МАТЕМАТИЧЕСКИЕ МЕТОДЫ КРИПТОГРАФИИ

УДК 519.23

АТАКИ РАЗЛИЧЕНИЯ НА БЛОЧНЫЕ ШИФРСИСТЕМЫ ПО РАЗНОСТЯМ ДВУБЛОЧНЫХ ТЕКСТОВ

О. В. Денисов

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

Предложена модель наблюдений (случайные двублочные тексты, шифруемые на независимых случайных ключах), в которой разностные атаки различения полностью соответствуют общепринятым схемам их статистического расчёта. В этой модели получены нижние границы и асимптотические оценки объёма материала мультиразностных атак различения. Показано, что материала объёма $O(1/p_{\max})$ недостаточно для успешной атаки при малых значениях $p_{\max}$ — максимальной вероятности перехода разностей. Проведены вычислительные и статистические эксперименты для марковских моделей шифрсистемы SmallPresent с длиной блока до 28 бит.

Ключевые слова: *мультиразностный анализ, атака различения, марковский шифр, SmallPresent.*

DOI 10.17223/20710410/48/5

DISTINGUISHING ATTACKS ON BLOCK CIPHERS BY DIFFERENTIALS OF TWO-BLOCK TEXTS

O. V. Denisov

Innovative Telecommunication Technologies, LLC, Moscow, Russia

E-mail: denisovOleg@yandex.ru

We present the observation model (random two-block texts encrypted on random independent keys) such that differential distinguishing attacks completely correspond to the generally accepted schemes of their statistical calculation. In this model, we get low bounds for data complexity of multiple differential distinguishing attacks. Let n be the block size, $M = 2^n - 1$, $D(a)$ be a set of high-probabilily output differences at a fixed input difference $a \in \mathbb{Z}_2^n$ in R-round encryption, $P_1 = \dfrac{|D(a)|}{M} < P_2 = \sum\limits_{b \in D(a)} p_{a,b}^{(R)} \leqslant \dfrac{1}{2}$. Let $\nu(D(a))$ be the number of these differences appearances. Suppose the attack is based on statistics $\nu(D(a))$ and have error probabilities $\alpha, \beta \in (0, 1/2)$. Then attack needs $N_{1,a*} > N_{2,a*} = \dfrac{4P_1(1-P_1)(1-\alpha-\beta)^2}{(P_2-P_1)^2}$ two-block texts. In particulary, in the case of $D(a) = \{b\}$, $1/M < p_{a,b}^{(R)} \leqslant 1/2$, we have low bound $N_{2,a,b*} = \dfrac{4\,(1-1/M)\,(1-\alpha-\beta)^2}{M\left(p_{a,b}^{(R)} - 1/M\right)^2}$. Consequently, the frequently used estimate $O(1/p_{\max})$ for data complexity is not enough for successful attack at small values

of $(p_{\max} - 1/M)$, where $p_{\max}$ is the maximal transition probability of differentials. We also get asymptotic estimates for data complexity of the most powerful criterion (MPC) in the case of converging hypotheses. Let $\rho(a,R) = \left|\mathbb{P}_a^{(R)} - \frac{1}{M}\mathbf{1}\right|$ is the Euclidean distance from the row $\mathbb{P}_a^{(R)}$ of transition probabilities matrix to the uniform distribution vector. Suppose $\max\limits_{b\neq 0}|p_{a,b}^{(R)}M - 1| \to 0$ in the series scheme with a growing R, $N(R) \sim N_{\mathrm{MPC}}(R,a) = \dfrac{(\varkappa_{1-\alpha} + \varkappa_{1-\beta})^2}{M\rho^2(a,R)}$. Then MPC errors tend to α, β (for some criteria bounds). We make experiments with Markov models of SmallPresent cipher for block size up to $n = 28$ bits and $R = 10$ rounds: we find inpute differences that minimize $N_{\mathrm{MPC}}(R,a)$ and we calculate empirical error probabilities for this number of texts.

Keywords: *multiple differential cryptanalysis, distinguishing attack, two-block texts model, Kullback — Leibler divergence, converging hypotheses, capacity, Markov cipher, SmallPresent.*

Введение

Разностный анализ является одним из основных методов криптоанализа блочных шифрсистем. Предложенный в [1], он получил теоретическое развитие в работах [2, 3] для обоснования стойкости марковских шифров и в настоящее время является стандартом при анализе стойкости шифрсистем к атакам по известным входным и выходным блокам. В [4] развиваются мультиразностные атаки восстановления ключа, учитывающие статистическое поведение многих пар «входная/выходная разность». В [5] для шифров с малой длиной блока для ранжирования вариантов последнего раундового ключа при их опробовании используется статистика логарифма отношения правдоподобий (ЛОП); этот подход позволяет учитывать потенциально все выходные разности при фиксированной входной.

Часто исследуются атаки на ключ k_R последнего (R-го) раунда шифрования, осуществляемые так: пусть известно значение $p = \mathsf{E}_k p(k)$ вероятности некоторого события, усреднённой по всем ключам шифрования k шифрсистемы с $R-1$ раундами. Пусть известно также значение q вероятности этого события, рассчитанное в ситуации применения случайной равновероятной подстановки на блоках. Тогда опробуются все варианты k_R, имеющиеся выходные блоки расшифровываются на текущем варианте и считается, что вероятность события близка к p при истинном варианте и близка к q при ложном. В этой модели объём требуемого материала рассчитывается как объём материала критерия проверки гипотез о параметре распределения Бернулли:

$$H_2 : \xi \sim \mathrm{Be}(p) \;\text{ против }\; H_1 : \xi \sim \mathrm{Be}(q) \tag{1}$$

при заданных вероятностях α, β ошибок критерия. Таким образом, для атаки восстановления ключа k_R достаточно построить *атаку различения* на $R-1$ раунд шифрсистемы, т. е. критерий для проверки гипотезы, что при шифровании применялась данная шифрсистема, против гипотезы о том, что применялась случайная подстановка.

Отметим, что при таком расчёте используются эвристические предположения [2, 4] о близости, а фактически о совпадении: 1) вероятностей $p(k)$ события при всех k с величиной p (гипотеза стохастической эквивалентности); 2) вероятностей события при R-раундовом шифровании и одноразундовом расшифровании на ложном раундовом ключе с величиной q (гипотеза о ложных ключах). Заметим, что [5] — единственная

известная автору работа, в которой при расчёте атаки отказались от предположения 2, использовав вместо него аналог предположения 1.

Кроме того, при расчёте разностных атак различения получили широкое распространение нестрогие оценки объёма материала вида $O(1/p_{\max})$,

где $$p_{\max} = \max_{a,b \in \{1,\dots,M\}} p_{a,b}^{(R)}, \quad M = 2^n - 1;$$

$p_{a,b}^{(R)}$ — средняя (по ключам) вероятность перехода входной разности a в выходную разность b; n (битов) — длина блока.

Возникает вопрос: а существует ли непротиворечивая вероятностная модель наблюдений, в которой разностная атака различения будет в точности приводить к статистической модели (1)? В данной работе предлагается такая модель наблюдений — «независимые двублочные тексты». Условия модели отличаются от традиционных и фактически означают, что каждая пара входных блоков шифруется на своём случайно выбираемом ключе k либо своей случайно равновероятно выбираемой подстановке при альтернативной гипотезе. Это близко к ситуации, когда криптоаналитик наблюдает много коротких открытых сообщений и результатов их шифрования. Тогда непосредственно с помощью атаки различения может проверяться гипотеза о том, что шифрование осуществляется заданной системой. Заметим, что ограничение атакой различения позволяет пока оставить в стороне сложный вопрос о соответствии реальности гипотезы о ложных ключах, возникающей в атаке восстановления раундового ключа.

В предлагаемой модели без использования эвристических предположений возможно получение оценок вероятностей ошибок разностных атак различения, основанных на статистиках [2, 4, 5]. При этом, как и в перечисленных работах, предполагаются известными некоторые строки матрицы $\mathbb{P}^{(R)} = \|p_{a,b}^{(R)}\|$ вероятностей переходов разностей за R раундов.

Далее структура работы следующая. В п. 1 вводится модель двублочных текстов. В п. 2 с помощью дивергенции Кульбака — Лейблера получены нижние оценки объёма материала в статистической модели (1). Показано, что при значениях $p_{\max}$, близких к $1/M$, использование оценок вида $O(1/p_{\max})$ как верхних оценок ошибочно. В п. 3 полученные результаты обсуждаются с точки зрения развития статистических вопросов разностного анализа.

В п. 4 на основе предельной теоремы А. С. Амбросимова [6] получены асимптотические формулы для вероятностей ошибок и объёма материала $N_{\text{КОП}}(\cdot)$ критерия отношения правдоподобий (КОП), основанного на частотах всех выходных разностей. Показано, что асимптотически при сближении гипотез для атак различения с фиксированной входной разностью a значение $N_{\text{КОП}}(\cdot)$ зависит от шифрсистемы через величину $\rho(a, R)$ — евклидово расстояние от строки $\mathbb{P}_a^{(R)}$ матрицы переходных вероятностей до вектора вероятностей равномерного распределения.

