i, \;\; i = 0, \ldots, N-1.$$

Метод восстановления ключа состоит из двух этапов. На первом этапе последовательно просматриваются траектории плоскостной аппроксимации и выбираются те из них, в начальной плоскости которых вероятнее всего находится искомый ключ. Для проверки одной траектории $\mathrm{Traj} = (m, \mathbb{L}, \mathbb{T})$ необходимо сравнить биты характеристики траектории $b_1, \ldots, b_m$ и биты выходной последовательности генератора $z_{t_1}, \ldots, z_{t_m}$, где $b_i \in \mathbb{B}$ и $t_i \in \mathbb{T}$ для $i = 1, \ldots, m$. Всюду далее будем полагать, что для любой траектории из плоскостной аппроксимации выполняется неравенство $t_m < N$.

Пусть p — минимальное значение p_i среди вероятностей характеристик всех траекторий. При расчете характеристик метода будем заменять вероятности p_i на p. Отметим, что оценки, полученные в рамках такого допущения, хотя и будут более грубыми, но останутся справедливыми. Проверка принадлежности ключа начальной плоскости траектории сводится к различению двух распределений Бернулли по выборке объёма m. Благодаря тому, что все траектории имеют одинаковую длину m и одинаковые вероятности p появления преобладающих значений на плоскостях, на первом этапе достаточно построить единое для всех траекторий решающее правило, отвечающее за принятие начальной плоскости траектории в качестве возможного расположения искомого ключа генератора. В силу введённых в [10] статистических гипотез данному решающему правилу соответствуют ошибки первого и второго рода α, β, где α — вероятность принять ложную траекторию, начальной плоскости которой ключ не принадлежит; β — вероятность отклонить истинную траекторию.

Второй этап метода заключается в полном опробовании векторов из начальных плоскостей тех траекторий, что успешно прошли первый этап. Для этого необходимо для каждого вектора плоскости построить отрезок выходной последовательности генератора длины N. При точном совпадении построенного и известного отрезков выходной последовательности рассматриваемый ключ объявляется искомым. При отсутствии такого совпадения метод заканчивает свою работу, не найдя ключа.

Как и в работе [10], будем предполагать, что основные параметры метода восстановления ключа далеки от своих крайних значений (например, значение параметра k не приближается к 0 или к n), так как именно в этом предположении предлагаемый метод имеет существенное преимущество относительно полного опробования ключевого пространства.

2.3. Характеристики метода восстановления ключа

Трудоёмкость

Оценка средней трудоёмкости метода восстановления ключа для общего случая приведена в работе [10]. Для получения значения средней трудоёмкости для случая однородной аппроксимации воспользуемся следствием 1 из упомянутой работы.

Следствие 1 [10]. Пусть s — количество траекторий в плоскостной аппроксимации, $C_\alpha = \sum\limits_{j=1}^{s} |L_0^{(j)}|\alpha_j$. Обозначим через M список, состоящий из векторов, которые не принадлежат ни одной начальной плоскости: $M = V_n \setminus \bigcup\limits_{L\in\mathbb{L}_{\text{start}}} L$. Если плоскости из $\mathbb{L}_{\text{start}}$ не имеют пересечений, то справедливо следующее соотношение:

$$\mathrm{E}D(u,\gamma_u) = s + C_\alpha + \frac{|M|}{2^n}\left(|M| + \sum_{i=1}^{s}|L_0^{(i)}|\beta_i\right) + \frac{1}{2^n}\sum_{i=1}^{s}|L_0^{(i)}|^2(1-\alpha_i-\beta_i).$$

Для случая однородной аппроксимации формула имеет следующий вид:

$$\mathrm{E}D(u,\gamma_u) = 2^{n-k} + 2^{n-k}|L|\,\alpha + \frac{1}{2^n}|L|^2(1-\alpha-\beta) = 2^{n-k} + 2^n\alpha + 2^k(1-\alpha-\beta).$$

Будем рассматривать случаи, когда $\alpha \ll \beta$ (наиболее практически значимые). Таким образом, будем пренебрегать α в последнем слагаемом:

$$\mathrm{E}D(u,\gamma_u) = 2^{n-k} + 2^n\alpha + 2^k(1-\beta).$$

Надёжность

Надёжность метода в общем случае описывается формулой

$$\pi \geqslant 1 - \frac{1}{2^n}\sum_{i=1}^{s}|L_0^{(i)}|\,\beta_i,$$

где s — число траекторий в плоскостной аппроксимации.

