

МАТЕМАТИЧЕСКИЕ МЕТОДЫ КРИПТОГРАФИИ

УДК 519.23

DOI 10.17223/20710410/69/4

ХАРАКТЕРИСТИКИ АТАК РАЗЛИЧЕНИЯ НА 3 И 4 РАУНДА
СХЕМЫ ЛУБИ — РАКОВА В МОДЕЛИ НЕЗАВИСИМЫХ
ПОДСТАНОВОК

О. В. Денисов*, Е. Д. Андреев*, М. А. Батаев**

*ООО «Инновационные телекоммуникационные технологии», г. Москва, Россия

**ФГУП «НИИ „Квант“», г. Москва, Россия

E-mail: denisovOleg@yandex.ru, andreev5902@mail.ru, misha.bat72@gmail.com

Вычислены математические ожидания статистик Патарина, используемых в атаках различения на 3 и 4 раунда схемы Луби — Ракова по выбранным открытым текстам. В модели независимых подстановок, к каждой из которых производится по два запроса, получены оценки вероятностей ошибок и явные выражения для объёмов материала атак, построенных на аналогичных статистиках. В случае четырёх раундов при длинах блока 16–52 получены эмпирические вероятности ошибок в модели независимых подстановок и в модели запроса значений одной подстановки.

Ключевые слова: *схема Луби — Ракова, статистики Патарина, атака различения.*

CHARACTERISTICS OF DISTINGUISHING ATTACKS ON 3 AND
4 ROUNDS OF THE LUBY — RACKOFF SCHEME IN INDEPENDENT
PERMUTATIONS MODEL

O. V. Denisov*, E. D. Andreev*, M. A. Bataev**

*Innovative Telecommunication Technologies, LLC, Moscow, Russia

**Federal State Unitary Enterprise “Scientific Research Institute ‘Kvant’”, Russia

We calculate the means of Patarin statistics that are used in distinguishing CPA-attacks on 3 and 4 rounds of the Luby — Rackoff scheme. We study a model of independent permutations and make two queries for each. In this model, we find estimates of error probabilities and explicit expressions for the data complexities of attacks based on similar statistics. In case of 4 rounds and block lengths 16–52 we have got empirical error probabilities in the model of independent permutations and in the model of queries for a single permutation.

Keywords: *Luby — Rackoff scheme, Patarin statistics, distinguishing attack.*

Введение

Рассмотрим R -раундовую схему Фейстеля с алфавитом полублоков $\mathbb{Z}_2^m$, операций $\oplus$ покомпонентного сложения полублоков по модулю 2 и раундовыми преобразованиями

$$(x^{r-1}, x^r) \rightarrow (x^r, x^{r-1} \oplus f_r(x^r)), \quad 1 \leq r \leq R.$$

В 1988 г. американские криптографы Луби (М. Luby) и Раков (С. Rackoff) ввели [1] вероятностную модель этой схемы, в которой раундовые функции усложнения выбираются независимо случайно равновероятно из множества всех двоичных вектор-функций от m переменных:

$$f_1, \dots, f_R \sim U(\mathcal{F}_m), \quad \mathcal{F}_m = \{f \mid f: \mathbb{Z}_2^m \rightarrow \mathbb{Z}_2^m\}.$$

Здесь и далее $U(\mathbb{X})$ — равномерное распределение вероятностей на множестве $\mathbb{X}$; символ $\sim$ означает «имеет распределение»; $\mathcal{S}(\mathbb{X})$ — множество подстановок на $\mathbb{X}$; $\mathbb{I}(A)$ — индикаторная функция условия A . Кратко эту вероятностную модель будем называть *ЛР-схемой* (Луби, Раков).

В течение следующих лет примерно до 2008 г. французским криптографом Жаком Патариным (J. Patarin) и другими авторами был получен ряд результатов: выведены нижние оценки объёмов материала атак различения на 3–4 раунда в разных моделях запросов (наблюдений), предложены атаки на R раундов с эвристическими асимптотическими ($m \rightarrow \infty$) оценками их объёма материала. Новый импульс к развитию эта тематика получила при разработке блочных *FPE-схем* (Format Preserving Encryption), т. е. схем шифрования, сохраняющих исходный формат шифруемых данных [2]. FPE-схемы FF1 и FF3, принятые как стандарт NIST, представляют собой схемы Фейстеля, раундовые функции усложнения которых при равновероятном выборе раундовых ключей близки к случайным равновероятным функциям. Поэтому свойства ЛР-схем используются для обоснования безопасности таких FPE-схем. Согласно [3, с. 446], идеи Патарина были использованы в [4] для построения атак восстановления сообщений. Методика Патарина использовалась также для расчёта нижних оценок стойкости схемы FEA-2 [5, с. 145–147], принятой в качестве стандарта FPE-шифрования в Южной Корее.

В работах [6; 7; 8, с. 68] приведены, в частности, статистики для атак различения на схему с $R = 3, 4$ раундами в модели запросов CPA (Chosen Plaintext Attack) — атаки на основе выбранных открытых текстов, однако критерии не были явно сформулированы и рассчитаны. В п. 1 данной работы получены точные значения математических ожиданий этих статистик при каждой из двух проверяемых гипотез.

В п. 2 в модели выбора независимых подстановок, к каждой из которых производится по два запроса, сформулированы аналогичные гипотезы, найдены математические ожидания и дисперсии аналогичных статистик. Построены критерии, размеры которых близки к заданному значению α . Даны оценки $N_1(\alpha, \beta)$ объёмов материала, при которых вероятности ошибок 2-го рода критериев близки к заданному значению β .

В п. 3 для атак на четыре раунда при длинах блоков $16 \leq 2m \leq 52$ проведены статистические эксперименты — получены эмпирические оценки вероятностей ошибок двух родов: 1) в модели запросов пар значений независимых подстановок; 2) в модели запроса значений одной подстановки на наборе из q заранее выбранных аргументов (т. е. в исходной модели Патарина) с таким выбором, что число $\binom{q}{2}$ слагаемых в статистике близко к $N_1(\alpha, \beta)$.

В п. 4 обсуждается связь полученных результатов с другими работами по разностному анализу. Пункт 1 написан авторами совместно, пп. 2–4 — первым автором. Авторы выражают признательность рецензенту за ряд полезных замечаний, способствовавших улучшению статьи.

