УДК 681.3.06

Р.В. ОЛЕЙНИКОВ, канд. техн. наук, Д.С. КАЙДАЛОВ

УТОЧНЕНИЕ ЭФФЕКТИВНОСТИ РАЗЛИЧЕНИЯ ЦЕПИ ФЕЙСТЕЛЯ И СЛУЧАЙНОЙ ПЕРЕСТАНОВКИ

Введение

Цепь Фейстеля является одной из наиболее широко распространенных высокоуровневых конструкций блочных симметричных шифров. На ее основе построены алгоритмы ГОСТ 28147-89 [1], DES [2], Camellia [3] и др. Основным преимуществом цепи Фейстеля является возможность построения инволютивного преобразования, т.е. расшифрование реализовано аналогично зашифрованию при использовании обратного порядка раундовых подключей. Дополнительным преимуществом этой конструкции является отсутствие требований к биективности раундовой функции (как у SPN-структур), что упрощает разработку и реализацию.

Шифрующее преобразование симметричного n-раундового ($n = 2k$) блочного шифра с размером блока $2l$ битов, построенного на базе цепи Фейстеля, может быть описано следующим образом:

$$F_K = \varphi_{n,k_n} \circ \varphi_{n-1,k_{n-1}} \circ \ldots \circ \varphi_{1,k_1},$$
$$\varphi_{i,k_i}\left(x_i^L, x_i^R\right) = \left(x_i^R, f\left(x_i^R, k_i\right) \oplus x_i^L\right),$$

где $K = \left(k_1, k_2, \ldots, k_n\right)$ – развернутый ключ шифрования, состоящий из раундовых подключей k_i, $i \in \{1,2,\ldots,n\}$; x_i^L и x_i^R – значения левого и правого полублока (длиной l битов) на входе i-го раунда; $f\left(x_i^R, k_i\right)$ – раундовая функция блочного шифра.

В большинстве публикаций при разработке и оценке стойкости блочных шифров, построенных на цепи Фейстеля, основное внимание уделяется свойствам раундовой функции или схемы разворачивания подключей. В то же время выбору и обоснованию свойств высокоуровневой конструкции посвящено сравнительно мало публикаций.

Эффективность цепи Фейстеля целесообразно оценивать через сложность различения перестановки, сформированной этой структурой, от случайной соответствующей степени, поскольку именно множество случайных перестановок степени является моделью идеального блочного шифра [4]. Сложность выполнения алгоритма-различителя и достигаемая вероятность успеха являются численными показателями эффективности высокоуровневой конструкции. Для исключения влияния свойств конкретной раундовой функции необходимо использовать идеализированное раундовое преобразование, такое как случайная функция или случайная перестановка.

Обзор известных результатов

В работе [5] предложен способ построения генератора псевдослучайных перестановок степени 2^{2n} на основе псевдослучайных функций, выполняющих отображение на множество степени 2^n. Показана возможность построения симметричного блочного шифра, при выполнении ряда условий стойкого к атакам на основе выбранных открытых текстов. Определено, что верхняя граница эффективности алгоритма-различителя для $k \geq 3$ раундов не превышает значения $|P_1 - P_1^*| \leq \frac{m^2}{2^n}$. Более короткое доказательство этих результатов представлено в [6], там же описана взаимосвязь между вероятностно-теоретическими и сложностно-теоретическими оценками в криптографии.

В статье [7] описаны атаки на схемы Фейстеля с 1,2,3 и 4 раундами. Кроме того, автор показывает, что для 5 раундов возможно различение рассматриваемой структуры от случай-

ной перестановки со сложностью не более чем $O(2^{\frac{3n}{2}})$ при наличии $O(2^{\frac{3n}{2}})$ пар выбранного открытого и соответствующего шифрованного текста. В работах [8, 9] рассмотрены варианты атак на основе известных и выбранных открытых текстов для 4-6 раундов и более. В [10] сложность различения 3 или 4 раундов уменьшена с $O(\frac{m^2}{2^n})$ до не более чем $O(\frac{m^4}{2^{3n}}+\frac{m^2}{2^{2n}})$.

Уменьшенная сложность доказательства результатов [5] и достаточности комбинации 2-раундовой цепи Фейстеля и конечных попарно независимых перестановок приведена в [11].

В работе [12] описаны новые результаты применительно к модифицированной цепи Фейстеля (L-схема), использованной в ряде блочных шифров, в том числе MISTY.

Исследование стойкости цепи Фейстеля с раундовыми функциями, не являющимися псевдослучайными, к атакам с выбранными открытыми текстами, приведено в [13].

Цепь Фейстеля со случайными перестановками вместо случайных функций в раундовых преобразованиях рассмотрена в [14]. Показано, что такая конструкция является стойкой при количестве запросов $m << 2^{\frac{n}{2}}$, где $2n$ – размер блока.

Сравнение высокоуровневых схем различных алгоритмов, представленных на конкурс AES, приведено в [15], где показано преимущество цепи Фейстеля перед другими конструкциями.

Используемые обозначения

При дальнейшем изложении будут использоваться следующие обозначения:

I_n – множество битовых векторов длины n;

I_{2n} – множество битовых векторов длины $2n$;

F_n – множество функций $F: I_n \to I_n$;

F_{2n} – множество функций $F: I_{2n} \to I_{2n}$;

σ_n – множество перестановок степени n;

σ_{2n} – множество перестановок степени $2n$;

ψ – биективное отображение на основе цепи Фейстеля;

P_1 – вероятность, с которой алгоритм-различитель определяет цепь Фейстеля;

P_1^* – вероятность определения случайной функции (перестановки) алгоритмом-различителем;

$x \bullet y$ – конкатенация двух векторов x и y.

Модель алгоритма-различителя

Алгоритм-различитель (рис. 1) получает на входе множество открытых текстов $\{(L_i, R_i), 1 \le i \le m\}$ и множество соответствующих шифрованных текстов $\{(S_i, T_i), 1 \le i \le m\}$.

Блочный шифр реализует определенное подмножество перестановок, количество которых равняется количеству возможных ключей шифрования, так как выбор такого подмножества определяется структурой шифра и не является случайным, то возможно построение алгоритма-различителя, который мог бы определить, является ли конкретная перестановка случайно выбранной из общего множества, либо полученной в результате действия блочного шифра. Таким образом, анализируя входные данные, поданные на вход алгоритма-различителя, на выходе алгоритма будем иметь «1» либо «0». «1» в том случае, если считается, что входные данные были получены с помощью блочного шифра (цепи Фейстеля в данном случае) и «0», если считается, что это результат действия случайной функции. Для цепи Фейстеля вероятность появления «1» будет иметь определенное значение. Однако и для слу-

чайной функции вероятность появления «1» на выходе алгоритма-различителя не равна нулю, поскольку возможен случайный выбор произвольной перестановки, аналогичной сформированной цепью Фейстеля.