В п. 5 проведены экспериментальные расчёты атак на марковские модели шифрсистем семейства SmallPresent с размером блока $n \leqslant 28$ битов. Найдены минимальные и максимальные значения $\rho(a, R)$ при $R \leqslant 10$ по некоторым $n/2$ входным разностям a. Проведены статистические эксперименты при вероятностях ошибок $\alpha = \beta = 0{,}1$ и объёме материала КОП не более 2^{30} разностей, что соответствует $R = 9$ в большинстве случаев.

1. Модель двублочных текстов, шифруемых на независимых ключах

Пусть на алфавите $\mathcal{X}$ входных и выходных блоков задана групповая операция $\odot$; $0 \in \mathcal{X}$ — нейтральный элемент относительно неё; $M = |\mathcal{X}| - 1$ — мощность множества $\mathcal{X}' = \mathcal{X} \setminus \{0\}$. Через $S(\mathcal{X})$ обозначим множество всех подстановок на множестве $\mathcal{X}$.

Рассматривается итеративный R-раундовый алгоритм блочного шифрования с независимым выбором случайных ключей (при известном и фиксированном их распределении) для каждого двублочного текста. Наблюдаются *входные разности* $\Delta X_t = X_t^{-1} \odot X_t^*$ случайных независимых входных пар блоков (X_t, X_t^*) и *выходные разности* $\Delta Y_t = Y_t^{-1} \odot Y_t^*$ соответствующих им шифртекстов (Y_t, Y_t^*), причём

$$\text{пары } (X_t, X_t^*) \text{ независимы и одинаково распределены на } \mathcal{X}^2 \tag{2}$$

при $1 \leqslant t \leqslant N$.

Гипотеза H_1 заключается в том, что шифртексты получены в результате применения подстановок, выбираемых независимо равновероятно из $S(\mathcal{X})$. Легко показать (см., например, [7]), что

$$\mathsf{P}_1\left\{\Delta Y_1 = b \,\middle|\, \Delta X_1 = a\right\} = \frac{1}{M}$$

для всех ненулевых разностей $a, b \in \mathcal{X}'$. Здесь и далее через P_i, E_i, D_i, $\alpha_i(N)$ обозначаем соответственно вероятностное распределение, математическое ожидание и дисперсию статистик, вероятности ошибок критериев при гипотезе H_i, $i = 1, 2$.

Гипотеза H_2 заключается в том, что шифртексты получены в результате применения случайных подстановок, распределение которых соответствует алгоритму шифрования и распределению ключа. Здесь считаются известными элементы

$$p_{a,b}^{(R)} = \mathsf{P}_2\left\{\Delta Y_1 = b \,\middle|\, \Delta X_1 = a\right\}$$

матрицы вероятностей переходов ненулевых разностей $a, b \in \mathcal{X}'$. Нулевой входной разности при обеих гипотезах соответствует нулевая выходная разность, поэтому этот переход не рассматривается.

Требуется построить критерий для проверки гипотез H_1 и H_2, вероятности ошибок $\alpha_1(N)$ и $\alpha_2(N)$ которого близки к заданным величинам α и β соответственно.

Из условий гипотез и предположения (2) вытекает, что при каждой гипотезе для всех $1 \leqslant t \leqslant N$

$$\text{пары } (\Delta X_t, \Delta Y_t) \text{ независимы и одинаково распределены,} \tag{3}$$

причём их распределение не зависит от ключа, поскольку получено усреднением по ключам. Это главный момент, позволяющий уйти от использования гипотезы о стохастической эквивалентности. В частности, статистика ЛОП [5] здесь является суммой независимых одинаково распределённых случайных величин, распределение которых при гипотезе H_2 получено усреднением по ключам шифрсистемы. Из (3) также следует, что распределение вектора частот переходов разностей будет полиномиальным. В стандартной модели наблюдений это обычно является эвристическим предположением (как, например, в [5, п. 3.1]).

Далее, распределение пары блоков (2) определяет вероятности распределения входных разностей

$$p_a = \mathsf{P}\{\Delta X_t = a\}, \quad a \in \mathcal{X}'.$$

Обозначим через

$$\Delta_0 = \{a \in \mathcal{X}' : p_a > 0\}$$

множество всех возможных входных разностей. Для каждой входной разности $a \in \Delta_0$ выделим множество $D(a) \subset \mathcal{X}'$, состоящее из некоторых высоковероятных (т. е. $p^{(R)}_{a,b} > 1/M$) выходных разностей, и положим

$$\mathbf{D} = (D(a),\, a \in \Delta_0)\,.$$

В этих обозначениях основная статистика работы [4], равная числу выделенных высоковероятных переходов, имеет вид

$$\nu(\mathbf{D}) = \sum_{1 \leqslant t \leqslant N} \mathbb{I}\{\Delta Y_t \in D(\Delta X_t)\}\,.$$

Статистика $\nu(b) = \sum\limits_{1 \leqslant t \leqslant N} \mathbb{I}\{\Delta Y_t = b\}$ работы [2] является частным случаем этой статистики в ситуации, когда $\Delta_0 = \{a\}$, т. е. входная разность фиксирована ($p_c = \mathbb{I}\{c = a\}$, $c \in \mathcal{X}'$), $D(a) = \{b\}$.

Введём величины

$$P_2 = P_2(\mathbf{D}) = \sum_{a \in \mathcal{X}',\, b \in D(a)} p_a p^{(R)}_{a,b}\,, \quad P_1 = P_1(\mathbf{D}) = \sum_{a \in \mathcal{X}'} p_a \frac{|D(a)|}{M};$$

из построения $\mathbf{D}$ следует, что $P_2 > P_1$. Если $\Delta_0 = \{a\}$, то

$$P_2 = \sum_{b \in D(a)} p^{(R)}_{a,b}\,, \quad P_1 = \frac{|D(a)|}{M}.$$

Из (3) следует, что события $A_t = \{\Delta Y_t \in D(\Delta X_t)\}$ независимы, $1 \leqslant t \leqslant N$, и в наших обозначениях для параметров распределения входных разностей их вероятности равны

$$\begin{aligned} \mathsf{P}_i(A_t) &= \mathsf{P}_i\left\{\exists a \in \mathcal{X}'\ (\Delta X_t = a,\ \Delta Y_t \in D(a))\right\} = \\ &= \sum_{a \in \mathcal{X}',\, b \in D(a)} \mathsf{P}_i\left\{\Delta X_t = a\right\} \mathsf{P}_i\left\{\Delta Y_t = b \,\middle|\, \Delta X_t = a\right\}, \end{aligned}$$

что равно P_i при гипотезе H_i, $i = 1, 2$. Следовательно, при гипотезе H_i статистика числа высоковероятных переходов имеет биномиальное распределение с соответствующей вероятностью успеха

$$\nu(\mathbf{D}) \sim \mathrm{Bin}(N, P_i),\ i = 1, 2.$$

В других моделях наблюдений это обычно является явно или неявно вводимым эвристическим предположением.

В частности, при $\Delta_0 = \{a\}$ и $D(a) = \{b\}$ отсюда вытекает, что в модели двублочных текстов (2) задача различения гипотез H_1, H_2 по разностям эквивалентна проверке гипотез (1) при $p = P_2 = p^{(R)}_{a,b}$, $q = P_1 = 1/M$.

2. Нижние оценки объёма материала мультиразностных критериев

Получим нижнюю оценку объёма материала любых критериев, основанных на функциях от индикаторов попадания разностей в выделенные множества. Это будет сделано с помощью *дивергенции Кульбака — Лейблера* (её также называют информацией в пользу гипотезы $\mathrm{Be}(p)$ против $\mathrm{Be}(q)$ [8, с. 15]):

$$K(p, q) = p \ln \frac{p}{q} + (1 - p) \ln \frac{1 - p}{1 - q}, \quad p, q \in (0, 1).$$

Из определения следует, что $K(p,q) = K(1-p, 1-q)$. Известно (см., например, [9, с. 80]), что

$$K(p,q) \geqslant 2(p-q)^2 \text{ при всех } p, q \in (0,1), \tag{4}$$

поэтому $K(p,q) \geqslant 0$ и равенство достигается только при $p = q$.

Получим неравенства для дивергенции (заметим, что нижняя оценка в (5) улучшает (4) при всех $0 < q < p \leqslant 1/2$) и нижнюю оценку объёма материала критериев проверки гипотез о параметре распределения Бернулли.

Лемма 1.

1. Если при объёме выборки N вероятности ошибок критерия проверки гипотез (1) равны $\alpha_1(N) = \alpha$ и $\alpha_2(N) = \beta$, то

$$N \geqslant N_*(p,q,\alpha,\beta) = \max\left\{\frac{K(1-\alpha,\beta)}{K(p,q)},\ \frac{K(1-\beta,\alpha)}{K(q,p)}\right\}.$$

2. При всех $0 < q < p \leqslant 1/2$

$$\frac{(p-q)^2}{2p(1-p)} < K(p,q) < \frac{(p-q)^2}{2q(1-q)}; \tag{5}$$

$$K(p,q) > K(q,p). \tag{6}$$

Доказательство.

1. Воспользуемся оценкой Кульбака: из [8, теорема 3.1, с. 86] имеем

$$N\,K(p,q) \geqslant K(\alpha, 1-\beta) = K(1-\alpha,\beta),$$
$$N\,K(q,p) \geqslant K(\beta, 1-\alpha) = K(1-\beta,\alpha),$$

откуда сразу вытекает оценка п. 1.