1. Статистики Патарина в атаках на 3 и 4 раунда

1.1. Необходимые леммы

Сформулируем и докажем две леммы, необходимые для расчёта атак различения. Здесь и далее через $(x)_k = x(x-1)\dots(x-k+1)$, $k \in \mathbb{N}$, обозначаем k -ю факториальную степень числа $x \in \mathbb{R}$, а через $v^{\text{left}}, v^{\text{right}} \in \mathbb{Z}_2^m$ — левый и правый полублоки вектора $v \in \mathbb{Z}_2^{2m}$.

Лемма 1. Если пара блоков (σ_1, σ_2) выбирается случайно равномерно без вращения из $\mathbb{Z}_2^{2m}$, то

$$\Pr[(\sigma_1 \oplus \sigma_2)^{\text{left}} = x] = \begin{cases} 2^m / (2^{2m} - 1) & \text{при } x \in \mathbb{Z}_2^m \setminus \{0\}, \\ 1 / (2^m + 1) & \text{при } x = 0. \end{cases}$$

Доказательство. Так как $\Pr[\sigma_1 = a, \sigma_2 = b] = \frac{1}{(2^{2m})_2}$ для любых различных $a, b \in \mathbb{Z}_2^{2m}$, то при $x \neq 0$

$$\begin{aligned} \Pr[(\sigma_1 \oplus \sigma_2)^{\text{left}} = x] &= \sum_{y \in \mathbb{Z}_2^{2m}} \Pr[\sigma_1 = y, \sigma_2 \in \{(y^{\text{left}} \oplus x, z) : z \in \mathbb{Z}_2^m\}] = \\ &= \sum_{y \in \mathbb{Z}_2^{2m}} \sum_{z \in \mathbb{Z}_2^m} \Pr[\sigma_1 = y, \sigma_2 = (y^{\text{left}} \oplus x, z)] = \sum_{y \in \mathbb{Z}_2^{2m}} \sum_{z \in \mathbb{Z}_2^m} \frac{1}{(2^{2m})_2} = \\ &= 2^{2m} \cdot 2^m \cdot \frac{1}{2^{2m}(2^{2m} - 1)} = \frac{2^m}{2^{2m} - 1}. \end{aligned}$$

Следовательно, $\Pr[(\sigma_1 \oplus \sigma_2)^{\text{left}} = 0] = 1 - (2^m - 1) \frac{2^m}{2^{2m} - 1} = \frac{2^m - 1}{2^{2m} - 1} = \frac{1}{2^m + 1}$. ■

Известен следующий факт: если $\xi = (\xi_1, \dots, \xi_n) \sim U(A^n)$, A конечно, то $\Pr[\xi_i = \xi_j] = |A|^{-1}$ для любых фиксированных i, j . В следующей лемме этот факт обобщается на ситуацию случайного выбора пары индексов.

Лемма 2. Пусть $|A| < \infty$, $\mathcal{F} = \{F \mid F : A \rightarrow A\}$, пара (x, y) выбирается случайно из A^2 так, что $\Pr[x = y] = 0$, случайная функция $f \sim U(\mathcal{F})$ не зависит от (x, y) . Тогда $\Pr[f(x) = f(y)] = |A|^{-1}$.

Доказательство. Условие равномерности распределения случайной функции f на $\mathcal{F}$ эквивалентно условию независимости и равномерной распределённости её значений на A . Поэтому для любых различных $a, b \in A$

$$\begin{aligned} \Pr[f(a) = f(b)] &= \sum_{c \in A} \Pr[f(a) = c, f(b) = c] = \\ &= \sum_{c \in A} \Pr[f(a) = c] \Pr[f(b) = c] = |A| \frac{1}{|A|^2} = \frac{1}{|A|}. \end{aligned} \tag{1}$$

Рассмотрим множество $E = \{(a, b) \in A^2 : a \neq b\}$. Так как $\Pr[(x, y) \in E] = 1$, то из аддитивности вероятности и независимости событий $\{x = a, y = b\}$ и $\{f(a) = f(b)\}$ с учётом (1) получаем

$$\begin{aligned}
\Pr[f(x) = f(y)] &= \sum_{(a,b) \in E} \Pr[(x, y) = (a, b), f(a) = f(b)] = \\
&= \sum_{(a,b) \in E} \Pr[x = a, y = b] \Pr[f(a) = f(b)] = \sum_{(a,b) \in E} \Pr[x = a, y = b] \frac{1}{|A|} = \\
&= \frac{1}{|A|} \Pr[(x, y) \in E] = \frac{1}{|A|}.
\end{aligned}$$

Лемма 2 доказана. ■

1.2. Средние значения статистик Патарина

Атаки различения на R раундов — это критерии проверки гипотез о случайной подстановке F на $\mathbb{Z}_2^{2m}$:

$$\begin{aligned}
H_1 : F &\sim U(\mathbb{S}(\mathbb{Z}_2^{2m})), \\
H_2 : F &\text{ получена по } R\text{-раундовой ЛР-схеме.}
\end{aligned}$$

При входном блоке (x^0, x^1) в первой строке табл. 1 выписана последовательность полублоков схемы Фейстеля; выходом R -раундовой схемы является блок (x^R, x^{R+1}) . Вывод строк 2, 3 обсуждается в п. 2 доказательства теоремы 1.

Т а б л и ц а 1

Раундовые последовательности в схеме Фейстеля: полублоки, полуразности при $\Delta x^0 = 0$ либо при $\Delta x^1 = 0$

r	0	1	2	3	4
x^r	x^0	x^1	$x^0 \oplus f_1(x^1)$	$x^1 \oplus f_2(x^0 \oplus f_1(x^1))$	$x^0 \oplus f_1(x^1) \oplus f_3(x^1 \oplus f_2(x^0 \oplus f_1(x^1)))$
Δx^r	0	Δx^1	$\Delta f_1(\Delta x^1)$	$\Delta x^1 \oplus \Delta f_2(\Delta f_1(\Delta x^1))$	$\Delta f_1(\Delta x^1) \oplus \Delta f_3(\Delta x^1 \oplus \Delta f_2(\Delta f_1(\Delta x^1)))$
Δx^r	Δx^0	0	Δx^0	$\Delta f_2(\Delta x^0)$	$\Delta x^0 \oplus \Delta f_3(\Delta f_2(\Delta x^0))$

При $R = 3$ атака выглядит так [7, с. 225]:

- 1) для произвольного фиксированного x^0 выбираем (не обязательно равновероятно) q попарно различных полублоков x_i^1 , $1 \leq i \leq q$;
- 2) запрашиваем значения $Y_i = F(x_i^0, x_i^1)$, $1 \leq i \leq q$;
- 3) вычисляем статистику $\nu_3 = \sum_{1 \leq i < j \leq q} \mathbb{I}\{Y_i^{\text{left}} \oplus x_i^1 = Y_j^{\text{left}} \oplus x_j^1\}$.