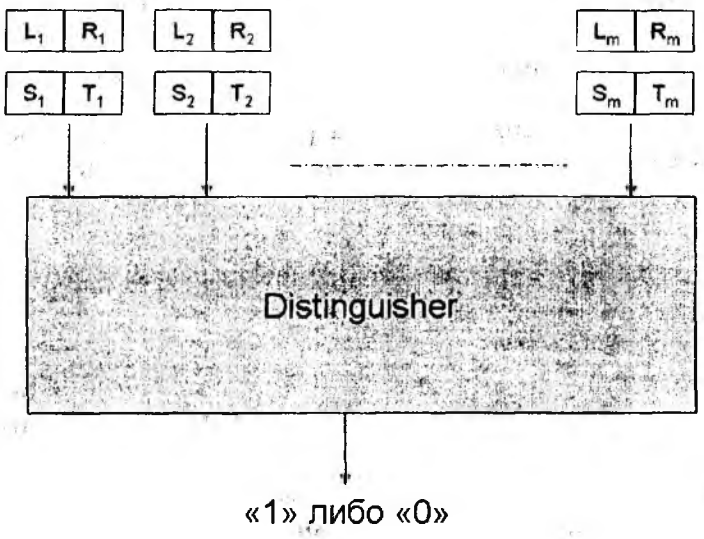

Рис. 1

Преимуществом алгоритма-различителя считается модуль разности вероятностей определения (различения) цепи Фейстеля и случайной функции:

$$Adv(\psi, F^*) = | P_\psi - P_{F^*} | ,$$

где P_ψ – вероятность определения (различения) алгоритмом цепи Фейстеля; P_{F^*} – вероятность определения (различения) алгоритмом случайной функции (перестановки).

Для одно- и двухраундовой цепи Фейстеля построение алгоритма-различителя является тривиальной задачей, поэтому в дальнейшем все исследования будут проводиться для трех или четырех раундов шифрования. Для большего количества раундов сложность различения будет только увеличиваться, поэтому полученные результаты являются нижней границей сложности (верхней границей вероятности) различения для пяти- и более раундовых структур на основе цепи Фейстеля.

Верхняя граница вероятности различения 3-раундовой цепи Фейстеля на основе случайных функций

На рис. 2 приведена 3-раундовая цепь Фейстеля. Это преобразование будет обозначено как $f = \psi(F^n, F^n, F^n)$. Принимается, что раундовые функции f_1^*, f_2^*, f_3^* случайно выбраны из множества F_n. Аргументами функции f являются блоки данных $x_i = L_i \bullet R_i$ длиной $2n$ *бит*. L_i и R_i являются соответственно левой и правой частью входного аргумента длиной по n бит каждая. Выходные значения будут обозначены как $f[L_i \bullet R_i] = [V_i \bullet T_i]$.

Пусть g – это некоторая функция, реализующая алгоритм-различитель. На вход функции g подаются $k \geq 2$ аргументов $x_1 ..., x_k$. Для каждого входного аргумента вычисляется выходное значение $f[x_i] = y_i$. Значение y_i также имеет длину $2n$ *бит*. Далее каждая пара значений $\{[x_i, y_i], [x_j, y_j]\}$, где $1 \leq i < j \leq k$, обрабатывается алгоритмом-различителем, результатом работы которого является значение «0» или «1». Если хотя бы одна пара значений привела к «1», то и возвращаемое значение функции g также равно «1», иначе возвращаемое значение равно «0».

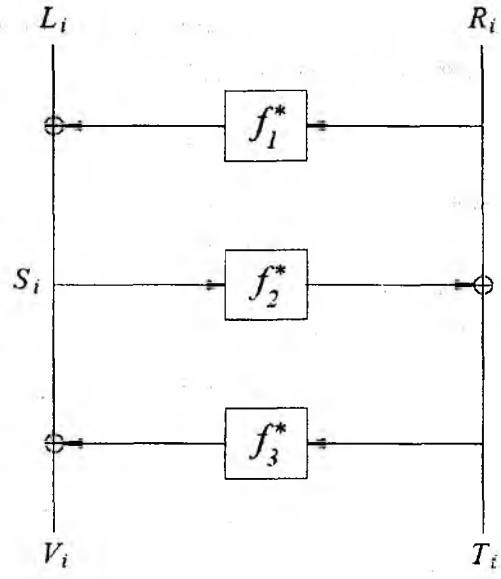

Рис. 2

Известный результат [5,6] сформулирован в виде следующей леммы.

Лемма 1 (Маурер, Лаби – Раков). Для каждой функции $g:(\{0,1\}^{2n})^k \to \{0,1\}$ и для каждого набора из k аргументов $x_1,...,x_k$ верно следующее неравенство:

$$|P[g(f(x_1),...,f(x_k))=1 : f \in_R \psi(F_n,F_n,F_n)] - P_g| \le \frac{k^2}{2^n}, \quad (1)$$

$$P_g = P[g(f(x_1),...,f(x_k))=1 : f \in_R F_{2n}].$$

В доказательстве этой леммы [6] используется предположение о равновероятности событий $\{S_i = S_j\}$ и $\{T_i = T_j\}$ (см. рис.1). Кроме того, исходя из суммирования вероятностей в доказательстве использовано допущение о несовместности событий, что корректно для определения верхней границы вероятности, но неверно для точного значения.

Точный результат может быть получен следующим образом.

Определим значения S_i, T_i и V_i для всех $1 \le i \le k$ следующим образом: $S_i = L_i \oplus f_1^*(R_i)$, $T_i = R_i \oplus f_2^*(S_i)$, $V_i = S_i \oplus f_3^*(T_i)$.

Для функции f с аргументами x_i, рассматриваемой как трехраундовое преобразование, выходы первого, второго и третьего раундов обозначаются как $R_i \bullet S_i$, $S_i \bullet T_i$ и $T_i \bullet V_i = f(L_i \bullet R_i)$ соответственно. В последующем предполагается, что все x_i при $1 \le i \le k$ различны. Выбор повторяющихся аргументов не предоставит никакой новой информации о структуре преобразования.

Пусть ε_S и ε_T обозначают такие события, что все $S_1,...,S_k$ и $T_1,...,T_k$ различны. Также пусть ε обозначает событие, когда произошли оба события ε_S и ε_T. Вероятность появления коллизии ($P[\bar{\varepsilon}]$) может быть оценена следующим образом.