2. Докажем оценку (5). Рассмотрим дивергенцию как функцию первого аргумента: $K(x,q) = x\ln\dfrac{x}{q} + (1-x)\ln\dfrac{1-x}{1-q}$; заметим, что (см., например, [10, с. 98])

$$K'(x,q) = \ln\frac{x}{q} - \ln\frac{1-x}{1-q},$$
$$K(q,q) = 0,\ K'(x,q)\big|_{x=q} = 0,\ K''(x,q) = \frac{1}{x(1-x)}.$$

Раскладывая $K(x,q)$ по формуле Тейлора в точке $x_0 = q$ с остаточным членом в форме Лагранжа, имеем

$$K(x,q) = \frac{(x-q)^2}{2\xi(1-\xi)}$$

при всех $x \in (0,1)$ для некоторой точки ξ, лежащей между x и q (заметим, что отсюда с учётом неравенства $\xi(1-\xi) \leqslant 1/4$ вытекает (4)). Из этого представления с учётом того, что функция $\xi(1-\xi)$ возрастает при $0 < \xi \leqslant 1/2$, получаем двустороннюю оценку (5).

Для доказательства неравенства (6) заметим, что

$$K(q,x) = q\ln\frac{q}{x} + (1-q)\ln\frac{1-q}{1-x} = -q\ln\frac{x}{q} + (1-q)\ln\frac{1-x}{1-q},$$

и рассмотрим функцию

$$f(x) = K(x,q) - K(q,x) = (x+q)\ln\frac{x}{q} + (2-x-q)\ln\frac{1-x}{1-q}; \quad f(q) = 0.$$

Так как

$$f'(x) = \ln\frac{x}{q} + \frac{x+q}{x} - \ln\frac{1-x}{1-q} - \frac{2-x-q}{1-x} = \ln\frac{x}{q} + \frac{q}{x} - \ln\frac{1-x}{1-q} - \frac{1-q}{1-x},$$
$$f''(x) = \frac{1}{x} - \frac{q}{x^2} + \frac{1}{1-x} - \frac{1-q}{(1-x)^2} = (x-q)(x^{-2} - (1-x)^{-2}),$$

то $f'(q) = 0$, $f''(x) > 0$ при $q < x < 1/2$, и из формулы Тейлора при $q < x \leqslant 1/2$ имеем $f(x) > 0$. ■

Ниже нам потребуется необременительное ограничение

$$P_1 < P_2 \leqslant 1/2, \quad \max\{\alpha,\beta\} < 1/2. \tag{7}$$

Теорема 1. Пусть при условии (2) статистика некоторого критерия проверки гипотез H_1 и H_2 является функцией от случайных величин $\mathbb{I}\{\Delta Y_t \in D(\Delta X_t)\}$, $1 \leqslant t \leqslant N$, а вероятности ошибок критерия равны соответственно α и β. Тогда

$$N \geqslant N_{1*} = \max\left\{\frac{K(1-\alpha,\beta)}{K(P_1,P_2)}, \frac{K(1-\beta,\alpha)}{K(P_2,P_1)}\right\}. \tag{8}$$

Если при этом выполнено (7), то $N_{1*} = \dfrac{K(1-\alpha,\alpha)}{K(P_1,P_2)}$ при $\alpha = \beta$,

$$N_{1*} > N_{2*} = \frac{4P_1(1-P_1)(1-\alpha-\beta)^2}{(P_2-P_1)^2}. \tag{9}$$

Доказательство.

1. Случайные величины $\mathbb{I}\{\Delta Y_t \in D(\Delta X_t)\}$, $1 \leqslant t \leqslant N$, независимые в силу условия (3), могут рассматриваться как элементы выборки, на которой работает критерий. Тогда из п. 1 леммы 1 при $p = P_2$, $q = P_1$ вытекает оценка (8).

2. При дополнительном условии (7) и $\alpha = \beta$ из (8) с учётом оценки (6) следует равенство для N_{1*}. Далее, из (4) и (5) соответственно вытекают неравенства

$$K(1-\beta,\alpha) > 2(1-\beta-\alpha)^2, \quad K(P_2,P_1) < \frac{(P_2-P_1)^2}{2P_1(1-P_1)},$$

с учётом которых из (8) получаем оценку (9):

$$N_{1*} \geqslant \frac{K(1-\beta,\alpha)}{K(P_2,P_1)} > \frac{4P_1(1-P_1)(1-\alpha-\beta)^2}{(P_2-P_1)^2}.$$

Теорема 1 доказана. ■

Замечание 1. Статистика $\nu(\mathbf{D})$ удовлетворяет условию теоремы 1, поскольку является суммой указанных в условии индикаторов.

Замечание 2. Оценка (9) мало меняется при изменении малых α, β. Она обладает следующими вполне естественными свойствами: если $P_2 - P_1 \to 0$ (гипотезы сближаются), то $N_{2*} \to \infty$; если $\alpha + \beta \to 1$ (критерий близок к критерию, заключающемуся в случайном выборе гипотез независимо от выборки), то $N_{2*} \to 0$. В последнем случае результат можно интерпретировать так: для критерия с такими вероятностями ошибок наблюдения не обязательны, поэтому нижняя граница близка к нулю.

В случае фиксированной входной разности исследуемая статистика $\nu(\mathbf{D})$ принимает следующий вид:

$$\nu(D(a)) = \sum_{1 \leqslant t \leqslant N} \mathbb{I}\{\Delta Y_t \in D(a)\}.$$

Здесь оценки теоремы 1 можно немного упростить.

Следствие 1. Пусть при условии (2) $\Delta_0 = \{a\}$. Тогда для любого критерия проверки H_1 и H_2 с вероятностями ошибок α, β, основанного на статистике $\nu(D(a))$, имеем

$$N \geqslant N_{1,a*} = \max\left\{ \frac{K(1-\alpha,\beta)}{K\left(\frac{|D(a)|}{M}, P_2\right)}, \frac{K(1-\beta,\alpha)}{K\left(P_2, \frac{|D(a)|}{M}\right)} \right\}.$$

Если при этом выполнено (7), то

$$N_{1,a*} > N_{2,a*} = \frac{4|D(a)|\left(1 - \frac{|D(a)|}{M}\right)(1-\alpha-\beta)^2}{M\left(P_2 - \frac{|D(a)|}{M}\right)^2}. \tag{10}$$

При $D(a) = \{b\}$, $1/M < p_{a,b}^{(R)} \leqslant 1/2$ величина $N_{2,a*}$ равна

$$N_{2,a,b*} = \frac{4\left(1 - \frac{1}{M}\right)(1-\alpha-\beta)^2}{M\left(p_{a,b}^{(R)} - \frac{1}{M}\right)^2}. \tag{11}$$

Доказательство. Так как здесь $\Delta X_t \equiv a$, то $\nu(D(a))$ является суммой индикаторных случайных величин, указанных в теореме 1. Поэтому условия теоремы 1 выполнены и из (8) и равенства $P_1 = |D(a)|/M$ вытекает нижняя оценка для N. При дополнительном условии (7) из (9) следует (10). Равенство (11) очевидно. ■

Заметим, что при малых значениях P_1, α, β значение оценки N_{2*} из (9) близко к величине $\dfrac{4}{P_1\left(P_2/P_1 - 1\right)^2}$, которая при условиях следствия 1 равна $\dfrac{4M}{|D(a)|\left(P_2/P_1 - 1\right)^2}$. Поэтому объём материала любого критерия, основанного на статистике $\nu(D(a))$, может быть существенно меньше количества M всех ненулевых блоков только при большом множестве $D(a)$ или большом отношении вероятностей P_2/P_1. Последние два условия противоречивы в том смысле, что если расположить все разности $b \in \mathcal{X}'$ в порядке убывания вероятностей $p_{a,b}^{(R)}$ и рассматривать в качестве $D(a)$ первые $1, 2, \ldots$ разностей, то отношение P_2/P_1 будет убывать.

Фактически в [2, теорема 1] рассматривалась атака различения, и применяемый при доказательстве теоремы способ оценки объёма материала можно описать так:

1) Выделение наиболее вероятного (при гипотезе H_2) перехода (a, b), т. е. определение величины $p_{\max} = p_{a,b}^{(R)} = \max\limits_{u,v \in \mathcal{X}'} p_{u,v}^{(R)} > 1/M$.

2) Переход к модели наблюдений с фиксированной входной разностью a и статистике $\nu(b) = \sum\limits_{1 \leqslant t \leqslant N} \mathbb{I}\{\Delta Y_t = b\}$.
3) Из эвристического предположения о необходимости условия

$$\mathsf{E}_2\, \nu(b) \geqslant \mathsf{E}_1\, \nu(b) + 1 \tag{12}$$

для возможности успешного применения критерия вида $(\nu(b) \geqslant C) \Rightarrow H_2$ авторы получают нижнюю оценку для числа разностей

$$N \geqslant N_{\text{heu}} = \frac{1}{p_{a,b}^{(R)} - 1/M}. \tag{13}$$

Если оценивается число блоков, то числитель последней дроби следует взять равным 2. Заметим, что эта оценка не вполне аккуратна, поскольку не зависит от α, β.