При $R = 4$ атака описывается аналогично [7, с. 225]:

- 1) для произвольного фиксированного x^1 выбираем (не обязательно равновероятно) q попарно различных полублоков x_i^0 , $1 \leq i \leq q$;
- 2) запрашиваем значения $Y_i = F(x_i^0, x^1)$, $1 \leq i \leq q$;
- 3) вычисляем статистику $\nu_4 = \sum_{1 \leq i < j \leq q} \mathbb{I}\{Y_i^{\text{left}} \oplus x_i^0 = Y_j^{\text{left}} \oplus x_j^0\}$.

Далее через P_s и E_s обозначаем соответственно вероятностную меру и математическое ожидание при гипотезе H_s , $s = 1, 2$. В [7, с. 225] для $R = 3, 4$ без доказательства отмечается, что $E_2 \nu_R \approx 2E_1 \nu_R$. Найдём точные значения величин $E_s \nu_R$ и докажем, что $E_s \nu_3 = E_s \nu_4$, $s \in \{1, 2\}$.

Теорема 1. При $s \in \{1, 2\}$ и $R \in \{3, 4\}$

$$E_s \nu_R = \binom{q}{2} p_s, \text{ где } p_1 = \frac{2^{-m}}{1 - 2^{-2m}}, p_2 = 2^{-m}(2 - 2^{-m}).$$

Доказательство. При фиксированных $i < j$ обозначим через $\Delta x^r = x_i^r \oplus x_j^r$ полуразности, $r \geq 0$. Тогда $(\Delta x^0, \Delta x^1)$ — входная разность подстановки F ; через $\Delta Y = Y_i \oplus Y_j$ обозначим её выходную разность. С учётом линейности математического ожидания достаточно найти вероятности событий A_{ij} под знаком индикатора в сумме для ν_R при каждой из гипотез.

С л у ч а й $s = 1$. При $R = 3$ событие под знаком индикатора можем записать в виде $A_{ij} = \{\Delta Y^{\text{left}} = \Delta x^1\}$.

Для любого такого распределения (x_i^1, x_j^1) , что $\Pr[x_i^1 = x_j^1] = 0$, при гипотезе H_1 случайные векторы (Y_i, Y_j) и (x_i^1, x_j^1) независимы. Действительно, в силу независимости набора из двух фиксированных значений равновероятной случайной подстановки F от пары входных блоков для любых $\{y_1, y_2\} \subset \mathbb{Z}_2^{2m}$, $\{x_1, x_2\} \subset \mathbb{Z}_2^m$ имеем

$$\begin{aligned} & \Pr\{(Y_i, Y_j) = (y_1, y_2), (x_i^1, x_j^1) = (x_1, x_2)\} = \\ &= \Pr\{(F(x^0, x_1), F(x^0, x_2)) = (y_1, y_2), (x_i^1, x_j^1) = (x_1, x_2)\} = \\ &= \Pr\{(F(x^0, x_1), F(x^0, x_2)) = (y_1, y_2)\} \Pr\{(x_i^1, x_j^1) = (x_1, x_2)\} = \\ &= \frac{1}{(2^{2m})_2} \Pr\{(x_i^1, x_j^1) = (x_1, x_2)\}, \end{aligned}$$

и (Y_i, Y_j) имеет распределение случайной равновероятной выборки без возвращения из $\mathbb{Z}_2^{2m}$. Поэтому функции ΔY и Δx^1 от этих случайных векторов тоже независимы, и с учётом леммы 1 получаем

$$\begin{aligned} \Pr(A_{ij}) &= \sum_{x \neq \mathbf{0}} \Pr\{(\Delta Y)^{\text{left}} = x, \Delta x^1 = x\} = \sum_{x \neq \mathbf{0}} \Pr\{(\Delta Y)^{\text{left}} = x\} \Pr\{\Delta x^1 = x\} = \\ &= \sum_{x \neq \mathbf{0}} \frac{2^m}{2^{2m} - 1} \Pr\{\Delta x^1 = x\} = \frac{2^m}{2^{2m} - 1} = \frac{1}{2^m} \frac{1}{1 - 2^{-2m}} = p_1. \end{aligned}$$

При $r = 4$ аналогично доказывается, что событие $A_{ij} = \{\Delta Y^{\text{left}} = \Delta x^0\}$ имеет вероятность p_1 .

С л у ч а й $s = 2$. Из рекуррентного соотношения $x^{r+1} = x^{r-1} \oplus f_r(x^r)$ следует $\Delta x^{r+1} = \Delta x^{r-1} \oplus \Delta f_r(\Delta x^r)$ для раундовых полуразностей схемы Фейстеля, $r \geq 1$; для краткости в записи приращений раундовых функций

$$\Delta f_r(x_i^r, \Delta x_i^r) = f_r(x_i^r) \oplus f_r(x_j^r) = f_r(x_i^r) \oplus f_r(x_i^r \oplus \Delta x_i^r)$$

первый аргумент опускаем.

Во второй строке табл. 1 последовательность раундовых полуразностей выписана при условии $\Delta x^0 = \mathbf{0}$, а в третьей — при $\Delta x^1 = \mathbf{0}$. При $R = 3$ с учётом второй строки имеем

$$A_{ij} = \{\Delta x^3 = \Delta x^1\} = \{\Delta f_2(\Delta f_1(\Delta x^1)) = \mathbf{0}\}.$$

Обозначая здесь для краткости $B = \{\Delta f_1(\Delta x^1) = \mathbf{0}\}$, применяя формулу полной вероятности и дважды лемму 2 (возможность применения леммы обосновывается независимостью случайных функций f_1 и f_2), находим

$$\begin{aligned} \Pr(A_{ij}) &= \Pr_2(B) + \Pr_2\{\bar{B}\} \Pr_2\{\Delta f_2(\Delta f_1(\Delta x^1)) = \mathbf{0}\} \mid \bar{B} = \\ &= 2^{-m} + (1 - 2^{-m})2^{-m} = 2^{-m}(2 - 2^{-m}) = p_2. \end{aligned}$$

При $R = 4$, согласно третьей строке табл. 1, получаем

$$A_{ij} = \{\Delta x^4 = \Delta x^0\} = \{\Delta f_3(\Delta f_2(\Delta x^0)) = \mathbf{0}\}.$$

Аналогично случаю $R = 3$ доказывается, что вероятность этого события равна p_2 .