Лемма 2. Вероятность события $\bar{\varepsilon}$ для каждого набора из $k \ge 2$ аргументов $x_1,...,x_k$

$$\mathrm{P}[\bar{\varepsilon}] \le 1-(1-\frac{1}{2^n})^{2(k-1)}(1-\frac{1}{2^n-1})^{2(k-2)}(1-\frac{1}{2^n-2})^{2(k-3)}...(1-\frac{1}{2^n-(k-2)})^{2(k-(k-1))} =$$

$$= 1-\prod_{i=0}^{k-2}(1-\frac{1}{2^n-i})^{2(k-(i+1))} \quad (2)$$

Доказательство. Рассмотрим сначала вероятность события $\bar{\varepsilon}_S$ – вероятность нахождения среди k аргументов такой пары, что $\{S_i = S_j\}$ при $i \neq j$. Количество возможных пар составляет $C_k^2 = \frac{k(k-1)}{2}$. При последовательном рассмотрении вероятность отсутствия коллизии в паре будет уменьшаться при добавлении очередного элемента.

Рассмотрим первые $k-1$ пар значений $\{(S_1;S_2),(S_1;S_3),...,(S_1;S_k)\}$. S_1 и S_i, $2 \le i \le k$ являются случайными, т.к. функция f_1^* является случайной (мы принимаем условие, что $R_i \neq R_j$, иначе $P[S_i = S_j] = 0$). Вероятность события $\{S_1 = S_i\}$ в таком случае составляет $P[S_1 = S_i] = 2^{-n}$. Вероятность того, что ни одно из событий $\{S_1 = S_i\}$, $2 \le i \le k$, не произойдет, равна $(1-\frac{1}{2^n})^{(k-1)}$.

Рассмотрим следующие $k-2$ пар значений $\{(S_2;S_3),(S_2;S_4),...,(S_2;S_k)\}$. S_2 и S_i, $3 \le i \le k$ также являются случайными, однако точно известно, что $S_2 \neq S_1$ и $S_i \neq S_1$, т.к. такие события уже были рассмотрены на предыдущем этапе. Соответственно вероятность события $\{S_2 = S_i\}$ в таком случае составляет $P[S_2 = S_i] = \frac{1}{2^n - 1}$. Общая вероятность того, что ни одно из событий $\{S_2 = S_i\}$, $3 \le i \le k$, не произойдет, равна $(1-\frac{1}{2^n-1})^{(k-2)}$.

Аналогично можно посчитать вероятности для остальных наборов пар. Всего таких наборов может быть $k-1$. Каждый из них будет иметь на один элемент (одну пару) меньше. Последний набор будет состоять из одной пары $\{S_{k-1} = S_k\}$ и вероятность такого события будет составлять $P[S_{k-1} = S_k] = \frac{1}{2^n - (k-2)}$.

Отсюда вероятность $P[\bar{\varepsilon}_S]$ имеет следующее значение:

$$P[\bar{\varepsilon}_S] \le 1 - (1-\frac{1}{2^n})^{(k-1)}(1-\frac{1}{2^n-1})^{(k-2)}(1-\frac{1}{2^n-2})^{(k-3)}...(1-\frac{1}{2^n-(m-2)})^{(k-(k-1))} =$$

$$= 1 - \prod_{i=0}^{k-2}(1-\frac{1}{2^n-i})^{(k-(i+1))}.$$

Вероятность события $\bar{\varepsilon}_T$ – это вероятность нахождения среди k аргументов такой пары, что $\{T_i = T_j\}$ при $i \neq j$. $\bar{\varepsilon}_T$ находится аналогично $\bar{\varepsilon}_S$. Фактически, если рассматривать событие $\bar{\varepsilon}$, при котором выполняется хотя бы одно из равенств $\{T_i = T_j\}$ либо $\{S_i = S_j\}$, то каждый из наборов, рассматриваемых выше, будет иметь в 2 раза больше возможных пар. Т.е. первый набор будет состоять из пар $[\{S_1;S_2\},\{S_1;S_3\},...,\{S_1;S_k\}]$ и $[\{T_1;S_2\},\{T_1;T_3\},...,\{T_1;T_k\}]$. Тогда общая вероятность того, что ни одно из событий $\{S_1 = S_i\}$ и $\{T_1 = T_i\}$, $2 \le i \le k$, не произойдет, равняется $(1-\frac{1}{2^n})^{2(k-1)}$. Аналогично можно найти вероятности для других наборов пар.

Тогда вероятность события $\bar{\varepsilon}$ будет иметь значение, представленное в формуле (2).

Доказательство окончено.

Формула (2) является достаточно сложной. Ее аппроксимацию можно получить при допущении, что вероятность коллизии среди выходных значений раундовой функции не увеличивается с ростом размера выборки. Если этот размер значительно меньше мощности множества открытых текстов, то аппроксимация будет достаточно точной.

Лемма 3. Аппроксимированное значение вероятности события $\bar{\varepsilon}$ для каждого набора из $k \geq 2$ аргументов $x_1,...,x_k$ может быть получено следующим образом:

$$P[\bar{\varepsilon}] \leq 1-(1-\frac{1}{2^n})^{2C_k^2} = 1-(1-\frac{1}{2^n})^{2^{\frac{k(k-1)}{2}}} = 1-(1-\frac{1}{2^n})^{k(k-1)}. \quad (3)$$

Доказательство. В данном случае рассматривается $2C_k^2 = k(k-1)$ событий $\{T_i = T_j\}$ и $\{S_i = S_j\}$. Вероятность каждого, как и в лемме 1, принята за $P[S_i = S_j] = P[T_i = T_j] \leq 2^{-n}$.

Доказательство окончено.

Теорема 1. Верхняя граница вероятности различения 3-раундовой цепи Фейстеля $\psi(F_n, F_n, F_n)$ на основе случайных функций $\{F_n\}$ и случайной функции F_{2n} для k запросов на входе алгоритма-различителя

$$|P[g(f(x_1),...,f(x_k)) = 1 : f \in_R \psi(F_n, F_n, F_n)] - P_g| \leq 1-\prod_{i=0}^{k-2}(1-\frac{1}{2^n-i})^{2(k-(i+1))} \leq 1-(1-\frac{1}{2^n})^{k(k-1)}.$$

Доказательство. Результат следует из леммы 2 и 3.

Сравнительная оценка значений преимущества цепи Фейстеля над случайной функцией, полученных с помощью формул (1), (2) и (3) для блоков длины 16 ($n = 8$) и 32 ($n = 16$) *бит*, показана на рис. 3.