Замечание 3. Сравним оценку (13) с нижней оценкой (11) при $p_{a,b}^{(R)} = c/M$, $c > 1$:

$$\frac{N_{\text{heu}}}{N_{2,a,b*}} = \frac{M(p_{a,b}^{(R)} - 1/M)}{4(1 - 1/M)(1 - \alpha - \beta)^2} = \frac{c - 1}{4(1 - 1/M)(1 - \alpha - \beta)^2}.$$

Отношение меньше 1 при всех $1 < c < 1 + 4(1 - 1/M)(1 - \alpha - \beta)^2$, и в такой ситуации любой критерий на материале объёма N_{heu} не может иметь вероятности ошибок α, β; следовательно, здесь использование (13) как верхней оценки объёма материала будет ошибочным.

Замечание 4. С теоретической точки зрения невозможность существования критерия с вероятностями ошибок α, β, где $\alpha + \beta < 1$, и верхней оценкой материала вида $N = C(\alpha, \beta)/p_{\max}$, обосновывается следующим соображением. Рассмотрим случай, когда $p_{\max} = 1/M$, т. е. все элементы матрицы переходных вероятностей равны $1/M$, гипотезы совпадают. Тогда при $N \geqslant C(\alpha, \beta)M$ разностях

$$\alpha_1(N) + \alpha_2(N) \leqslant \alpha + \beta < 1.$$

С другой стороны, так как гипотеза H_2 совпадает с H_1, то

$$\alpha_1(N) + \alpha_2(N) = \mathsf{P}_1(\mathcal{X}_2) + \mathsf{P}_1(\overline{\mathcal{X}_2}) = 1,$$

где $\mathcal{X}_2$ — область принятия H_2.

Замечание 5. Наверное, невозможно полностью отказаться от ограничений на выбор ключей при расчёте атаки различения, основанной на высоковероятном переходе разностей. Дело в том, что вероятность $p_{\max} = p_{a,b}^{(R)}$ получена как результат усреднения по всем ключам k вероятностей $p_{a,b}^{(R)}(k)$ переходов при фиксированном ключе k. Поэтому вполне возможно наличие такого ключа k, при котором вероятность $p_{a,b}^{(R)}(k)$ меньше вероятности $1/M$ равновероятного перехода. Тогда для материала, полученного на таком ключе, критерий с большой вероятностью будет принимать неверное решение.

3. Обсуждение результатов

Обсудим полученные результаты с точки зрения развития статистических вопросов разностного анализа.

Впервые разностный анализ предложили E. Biham и A. Shamir при атаке на DES на конференции CRYPTO'90 и более подробно изложили материал в [1], где введены разности при покоординатном сложении по модулю 2 (названные XOR-значениями) и переходные вероятности разностей. Для блочных схем Фейстеля с независимыми ключами была получена мультипликативная формула для R-раундовой характеристики, напоминающая формулу для вероятности заданной цепочки состояний цепи Маркова.

Этот подход был сразу же подхвачен в работах X. Lai, J. Massey и S. Murphy: в [2] введены понятия разности относительно произвольной операции на группе блоков $(\mathcal{X}, \odot)$ марковского шифра. Там же предложена атака восстановления последнего раундового ключа марковского шифра, основанная на атаке различения, и получена оценка (13) объёма материала. Целью этой оценки было обоснование хороших свойств предложенной авторами шифрсистемы IPES (позже названной IDEA), вероятностная модель которой является марковским шифром [2, с. 30; 3, с. 58]. В случае, если соответствующая цепь Маркова является эргодической (т. е. неприводимой и ациклической), что проверено в [2] для моделей с длинами блоков $n = 8$ и 16, то в силу двойной стохастичности матрицы переходных вероятностей марковского шифра предельное распределение разностей за R раундов сходится к равномерному при $R \to \infty$. Таким образом, для эргодического марковского шифра оценка N_{heu} будет сколь угодно велика при достаточно больших R. Критерии того, что шифрсистема является марковской, и обобщение класса [3, с. 58] марковских шифров приведены в [7].

Но при выводе оценки N_{heu} условия наблюдений не были четко сформулированы, выбор границы 1 для разности средних в базовом предположении (12) не обоснован, полученная оценка никак не связана с вероятностями ошибок предполагаемой атаки. Отметим, что этот недостаток ни в коей мере не умаляет значения большой пионерской работы [2], основные направления исследований которой не статистические.

Далее в течение почти двух десятков лет эта оценка часто использовалась необоснованно как значение достаточного объёма материала, с игнорированием области применимости, т. е. марковость шифра не проверялась. Позже (возможно, из соображений малости величины $1/M$) её преобразовали к виду $O(1/p_{\max})$; об этом критически говорится в [4]: «С 1991 г. общепринято, что объём материала в разностном криптоанализе имеет порядок $O(1/p_{\max})$». Заметим, что такое преобразование оправдано лишь при $p_{\max} \gg 1/M$, но первоначальная концепция марковских шифров предполагает, напротив, близость этих вероятностей при достаточно больших R.

Дополнительный вклад в обоснование верхней оценки $O(1/p_{\max})$ внесла работа [11], автор которой получил оценку такого вида для метода [1] в следствии 2. Как признаётся в [11] и отмечается в [4], недостатком оценки является большая погрешность нормального приближения для распределения статистики $\nu(b) \sim \text{Bin}(n, p)$, $p = p_{\max}$. Но, вероятно, главной причиной появления такой оценки является даже не использование нормального приближения, а недостаточно обоснованное равенство $P_2 = p + (1-p)P_1 \approx p + P_1$ [11, п. 3.1] (в наших обозначениях), фактически увеличивающее p на величину $P_1 = 1/M$. Критика оценки $O(1/p_{\max})$ проведена в замечаниях 3 и 4.

Более аккуратные оценки параметров атак получены в работах C. Blondeau и соавт. [4, 12]; в [4] развивается также мультиразностный анализ. Напомним, что для эргодических марковских шифров при увеличении R значения вероятностей перехо-

дов неограниченно приближаются к $1/M$ и вероятность $p_{\max}$ всё меньше отличается от остальных значений. Поэтому существенно возрастает роль мультиразностных статистических методов, учитывающих эмпирические вероятности не одного перехода из M^2 возможных, а многих. В [4] явно выписаны все четыре эвристических предположения, применяемых авторами: 1) упоминавшееся выше предположение о ложных ключах «wrong key assumption»; 2) независимость слагаемых-индикаторов в статистиках $\nu(D(a), K)$ и независимость статистик между собой при всех $a \in \Delta_0$; 3) независимость раундовых ключей; 4) одинаковое значение сумм $\sum\limits_{a\in\Delta_0, b\in D(a)} p_{a,b}^{(R)}(K)$ при всех опробуемых K — k-битных частей раундовых ключей. Тогда вероятность того, что статистика $\nu(\mathbf{D}, K^*)$ попадёт в первые l членов вариационного ряда, будет близка к 0,5 при числе блоков $N = \dfrac{2\ln(2^{k+1}/(l\sqrt{\pi}))}{|\Delta_0| K(P_2, P_1)}$ [4, следствие 1]. Заметим, что здесь, несмотря на различие задач, оценка тоже зависит от дивергенции Кульбака, как и нижняя оценка (8).

4. Разностная атака различения на основе КОП

В нашей модели гипотезы H_1 и H_2 являются простыми, поэтому возможно построение оптимального (наиболее мощного) критерия, основанного на статистике логарифма отношения правдоподобия. Выделение набора $\mathbf{D}$ здесь уже не требуется, поскольку статистика ЛОП учитывает вероятности всех выходных разностей при всех допустимых входных разностях.

Рассмотрим более общую по сравнению с п. 3 вероятностную схему. Пусть $\mathbf{p} = (p(a),\, a \in A)$, $\mathbf{q} = (q(a),\, a \in A)$ — наборы вероятностей распределения на произвольном конечном множестве A. Тогда *дивергенцией Кульбака — Лейблера (информацией в пользу* $\mathbf{p}$ *против* $\mathbf{q}$*)* называется величина $K(\mathbf{p}, \mathbf{q}) = \sum\limits_{a\in A} p(a) \ln \dfrac{p(a)}{q(a)}$ [8, с. 15]. Использованная выше функция $K(p, q)$ является частным случаем этой величины при $A = \{0, 1\}$:

$$K(p, q) = K((p, 1-p), (q, 1-q)).$$

Вероятностно-статистический смысл дивергенции состоит в следующем: если случайная величина ξ распределена на A в соответствии с вектором вероятностей θ, $\eta = \ln \dfrac{p(\xi)}{q(\xi)}$ — логарифм отношения правдоподобий гипотез $H_1 : \theta = \mathbf{q}$ и $H_2 : \theta = \mathbf{p}$, то $\mathsf{E}_2\eta = K(\mathbf{p}, \mathbf{q})$, $\mathsf{E}_1\eta = -K(\mathbf{q}, \mathbf{p})$.

Расчёт критерия произведём на основе оценок А. С. Амбросимова для распределений ЛОП в полиномиальной схеме в случае сближающихся гипотез. Приведём для удобства читателя преамбулу, формулировку и доказательство предельной теоремы, опубликованной в [6, с. 24–30], с небольшими изменениями для приближения к нашим обозначениям.