В случае однородной аппроксимации $s = 2^{n-k}$, $\beta_i = \beta$ для всех i, поэтому данная оценка преобразуется следующим образом:

$$\pi \geqslant 1 - \frac{s}{2^n}|L|\,\beta = 1 - \beta.$$

3. Построение плоскостных аппроксимаций

В данном разделе описан переборный метод построения однородной аппроксимации фильтрующего генератора (п. 3.1) и предложена модель на основе случайных множеств (п. 3.2), в которой получены оценки его характеристик (п. 3.3). Приведено экспериментальное подтверждение релевантности предложенной модели (п. 3.4).

Пусть необходимо построить такую однородную аппроксимацию фильтрующего генератора, чтобы метод восстановления ключа на её основе имел трудоёмкость не больше Q, надёжность не меньше π_0 и требовал для своей работы минимально возможный объём выходной последовательности генератора.

3.1. Описание метода

Каждая траектория в однородной аппроксимации обладает следующими параметрами:

1) $k \in \{0, 1, \ldots, n\}$ — размерность плоскости в траектории, мощность плоскости будем обозначать через $\Omega = 2^k$;
2) m — длина траектории;
3) p — нижняя граница доли преобладающих значений функции на плоскости.

При построении траектории будем использовать параметр $\delta \in \{0, 1, \ldots, \Omega\}$, определяющий минимально приемлемое преобладание той или иной константы на плоскости. Пусть $T_0 = \Omega/2 - \delta/2$, $T_1 = \Omega/2 + \delta/2$, S — вес функции f на плоскости, входящей в однородную аппроксимацию. Тогда либо $S < T_0$ (плоскость содержит достаточное количество нулей функции), либо $S > T_1$ (плоскость содержит достаточное количество единиц функции). Таким образом, вероятность появления преобладающего значения функции на плоскости выражается через параметр δ следующим образом: $p = \dfrac{1}{2} + \dfrac{\delta}{2\Omega}$.

Метод построения однородной аппроксимации с параметрами k, m и p описывается следующим образом. Фиксируем произвольным образом подпространство L размерности k. Для каждого из 2^{n-k} смежных классов строится своя траектория. Для конкретного смежного класса, обозначаемого через L_0, выполняем следующие действия для каждого $i = 0, 1, 2, \ldots$:

1) $L_i = A^i(L_0)$;
2) обозначим через S_i вес функции f на плоскости L_i. Тогда если $|S_i - \Omega/2| > \delta/2$, то добавляем плоскость в строящуюся траекторию с соответствующей преобладающей константой; иначе переходим к п. 1;
3) если длина траектории равна m, то закончить построение траектории для текущего смежного класса.

Число проверок веса функции на плоскости при построении одной траектории определяет длину выходной последовательности генератора, необходимую для проверки принадлежности ключа начальной плоскости этой траектории на первом этапе работы алгоритма восстановления ключа. Таким образом, чем больше максимальное число проверок веса функции на плоскости среди траекторий однородной аппроксимации, тем больше и целевой параметр — объём выходной последовательности, требуемый для работы метода.

3.2. Математическая модель

При оценке характеристик метода будем считать, что все плоскости L_i, участвующие в проверке веса функции, являются множествами мощности 2^k, выбранными из множества всех подмножеств V_n случайно и равновероятно. Допустимость данного предположения обосновывается точностью получаемых с её помощью оценок параметров метода (см. п. 3.4).

Обозначим через $\widetilde{p}(\delta, k)$ вероятность принять случайное множество мощности 2^k в траекторию. Тогда $\widetilde{N}(\delta, k) = \dfrac{m}{\widetilde{p}(\delta, k)}$ — среднее количество шагов метода, которое необходимо выполнить для построения траектории длины m. Отметим, что величина $\widetilde{N}(\delta, k)$ равна среднему объёму выходной последовательности генератора, необходимому для восстановления ключа с заданными параметрами; именно её требуется минимизировать, согласно постановке задачи.