Как видно из графика, формула (3) задает хорошую аппроксимацию для точного значения вероятности преимущества, полученного с помощью формулы (2). Следует отметить, что граница вероятности, полученная с помощью известной из [5] формулы (1), становится больше 1 примерно при $\sqrt{2^n}$ аргументах, что делает невозможным ее использование для оценки отличия цепи Фейстеля от случайной функции при $|\{x_i\}| > \sqrt{2^n}$.

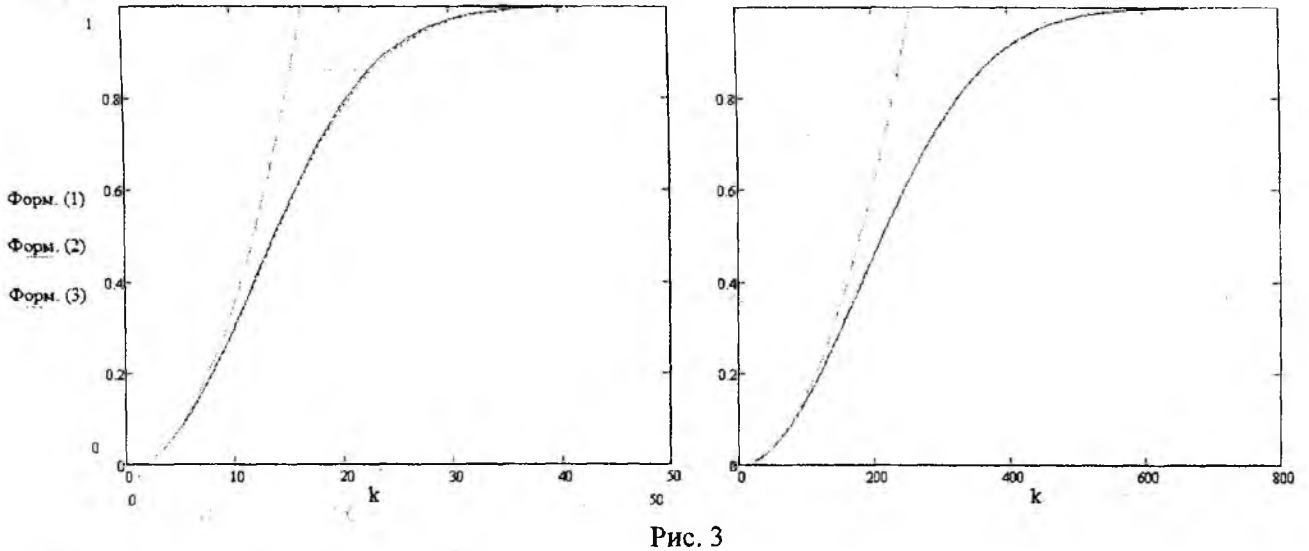

Рис. 3

Оценка эффективности алгоритмов-различителей для 3-раундовой цепи Фейстеля на основе случайных функций

Оценка вероятности различения, приведенная в предыдущем разделе, является теоретически максимально возможной вероятностью различения. Однако для реальных алгоритмов различения данная вероятность будет меньше. Кроме того, с ростом числа раундов вероятность различения также будет уменьшаться. Стоит отметить, что высокоуровневая конструкция может иметь несколько различителей одновременно, каждый из которых позволяет

получить определенную вероятность. В частности, для 3-раундовой цепи Фейстеля со случайными функциями в качестве раундовых преобразований будет рассмотрено два алгоритма-различителя. Вполне возможно, что существуют другие алгоритмы-различители, которые позволяют получить более высокую вероятность различения (но не превышающую порог из теоремы 1).

Применение конкретных алгоритмов позволяет изменить модель различения от случайной функции к случайной перестановке, что дает более точный результат.

Алгоритм-различитель №1 для 3-раундовой цепи Фейстеля со случайными функциями в качестве раундовых преобразований [9].

Для k входных аргументов с различными левыми и одинаковыми правыми половинами проверяется выполнение двух равенств: $\mathrm{L}_i \oplus V_i = \mathrm{L}_j \oplus V_j$ и $T_i = \mathrm{T}_j$ (см. рис. 2). В случае выполнения равенств хотя бы для одного аргумента возвращаемое значение будет «1», иначе «0».

Лемма 4. Вероятность выполнения равенств $\mathrm{L}_i \oplus V_i = \mathrm{L}_j \oplus V_j$ и $T_i = \mathrm{T}_j$ для набора из $k \geq 2$ аргументов $x_1, ..., x_k$ с различными левыми и одинаковыми правыми половинами (алгоритм-различитель №1) для 3-раундовой цепи Фейстеля (рис. 2)

$$\mathrm{P}_1 = 1 - \prod_{i=0}^{k-2} \left(1 - \frac{1}{2^n - i}\right)^{k-(i+1)} \leq 1 - \left(1 - \frac{1}{2^n}\right)^{\frac{k(k-1)}{2}}. \qquad (4)$$

Доказательство. Поскольку $R_i = R_j$ и $L_i \neq L_j$, то $S_i \neq S_j$ (см. рис. 2). В таком случае, оба условия $\mathrm{L}_i \oplus V_i = \mathrm{L}_j \oplus V_j$ и $T_i = \mathrm{T}_j$ будут выполнены, если $f_2^*(S_i) = f_2^*(S_j)$. Вероятность такого события для одной пары будет равна $\frac{1}{2^n}$. Для k аргументов можно составить C_k^2 пар, соответственно суммарная вероятность выполнения условия $f_2^*(S_i) = f_2^*(S_j)$ равна $\mathrm{P} = 1 - \prod_{i=0}^{k-2} \left(1 - \frac{1}{2^n - i}\right)^{k-(i+1)}$. Способ вычисления такой суммарной вероятности рассматривался при доказательстве леммы 2. Кроме того, в формуле (4) приведено аппроксимированное значение вероятности, которое рассматривалось выше.

Доказательство окончено.

Лемма 5. Вероятность выполнения условий $\mathrm{L}_i \oplus V_i = \mathrm{L}_j \oplus V_j$ и $T_i = \mathrm{T}_j$ для случайной перестановки при произвольных различных входах

$$\mathrm{P}_1^* = 1 - \prod_{i=0}^{k-2} \left(1 - \frac{1}{2^{2n} - 1 - i}\right)^{k-(i+1)} \leq 1 - \left(1 - \frac{1}{2^{2n} - 1}\right)^{\frac{k(k-1)}{2}}. \qquad (5)$$

Доказательство. Для случайной перестановки при произвольных различных входах выходное значение $(V_i \bullet T_i)$ является равномерно распределенной случайной величиной.

Для произвольно выбранной пары значений равномерно распределенной случайной величины $h_l, h_k \in_R I_{2n}$ вероятность совпадения равна $\frac{1}{2^{2n} - 1}$.