Проводятся N независимых испытаний $\{x_t, 1 \leqslant t \leqslant N\}$ в схеме с исходами $\{1, \ldots, M\}$. Для проверки двух простых гипотез $H_i : \mathbf{p} = \mathbf{p}_i$ о вероятностях исходов, где вероятностные векторы $\mathbf{p}_i = (p_i(1), \ldots, p_i(M))$ положительные, используется статистика ЛОП $S_N = \sum\limits_{1\leqslant t\leqslant N} \ln \dfrac{p_2(x_t)}{p_1(x_t)}$.

Определение 1. Пусть в схеме серий параметры N, M и векторы $\mathbf{p}_i$ зависят от номера серии. Гипотезы H_1 и H_2 называются *сближающимися*, если

$$\delta = \max_{1\leqslant j\leqslant M} |\delta_j| \to 0, \text{ где } \delta_j = \frac{p_2(j)}{p_1(j)} - 1.$$

Теорема 2. Пусть в схеме серий гипотезы H_1 и H_2 сближаются так, что $N \to \infty$, $\sigma^2(N) = N \sum\limits_{1\leqslant j\leqslant M} p_1(j)\delta_j^2 \to \sigma^2 > 0$. Тогда:

1) $\mathsf{E}_i S_N \to (-1)^i\sigma^2/2$, $\mathsf{D}_i\eta \to \sigma^2$ при $i = 1, 2$;
2) $\dfrac{S_N - \mathsf{E}_i S_N}{\sqrt{\mathsf{D}_i S_N}} \xrightarrow{w} \mathcal{N}(0,1)$ при гипотезе H_i, $i = 1, 2$.

Доказательство.

1. Случайные величины $\eta_t = \ln\dfrac{p_2(x_t)}{p_1(x_t)}$ независимы и при гипотезе H_i имеют распределение такой случайной величины η, что

$$\mathsf{P}\left\{\eta = \ln\frac{p_2(j)}{p_1(j)}\right\} = p_i(j), \quad i = 1, 2.$$

Из асимптотических равенств $\ln(1+x) = x - x^2/2 + O(x^3) = x + O(x^2)$, $x \to 0$, и условия сближения гипотез получаем

$$\ln\frac{p_2(j)}{p_1(j)} = \ln(1+\delta_j) = \delta_j - \delta_j^2/2(1+O(\delta)) = \delta_j + O(\delta^2). \tag{14}$$

Заметим также, что

$$\sum_j \delta_j p_1(j) = \sum_j (p_2(j) - p_1(j)) = 1 - 1 = 0; \tag{15}$$

$$\sum_j \delta_j p_2(j) = \sum_j \delta_j(p_1(j) + \delta_j p_1(j)) = \sum_j \delta_j^2 p_1(j) = \sigma^2(N)/N. \tag{16}$$

С учётом (14) и (16) имеем

$$\begin{aligned}\mathsf{E}_1\eta = \sum_j p_1(j)\ln(1+\delta_j) = \sum_j p_1(j)\left(\delta_j - \delta_j^2/2(1+O(\delta))\right) = \\ = \sum_j (-1/2)\delta_j^2 p_1(j)\,(1+O(\delta)) \sim -\frac{\sigma^2(N)}{2N},\end{aligned} \tag{17}$$

и поэтому $\mathsf{E}_1 S_N = N\mathsf{E}_1\eta \to -\sigma^2/2$.

С учётом (14) и (15) имеем

$$\mathsf{E}_2\eta = \sum_j p_2(j)\ln(1+\delta_j) = \sum_j (p_1(j) + \delta_j p_1(j))\left(\delta_j - \delta_j^2/2(1+O(\delta))\right) =$$

$$= \sum_j \left(\delta_j^2 p_1(j) - \frac{1}{2}\delta_j^2 p_1(j)(1+O(\delta)) + \delta_j^2 p_1(j)O(\delta)\right) = \sum_j \frac{1}{2}\delta_j^2 p_1(j)\,(1+O(\delta)) \sim \frac{\sigma^2(N)}{2N},$$

откуда $\mathsf{E}_2 S_N = N\mathsf{E}_2\eta \to \sigma^2/2$. Предельные равенства для математических ожиданий п. 1 доказаны, переходим к равенствам для дисперсий.

С учётом (14) и (17) имеем

$$\mathsf{D}_1\eta = \sum_j p_1(j)\ln^2(1+\delta_j) - (\mathsf{E}_1\eta)^2 = \sum_j p_1(j) + \delta_j^2(1+O(\delta)) + O(\sigma^4(N)/N^2) \sim \frac{\sigma^2(N)}{N},$$

и поэтому $\mathsf{D}_1 S_N = N\mathsf{D}_1\eta \to \sigma^2$. Аналогично

$$\mathsf{D}_2\eta = \sum_j p_2(j)\ln^2(1+\delta_j) - (\mathsf{E}_2\eta)^2 = \sum_j p_1(j)(1+O(\delta))\left(\delta_j + O(\delta_j^2\right)^2 + O(\sigma^4(N)/N^2) =$$

$$= \sum_j \delta_j^2 p_1(j)(1+O(\delta)) + O(N^{-2}) \sim \frac{\sigma^2(N)}{N},$$

откуда $\mathsf{D}_2 S_N = N\mathsf{D}_2\eta \to \sigma^2$. Пункт 1 теоремы 2 доказан.

2. Достаточно доказать, что при $i = 1, 2$ выполнено условие Линдеберга [9, с. 351; 10, с. 174]: в схеме серий для любого фиксированного $\varepsilon > 0$

$$L_N(\varepsilon) = \sum_{1\leqslant t\leqslant N} \int_{|\eta_t - \mathsf{E}_i\eta_t| \geqslant \varepsilon\sqrt{\mathsf{D}_i S_N}} \frac{(\eta_t - \mathsf{E}_i\eta_t)^2}{\mathsf{D}_i S_N}\, d\,\mathsf{P}_i \to 0.$$

Так как $|\ln(1+x)| \leqslant |x|$, то $\left|\ln\dfrac{p_2(j)}{p_1(j)}\right| = |\ln(1+\delta_j)| \leqslant \delta(N)$, т. е. $|\eta| \leqslant \delta(N)$, $|\mathsf{E}_i\eta| \leqslant$ $\leqslant \mathsf{E}_i|\eta| \leqslant \delta(N)$. Отсюда следует, что

$$\frac{(\eta - \mathsf{E}_i\eta)^2}{\mathsf{D}_i S_N} \leqslant \frac{4\delta^2(N)}{\sigma^2(N)(1+o(1))} = O(\delta^2) = o(1) < \varepsilon^2,$$

начиная с некоторого $N = N(\varepsilon)$, и область интегрирования в слагаемых $L_N(\varepsilon)$ является пустым множеством, т. е. $L_N(\varepsilon) = 0$. ■

Вернёмся к нашей задаче и рассмотрим схему серий, в которой номер серии R равен числу раундов. Будем считать, что алфавит блоков $\mathcal{X}$ и множество возможных входных разностей $\Delta_0 \subset \mathcal{X}'$ фиксированы, а распределение $\{p_a\}$, набор $\mathbf{D}$ и число пар разностей N зависят от R. Введённые выше величины и условия принимают вид: при росте R

$$\begin{gathered}\delta(R) = \max_{a\in\Delta_0,\, b\in\mathcal{X}'} |p_{a,b}^{(R)} M - 1| \to 0,\\ N(R)\to\infty \text{ так, что } \sigma^2(R) = N(R)\sum_{a\in\Delta_0} p_a M \sum_{b\in\mathcal{X}'}\left(p_{a,b}^{(R)} - 1/M\right)^2 \to \sigma^2 > 0.\end{gathered} \tag{18}$$

Обозначим через $\varkappa_\gamma$ квантиль уровня γ стандартного нормального закона $\mathcal{N}(0,1)$, через $\Phi(x)$ — его функцию распределения, через $\mathbf{1}$ — вектор-строку из M единиц.

Сначала рассмотрим ситуацию нескольких входных разностей. Здесь при каждом R критерий отношения правдоподобия зададим так:

$$S_N = \sum_{1\leqslant t\leqslant N} \ln\left(p_{\Delta X_t,\Delta Y_t}^{(R)} M\right) > -\sigma^2/2 + \varkappa_{1-\alpha}\sigma \Rightarrow \text{ принимаем } H_2. \tag{19}$$

Теорема 3. Пусть при условии (2) рассматривается схема серий, в которой выполнены условия (18). Тогда:

1) $\alpha_1(N) \to \alpha$ для последовательности критериев (19). Если $\sigma = \varkappa_{1-\alpha} + \varkappa_{1-\beta}$, то $\lim\limits_{R\to\infty} \alpha_2(N) = \beta$;

2) $K(\mathbf{p}(R), \mathbf{q}(R)) \sim K(\mathbf{q}(R), \mathbf{p}(R)) \sim \dfrac{\sigma^2}{2N(R)}$, где $\mathbf{p}(R) = \left(p_a \mathbb{P}_a^{(R)},\, a\in\Delta_0\right)$, $\mathbf{q}(R) = \left(\dfrac{p_a}{M}\mathbf{1},\, a\in\Delta_0\right)$ — векторы распределений вероятностей случайного набора $(\Delta X_t, \Delta Y_t)$ при гипотезах H_2 и H_1.

Доказательство.