Перед началом работы метода построения однородной аппроксимации должны быть зафиксированы параметры k и δ.

Отметим, что стремление параметра k к нулю означает уменьшение мощности Ω плоскостей в траектории, а значит, приближение метода восстановления ключа к полному перебору векторов ключевого пространства на первом этапе работы алгоритма. При $k \to n$, то есть при увеличении мощности плоскости в траектории, трудоёмкость второго этапа алгоритма приближается к трудоёмкости полного перебора.

Рассмотрим подробнее влияние параметра δ на характеристики метода восстановления ключа при фиксированном значении k. Согласно математической модели, используемой для расчёта характеристик в [10], для принятия решения о возможном расположении ключа генератора в начальной плоскости траектории необходимо различать две простые статистические гипотезы по выборке объёма m. Для однородной аппроксимации эта задача сводится к построению оптимального критерия в схеме Бернулли с вероятностями $p_0 = \dfrac{1}{2}$ и $p_1 = \dfrac{1}{2} + \dfrac{\delta}{2\Omega}$. Чем больше δ, тем меньше объём выборки, необходимый для различения двух гипотез с заданными вероятностями ошибок первого и второго рода.

Таким образом, для фиксированного k при $\delta \to 0$ вероятность принятия случайной плоскости в траекторию увеличивается, но растёт и значение длины траектории m, требуемой для достижения заданной надёжности метода восстановления ключа. При $\delta \to \Omega$ длина траектории m уменьшается, но уменьшается и вероятность принять плоскость в траекторию, так как растёт требуемое преобладание на плоскости.

Возникает следующая задача: найти такие значения параметров δ и k, при которых величина $\widetilde{N}(\delta, k)$ минимальна для заданных ограничений трудоёмкости и надёжности метода восстановления ключа.

3.3. Характеристики метода

Математическая формулировка задачи минимизации

Как отмечено ранее, среднее количество проверок веса функции на плоскости, осуществляемых при построении траекторий, входящих в однородную аппроксимацию, равно $\widetilde{N}(\delta, k) = \dfrac{m}{\widetilde{p}(\delta, k)}$. Оценим числитель и знаменатель этой дроби.

Для оценки m найдём сначала вероятности ошибок первого и второго рода, возникающих при восстановлении ключа на основе однородной аппроксимации, построенной описанным методом с параметрами k и δ. Напомним, что для случая однородной аппроксимации эти вероятности одинаковы для всех траекторий:

- $\beta = 1 - \pi_0$ — вероятность ошибки второго рода;
- так как $Q = 2^{n-k} + \alpha\, 2^n + (1 - \beta)\, 2^k$, то вероятность ошибки первого рода $\alpha = 2^{-n}\left(Q - 2^{n-k} - \pi_0\, 2^k\right)$.

Получим выражение для длины траектории m, необходимой для работы метода восстановления ключа с заданными вероятностями ошибок первого и второго рода. На первом этапе метода для каждой траектории в однородной аппроксимации строится вектор $w = (c_1 \oplus \widetilde{z}_1, \ldots, c_m \oplus \widetilde{z}_m)$, $\widetilde{z}_i = z_{t_i}$, $i = 1, \ldots, m$. При этом если в начальном множестве данной траектории находится ключ, то $\mathsf{P}[w_i = 0] \geqslant q_1 = \dfrac{1}{2} + \dfrac{\delta}{2\Omega}$, в противном случае $\mathsf{P}[w_i = 0] = q_0 = \dfrac{1}{2}$. Для статистического различения двух распределений Бернулли с вероятностями успеха q_0 и q_1, вероятностями ошибок первого и второго

рода α и β соответственно потребуется объём материала [9]

$$m \approx \frac{\left(u_\alpha\sqrt{q_0(1-q_0)}+u_\beta\sqrt{q_1(1-q_1)}\right)^2}{(q_1-q_0)^2},$$

где u_α, u_β — квантили стандартного нормального распределения.