Теорема 1 доказана. ■

Итак, в модели наблюдений Патарина, где запрашиваются значения одной подстановки, вычислены средние значения случайных величин ν_R при гипотезах H_1 и H_2 . Но пока не вычислены значения дисперсий, из-за чего теоретический расчёт вероятностей ошибок критериев не представляется возможным. Однако эта проблема разрешима в модели наблюдения независимых двублочных текстов [9]. Условия этой модели дают независимость слагаемых в суммах для вводимых далее статистик $\nu'_{3,4}$, соответствующих статистикам $\nu_{3,4}$. Условие независимости слагаемых используется при оценке дисперсий статистик $\nu_{3,4}$ во всех известных работах Патарина, хотя оно не соответствует его модели наблюдений.

Поэтому, с одной стороны, переход к более простой модели независимых двублочных текстов позволяет более обоснованно строить критерии, рассчитывать их пороги и получать математически строгие результаты о вероятностях ошибок. Затем на этой основе строятся аналогичные критерии в исходной модели наблюдений. Соответствие параметров распределений статистик $\nu'_{3,4}$ и статистик $\nu_{3,4}$, определяющих вероятности ошибок критериев, проверяется экспериментально.

С другой стороны, в криптографии эта модель и более общая модель независимых q -блочных текстов возникает, например, когда для каждого шифртекста, полученного в режиме гаммирования, аналитик знает лишь q блоков исходного открытого текста, поэтому может вычислить лишь $\binom{q}{2}$ пар «входная/выходная разность» (подробнее см. [10, с. 98–100]).

2. Атаки на основе статистик Патарина в модели независимых двублочных текстов

Рассматривается модель наблюдений, в которой запросы делаются к случайным независимым подстановкам $F_1, \dots, F_N$, одинаково распределённым на $\mathbb{S}(\mathbb{Z}_2^{2m})$. Требуется проверить следующие гипотезы:

$$\begin{aligned} H_1 : F_1, \dots, F_N \text{ выбраны равновероятно из } \mathbb{S}(\mathbb{Z}_2^{2m}), \\ H_2 : F_1, \dots, F_N \text{ — композиции } R \text{ раундов схемы Луби — Ракова.} \end{aligned} \quad (2)$$

При $R = 3, 4$ построим статистики $\nu'_{3,4}$, аналогичные статистикам Патарина, и на их основе — критерии проверки.

При $R = 3$: для $t = 1, \dots, N$ запрашиваем значения $Y_{t1} = F_t(x_t^0, x_{t1}^1)$ и $Y_{t2} = F_t(x_t^0, x_{t2}^1)$, где случайные векторы $(x_t^0, x_{t1}^1, x_{t2}^1)$, $t = 1, \dots, N$, независимы, $x_{t1}^1 \neq x_{t2}^1$; вычисляем статистику

$$\nu'_3 = \sum_{1 \leq t \leq N} \mathbb{I}\{(\Delta Y_t)^{\text{left}} = (\Delta X_t)^{\text{right}}\},$$

где $\Delta X_t = (x_t^0, x_{t1}^1) \oplus (x_t^0, x_{t1}^1)$ и $\Delta Y_t = Y_{t1} \oplus Y_{t2}$ — разности между аргументами и значениями подстановок.

При $R = 4$: для $t = 1, \dots, N$ запрашиваем значения $Y_{t1} = F_t(x_{t1}^0, x_t^1)$ и $Y_{t2} = F_t(x_{t2}^0, x_t^1)$, где случайные векторы $(x_{t1}^0, x_{t2}^0, x_t^1)$ независимы, $x_{t1}^0 \neq x_{t2}^0$, вычисляем статистику

$$\nu'_4 = \sum_{1 \leq t \leq N} \mathbb{I}\{(\Delta Y_t)^{\text{left}} = (\Delta X_t)^{\text{left}}\}.$$

В силу независимости выбора блоков и подстановок в обоих случаях случайные величины ν'_R имеют биномиальное распределение $\text{Bin}(N, p_i)$ при гипотезе H_i , $i = 1, 2$, где значения p_i определены в теореме 1. Поэтому критерии проверки гипотез (2) записываем единообразно:

$$d'_R : \nu'_R > c \implies \text{принимаем } H_2, \quad (3)$$

где $c \in [0, \infty)$ — параметр критерия.

Пусть $\alpha_i(d)$ — вероятность ошибки i -го рода критерия d , т. е. вероятность принятия гипотезы H_{3-i} при распределении наблюдений, соответствующих гипотезе H_i ;

$$\Pi_\lambda(k) = \sum_{0 \leq i \leq k} \frac{\lambda^k e^{-\lambda}}{k!}$$

— функция распределения закона Пуассона $\text{Pois}(\lambda)$; $\Phi(x)$ — функция распределения стандартного нормального закона $\mathcal{N}(0, 1)$; $\varkappa_\gamma$ — квантиль уровня γ распределения $\mathcal{N}(0, 1)$.

Обозначим $\lambda = N/2^m$, $\lambda_i = Np_i$ при $i = 1, 2$. Значения λ и $\lambda_{1,2}$ связаны неравенствами

$$\lambda = \lambda_1(1 - 2^{-2m}) < \lambda_1 < \lambda_2 = 2(1 - 2^{-m-1})\lambda < 2\lambda,$$

т. е. λ_1 практически совпадает с λ при не очень малых m , а λ_2 примерно в 2 раза больше значений λ_1 и λ .

В следующей теореме получены допредельные оценки вероятностей ошибок критериев с использованием пуассоновского приближения для распределений ν'_R (п. 1), нормального приближения (п. 2). Точность первого приближения лучше при малых значениях λ , второго — при больших. В п. 2 выписана явная оценка объёма материала.