1. При обеих гипотезах имеем схему с N независимыми испытаниями и $|\Delta_0|M$ исходами вида $j=(a,b)\in\Delta_0\times\mathcal{X}'$, вероятности которых равны $p_{j,1}=p_a/M$ и $p_{j,2}=$ $=p_a p_{a,b}^{(R)}$ при гипотезах $H_1=H_1$ и $H_2=H_2$ соответственно. Тогда $\delta_j=p_{j,2}/p_{j,1}-1=$ $=p_{a,b}^{(R)}M-1$ и первое условие в (18) означает, что гипотезы H_1 и H_2 сближаются. Второе условие соответствует условию теоремы 2, поскольку

$$\sum_j p_{j,1}\delta_j^2=\sum_{a\in\Delta_0,b\in\mathcal{X}'}\frac{p_a}{M}\left(p_{a,b}^{(R)}M-1\right)^2.$$

С помощью теоремы о равномерной сходимости к непрерывной функции распределения (см., например, [10, с. 54]) легко доказать, что если последовательность функций распределения F_R сходится к непрерывной функции распределения F и $x_R\to x$, то $F_R(x_R)\to F(x)$. С учётом этого из п. 1 и 2 теоремы 2 получаем, что $S_N\xrightarrow{w}\mathcal{N}\left((-1)^i\sigma^2/2,\sigma^2\right)$ при гипотезе H_i, $i=1,2$.

Отсюда следует, что для любого фиксированного $\alpha\in(0,1)$

$$\alpha_1(N)=\mathsf{P}_1\{S_N>-\sigma^2/2+\varkappa_{1-\alpha}\sigma\}=\mathsf{P}_1\{(S_N-(-\sigma^2/2))/\sigma>\varkappa_{1-\alpha}\}\to 1-\Phi(-\varkappa_{1-\alpha})=\alpha.$$

Если $\sigma=\varkappa_{1-\alpha}+\varkappa_{1-\beta}$, то

$$\alpha_2(N)=\mathsf{P}_2\left\{S_N\leqslant-\sigma^2/2+\varkappa_{1-\alpha}\sigma\right\}=$$
$$=\mathsf{P}_2\left\{(S_N-\sigma^2/2)/\sigma\leqslant-\sigma+\varkappa_{1-\alpha}\right\}\to\Phi(-\varkappa_{1-\beta})=\beta.$$

2. С учётом п. 1 теоремы 2 получаем

$$K(\mathbf{p}(R),\mathbf{q}(R))=\mathsf{E}_2\ln\left(p_{\Delta X_1,\Delta Y_1}^{(R)}M\right)=\frac{\mathsf{E}_2S_N}{N(R)}\sim\frac{\sigma^2}{2N(R)},$$
$$-K(\mathbf{q}(R),\mathbf{p}(R))=\mathsf{E}_1\ln\left(p_{\Delta X_1,\Delta Y_1}^{(R)}M\right)=\frac{\mathsf{E}_1S_N}{N(R)}\sim\frac{-\sigma^2}{2N(R)}.$$

Теорема 3 доказана. ■

Заметим, что если $\sigma=\varkappa_{1-\alpha}+\varkappa_{1-\beta}$ и $\alpha=\beta$, то $\sigma=2\varkappa_{1-\alpha}$ и граница критерия (19) равна нулю.

Рассмотрим ситуацию фиксированной входной разности a. Здесь критерий (19) принимает вид

$$S_N=\sum_{1\leqslant t\leqslant N}\ln\left(p_{a,\Delta Y_t}^{(R)}M\right)>-\sigma^2/2+\varkappa_{1-\alpha}\sigma\Rightarrow\text{ принимаем }H_2.\tag{20}$$

Пусть $\rho(a,R)=\left|\mathbb{P}_a^{(R)}-\frac{1}{M}\mathbf{1}\right|$ — евклидово расстояние от строки $\mathbb{P}_a^{(R)}$ матрицы переходных вероятностей до вектора вероятностей равномерного распределения. Из теорем 2 и 3 вытекают следующие результаты.

Следствие 2. Пусть $\Delta_0=\{a\}$ при условии (2) и в схеме серий при росте R

$$\max_{b\in\mathcal{X}'}|p_{a,b}^{(R)}M-1|\to 0,\quad\sigma^2(R)=N(R)M\rho^2(a,R)\to\sigma^2>0.$$

Тогда:

1) $S_N \xrightarrow{w} \begin{cases} \mathcal{N}(-\sigma^2/2, \sigma^2) \text{ при гипотезе } H_1, \\ \mathcal{N}(\sigma^2/2, \sigma^2) \text{ при гипотезе } H_2, \end{cases}$

$$K\left(\mathbb{P}_a^{(R)}, \frac{1}{M}\mathbf{1}\right) \sim K\left(\frac{1}{M}\mathbf{1}, \mathbb{P}_a^{(R)}\right) \sim \frac{\sigma^2}{2N(R)};$$

2) если $\sigma = \varkappa_{1-\alpha} + \varkappa_{1-\beta}$, т. е.

$$N(R) \sim N_{\text{КОП}}(R, \alpha, \beta) = \frac{(\varkappa_{1-\alpha} + \varkappa_{1-\beta})^2}{M\rho^2(a, R)}, \tag{21}$$

то $\alpha_1(N) \to \alpha$, $\alpha_2(N) \to \beta$ для последовательности критериев (20).

Заметим, что в полученной асимптотической оценке (21) знаменатель пропорционален величине $\rho^2(a, R)$, которая исследовалась в ряде работ отечественных криптографов; в частности, в [13] получена явная верхняя оценка таких величин в несколько более общей ситуации шифрования (произведения подстановок, управляемых цепью Маркова). В зарубежной криптографической литературе величина $M\rho^2(a, R)$ называется *мощностью* (*capacity*) вероятностного вектора $\mathbb{P}_a^{(R)}$ (см., например, [14, с. 211]).

5. Эксперименты с марковскими моделями шифрсистемы SmallPresent

5.1. Выбор шифрсистемы для экспериментов

В наших условиях атаки предполагаются потенциально известными все элементы матрицы $\mathbb{P}^{(R)}$ вероятностей переходов разностей за R раундов. Заметим, что временная сложность вычисления даже одного элемента путём перебора ключей шифрования и вычисления переходов за R раундов всех пар блоков с фиксированной разностью a обычно превосходит сложность тотального опробования. Поэтому в ряде работ по разностному анализу (в основном начиная с [1]) рассматриваются модели шифрсистем, где алгоритм развёртывания раундовых ключей отсутствует, а ключи считаются случайными независимыми одинаково распределёнными. Если при этом последовательность раундовых разностей образует простую однородную цепь Маркова, то такая вероятностная модель называется *марковским шифром* [2]. В этих марковских моделях $\mathbb{P}^{(R)} = \mathbb{P}^R$, где $\mathbb{P} = \mathbb{P}^{(1)}$ — матрица вероятностей переходов разностей за 1 раунд, и вычисление $\mathbb{P}^{(R)}$ можно провести с помощью перебора ключей только одного раунда, а не всех R, что значительно сложнее.

Вычисление (и даже хранение) элементов $\mathbb{P}^{(R)}$ является отдельной задачей, требующей значительных ресурсов ЭВМ; только хранение одной строки матрицы $\mathbb{P}^{(R)}$ требует порядка 2^{n+2} байт памяти. Поэтому для экспериментов было выбрано семейство шифрсистем SmallPresent(n) [15] с переменной длиной блока n, кратной 4, предложенное одним из авторов шифрсистемы PRESENT. Эта система принята в 2012 г. как международный стандарт ISO/IEC блочных шифрсистем для устройств с ограниченными ресурсами [16]. Мы переходим к модели шифрсистемы SmallPresent(n) с независимым выбором R раундовых ключей и обозначаем её далее Present(n, R).

Раунд шифрования модели Present(n, R) состоит из применения набора из $s = n/4$ одинаковых параллельных S-боксов, покоординатного аддитивного наложения раундового ключа, перестановки бит π. Способ наложения раундовых ключей даёт марковость модели, согласно [3, с. 58], причём, согласно [7, следствие 2], матрица вероятностей переходов является тензорной степенью матрицы $\mathbb{P}_1$ вероятностей перехода разностей полубайт S-бокса с переставленными столбцами:

$$\mathbb{P} = \mathbb{P}_1^{[s]}\Pi, \tag{22}$$

где Π — матрица подстановки на двоичных векторах размерности n, индуцированной перестановкой π.

Заметим, что в [5, п. 4.1] производится атака восстановления ключа последнего раунда шифрсистемы Present(16,R), $7 \leqslant R \leqslant 9$, также фактически использующая формулу (22) при матрично-векторном умножении для расчёта элементов $\mathbb{P}^R$, как это указано в [17, п. 3.3] — расширенной версии работы [5].

5.2. Формулы для расчета строки матрицы вероятностей переходов за R раундов

Для ограничения емкостной сложности эксперимента выбрана атака с фиксированной входной разностью a. Вычисление строки матрицы вероятностей переходов за R раундов производится путём умножения орта $\mathbf{e}_a$ на $\mathbb{P}$, умножения результата на $\mathbb{P}$ и т. д., всего R раз:

$$\mathbb{P}_a^{(R)} = \mathbf{e}_a \underbrace{\mathbb{P} \ldots \mathbb{P}}_{R}.$$

Для этого требуются два массива размера 2^n и память для хранения $\mathbb{P}$. Поэтому актуален вопрос о способе представления, сложности вычисления и хранения матрицы $\mathbb{P}$.