Оценим величину $\widetilde{p}(\delta,k)$ — вероятность принять случайное множество мощности Ω в траекторию. Для этого оценим распределение веса функции f на плоскостях L_i, выбираемых, согласно модели, случайно и равновероятно. Выбор таких множеств можно представить в виде случайной выборки без возвращения мощности Ω из конечной совокупности 2^n элементов, при этом ровно 2^{n-1} из них обладают тем свойством, что значение функции f на них равно единице. Рассмотрим случайную величину S, равную количеству точек выбранного множества, на которых функция f равна единице. Случайная величина S имеет гипергеометрическое распределение с параметрами $2^n, 2^{n-1}, \Omega$, то есть $S \sim \mathrm{HG}(2^{n-1}, 2^n, \Omega)$ [12].

Для удобства вычислений перейдём от гипергеометрического распредления к нормальному $\mathrm{N}(\Omega/2, \Omega/4)$ следующим образом. Так как $\mathrm{HG}(D,N,n) \approx \mathrm{Bin}(n, D/N)$ при $N \to \infty$ [12], то будем считать, что $S \sim \mathrm{Bin}(\Omega, 1/2)$. Поскольку $\mathrm{Bin}(n,p) \approx \mathrm{N}(np, npq)$ при больших n, где $\mathrm{N}(np, npq)$ — нормальное распределение с математическим ожиданием np и дисперсией npq, то $\mathrm{Bin}(\Omega, 1/2) \approx \mathrm{N}(\Omega/2, \Omega/4)$, $p = q = 1/2$.

Пусть $\Phi(y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-x^2/2}dx$ — функция распределения стандартной нормальной случайной величины. Тогда

$$\begin{aligned}\widetilde{p}(\delta,k) &= 1 - \mathsf{P}\left[T_0 \leqslant S \leqslant T_1\right] = 1 - \mathsf{P}\left[\frac{T_0-\Omega p}{\sqrt{\Omega pq}} \leqslant \frac{S-\Omega p}{\sqrt{\Omega pq}} \leqslant \frac{T_1-\Omega p}{\sqrt{\Omega pq}}\right] = \\ &= 1 - Pr\left[-\frac{\delta}{\sqrt{\Omega}} \leqslant \frac{S-\Omega p}{\sqrt{\Omega pq}} \leqslant \frac{\delta}{\sqrt{\Omega}}\right] = 1 - \left(\Phi\left(\frac{\delta}{\sqrt{\Omega}}\right) - \Phi\left(-\frac{\delta}{\sqrt{\Omega}}\right)\right) = \\ &= 2\left(1 - \Phi\left(\frac{\delta}{\sqrt{\Omega}}\right)\right).\end{aligned}$$

Таким образом, объём выходной последовательности генератора, необходимый для построения траектории, равен

$$\widetilde{N}(\delta,k) = \frac{m}{\widetilde{p}(\delta,k)} \approx \frac{\left(u_\alpha + u_\beta\sqrt{1-(\delta/\Omega)^2}\right)^2 (\Omega/\delta)^2}{2\left(1-\Phi\left(\delta/\sqrt{\Omega}\right)\right)}.$$

Необходимо минимизировать величину $\widetilde{N}(\delta,k)$ при возможных значениях параметров $\delta \in \{0,1,\ldots,\Omega\}$ и $k \in \{0,1,\ldots,n\}$.

Решение задачи минимизации

Перейдём к непрерывной переменной $t = \delta/\sqrt{\Omega}$, $t \in (0;\sqrt{\Omega}]$. Тогда

$$\widetilde{N}(\delta,k) \approx \frac{\Omega}{2}\left(u_\alpha + u_\beta\sqrt{1-t^2/\Omega}\right)^2 \frac{1}{t^2(1-\Phi(t))}.$$

Рассмотрим наиболее интересный для решаемой задачи случай, когда $Q \ll 2^n$, а π_0 принимает значения из области, естественной для практических применений (например, $\pi_0 = 1/2$ или $\pi_0 = 1/10$). При таком условии $\alpha \ll \beta$ и, следовательно, $u_\alpha \gg u_\beta$. Поэтому далее будем полагать $\widetilde{N}(\delta,k) \approx \frac{\Omega}{2}\frac{u_\alpha^2}{t^2(1-\Phi(t))}$.

