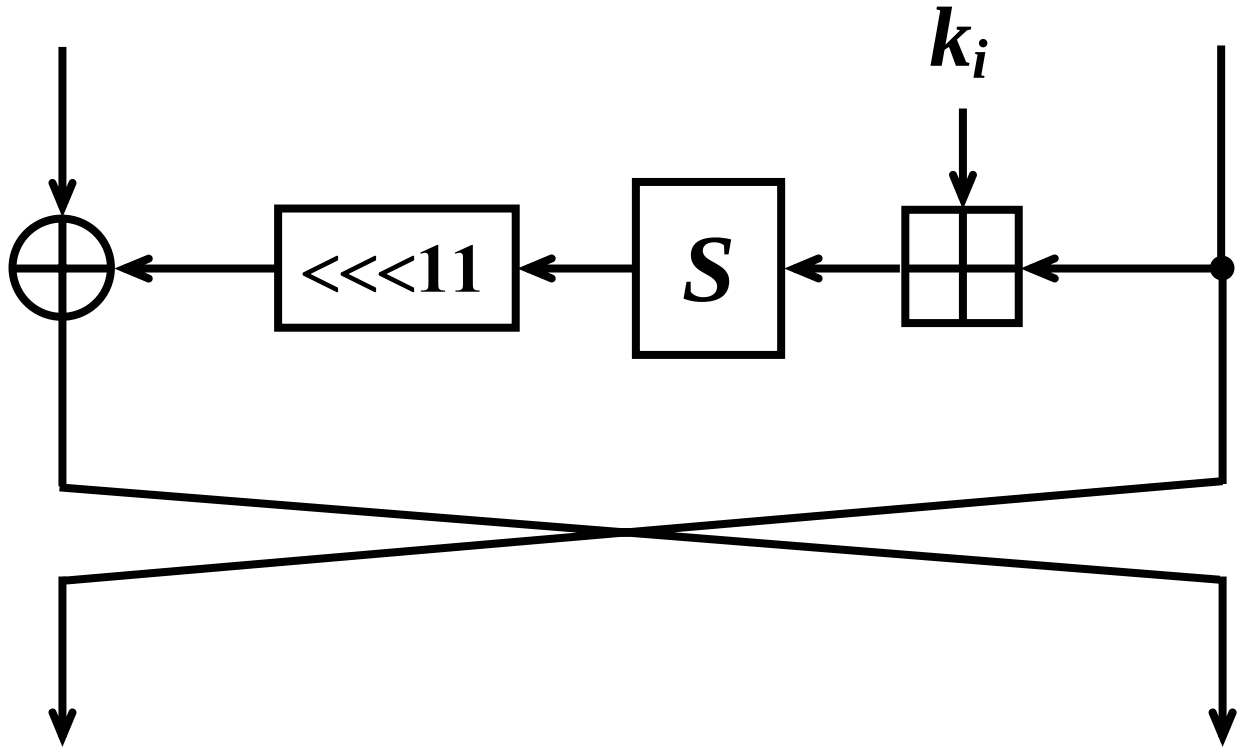
Тришин А.Е.

О дифференциальных атаках Н. Куртуа на алгоритм шифрования ГОСТ 28147-89



$i = 1, ..., 32$

$$k_j = k_{j+8} = k_{j+16} = k_{33-j},\ j = 1, ..., 8$$

$$S : (\pi_7, ..., \pi_0),\ \pi_j : \{0,1\}^4 \to \{0,1\}^4,\ j = 1, ..., 8$$

ГОСТ 28147-89

ГОСТ Р 34.12-2015

Цикл работ авторов

1. Courtois N., Misztal M. First differential attack on full 32-round GOST. – ICICS'11, Lect. Notes Comput. Sci., 2011, v. 7043, p.216 – 227.

2. Courtois N., Misztal M. Aggregated differentials and cryptanalysis of PP-1 and GOST. – CECC 2011, Periodica Mathematica Hungarica, 2012, vol. 65 (2), p. 177–192.