Теорема 2. Пусть $R \in \{3, 4\}$. Тогда:

1) для любого $c \in \mathbb{N}_0$ справедливы оценки

$$|\alpha_1(d'_R) - (1 - \Pi_{\lambda_1}(c))| \leq \lambda_1^2/N, \quad |\alpha_2(d'_R) - \Pi_{\lambda_2}(c)| \leq \lambda_2^2/N \leq 4\lambda^2/N;$$

2) при пороге

$$c = c_\alpha = \lambda_1 + \varkappa_{1-\alpha} \sqrt{\lambda_1(1 - p_1)} = \lambda_1 + \varkappa_{1-\alpha} \sqrt{\frac{\lambda_1(1 - 2^{1-m} - 2^{-2m})}{1 - 2^{-2m}}}$$

справедлива следующая оценка размера:

$$|\alpha_1(d'_R) - \alpha| \leq \frac{0,4693(1 - 2^{-m})}{\sqrt{\lambda_1(1 - 2^{-2m})(1 - 2^{1-m})}}, \quad 0 < \alpha < 1;$$

3) если при этом пороге

$$N = N_1(\alpha, \beta) = \left(\frac{\varkappa_{1-\alpha} \frac{\sqrt{1 - 2^{1-m}}}{1 - 2^{-m}} + \varkappa_{1-\beta}(1 - 2^{-m})\sqrt{2 - 2^{-m}}}{1 - 2^{-m} - \frac{1}{2^{2m} - 1}} \right)^2 2^m,$$

$$\text{то } |\alpha_2(d'_R) - \beta| \leq \frac{0,4693}{(1 - 2^{-m})\sqrt{2\lambda_1(1 - 2^{-2m})(1 - 2^{-m-1})}}.$$

Доказательство.

1) Неравенства п. 1 вытекают из оценки [11, с. 105] точности пуассоновской аппроксимации для распределения сумм независимых одинаково распределённых индикаторов и того, что величина 2λ чуть больше, чем λ_2 .

2) Воспользуемся следующим результатом, вытекающим из оценки [12] константы в неравенстве Берри — Эссеена: если $S_N = \xi_1 + \dots + \xi_N$ — сумма независимых одинаково распределённых случайных величин с нулевым средним и конечным третьим абсолютным моментом и $F(x)$ — функция распределения случайной величины $S_N/\sqrt{DS_N}$, то для отклонения $\delta_N = \sup_{x \in \mathbb{R}} |F(x) - \Phi(x)|$ выполнено неравенство

$$\delta_N \sqrt{N} \leq 0,4693 \frac{E|\xi_1|^3}{(D\xi_1)^{3/2}}.$$

Если при этом $\xi_1 \sim \eta - p$, где $\eta \sim \text{Be}(p)$, $0 < p < 1$, то $D\xi_1 = D\eta = pq$, $q = 1 - p$, $E|\xi_1|^3 = pq^3 + qp^3 = (p^2 + q^2)D\xi_1$, и тогда

$$\delta_N \leq 0,4693 \frac{1 - 2pq}{\sqrt{Npq}} < \frac{0,4693}{\sqrt{Npq}}. \quad (4)$$

При гипотезе H_1 выражение под корнем в оценке (4) для распределения ν_R равно

$$Np_1(1 - p_1) = \frac{\lambda}{1 - 2^{-m}} \left(1 - \frac{2^{-m}}{1 - 2^{-m}} \right) = \frac{\lambda}{(1 - 2^{-m})^2} (1 - 2^{1-m}), \quad (5)$$

откуда, согласно (4), получаем оценку близости для вероятности

$$\alpha_1(d'_R) = P_1 \left\{ \frac{\nu'_R - \lambda_1}{\sqrt{Np_1(1 - p_1)}} > \varkappa_{1-\alpha} \right\}$$

и её нормального приближения $1 - \Phi(\varkappa_{1-\alpha}) = \alpha$.

3) При гипотезе H_2 выражение под корнем в оценке (4) равно

$$Np_2(1 - p_2) = \lambda 2(1 - 2^{-m-1})(1 - 2^{-m+1} + 2^{-2m}) = 2\lambda(1 - 2^{-m-1})(1 - 2^{-m})^2, \quad (6)$$

что соответствует оценке близости вероятности

$$\alpha_2(d'_R) = P_2 \left\{ \frac{\nu'_R - \lambda_2}{\sqrt{Np_2(1 - p_2)}} \leq x \right\}, \quad x = \frac{c_\alpha - \lambda_2}{\sqrt{Np_2(1 - p_2)}}$$

и её нормального приближения $\Phi(x)$.

Осталось показать, что $x = \varkappa_\beta = -\varkappa_{1-\beta}$. Обозначая здесь для краткости $t = 2^{-m}$ и учитывая (5) и (6), получаем

$$-x = \frac{2\lambda(1 - t/2) - \frac{\lambda}{1-t^2} - \varkappa_{1-\alpha} \frac{\sqrt{\lambda(1-2t)}}{1-t}}{(1-t)\sqrt{2\lambda(1-t/2)}} = \frac{\sqrt{\lambda}(2 - t - \frac{1}{1-t^2}) - \varkappa_{1-\alpha} \frac{\sqrt{1-2t}}{1-t}}{(1-t)\sqrt{2-t}}.$$

Это выражение равно $\varkappa_{1-\beta}$ при $\sqrt{\lambda} = \frac{\varkappa_{1-\alpha}\sqrt{1-2t}/(1-t) + \varkappa_{1-\beta}(1-t)\sqrt{2-t}}{1-t-t^2/(1-t^2)}$, что выполнено при $N = N_1(\alpha, \beta)$. ■

Заметим, что абсолютная точность оценок погрешностей в пп.2–3 обратно пропорциональна величине $\sqrt{\lambda}$, которая при малых значениях 2^{-m} близка к величине $\varkappa_{1-\alpha} + \varkappa_{1-\beta}\sqrt{2}$. При $m \rightarrow \infty$ достаточное для атаки количество запрашиваемых пар выходных блоков асимптотически равно

$$N_1^*(\alpha, \beta) = (\varkappa_{1-\alpha} + \varkappa_{1-\beta}\sqrt{2})^2 2^m.$$

Это соответствует оценкам $q(m) = O(2^{m/2})$ числа запросов при атаках на $R = 3, 4$ раунда [7, с. 225–226], поскольку при таких $q(m)$ общее количество слагаемых в суммах статистик $\nu_{3,4}$ равно $\binom{q(m)}{2} = O(2^m)$.