Найдём значение переменной t, при котором достигается максимум выражения $f(t) = t^2(1 - \Phi(t))$, $t \in (0; \sqrt{\Omega}]$. Приравнивая производную $f'(t)$ к нулю, получим уравнение

$$2(1 - \Phi(t)) = \frac{t}{\sqrt{2\pi}} e^{-t^2/2},$$

которое эквивалентно

$$\frac{t}{2} = \frac{1 - \Phi(t)}{\varphi(t)},$$

где $\varphi(t) = \dfrac{e^{-t^2/2}}{\sqrt{2\pi}}$.

Выражение $R(t) = \dfrac{1 - \Phi(t)}{\varphi(t)}$ называется отношением Миллса [13]. Функция $R(t)$ монотонно убывает [14], значит, уравнение имеет в точности одно решение. Получить аналитическое выражение для корня данного уравнения не представляется возможным, но можно его посчитать с любой наперёд заданной точностью. Приведём значение корня с точностью до 0,00001: $t_0 = 1{,}19061$.

Произведём обратную замену и тем самым вернёмся к дискретным величинам: $\delta_0 \approx \lceil t_0\sqrt{\Omega} \rceil$. Значение $\widetilde{N}(\delta_0, k)$ в этой точке равно

$$\widetilde{N}(\delta_0, k) \approx \frac{\Omega}{2} \cdot u_\alpha^2 \cdot C_\Phi,$$

где $C_\Phi = t_0^{-2}(1 - \Phi(t_0))^{-1} \approx 6{,}03442$.

Так как для малых α справедливо $u_\alpha \approx \sqrt{-\ln(2\pi\alpha^2)}$, то u_α изменяется не более чем на n при допустимых k. В то же время Ω растет по k экспоненциально. Таким образом, $\widetilde{N}(\delta_0, k)$ достигает минимума при минимально допустимом k. Допустимыми являются те $k \in \{1, 2, \ldots, n\}$, для которых $\alpha = 2^{-n}\left(Q - 2^{n-k} - \pi_0 2^k\right) > 0$. Минимально допустимое значение k определяется равенством

$$k = \left\lceil \log_2 \left(\frac{Q - \sqrt{Q^2 - \pi_0 2^{n+2}}}{2\pi_0} \right) \right\rceil.$$

В табл. 1 указаны параметры k, δ и величины m, $\widetilde{N}$ при $n = 128$, $\pi_0 = 1/2$ для различных значений трудоёмкости.

Таблица 1

Q	k	δ	m	$\widetilde{N}$
2^{70}	59	2^{30}	2^{73}	2^{73}
2^{80}	49	2^{25}	2^{63}	2^{63}
2^{90}	39	2^{20}	2^{53}	2^{53}

3.4. Экспериментальная проверка релевантности математической модели

Рассуждения из п. 3.2 основаны на замене плоскостей, участвующих в однородной аппроксимации, случайными множествами. Проверим точность полученных оценок.

Для этого рассмотрим следующий фильтрующий генератор. Пусть $n = 32$, а преобразование A задаётся регистром сдвига с линейной обратной связью, характеристический многочлен которого является примитивным и равен $p(x) = x^{32} + x^7 + x^6 + x^2 + 1$. Перейдём к описанию фильтрующей функции. Пусть $\pi_i : V_4 \to V_4$ — перестановки, заданные в [15] ($i = 0, 1, \ldots, 7$); $\Psi : V_{32} \to V_{32}$ — отображение $\Psi(x) = \Psi(x_0 || \ldots || x_7) = \pi_0(x_0) || \ldots || \pi_7(x_7)$, где $x = x_0 || \ldots || x_7 \in V_{32}$, $x_i \in V_4$, $i = 0, 1, \ldots, 7$; $\mathcal{S}(x) : V_{32} \to V_{32}$ — циклический сдвиг вправо на 11 бит; $\mathcal{X}(x) : V_{32} \to \mathbb{F}_2$ — сумма по модулю 2 бит вектора x. Эксперименты проводились для функции $f(x) = \mathcal{X}(\Psi(\mathcal{S}(\Psi(\mathcal{S}(\Psi(x))))))$.