3. Courtois N., Misztal M. Differential cryptanalisys of GOST. – Cryptology ePrint Archive, 2011, http://eprint.iacr.org/2011/312.

4. Courtois N. An improved differential attack on full GOST. – Cryptology ePrint Archive, 2012, http://eprint.iacr.org/2012/138.

Разностные соотношения

$F : V_N \to V_M, V_M = GF(2)^M$

Разностное соотношение (differential): $(a, b) \in V_N \times V_M$,

$$p^F_{a,b} = \frac{1}{2^N} \cdot \left| \left\{ x \in V_N \mid F(x+a) + F(x) = b \right\} \right|.$$

Совокупное разностное соотношение (aggregated differential):

$$(A, B) \subseteq V_N \times V_M, \quad 0 \notin A,$$

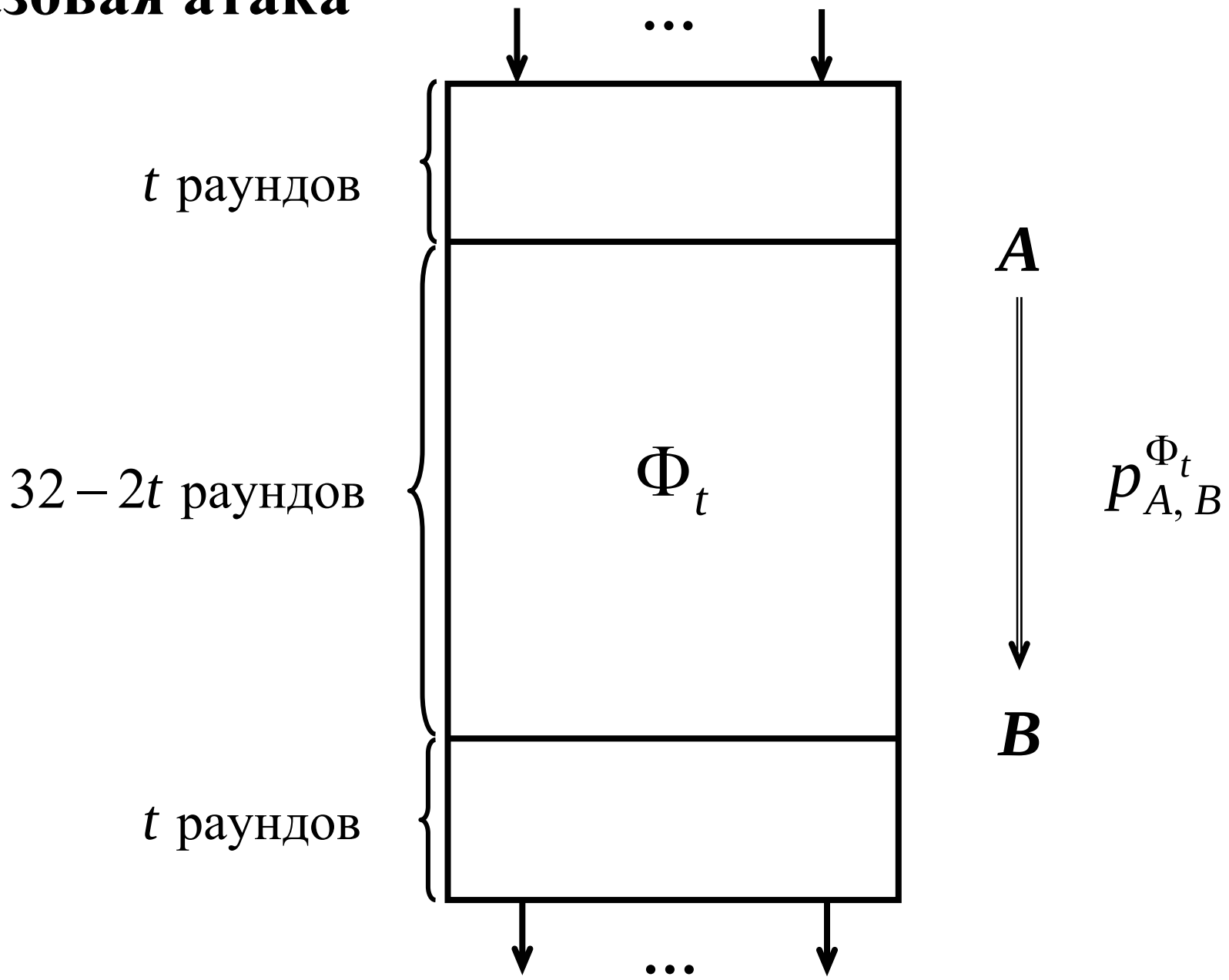
$$p^F_{A,B} = \mathrm{P}\left\{ F(x) + F(x') \in B \middle/ x + x' \in A \right\},$$