При малых значениях 2^{-m} порог в п. 3 теоремы 2 близок к $c^* = \lambda + \varkappa_{1-\alpha}\sqrt{\lambda}$. Тогда при $\alpha = \beta$, $N = N_1^*(\alpha, \alpha)$ имеем

$$\varkappa_{1-\alpha} = \frac{\sqrt{\lambda}}{1 + \sqrt{2}} = \sqrt{\lambda}(\sqrt{2} - 1), \quad \alpha = \Phi(-\sqrt{\lambda}(\sqrt{2} - 1)), \quad c^* = \lambda\sqrt{2}. \quad (7)$$

При пороге $c = \lfloor \lambda\sqrt{2} \rfloor$ (целое число, ближайшее к c^*), согласно п. 1 теоремы, получаем, что вероятности ошибок критерия (3) будут близки к $\alpha_1^*(d) = 1 - \Pi_\lambda(c)$ и $\alpha_2^*(d) = \Pi_{2\lambda}(c)$. В табл. 2 приведены примеры вычисления этих величин.

Т а б л и ц а 2

Грубая оценка α вероятностей ошибок атак и оценки вероятностей ошибок, близкие к истинным

λ	$\varkappa_{1-\alpha}$	α	c	$\alpha_1^*(d)$	$\alpha_2^*(d)$
4	0,82	0,204	6	0,11	0,31
9	1,24	0,107	13	0,073	0,142
16	1,65	0,048	23	0,026	0,06
25	2,07	0,019	35	0,022	0,016
36	2,48	0,006	51	0,007	0,005
49	2,89	0,0018	69	0,0027	0,0018
64	3,31	$4,6 \cdot 10^{-4}$	91	$5,8 \cdot 10^{-4}$	$3,5 \cdot 10^{-4}$

Из табл. 2 видно, что для критериев, получаемых посредством нормальной аппроксимации, формулы (7) позволяют неплохо оценивать среднее значение двух вероятностей ошибок.

3. Эксперименты при $R = 4$ в двух моделях наблюдений

Параметрами, определяющими проведение экспериментов, в обеих моделях наблюдений были длины полублока m и различные значения λ (первая часть табл. 3 — два столбца до двойной вертикали).

Во второй части табл. 3 представлены теоретические значения оценок $\alpha_1^* = 1 - \Pi_\lambda(c)$, $\alpha_2^* = \Pi_{2\lambda}(c)$ (при некоторых λ они были вычислены в табл. 2) вероятностей ошибок критерия (3) с порогом $c(\lambda) = \lfloor \lambda\sqrt{2} \rfloor$ (см. формулу (7)) в модели независимых двублочных текстов при объёме выборки $N_1^* = \lambda 2^m$. Количество A экспериментов выбиралось с учётом порядка малости величин α_i^* .

В третьей части табл. 3 представлены объём материала N_1^* в модели независимых двублочных текстов и полученные эмпирические вероятности ошибок $\hat{\alpha}_1(d'_4)$.

В четвёртой части табл. 3 приведены условие и результаты экспериментов в модели наблюдений Патарина, т.е. при q -блочной выборке из одной подстановки: количество запросов $q = q(\lambda) = \lceil \sqrt{2^{m+1}} \lambda \rceil = \lceil \sqrt{2} \lambda 2^{m/2} \rceil$, которое даёт общее число слагаемых $\binom{q}{2} \approx \lambda 2^m$ индикаторов в сумме для статистики ν_4 , близкое к N_1^* , и полученные эмпирические вероятности ошибок $\hat{\alpha}_1(d_4)$. Здесь применялся критерий с тем же порогом:

$$d_4 : \nu_4 > \lceil \lambda \sqrt{2} \rceil \implies \text{принимается } H_2. \quad (8)$$

Т а б л и ц а 3

Теоретические и эмпирические характеристики атак различения на 4 раунда схемы Луби — Ракова

$2m$	λ	α_1^*	α_2^*	A	N_1^*	$\hat{\alpha}_1(d_4)$	$\hat{\alpha}_2(d_4)$	$q(\lambda)$	$\hat{\alpha}_1(d_4)$	$\hat{\alpha}_2(d_4)$
16	25	0,022	0,016	10^3	6,4 E3	0,026	0,01	113	0,032	0,03
20	25	0,022	0,016	10^3	2,5 E4	0,034	0,009	226	0,039	0,01
24	4	0,11	0,31	10^2	1,6 E4	0,22	0,19	181	0,25	0,15
24	9	0,073	0,142	10^3	3,6 E4	0,132	0,105	272	0,13	0,09
28	4	0,11	0,31	50	6,5 E4	0,16	0,24	362	0,12	0,28
32	4	0,11	0,31	50	2,6 E5	0,22	0,16	724	0,18	0,20
36	4	0,11	0,31	50	1,0 E6	0,14	0,24	1,4 E3	0,16	0,20
40	4	0,11	0,31	50	4,2 E6	0,26	0,22	2,8 E3	0,10	0,14
40	9	0,073	0,142	10^2	9,4 E6	0,12	0,05	4,3 E3	0,17	0,10
44	4	0,11	0,31	20	1,6 E7	0,2	0,15	5,7 E3	0,3	0,25
48	4	0,11	0,31	20				1,1 E4	0,1	0,15
52	4	0,11	0,31	20				2,3 E4	0,2	0,25

Статистика ν_4 вычислялась так: при фиксированном $x^1 = 0$ (здесь и далее отождествляем векторы из $\mathbb{Z}_2^m$ с числами, двоичной записью которых являются векторы) полагаем $x_i^0 = i$. Поэтому здесь

$$\nu_4 = \sum_{0 \leq i < j \leq q-1} \mathbb{I}\{Y_i^{\text{left}} \oplus i = Y_j^{\text{left}} \oplus j\}, \quad Y_i = F(i, 0), \quad 0 \leq i \leq q-1.$$

Как отмечается в [7, с. 225], для вычисления ν_R за $O(q)$ операций можно сохранять значения $Y_i^{\text{left}} \oplus x_i^0$ и «считать коллизии». Подробнее: 1) инициализируем нулями все элементы массива $\text{coll}[]$ длины 2^m , полагаем $i = 0$; 2) увеличиваем на 1 содержимое элемента с индексом $Y_i^{\text{left}} \oplus i$; 3) если $i < q-1$, то увеличиваем i на 1 и повторяем п. 2, в противном случае завершаем подсчёт числа коллизий. Тогда верно равенство

$$\nu_4 = \sum_{0 \leq y \leq 2^m-1} |\{ \{i, j\} \subset \{0, \dots, q-1\} : Y_i^{\text{left}} \oplus i = Y_j^{\text{left}} \oplus j = y \}| = \sum_{0 \leq y \leq 2^m-1} \binom{\text{coll}[y]}{2}.$$

Временная сложность эксперимента $O(2^m)$ в первой модели быстро растёт, что ограничило длину блока 48 битами. Снижение сложности во второй модели наблюдений до $O(q) = O(2^{m/2})$ позволило провести эксперименты до длины блока 52 битов включительно.