Матрица $\mathbb{P}_1$ имеет 97 ненулевых элементов из 256 и, согласно формуле (22), доля ненулевых элементов в $\mathbb{P}$ равна $(97/256)^m < 0{,}38^m$ — экспоненциально убывает с ростом m. Таким образом, выгодно хранить матрицы $\mathbb{P}$ в разреженном виде: например, при $n = 16$ возможно хранение непосредственно всех $8{,}8 \cdot 10^7$ ненулевых элементов, что составляет 2 % от 2^{2n}.

Но увеличение n на 4 бита приводит к росту числа элементов в 97 раз, что превышает, например, ограничение в 10^9 элементов. Поэтому дальнейший рост размера блока до $n = 28$ был достигнут путем использования следующей формулы: если A, B — квадратные матрицы размеров M, N соответственно, $\otimes$ — произведение Кронекера, $X = (X_1, \ldots, X_M)$ — вектор-строка размерности MN, $X^{\downarrow} = X^{\top}$ — вектор-столбец, то

$$(A \otimes B)X^{\downarrow} = \begin{pmatrix} a_{11}B & \ldots & a_{1M}B \\ \ldots & \ldots & \ldots \\ a_{M1}B & \ldots & a_{MM}B \end{pmatrix} X^{\downarrow} = \begin{pmatrix} a_{11}BX_1^{\downarrow} + \ldots + a_{1M}BX_M^{\downarrow} \\ \ldots \\ a_{M1}BX_1^{\downarrow} + \ldots a_{MM}BX_M^{\downarrow} \end{pmatrix}.$$

Она похожа на формулу, применяемую в быстром преобразовании Фурье — Адамара, и показывает, что умножение тензорного произведения матриц на вектор можно осуществлять так: 1) заменяем каждый подвектор X_i на $BX_i^{\downarrow}$; 2) для всех $1 \leqslant i \leqslant M$ вычисляем суммы векторов $\sum\limits_{j:a_{ij}\neq 0} a_{ij}X_j^{\downarrow}$, из которых составляем результат.

Например, при $n = 28$, когда действуют $s = 7$ S-боксов, умножение вектора-строки $\mathbf{v}$ на $\mathbb{P}$, согласно формуле (22) и свойствам тензорных степеней

$$\left(A^{[s]}\right)^{\top} = \left(A^{\top}\right)^{[s]}, \quad A^{[s_1+s_2]} = A^{[s_1]} \otimes A^{[s_2]},$$

осуществляется вычислением сначала вектор-столбца

$$\left(Q^{[4]} \otimes Q^{[3]}\right) v^{\downarrow}, \quad Q = \mathbb{P}_1^{\top},$$

а затем умножением полученного вектора на Π путём перестановки его координат.

5.3. Выбор входных разностей и вычислительные эксперименты при $n = 16$

Поиск входной разности a, дающей наиболее неравновероятное (в смысле величины $\max\limits_{b\in\mathcal{X}'} p_{a,b}^{(R)}$ или $\rho(a, R)$) выходное распределение за R раундов, является сложной и не решённой теоретически задачей. Тотальный поиск a среди всех ненулевых разностей при фиксированном $R \approx 10$ при $n \geqslant 20$ также требует длительных вычислений. Можно предложить следующий выход из этой ситуации: будем использовать множество входных разностей (задаваемых далее как упорядоченные наборы из $s = n/4$ полубайт, соответствующих S-боксам) $A(n)$ мощности $n/2$, в которых ровно один полубайт ненулевой, при этом он равен 0x7 или 0xf (обозначения в 16-ричном формате). Разности такого вида называются разностями с одним активным S-боксом и рассматриваются в [5, п. 4.1] в экспериментах с Present$(16, R)$; при этом разность $a = 0\text{x}0007$ называется в некотором смысле «наилучшей» входной 6,7,8-раундовой разностью.

Для модели Present$(16, R)$ при $a = 0\text{x}0007$ сравним результаты экспериментов [5, п. 4.1] с нашими. В табл. 1 указаны точные значения μ_i, σ_i^2 (i — номер гипотезы) [5, п. 4.1] средних и дисперсий статистик ЛОП при проверке по полной кодовой книге ($N = 2^{15}$ пар блоков) гипотезы 2 о истинности опробуемого ключа раунда $R + 1$ (совпадает с H_2 в нашей модели) против гипотезы 1 о ложном опробовании, что при расчёте [5] предполагается совпадающей со строкой матрицы $\mathbb{P}^{R+1}\mathbb{P}'$, где $\mathbb{P}'$ — матрица переходных вероятностей при одноraундовом расшифровании. В предпоследнем столбце приведены рассчитанные, согласно следствию 2, значения $\sigma^2(N) = N\rho^2(a, R)$ при $N = 2^{15}$; они близки к дисперсии и удвоенному модулю математического ожидания статистики ЛОП при различении гипотез H_1 и H_2.

Таблица 1

R	μ_1	σ_1^2	μ_2	σ_2^2	$\sigma^2 = 2^{15}\rho^2(a, R)$	$\sigma^2/2$
6	53,821	124,087	−47,289	84,237	140,6	70,3
7	2,225	4,611	−2,149	4,152	4,8	2,4
8	0,057	0,113	−0,056	0,112	0,12	0,06

Из табл. 1 видно, что при увеличении числа раундов модули средних значений, а также дисперсии сближаются друг с другом, как и предсказывает теорема 2. При этом дисперсии и модули средних значений близки, но немного меньше значений σ^2 и $\sigma^2/2$ соответственно. Вероятно, это объясняется тем, что строки матрицы $\mathbb{P}^R$ ближе к строкам матрицы $\mathbb{P}^{R+1}\mathbb{P}'$, чем к вектору равномерного распределения.

5.4. Вычислительные и статистические эксперименты при $n \leqslant 28$

В табл. 2 при $6 \leqslant R \leqslant 10$ приведены значения $\log_2 \min\limits_{a\in A(n)} \rho^{-2}(a, R), \log_2 \max\limits_{a\in A(n)} \rho^{-2}(a, R)$ для марковских шифров Present(n, R) и входные разности $a \in A(n)$, на которых экстремумы достигаются. Заметим, что при $R = 4$ модели с длиной блока $n \geqslant 16$, а при $R = 5$ модели с $n \geqslant 24$ имеют запретные переходы разностей при некоторых $a \in A(n)$, т. е. нулевые вероятности в строке $\mathbb{P}_a^{(R)}$, и тогда критерий отношения правдоподобий не определён. Модель Present(28,5) имеет запретные переходы разностей при всех $a \in A(28)$.

При росте n на 4 бита количество исходов вероятностной схемы (выходных разностей) возрастает в 16 раз, т. е. в среднем вероятность исхода уменьшается в 16 раз.

Это при $R \geqslant 7$ ведёт к монотонному возрастанию соответствующих величин в столбцах табл. 2 везде, за исключением строк модели Present$(20,R)$: величины, нарушающие монотонность, выделены жирным шрифтом. Заметим, что в модели Present$(20,R)$ сходимость последовательности матриц к равномерной происходит быстрее по сравнению с остальными; интересно было бы установить причину этого явления.

Таблица 2

n	R				
	6	7	8	9	10
8	20,73, 0x07 22,62, 0x70	24,89, 0x0f 26,42, 0x70	28,62, 0x0f 30,28, 0xf0	32,64, 0x0f 34,45, 0xf0	36,49, 0x0f 38,70, 0x70
12	20,49, 0x007 25,45, 0xf00	25,19, 0x007 30,18, 0xf00	29,67, 0x007 34,56, 0xf00	34,61, 0x007 38,68, 0xf00	39,58, 0x007 43,24, 0xf00
16	23,84, 0x0007 25,32, 0xf000	28,73, 0x0007 30,11, 0xf000	33,95, 0x0070 35,52, 0xf000	38,04, 0x0f00 39,42, 0xf000	41,98, 0x0700 43,16, 0xf000
20	**29,73**, 0x00f00 31,61, 0x07000	35,25, 0x00f00 38,05, 0x00007	40,82, 0x00f00 44,38, 0x07000	46,56, 0x00f00 **50,51**, 0x07000	52,34, 0x00f00 **56,33**, 0x07000
24	29,35, 0xf00000 32,96, 0x00f000	35,26, 0xf00000 38,69, 0x00f000	41,62, 0x000007 44,41, 0x00f000	47,07, 0x000007 49,96, 0x00f000	52,38, 0xf00000 55,57, 0x00f000
28	34,05, 0x0000070 35,23, 0x000f000	39,61, 0x0000070 40,92, 0x000f000	45,39, 0x0000070 45,69, 0x000f000	51,20, 0x0000070 52,26, 0x000f000	56,97, 0x0000070 58,37, 0x000f000

В табл. 3 приведены результаты вычислительных и статистических экспериментов для марковских шифров Present(n,R) при $R = 9$: значения $\log_2 N_{\text{КОП}}(n,R,\alpha,\beta)$ объёма материала оптимальной атаки различения на марковские шифры Present(n,R) при $\alpha = \beta = 0{,}1$ и эмпирические вероятности ошибок в серии из 20 испытаний.