$x \sim U(V_N)$, $x' \sim U(V_N)$ – незав. Обозн.: $A \overset{F}{\Rightarrow} B$

$$p^F_{A,B} = \frac{1}{2^N} \cdot \frac{1}{|A|} \cdot \sum_{a \in A} \sum_{b \in B} \left| \left\{ x \in V_N \mid F(x+a) + F(x) = b \right\} \right|$$

Базовая атака

$t = 6$



Различитель

$G : V_{64} \to V_{64}$ – подстановка, гипотезы H_0, H_1.

H_1: $G = \Phi_t$;

H_0: G – случайная подстановка.

Решение принимается по известному набору пар

$$\{(x, y) \mid x \in V_{64},\ y = G(x)\} = \{(x_i, y_i) \mid i = 1, ..., T\},\ T = 2^{64}.$$

Статистика:

$$\eta = \left|\left\{(i, j) \mid 1 \le i < j \le T,\ x_i + x_j \in A,\ y_i + y_j \in B\right\}\right|.$$

Заметим, что если $0 \notin A$, то

$$\left|\left\{(i, j) \mid 1 \le i < j \le T,\ x_i + x_j \in A\right\}\right| = |A| \cdot T/2.$$

Напомним: $0 \notin A,\ 0 \notin B$.

Пусть $(i,\ j) \sim U\left(\left\{(i,\ j) \,|\, 1 \le i < j \le T,\ x_i + x_j \in A\right\}\right)$

G – случайная подстановка:

$$\mathrm{P}\{y_i + y_j \in B \,|\, H_0\} = \left|B\right| \Big/ (T-1)$$

$$Ep^{G}_{A,\,B} = p_0 = \left|B\right| \Big/ (T-1)$$

$$E\eta = \eta_0 = \left|A\right| \cdot \frac{T}{2} \cdot \frac{\left|B\right|}{T-1} = \frac{\left|A\right| \cdot \left|B\right|}{2} \cdot \frac{T}{T-1}$$

$G = \Phi_t$:

$$p^{G}_{A,\,B} = p_1 = p^{\Phi_t}_{A,\,B}$$

$$\eta \approx \eta_1 = \left|A\right| \cdot \frac{T}{2} \cdot p_1$$

$$\eta_1 > \eta_0 \quad \Leftrightarrow \quad (T-1)\,p_1 > |B|$$

$C \in \mathbb{N}_0$ – параметр;

$$H_1: \ \eta > C;$$

$$H_0: \ \eta \le C.$$

Положим $|A| \cdot T/2 = \mu / p_1 = \nu / p_0$.