Таким образом, можно сделать следующие выводы:

1. Части 2 и 3 табл. 3 показывают хорошее¹ согласие эмпирических вероятностей ошибок в модели независимых двублочных текстов и теоретических: отношения

¹При неверных расчётах эмпирические вероятности могут отличаться от теоретических на несколько порядков.

$\hat{\alpha}_i(d'_4)/\alpha_i^*$ в основном лежат в пределах от 0,5 до 2. Это подтверждает правильность расчётов в теоремах 1 и 2.

2. Аналогично части 2 и 4 табл. 3 показывают хорошее согласие эмпирических вероятностей ошибок в модели Патарина и теоретических оценок, рассчитанных в модели независимых двублочных текстов. Это говорит о том, что зависимость между слагаемыми статистики ν_4 , имеющаяся в модели Патарина, не привела к существенным отклонениям её распределения от распределения аналогичной статистики ν'_4 в значительно более простой модели наблюдений.

3. Разработанные подходы и полученные результаты дают основу для построения и расчёта атак на большее количество раундов ЛР-схемы.

4. Обзор связей с другими работами по разностному анализу

Статистики Патарина могут быть записаны в виде

$$\nu_3 = \sum_{1 \leq i < j \leq q} \mathbb{I}\{\Delta Y_{ij}^{\text{left}} = \Delta x_{ij}^{\text{right}}\}, \quad \nu_4 = \sum_{1 \leq i < j \leq q} \mathbb{I}\{\Delta Y_{ij}^{\text{left}} = \Delta x_{ij}^{\text{left}}\},$$

т.е. являются разностными. Точнее, они являются функциями от усечённых разностей [13], атаки производятся при нулевой левой (правой) входной полуразности при $R = 3$ ($R = 4$), подсчитывается число совпадений левой выходной полуразности с правой (левой) входной. Но, несмотря на одновременное развитие разностного анализа с 1990 г., даже замечаний о связи с ним в известных работах Патарина не обнаружено, в том числе в итоговой монографии [8]. Таким образом, эта ветвь разностного анализа развивалась отдельно от «основного дерева». Что касается его представителей, то нам известны лишь две работы [14, 15], в которых дан некоторый анализ статистик Патарина с точки зрения разностного метода.

В русскоязычной литературе, насколько известно авторам, разностный анализ схем Луби — Ракова впервые был начат в [16], где установлено, что схема является марковским шифром, т.е. последовательность раундовых разностей образует однородную цепь Маркова. Был найден блочный вид матрицы переходных вероятностей разностей за один раунд, обозначаемой далее через $\mathbb{P} = \mathbb{P}(M)$, $M = 2^m$. Из теоремы о блочном умножении матриц [17, с. 21] следует, что все степени $\mathbb{P}$ имеют такое же разбиение на M^2 клеток размера M , как и $\mathbb{P}(M)$. В [16] найден блочный вид матриц $\mathbb{P}^2$ и $\mathbb{P}^4$; доказательства этих результатов можно найти в [18]. Заметим, что из описания [16, формула (4)] элементов $\mathbb{P}^4$ при нулевой правой входной полуразности $\Delta X^{\text{right}} = 0$ может быть выведено равенство $p_2 = M \frac{2M^2 - M}{M^4} = M^{-1}(2 - M^{-1})$ теоремы 1 при $R = 4$.

В [16] также построена последовательная атака различения (критерий Вальда) на 4 раунда в модели независимых двублочных текстов с входными полуразностями $\Delta X_t^{\text{right}} = 0$. Получены оценки средней длины материала при гипотезах $H_{1,2}$ и больших M

$$T_1(M) = \frac{(1 - 2\alpha) \ln((1 - \alpha)/\alpha)}{1 - \ln 2} M, \quad T_2(M) = \frac{(1 - 2\alpha) \ln((1 - \alpha)/\alpha)}{2 \ln 2 - 1} M$$

для критерия с расчётными вероятностями ошибок α . В частности, при $\alpha = 0,1$ эти величины равны соответственно $5,72M$ и $4,55M$, что примерно в 2 раза меньше значения $9M$, при котором, согласно табл. 2, вероятности ошибок атаки на основе аналога статистики Патарина будут близки к 0,107 и 0,073.

В [16] эксперименты проводились для длин блока от 12 до 44 битов; в данной работе в модели экспериментов с одной подстановкой «потолок» длины блоков повышен

до 52 битов благодаря малой временной сложности вычисления статистик Патарина. Результаты каких-либо других статистических экспериментов с ЛР-схемой нам неизвестны; их отсутствие можно частично объяснить недостаточной теоретической проработанностью материала зарубежными криптографами — отсутствием формул для порога критерия, обеспечивающего заданный размер; отсутствием явных оценок объёма материала.