Пусть необходимо построить такую однородную аппроксимацию, чтобы метод восстановления ключа на её основе имел трудоёмкость $Q = 2^{24}$ и надёжность не меньше $\pi_0 = 1/2$. Согласно выражениям, представленным в п. 3.3, оптимальные значения параметров для построения однородной аппроксимации следующие:

$$k = 9, \quad \delta = 27.$$

При данных параметрах для достижения целевых значений надёжности и трудоёмкости метода восстановления ключа потребуется отрезок выходной последовательности длины $\widetilde{N} = 12861$, длина траекторий составляет $m = 3007$.

Для проверки релевантности предложенной математической модели разработана программа, реализующая построение траектории плоскостной однородной аппроксимации для заданных параметров метода при случайном выборе начальной плоскости L_0. Для подсчёта среднего значения $\widetilde{N}$ для каждой пары параметров (k, δ) процедура построения траектории была запущена 100 раз. В табл. 2 представлены значения целевого параметра $\widetilde{N}$, полученные с помощью формулы из п. 3.3 ($\widetilde{N}_t$) и в результате экспериментов ($\widetilde{N}_e$). Жирным шрифтом выделены столбцы, в которых достигается минимальное значение длины отрезка выходной последовательности.

Таблица 2

$k = 9$							
δ	24	25	26	**27**	28	29	30
$\widetilde{N}_t$	13121	12974	12892	**12866**	12898	12981	13120
$\widetilde{N}_e$	14094	13911	13880	**13849**	13975	14018	14174
$k = 10$							
δ	36	37	38	**39**	40	41	42
$\widetilde{N}_t$	23573	23487	23456	**23477**	23549	23667	23833
$\widetilde{N}_e$	26201	24736	23451	**24635**	26256	25026	23842
$k = 11$							
δ	51	52	53	**54**	55	56	57
$\widetilde{N}_t$	45406	45282	45213	**45198**	45224	45302	45429
$\widetilde{N}_e$	47048	48714	46999	**45278**	46935	48848	47203
$k = 12$							
δ	74	75	76	**77**	78	79	80
$\widetilde{N}_t$	89005	88925	88890	**88905**	88966	89075	89233
$\widetilde{N}_e$	89085	91210	93756	**91362**	89037	91424	94197

Данные эксперименты подтверждают, что теоретически предсказанные результаты действительно являются оптимальными, а величина целевого параметра $\widetilde{N} \approx 12861$.

Заключение

В настоящей работе получены выражения для оптимальных характеристик переборного метода построения плоскостных аппроксимаций одного специального вида. Проведённый численный эксперимент показал, что данные выражения, будучи получены в рамках довольно сильного модельного предположения, позволили получить относительно точные значения характеристик метода. Уточнение оценок за счёт конкретизации модели остаётся открытым вопросом для будущих исследований. Другим открытым вопросом является исследование проблемы построения плоскостных аппроксимаций более общего вида.