$$\mathrm{P}\{\eta = k / H_1\} = \binom{|A| \cdot T/2}{k} \cdot p_1^k \cdot (1-p_1)^{|A| \cdot T/2 - k} \approx \frac{\mu^k}{k!} \cdot e^{-\mu},$$

$$\mathrm{P}\{\eta = k / H_0\} = \binom{|A| \cdot T/2}{k} \cdot p_0^k \cdot (1-p_0)^{|A| \cdot T/2 - k} \approx \frac{\nu^k}{k!} \cdot e^{-\nu},$$

$$\alpha = \mathrm{P}\{\eta \le C / H_1\} \approx \sum_{k=0}^{C} \frac{\mu^k}{k!} \cdot e^{-\mu}, \ \beta = \mathrm{P}\{\eta > C / H_0\} \approx \sum_{k>C} \frac{\nu^k}{k!} \cdot e^{-\nu}$$

Сложность атаки на ключ (при условии, что $|A|$ невелико)

$$Q = 2^{32(\tau+2)} \cdot \frac{t}{16} + 2^{256} \cdot \beta \cdot \left(1 - \frac{t}{16}\right).$$

Объём памяти $5 \cdot 2^{70}$ бит.

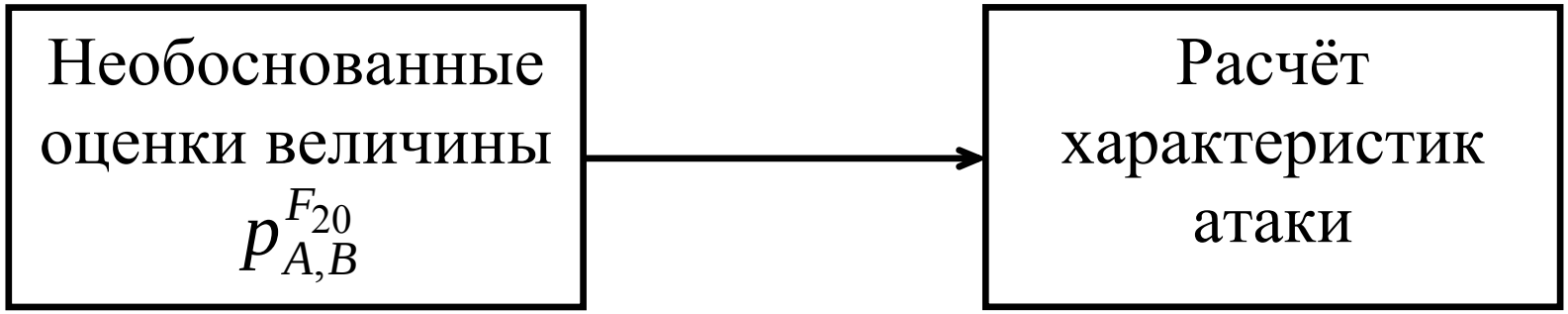
$\exists$ другие атаки авторов, развивающие идею базовой атаки, и являющиеся «более эффективными», чем эта атака.

Обоснование авторами характеристик конкретных реализаций данной атаки является нестрогим и содержит ряд **неточностей**.