Выходные значения перестановки при разных входных аргументах не могут повторяться, т.е. гарантированно $V_i \neq V_j$ и/или $T_i \neq T_j$. Учитывая, что $L_i \neq L_j$, такое условие исключает один из неподходящих вариантов для выполнения заданных равенств.

Для k аргументов можно составить C_k^2 пар, соответственно вероятность выполнения условий $L_i \oplus V_i = L_j \oplus V_j$ и $T_i = T_j$ равна $P = 1 - \prod_{i=0}^{k-2} (1 - \frac{1}{2^{2n} - 1 - i})^{k-(i+1)}$. Способ вычисления такой суммарной вероятности рассматривался при доказательстве леммы 2. В формуле (5) приведено аппроксимированное значение вероятности, которое также рассматривалось выше.

Доказательство окончено.

Теорема 2. Преимущество алгорима-различителя №1 при $k \geq 2$ запросах ограничено сверху величиной

$$\mathrm{Adv}_1(\psi, \sigma_{2n}) = | P_1 - P_1^* | \leq \left| \prod_{i=0}^{k-2} (1 - \frac{1}{2^n - i})^{k-(i+1)} - \prod_{i=0}^{k-2} (1 - \frac{1}{2^{2n} - 1 - i})^{k-(i+1)} \right| \leq$$

$$\leq \left| (1 - \frac{1}{2^n})^{\frac{k(k-1)}{2}} - (1 - \frac{1}{2^{2n} - 1})^{\frac{k(k-1)}{2}} \right|.$$

Доказательство. Результат следует из леммы 4 и 5.

Преимущество алгоритма-различителя №1 для $n = 8$ и $n = 16$ представлено на рис.4. По оси абсцисс приведено количество аргументов, которые подаются на вход, по оси ординат отмечены значения преимущества (вероятность различения случайной перестановки и цепи Фейстеля).

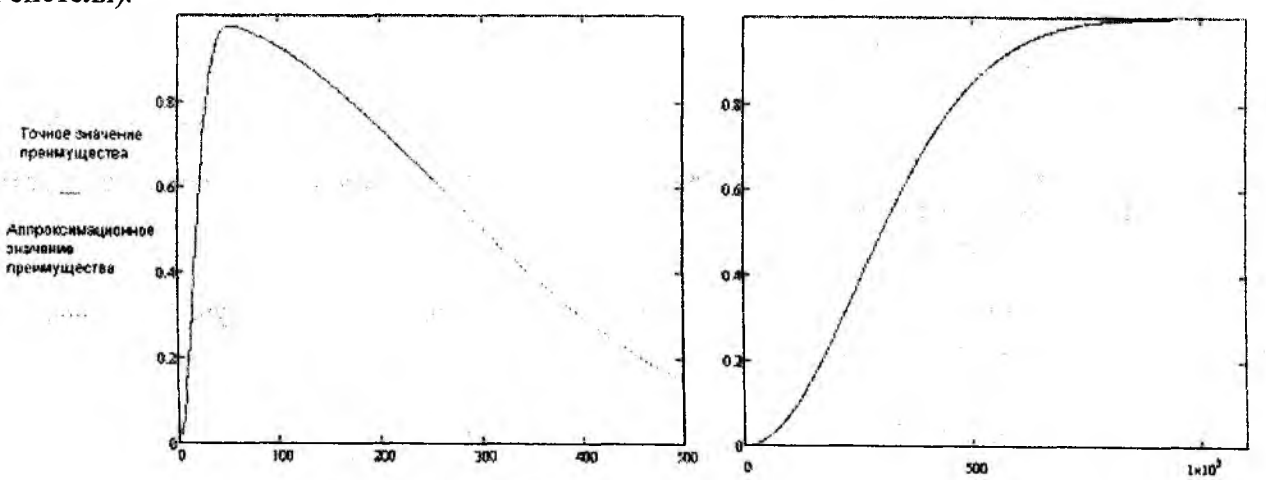

Рис. 4

Следует отметить, что функция на правом графике (рис. 4) не стремится к единице. При достаточно большом количестве аргументов функция пойдет на убывание и для большого количества аргументов будет асимптотически стремиться к нулю, как и на левом графике.

Существует аналитическое представление функции, определяющей количество запросов, максимизирующих преимущество алгоритма-различителя 1. Эта функция достаточно сложная, поэтому здесь не приводится. В табл.1 приведены вычисленные значения оптимального количества запросов для различных размеров блока и соответствующее преимущество.

Таблица 1

Размер полублока ($2n$)	Оптимальное число запросов	Максимальное преимущество
8	10	0.783876
16	54	0.97470815043
24	262	0.997727463
32	1206	0.99981553
40	5392	0.9999859

Алгоритм-различитель №2 для 3-раундовой цепи Фейстеля со случайными функциями в качестве раундовых преобразований [16].

Для k входных аргументов с различными правыми и одинаковыми левыми половинами проверяется выполнение равенства $\mathrm{R}_i \oplus T_i = \mathrm{R}_j \oplus T_j$ (см. рис. 2). В случае выполнения равенства хотя бы для одного аргумента возвращаемое значение будет «1», иначе «0».

Лемма 6. Вероятность выполнения равенства $R_i \oplus T_i = R_j \oplus T_j$ для набора из $k \geq 2$ аргументов $x_1 ..., x_k$ для 3-раундовой цепи Фейстеля (см. рис.2) при использовании алгоритма-различителя №2

$$\mathrm{P}_2 = 1 - \prod_{i=0}^{k-2} (1 - \frac{1}{2^n - i})^{2(k-(i+1))} \leq 1 - (1 - \frac{1}{2^n})^{k(k-1)}. \qquad (6)$$

Доказательство. Поскольку $R_i \oplus T_i = f_2(L_i \oplus f_1(R_i))$ (см. рис. 2), из равенства $R_i \oplus T_i = R_j \oplus T_j$ следует $f_2(L_i \oplus f_1(R_i)) = f_2(L_j \oplus f_1(R_j))$. Это условие выполнено, когда произошло совпадение выходного значения раундовой функции f_1 или f_2, т.е. $f_2(L_i \oplus f_1(R_i)) = f_2(L_j \oplus f_1(R_j))$ или $L_i \oplus f_1(R_i) = L_j \oplus f_1(R_j)$. Вероятность появления такой коллизии для одной пары равна $\frac{2}{2^n}$. Для k аргументов возможно создать $2C_2^k$ пар, соответственно суммарная вероятность выполнения условия $f_2(L_i \oplus f_1(R_i)) = f_2(L_j \oplus f_1(R_j))$ равна $\mathrm{P}_2 = 1 - \prod_{i=0}^{k-2} (1 - \frac{1}{2^n - i})^{2(k-(i+1))}$. Способ вычисления вероятности рассматривался при доказательстве леммы 2. В формуле (6) приведено аппроксимированное значение вероятности, вывод которого также рассматривалось выше.