ЛИТЕРАТУРА

1. `http://www.bluetooth.org` — Bluetooth^{TM}. Bluetooth Specification, version 1.2. 2003. P. 903–948.
2. *Briceno M., Goldberg I., and Wagner D.* A pedagogical implementation of A5/1. 1999. `https://cryptome.org/jya/a51-pi.htm`.
3. *Ekdahl P. and Johansson T.* A new version of the stream cipher SNOW // LNCS. 2003. V. 2595. P. 47–61.
4. *Hell M., Johansson T., and Meier W.* Grain: stream cipher for constrained environments // Int. J. Wireless Mobile Computing. 2007. No. 2(1). P. 86–93.
5. *Siegenthaler T.* Decrypting a class of stream cipher using ciphertext only // IEEE Trans. Comput. 1985. V. C-34(1). P. 81–85.
6. *Meier W. and Staffelbach O.* Fast correlation attacks on certain stream cipher // J. Cryptology. 1989. V. 1. No. 3. P. 159–176.
7. *Courtois N. and Meier W.* Algebraic attacks on stream ciphers with linear feedback // LNCS. 2003. V. 2656. P. 345–359.
8. *Courtois N.* Fast algebraic attacks on stream ciphers with linear feedback // LNCS. 2003. V. 2729. P. 176–194.
9. *Логачев О. А., Сальников А. А., Ященко В. В.* Корреляционная иммунность и реальная секретность // Математика и безопасность информационных технологий. Материалы конф. в МГУ 23–24 октября 2003 г. М.: МЦНМО, 2004. С. 165–171.
10. *Алексеев Е. К., Кущинская Л. А.* Обобщение одного метода восстановления ключа фильтрующего генератора // Дискретная математика. 2017. Т. 29. № 4. С. 3–27.
11. *Глухов М. М., Елизаров В. П., Нечаев А. А.* Алгебра: Учебник. В 2-х т. Т. 1. М.: Гелиос АРВ, 2003. 416 с.
12. *Феллер В.* Введение в теорию вероятностей и ее приложения. Т. 1. М.: Мир, 1984.
13. *Mills J. P.* Table of the ratio: area to bounding ordinate, for any portion of normal curve // Biometrika. 1986. V. 18. No. 3/4. P. 395–400.
14. *Gasull A. and Utzet F.* Approximating Mills ratio // J. Math. Anal. Appl. 2014. V. 420. Iss. 2. P. 1832–1853.
15. ГОСТ Р 34.12–2015. Информационная технология. Криптографическая защита информации. Блочные шифры. М.: Стандартинформ, 2015.

REFERENCES

1. `http://www.bluetooth.org` — Bluetooth^{TM}. Bluetooth Specification, version 1.2, 2003, pp. 903–948.

2. *Briceno M., Goldberg I., and Wagner D.* A pedagogical implementation of A5/1. 1999. `https://cryptome.org/jya/a51-pi.htm`.
3. *Ekdahl P. and Johansson T.* A new version of the stream cipher SNOW. LNCS, 2003, vol. 2595, pp. 47–61.
4. *Hell M., Johansson T., and Meier W.* Grain: stream cipher for constrained environments. Int. J. Wireless Mobile Computing, 2007, no. 2(1), pp. 86–93.
5. *Siegenthaler T.* Decrypting a class of stream cipher using ciphertext only. IEEE Trans. Comput., 1985, vol. C-34(1), pp. 81–85.
6. *Meier W. and Staffelbach O.* Fast correlation attacks on certain stream cipher. J. Cryptology, 1989, vol. 1, no. 3, pp. 159–176.
7. *Courtois N. and Meier W.* Algebraic attacks on stream ciphers with linear feedback. LNCS, 2003, vol. 2656, pp. 345–359.
8. *Courtois N.* Fast algebraic attacks on stream ciphers with linear feedback. LNCS, 2003, vol. 2729, pp. 176–194.
9. *Logachev O. A., Sal'nikov A. A., and Yashchenko V. V.* Korrelyatsionnaya immunnost' i real'naya sekretnost' [Correlation immunity and real secrecy]. Matematika i Bezopasnost' Informatsionnykh Tekhnologiy, Moscow, MCCME Publ., 2004, pp. 165–171. (in Russian)
10. *Alekseev E. K. and Kushchinskaya L. A.* Generalization of one method of a filter generator key recovery. Discrete Math. Appl., 2019, vol. 29, no. 2, pp. 69–87.
11. *Glukhov M. M., Elizarov V. P., and Nechaev A. A.* Algebra [Algebra]. V. 1, Moscow, Gelios ARV Publ., 2003. 416 p.
12. *Feller V.* Vvedenie v teoriyu veroyatnostey i ee prilozheniya [An Introduction to Probability Theory and its Applications]. V. 1, Moscow, Mir Publ., 1984.
13. *Mills J. P.* Table of the ratio: area to bounding ordinate, for any portion of normal curve. Biometrika, 1986, vol. 18, no. 3/4, pp. 395–400.
14. *Gasull A. and Utzet F.* Approximating Mills ratio. J. Math. Anal. Appl., 2014, vol. 420, iss. 2, pp. 1832–1853.
15. GOST R 34.12–2015. Informatsionnaya tekhnologiya. Kriptograficheskaya zashchita informatsii. Blochnye shifry [Information Technology. Information Security. Block Ciphers]. Moscow, Standartinform Publ., 2015. (in Russian)