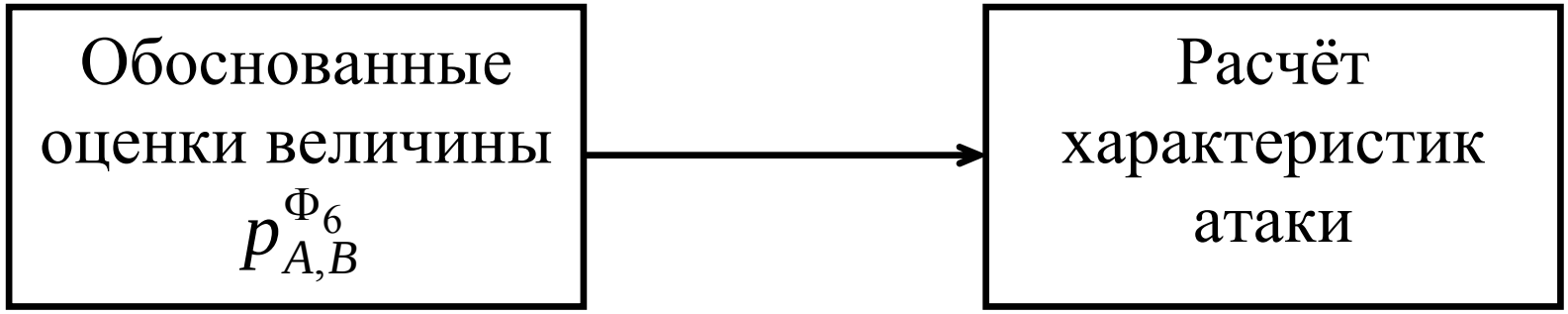
$$t = 6$$

$$A = [80000000, 00000000]^{\times}\ ,\ B = [00000100, 80600200]^{\times}$$

Авторы:



Надо:



F_r – подстановка, задаваемая r-раундовым алгоритмом ГОСТ 28147-89, в котором раундовые ключи выбираются независимо, случайно и равновероятно (Марковская модель)

Факт 7.1 [1]. Пусть пара открытых текстов $x, x' \in V_{64}$ выбирается случайным образом так, что

$$x + x' = a = (80000000, 00000000),$$

а x имеет равномерное распределение вероятностей на множестве V_{64}. Тогда усреднённая по всевозможным ключам вероятность события $F_{20}(x) + F_{20}(x') \in B$ не меньше величины, равной приблизительно $p_1 = 2^{-58}$.

Замечание 1

Неверно вычислено значение усреднённой вероятности.

А именно, усреднённая вероятность p_0 выполнения разностного соотношения

$$[80000000, 0]^{\times} \overset{F}{\Rightarrow} [00000100, 80600200]^{\times},$$

где среднее берётся по всем подстановкам $F : V_{64} \to V_{64}$, равна

$$31/2^{64},$$

тогда как у авторов (см. [1], с. 223) –

$$p_0 = 2^{-59} = 32/2^{64}.$$

Замечание 2

Неверно сформулировано условие принятия гипотезы H_1.

Утверждается, что среди всех 2^{63} неупорядоченных пар $\{x, x'\} \subset V_{64}$,

$$x + x' = a,$$

приблизительно для 2^4 пар будет выполнено включение

$$F(x) + F(x') \in B$$

в случае, когда $F : V_{64} \to V_{64}$ – случайная подстановка и приблизительно для $2^4 + 2^5$ пар будет выполнено включение

$$F_{20}(x) + F_{20}(x') \in B. \qquad (*)$$

Последнее объясняется тем, что будто бы за счёт выполнения $A \overset{F_{20}}{\Rightarrow} B$ к имеющимся 2^4 парам (получающимся, якобы, за счёт того, что отображение F_{20} обладает свойствами случайной подстановки) добавляется ещё 2^5 пар. Данное рассуждение ошибочно. Выполнение включения (*) для $2^4 + 2^5$ пар $\{x, x'\}$, $x + x' = a$, означает, что $p_1 = 2^{-58} + 2^{-59}$. Однако, в соответствии с фактом 7.1 усреднённая вероятность события $F_{20}(x) + F_{20}(x') \in B$ не менее 2^{-58}, следовательно, можно лишь утверждать, что включение (*) должно быть выполнено не менее чем для 2^5 пар $\{x, x'\}$ с разностью $x + x' = a$.

У нас $\left|A\right| = 1,\ p_1 = 2^{-58},\ p_0 = 31 \cdot 2^{-64}.$

Положим $T/2 = \nu/p_0 = \mu/p_1$, откуда $\mu = 32,\ \nu = 15{,}5.$

Пусть η – число всех пар $\{x,\ x'\} \subset V_{64}$, для которых

$$F(x) + F(x') \in B,$$

и $C \in \mathbb{N}$ – параметр критерия.

Гипотеза H_0 принимается, если $\eta \le C$,

а гипотеза H_1 принимается, если $\eta > C$.

Замечание 3

Неверно оценена вероятность ошибки второго рода.