Доказательство окончено.

Лемма 7. Вероятность выполнения условия $\mathrm{R}_i \oplus T_i = \mathrm{R}_j \oplus T_j$ для случайной перестановки при произвольных различных входах

$$\mathrm{P}_2^* = 1 - \prod_{i=0}^{k-2} (1 - \frac{2^n}{2^{2n} - 1 - i \cdot 2^n})^{k-(i+1)} \leq 1 - (1 - \frac{2^n}{2^{2n} - 1})^{\frac{k(k-1)}{2}}. \qquad (7)$$

Доказательство. Для случайной перестановки при произвольных различных входах выходное значение $(V_i \bullet T_i)$ является равномерно распределенной случайной величиной. Пусть $(L_1, R_1) => (V_1, R_1)$, и для перестановки имеем $(L_2, R_2) \neq> (V_1, R_1)$, т.е. множество допустимых выходных значений для второго запроса уменьшилось на один элемент и стало равным $2^{2n} - 1$. При этом $\mathrm{R}_1 \oplus T_1 \neq \mathrm{R}_2 \oplus T_1$, соответственно отбрасывается заведомо неудовлетворяющий уравнению вариант. Для подходящего значения правой половины на выходе T_2, соответствующей уравнению $\mathrm{R}_1 \oplus T_1 = \mathrm{R}_2 \oplus T_2$, существует 2^n вариантов левой половины (коллизия ищется на полублоке, а не на всем блоке). Из всего множества возможных $2^{2n} - 1$ значений уравнению будут удовлетворять 2^n, откуда для одной пары входных аргументов вероятность выполнения условия $\mathrm{R}_i \oplus T_i = \mathrm{R}_j \oplus T_j$ равна $P = \frac{2^n}{2^{2n} - 1}$.

Для k аргументов можно составить C_k^2 пар, соответственно суммарная вероятность выполнения условия $\mathrm{R}_i \oplus T_i = \mathrm{R}_j \oplus T_j$ равна $\mathrm{P} = 1 - \prod_{i=0}^{k-2} (1 - \frac{2^n}{2^{2n} - 1 - i \cdot 2^n})^{k-(i+1)}$. Способ опреде-

ления такой вероятности рассматривался при доказательстве леммы 2. В формуле (7) приведено аппроксимированное значение вероятности, вывод которого также рассматривалось выше.

Доказательство окончено.

Теорема 3. Преимущество алгорима-различителя №2 при $k \geq 2$ запросах ограничено сверху величиной

$$\mathrm{Adv}_2(\psi, \sigma_{2n}) = |\,\mathrm{P}_2 - \mathrm{P}_2^*\,| \leq \left| \prod_{i=0}^{k-2} \left(1 - \frac{1}{2^n - i}\right)^{2(k-(i+1))} - \prod_{i=0}^{k-2} \left(1 - \frac{2^n}{2^{2n} - 1 - i \cdot 2^n}\right)^{k-(i+1)} \right| \leq$$

$$\leq \left| \left(1 - \frac{1}{2^n}\right)^{k(k-1)} - \left(1 - \frac{2^n}{2^{2n} - 1}\right)^{\frac{k(k-1)}{2}} \right|.$$

Доказательство. Результат следует из леммы 6 и 7.

Преимущество алгоритма-различителя №2 для $n = 8$ и $n = 16$ представлено на рис.5. По оси абсцисс приведено количество аргументов, которые подаются на вход, по оси ординат отмечены значения преимущества (вероятность различения случайной перестановки и цепи Фейстеля).

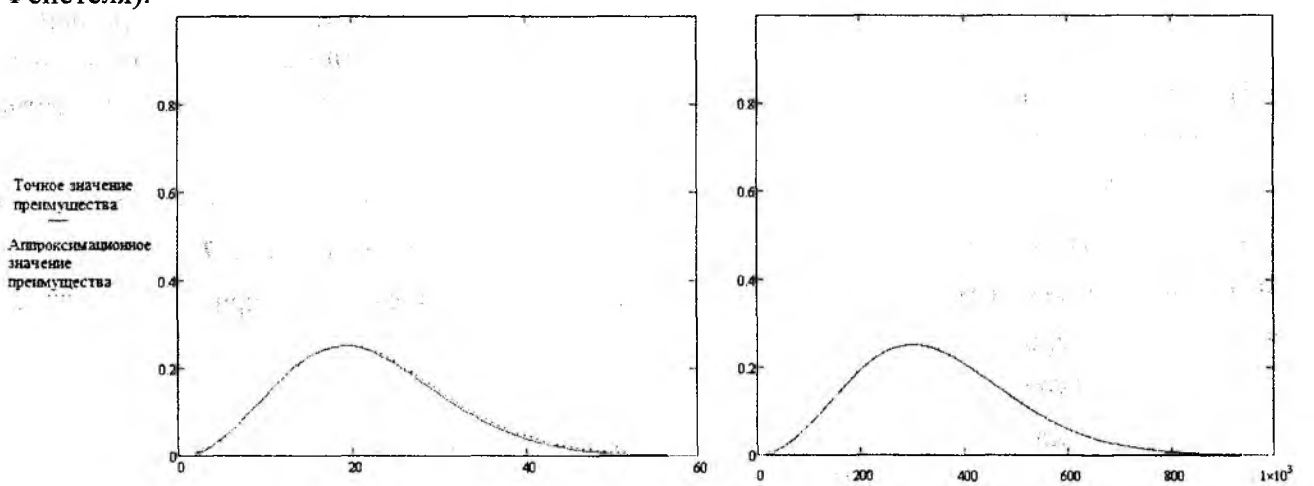

Рис. 5

Существует аналитическое представление функции, определяющей количество запросов, максимизирующих преимущество алгоритма-различителя 2. Эта функция достаточно сложная, поэтому не приводится здесь. В табл. 2 приведены вычисленные значения оптимального количества запросов для различных размеров блока и соответствующее преимущество.

Как видно из табл. 2, на оптимальном числе запросов преимущество алгоритма-различителя 2 асимптотически стремится к 0,25 с увеличением размера блока. Этот результат значительно хуже, чем для предыдущего алгоритма.

Таким образом, алгоритм-различитель №1 является значительно более эффективным, чем алгоритм-различитель №2.