Таблица 3

n	8	12	16	20	24	28
$\log_2 \min\limits_{a \in A(n)} 1/((2^n-1)\rho^2(a,R))$	24,6	22,6	22,0	26,5	23,0	23,2
a	0x0f	0x007	0x0f00	0x00f00	0x000007	0x0000070
$\log_2 N = \log_2 N_{\text{КОП}}(n,9,\alpha,\beta)$	27,3	25,3	24,7	29,2	25,7	25,9
$\hat{\alpha}_1(N)$	0,5	0	0	0	0	0
$\hat{\alpha}_2(N)$	0,2	0	0	0,2	0,25	0

Более удобной для расчёта материала КОП в атаке различения при фиксированной разности может быть максимальная мощность $M\rho^2(a,R)$ — величина знаменателя в асимптотической оценке (21). Например, при $R = 9$ во второй строке табл. 3 (полученной по данным табл. 2) значения логарифмов величин, обратных к максимальной мощности векторов распределений выходной разности, непосредственно характеризуют устойчивость шифрсистем Present$(n,9)$ к атакам различения. Эта строка показывает, что сначала увеличение размера блока n ведёт к небольшому снижению скорости «перемешивания», т. е. скорости сближения элементов матрицы $\mathbb{P}^R$ между собой. Но у модели Present(20,9) примерно на десятичный порядок больше устойчивость к атаке различения, чем у соседних с ней моделей. В третьей строке таблицы указана разность, на которой достигается максимальная мощность.

Для практического подтверждения правильности полученных формул (для предельных вероятностей ошибок обоих родов КОП и объёма материала), а также правильности вычислений строк $\mathbb{P}_a^{(R)}$ для шифров Present$(n,9)$ проведены экспериментальные атаки различения с генерацией текстов согласно каждой из двух гипотез.

Были выбраны сравнительно большие расчётные значения $\alpha = \beta = 0{,}1$, поскольку для каждого значения n из шести возможных проведение серии из 20 независимых атак требует до 2–3 часов работы ЭВМ. Последние две строки табл. 3 показывают большую дисперсию полученных значений эмпирических вероятностей ошибок обоих родов, но средние значения эмпирических вероятностей равны соответственно $0{,}5/6 = 0{,}083$ и $0{,}65/6 = 0{,}108$, что близко к расчётным значениям $\alpha = \beta = 0{,}1$.

Работа [5] наиболее близка к нашей по своему подходу: использованию всех выходных разностей при нескольких возможных входных, теоретическому вычислению их распределений с помощью матричных формул для марковской модели шифрсистемы SmallPresent с длиной блока 16. В нашей работе получены некоторые продвижения: использование модели двублочных текстов и теоремы Амбросимова позволило в модели сближающихся гипотез строго доказать асимптотическую нормальность ЛОП, предполагаемую в [5]; применение формул тензорного умножения дало возможность достигнуть длины блока 28 при использовании оперативной памяти обычной ЭВМ.

ЛИТЕРАТУРА

1. *Biham E. and Shamir A.* Differential cryptanalysis of DES-like cryptosystems // J. Cryptology. 1991. V. 4. No. 1. P. 3–72.
2. *Lai X., Massey J., and Murphy S.* Markov ciphers and differential cryptanalysis // Eurocrypt-1991. LNCS. 1991. V. 547. P. 17–38.
3. *Lai X.* On the Design and Security of Block Ciphers: dissertation for the degree of Doctor of Technical Sciences. Swiss Federal Institute of Technology, Zurich, 1992. 118 p.
4. *Blondeau C. and Gérard B.* Multiple differential cryptanalysis: theory and practice // FSE-2011. LNCS. 2011. V. 6733. P. 35–54.
5. *Albrecht M. and Leander G.* An all-in-one approach to differential cryptanalysis for small block ciphers // SAC-2012. LNCS. 2013. V. 7707. P. 1–15.
6. *Амбросимов А. С.* Предельные теоремы для вероятностей ошибок первого и второго родов наиболее мощного критерия проверки двух сближающихся гипотез относительно вероятностей исходов полиномиальной схемы в схеме серий // Дополнительные главы теории вероятностей. Учебно-методич. пособие / ред. А. С. Амбросимов, Ю. И. Громак, И. А. Круглов, Б. В. Столпаков. М., 1992. С. 24–34.
7. *Денисов О. В.* Критерии марковости алгоритмов блочного шифрования // Прикладная дискретная математика. 2018. № 41. С. 28–37.
8. *Кульбак С.* Теория информации и статистика. М.: Наука, 1967. 408 с.
9. *Ширяев А. Н.* Вероятность. М.: Наука, 1989. 640 с.
10. *Боровков А. А.* Теория вероятностей. М.: Эдиториал УРСС, 1999. 472 с.
11. *Selçuk A. A.* On probability of success in linear and differential cryptanalysis // J. Cryptology. 2008. No. 21(1). P. 131–147.
12. *Blondeau C., Gérard B., and Tillich J.* Accurate estimates of the data complexity and success probability for various cryptanalyses // Designs, Codes and Cryptography. 2011. V. 59. P. 3–34.
13. *Круглов И. А.* Оценка скорости сходимости к равномерному распределению для произведений элементов конечной группы, управляемых цепью Маркова // Матем. вопр. криптогр. 2014. Т. 5. Вып. 1. С. 85–94.
14. *Hermelin M., Cho J., and Nyberg K.* Multidimensional extension of Matsui's algorithm 2 // FSE-2009. LNCS. 2009. V. 5665. P. 209–227.
15. *Leander G.* Small Scale Variants of the Block Cipher PRESENT. Technical University of Denmark, 2010. `http://eprint.iacr.org/2010/143.pdf`.

16. www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=56552
17. *Albrecht M. and Leander G.* An All-in-one Approach to Differential Cryptanalysis for Small Block Ciphers. Cryptology ePrint Archive, Report 2012/401, 2012. http://eprint.iacr.org.

REFERENCES

1. *Biham E. and Shamir A.* Differential cryptanalysis of DES-like cryptosystems. J. Cryptology, 1991, vol. 4, no. 1, pp. 3–72.
2. *Lai X., Massey J., and Murphy S.* Markov ciphers and differential cryptanalysis. Eurocrypt-1991, LNCS, 1991, vol. 547, pp. 17–38.
3. *Lai X.* On the Design and Security of Block Ciphers: dissertation for the degree of Doctor of Technical Sciences. Swiss Federal Institute of Technology, Zurich, 1992. 118 p.
4. *Blondeau C. and Gérard B.* Multiple differential cryptanalysis: theory and practice. FSE-2011, LNCS, 2011, vol. 6733, pp. 35–54.
5. *Albrecht M. and Leander G.* An all-in-one approach to differential cryptanalysis for small block ciphers. SAC-2012, LNCS, 2013, vol. 7707, pp. 1–15.
6. *Ambrosimov A. S.* Predel'nye teoremy dlya veroyatnostej oshibok pervogo i vtorogo rodov naibolee moshchnogo kriteriya proverki dvuh sblizhayushchihsya gipotez otnositel'no veroyatnostej iskhodov polinomial'noj skhemy v skheme serij [Limit theorems for error probabilities of the first and second genera the most powerful test for two converging hypotheses regarding the probabilities of polynomial scheme outcomes in a series scheme]. Additional Chapters of Probability Theory. Educational and Methodical Manual, eds. A. S. Ambrosimov, Yu. I. Gromak, I. A. Kruglov, and B. V. Stolpakov. Moscow, 1992, pp. 24–34. (in Russian)
7. *Denisov O. V.* Kriterii markovosti algoritmov blochnogo shifrovaniya [Criteria for Markov block ciphers]. Prikladnaya Diskretnaya Matematika, 2018, no. 41, pp. 28–37. (in Russian)
8. *Kullback S.* Information Theory and Statistics. John Wiley and Sons, 1959.
9. *Shiryaev A. N.* Veroyatnost' [Probability]. Moscow, Nauka Publ., 1989. 640 p. (in Russian)
10. *Borovkov A. A.* Teoriya veroyatnostej [Probability Theory]. Moscow, URSS, 1999. 472 p. (in Russian)
11. *Selçuk A. A.* On probability of success in linear and differential cryptanalysis. J. Cryptology, 2008, no. 21(1), pp.131–147.
12. *Blondeau C., Gérard B., Tillich J.* Accurate estimates of the data complexity and success probability for various cryptanalyses. Designs, Codes and Cryptography, 2011, vol. 59, pp. 3–34.
13. *Kruglov I. A.* Ocenka skorosti skhodimosti k ravnomernomu raspredeleniyu dlya proizvedenij elementov konechnoj gruppy, upravlyaemyh cep'yu Markova [Estimation of convergence rate to uniform distribution for products of finite group elements controlled by Markov chain]. Matematicheskie Voprosy Kriptografii, 2014, vol. 5, iss. 1, pp. 85–94. (in Russian)
14. *Hermelin M., Cho J., and Nyberg K.* Multidimensional extension of Matsui's algorithm 2. FSE-2009, LNCS, 2009, vol. 5665, pp. 209–227.
15. *Leander G.* Small Scale Variants of the Block Cipher PRESENT. Technical University of Denmark, 2010. http://eprint.iacr.org/2010/143.pdf.
16. www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=56552.
17. *Albrecht M. and Leander G.* An All-in-one Approach to Differential Cryptanalysis for Small Block Ciphers. Cryptology ePrint Archive, Report 2012/401, 2012. http://eprint.iacr.org.