ЛИТЕРАТУРА

1. *Luby M. and Rackoff C.* How to construct pseudorandom permutations from pseudorandom functions // SIAM J. Comput. 1988. V. 17. P. 373–386.
2. *Tsaregorodtsev K. D.* Format-preserving encryption: a survey // Мат. вопр. криптогр. 2022. Т. 13. Вып. 2. С. 133–153.
3. *Bellare M., Hoang V. T., and Tessaro S.* Message-recovery attacks on Feistel-based format preserving encryption // Proc. CCS'16. Vienna, Austria, 2016. P. 444–455.
4. *Bellare M., Ristenpart T., Rogaway P., and Stegers T.* Format-preserving encryption // LNCS. 2009. V. 5867. P. 295–312.
5. *Lee J., Koo B., Roh D., et al.* Format-preserving encryption algorithms using families of tweakable blockciphers // LNCS. 2015. V. 8949. P. 132–159.
6. *Patarin J.* New results on pseudorandom permutation generators based on the DES scheme // LNCS. 1992. V. 576. P. 301–312.
7. *Patarin J.* Generic attacks on Feistel schemes // LNCS. 2001. V. 2248. P. 222–238.
8. *Nachev V., Patarin J., and Volte E.* Feistel Ciphers: Security Proofs and Cryptanalysis. Cham: Springer, 2017. 309 p.
9. *Денисов О. В.* Атаки различения на блочные шифрсистемы по разностям двублочных текстов // Прикладная дискретная математика. 2020. № 48. С. 43–62.
10. *Денисов О. В.* Многомерный спектральный критерий для проверки гипотез о случайных подстановках // Мат. вопр. криптогр. 2023. Т. 14. Вып. 3. С. 85–106.
11. *Боровков А. А.* Теория вероятностей. М.: Эдиториал УРСС, 1999. 472 с.
12. *Шевцова И. Г.* Об абсолютных константах в неравенстве Берри — Эссеена и его структурных и неравномерных уточнениях // Информатика и ее примен. 2013. Т. 7. Вып. 1. С. 124–125.
13. *Knudsen L.* Truncated and higher order differentials // LNCS. 1995. V. 1008. P. 196–211.
14. *Knudsen L.* The security of Feistel ciphers with six rounds or less // J. Cryptology. 2002. V. 15. P. 207–222.
15. *Dunkelman O., Kumar A., Lambooi E., and Sanadhya S. K.* Cryptanalysis of Feistel-Based Format-Preserving Encryption. Cryptology ePrint Archive. 2020. Report 2020/1311. <https://eprint.iacr.org/2020/1311>.
16. *Денисов О. В.* Атака различения на четыре раунда шифра Люби — Ракофф по разностям двублочных текстов // Прикладная дискретная математика. Приложение. 2023. № 16. С. 32–35.
17. *Ланкастер П.* Теория матриц. М.: Наука, 1978. 280 с.
18. *Денисов О. В.* Спектральные атаки различения на схемы Луби — Ракова по независимым двублочным текстам // Мат. вопр. криптогр. 2024. Т. 15. Вып. 4. С. 23–42.

REFERENCES

1. *Luby M. and Rackoff C.* How to construct pseudorandom permutations from pseudorandom functions. SIAM J. Comput., 1988, vol. 17, pp. 373–386.

2. *Tsaregorodtsev K. D.* Format-preserving encryption: a survey. *Matematicheskie Voprosy Kriptografii*, 2022, vol. 13, iss. 2, pp. 133–153.
3. *Bellare M., Hoang V. T., and Tessaro S.* Message-recovery attacks on Feistel-based format preserving encryption. *Proc. CCS'16*, Vienna, Austria, 2016, pp. 444–455.
4. *Bellare M., Ristenpart T., Rogaway P., and Stegers T.* Format-preserving encryption. *LNCS*, 2009, vol. 5867, pp. 295–312.
5. *Lee J., Koo B., Roh D., et al.* Format-preserving encryption algorithms using families of tweakable blockciphers. *LNCS*, 2015, vol. 8949, pp. 132–159.
6. *Patarin J.* New results on pseudorandom permutation generators based on the DES scheme. *LNCS*, 1992, vol. 576, pp. 301–312.
7. *Patarin J.* Generic attacks on Feistel schemes. *LNCS*, 2001, vol. 2248, pp. 222–238.
8. *Nachev V., Patarin J., and Volte E.* Feistel Ciphers: Security Proofs and Cryptanalysis. Cham, Springer, 2017. 309 p.
9. *Denisov O. V.* Ataki razlicheniya na blochnye shifrsistemy po raznostyam dvublochnykh tekstov [Distinguishing attacks on block ciphers by differentials of two-block texts]. *Prikladnaya Diskretnaya Matematika*, 2020, no. 48, pp. 43–62. (in Russian)
10. *Denisov O. V.* Mnogomernyy spektral'nyy kriteriy dlya proverki gipotez o sluchaynykh podstanovkakh [Multidimensional spectral criterion for testing hypotheses on random permutations]. *Matematicheskie Voprosy Kriptografii*, 2023, vol. 14, iss. 3, pp. 85–106. (in Russian)
11. *Borovkov A. A.* Teoriya veroyatnostey [Probability Theory]. Moscow, URSS, 1999. 472 p. (in Russian)
12. *Shevtsova I. G.* Ob absolyutnykh konstantakh v neravenstve Berri — Esseena i ego strukturnykh i neravnomernykh utochneniyakh [On absolute constants in the Berry-Esseen inequality and its structural and uneven refinements]. *Informatika i ee Primeneniya*, 2013, vol. 7, no. 1, pp. 124–125. (in Russian)
13. *Knudsen L.* Truncated and higher order differentials. *LNCS*, 1995, vol. 1008, pp. 196–211.
14. *Knudsen L.* The security of Feistel ciphers with six rounds or less. *J. Cryptology*, 2002, vol. 15, pp. 207–222.
15. *Dunkelman O., Kumar A., Lambooi E., and Sanadhya S. K.* Cryptanalysis of Feistel-Based Format-Preserving Encryption. *Cryptology ePrint Archive*, 2020, Report 2020/1311. <https://eprint.iacr.org/2020/1311>.
16. *Denisov O. V.* Ataka razlicheniya na chetyre raunda shifra Lyubi — Rakoff po raznostyam dvublochnykh tekstov [Distinguishing attack on four rounds of the Luby — Rackoff cipher by differentials of two-block texts]. *Prikladnaya Diskretnaya Matematika. Prilozhenie*, 2023, No. 16, pp. 32–35. (in Russian)
17. *Lankaster P.* Theory of Matrices. N.Y.; London, Academic Press, 1969.
18. *Denisov O. V.* Spektral'nye ataki razlicheniya na skhemy Lubi — Rakova po nezavisimym dvublochnym tekstam [Spectral attacks on Luby-Rackoff schemes based on independent two-block texts]. *Matematicheskie Voprosy Kriptografii*, 2024, vol. 15, iss. 4, pp. 23–42. (in Russian)