Таблица 2

Размер полублока ($2n$)	Оптимальное число запросов	Максимальное преимущество
8	5	0.24859824088
16	19	0.24999470129
24	76	0.24999997934
32	302	0.24999999991
40	1206	0.24999999999

Оценка эффективности алгоритма-различителя для 3-раундовой цепи Фейстеля на основе случайных перестановок в качестве раундовых функций

Рассмотренные выше алгоритмы-различители использовали возможность появления коллизии в раундовой функции. Если же в качестве раундовых функций использовать случайные перестановки, где коллизии невозможны, то можно достигнуть еще лучших результатов по различению [17]. Отметим также, что модель со случайными функциями в раундовом преобразовании применима к DES, а модель со случайными перестановками соответствует ГОСТ 28147-89, Camellia и др.

Пусть f([L,R]) = [V,T], где f – 3-раундовая цепь Фейстеля (рис.2), при этом преобразования f_i^* – это случайные перестановки.

Алгоритм-различитель №3 для 3-раундовой цепи Фейстеля со случайными подстановками в качестве раундовых преобразований [9].

Для k входных аргументов с различными левыми и одинаковыми правыми половинами проверяется выполнение равенства $T_i \neq T_j$. В случае невыполнения равенства хотя бы для одного аргумента возвращаемое значение будет «0» (случайная перестановка), иначе «1» (цепь Фейстеля).

Лемма 8. Вероятность выполнения равенства $T_i \neq T_j$ для набора из $k \geq 2$ аргументов $x_1,...,x_k$ с различными левыми и одинаковыми правыми половинами (алгоритм-различитель №3) для 3-раундовой цепи Фейстеля (рис. 2) со случайными перестановками в качестве раундовых преобразований $P_3 = 1$.

Доказательство. Поскольку $R_i = R_j$, то $S_i \neq S_j$. Так как f_2^* – перестановка, то $f_2^*(S_i) \neq f_2^*(S_j)$. Соответственно $R_i = R_j$ и $f_2^*(S_i) \oplus R_i \neq f_2^*(S_j) \oplus R_j$, т.е. $T_i \neq T_j$ для всех различных входов. Таким образом, для 3-раундовой цепи Фейстеля неравенство $T_i \neq T_j$ выполняется с вероятностью $P_3 = 1$.

Доказательство окончено.

Лемма 9. Для случайной перестановки вероятность выполнения равенства $T_i \neq T_j$ для набора из $k \geq 2$ различных аргументов

$$P_3^* = \frac{2^{n(k-1)}\left(2^n-1\right)\left(2^n-2\right)..\left(2^n-k+1\right)}{\left(2^{2n}-1\right)\left(2^{2n}-2\right)..\left(2^{2n}-k+1\right)} = \frac{2^{n(k-1)}\left(2^n-1\right)_k}{\left(2^{2n}-1\right)_k} \leq \left(\frac{2^n}{2^n+1}\right)^{\frac{k(k-1)}{2}}.$$

Доказательство. При различных входных аргументах выходные значения случайной подстановки являются равномерно распределенной случайной величиной, причем в выходной последовательности элементы не повторяются. Количество комбинаций неповторяющихся правых половин ($T_i \neq T_j$) можно определить следующим образом.

Первый элемент всегда уникален, второй элемент можно выбрать $2^n\left(2^n-1\right)$ способами, третий – $2^n\left(2^n-2\right)$, и т.д., k-й – $2^n\left(2^n-k+1\right)$. Всего существует $2^{n(k-1)}\left(2^n-1\right)\left(2^n-2\right)..\left(2^n-k+1\right) = 2^{n(k-1)}\left(2^n-1\right)_k$ вариантов выбора уникальных элементов со 2-го по k-й.

В то же время всего возможно $\left(2^{2n}-1\right)$ вариантов выбора второго элемента выходного значения случайной подстановки, $\left(2^{2n}-2\right)$ – третьего элемента и т.д., $\left(2^{2n}-k+1\right)$ – k-го. Соответственно, возможно $\left(2^{2n}-1\right)\left(2^{2n}-2\right)..\left(2^{2n}-k+1\right) = \left(2^{2n}-1\right)_k$ вариантов выбора элементов выходных значений.

Отсюда вероятность появления на выходе случайной подстановки последовательности, у которой все правые половины будут уникальными:

$$P_3^* = \frac{2^{n(k-1)}\left(2^n-1\right)_k}{\left(2^{2n}-1\right)_k}.$$

Для получения формулы аппроксимации целесообразно воспользоваться тем, что вероятность уникального выбора второго элемента выше, чем любого последующего, $\frac{2^n\left(2^n-1\right)}{\left(2^{2n}-1\right)} = \frac{2^n}{\left(2^n+1\right)}$, а всего существует $C_k^2 = \frac{k(k-1)}{2}$ различных пар в последовательности, откуда верхняя граница вероятности выбора уникальной последовательности $P_3^* \le \left(\frac{2^n}{2^n+1}\right)^{\frac{k(k-1)}{2}}$.

Доказательство окончено.

Теорема 4. Преимущество алгорима-различителя №3 при $k \ge 2$ запросах ограничено сверху величиной

$$\mathrm{Adv}_3(\psi, \sigma_{2n}) = \left| P_3 - P_3^* \right| = 1 - \frac{2^{n(k-1)}\left(2^n-1\right)_k}{\left(2^{2n}-1\right)_k} \le 1 - \left(\frac{2^n}{2^n+1}\right)^{\frac{k(k-1)}{2}}.$$

Доказательство. Результат следует из леммы 8 и 9.

Преимущество алгоритма-различителя №3 для $n = 8$ и $n = 16$ представлено на рис.6. По оси абсцисс приведено количество аргументов, которые подаются на вход, по оси ординат отмечены значения преимущества (вероятность различения случайной перестановки и цепи Фейстеля).

В отличие от алгоритмов-различителей №1 и №2 в данном случае не существует оптимального количества запросов (кроме перебора всех элементов множества I_{2n}). Увеличение количества запросов максимизирует вероятность различения.