В работе [1] приведены следующие значения:

$$\beta = 2^{-50},\ \alpha = 0{,}5.$$

Непосредственными же вычислениями получаем, что значение α, наиболее близкое к 1/2, получается при $C = 31$, и в этом случае

$$\alpha \approx 0{,}48,\ \beta \approx 1{,}6 \cdot 10^{-4}.$$

Заметим, правда, что на итоговую трудоёмкость атаки, оцениваемую величиной $\frac{3}{8}\cdot 2^{256}$ операций зашифрования ГОСТ, последнее замечание не влияет. Данная оценка остаётся справедливой, если только имеет место неравенство $p_1 \geq 2^{-58}$. Вместе с тем, используемое для p_1 значение 2^{-58}, несмотря на имеющиеся «обоснования», следует признать взятым «с потолка». При этом не представляется возможным ни обосновать правильность этого выбора, ни опровергнуть его.

Значения основной характеристики p_1 разностного метода в работах авторов выбираются исходя из некоторых недоказанных фактов. В этих «фактах» указываются усреднённые оценки вероятностей выполнения совокупных разностных соотношений для алгоритма ГОСТ с уменьшенным числом раундов. Как связаны эти усреднённые оценки со значением p_1, вообще говоря, неясно.

Кроме того, приводимые оценки не сопровождаются конкретными сведениями о том наборе (или наборах) полубайтовых подстановок, для которых они справедливы.

К наименее обоснованным относятся результаты экспериментальной проверки содержащихся в работах авторов оценок основной характеристики p_1 дифференциальной атаки. Это так называемые теоретические «факты», в частности, факт 7.1. Результаты экспериментальной проверки некоторых из этих фактов показали следующее. Для алгоритма ГОСТ 28147-89 с полубайтовыми подстановками из стандарта хеш-функции ГОСТ Р 34.11-94 и полубайтовыми подстановками из стандарта блочного шифрования ГОСТ Р 34.12-2015 полученные нами оценки не подтверждают оценки вероятностей выполнения разностных соотношений $A \overset{F_r}{\Rightarrow} B$, рассмотренных в работах авторов. Оценки из рассмотренных работ являются существенно завышенными.

Таким образом, оценивая характеристики предлагаемых дифференциальных атак в рамках Марковской модели, авторы завышают возможности рассматриваемого варианта дифференциального метода применительно к алгоритму блочного шифрования ГОСТ 28147-89.

Выводы

В рассмотренном цикле работ содержатся интересные подходы к построению атак на ключ алгоритма шифрования ГОСТ 28147-89 дифференциальным методом, в числе которых использование

1) обобщённых разностных соотношений нового типа, названных совокупными разностными соотношениями;

2) использование «различителя» для средней части блочного преобразования ГОСТ – статистического критерия, с помощью которого подстановку, задаваемую 20 средними раундами алгоритма зашифрования, можно «отличить» от случайной подстановки на множестве V_{64}.

Основными недостатками предлагаемых подходов являются:

1) необходимость наличия полной кодовой книги шифра;

2) получение оценок характеристик криптографических атак на основе Марковской модели, в рамках которой основным параметром метода является усреднённая оценка разностной характеристики блочного преобразования по всему множеству ключей, причём считается, что раундовые ключи выбираются независимо, случайно и равновероятно;

3) получение оценок основной характеристики метода – вероятности выполнения глобального совокупного разностного соотношения – путём перемножения определённых усреднений эмпирических оценок вероятностей выполнения некоторых событий, не являющихся, вообще говоря, независимыми; отсутствие данных о практических проверках правомерности подобных перемножений.

4) завышение в пользу метода ряда оценок криптографических характеристик.

Предложенные в рассмотренных работах дифференциальные атаки на алгоритм блочного шифрования ГОСТ 28147-89 представляют несомненный теоретический интерес, однако на практическую стойкость алгоритма ГОСТ 28147-89 и, в частности, алгоритма блочного шифрования «Магма» из стандарта ГОСТ Р 34.12-2015, не влияют.