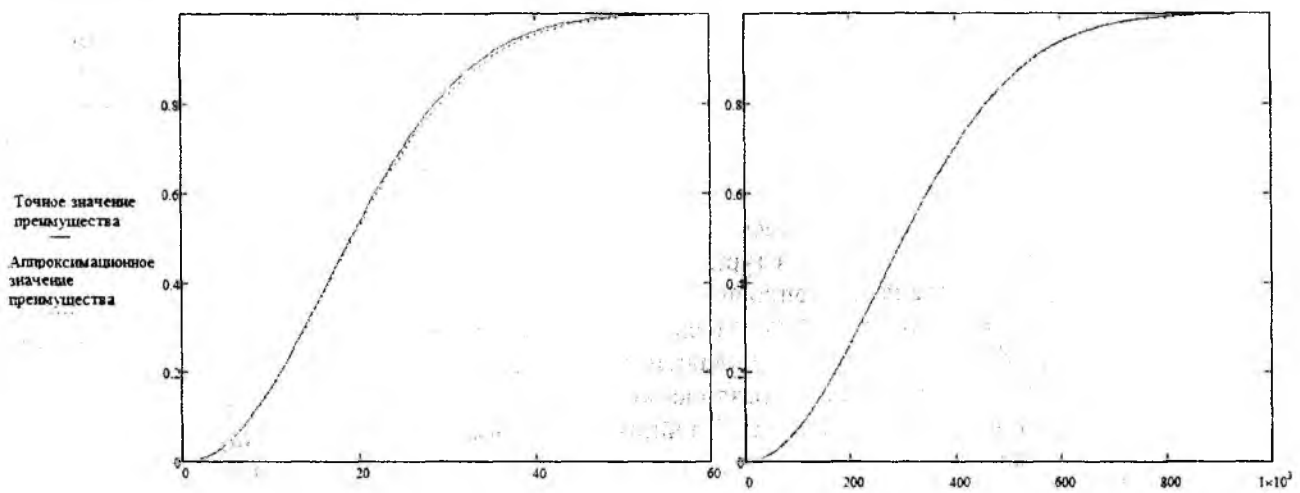

Рис. 6

Выводы

Использование модели идеального блочного шифра как случайной перестановки позволяет получить численную оценку эффективности высокоуровневой конструкции алгоритма шифрования, в рассмотренном случае – цепи Фейстеля.

Для исключения влияния свойств конкретной цикловой функции в качестве раундового преобразования целесообразно брать случайную функцию или случайную перестановку.

$$P_3^* = \frac{2^{n(k-1)}\left(2^n-1\right)_k}{\left(2^{2n}-1\right)_k}.$$

Для получения формулы аппроксимации целесообразно воспользоваться тем, что вероятность уникального выбора второго элемента выше, чем любого последующего, $\frac{2^n\left(2^n-1\right)}{\left(2^{2n}-1\right)} = \frac{2^n}{\left(2^n+1\right)}$, а всего существует $C_k^2 = \frac{k(k-1)}{2}$ различных пар в последовательности, откуда верхняя граница вероятности выбора уникальной последовательности $P_3^* \le \left(\frac{2^n}{2^n+1}\right)^{\frac{k(k-1)}{2}}$.

Доказательство окончено.

Теорема 4. Преимущество алгорима-различителя №3 при $k \ge 2$ запросах ограничено сверху величиной

$$\mathrm{Adv}_3(\psi, \sigma_{2n}) = \left| P_3 - P_3^* \right| = 1 - \frac{2^{n(k-1)}\left(2^n-1\right)_k}{\left(2^{2n}-1\right)_k} \le 1 - \left(\frac{2^n}{2^n+1}\right)^{\frac{k(k-1)}{2}}.$$

Доказательство. Результат следует из леммы 8 и 9.

Преимущество алгоритма-различителя №3 для $n = 8$ и $n = 16$ представлено на рис.6. По оси абсцисс приведено количество аргументов, которые подаются на вход, по оси ординат отмечены значения преимущества (вероятность различения случайной перестановки и цепи Фейстеля).

В отличие от алгоритмов-различителей №1 и №2 в данном случае не существует оптимального количества запросов (кроме перебора всех элементов множества I_{2n}). Увеличение количества запросов максимизирует вероятность различения.

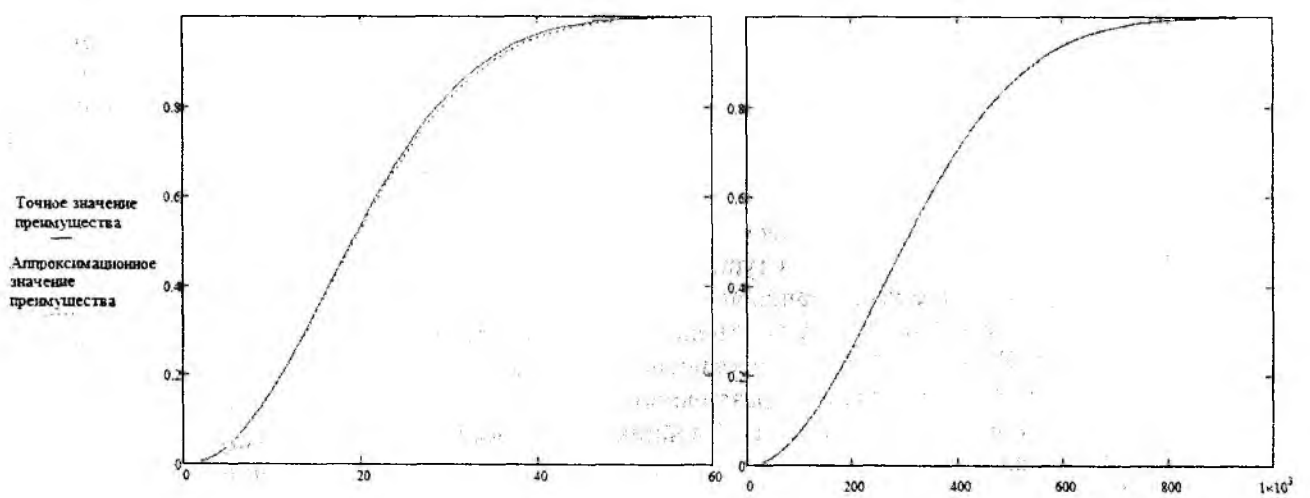

Рис. 6

Выводы

Использование модели идеального блочного шифра как случайной перестановки позволяет получить численную оценку эффективности высокоуровневой конструкции алгоритма шифрования, в рассмотренном случае – цепи Фейстеля.

Для исключения влияния свойств конкретной цикловой функции в качестве раундового преобразования целесообразно брать случайную функцию или случайную перестановку.

Первый вариант соответствует DES-подобной, второй – ГОСТ-подобной высокоуровневой конструкции блочного шифра.

Полученные результаты позволяют точно оценить верхнюю границу эффективности (преимущества) произвольного алгоритма-различителя для 3-раундовой цепи Фейстеля.

Для рассмотренных конкретных алгоритмов выведены точные значения преимущества, определен метод расчета оптимального количества запросов, при котором преимущество будет максимальным.

Для алгоритмов различения цепи Фейстеля на основе случайных функций существует конкретное значение количества запросов, на котором преимущество будет максимальным. В то же время для алгоритма различения цепи Фейстеля на основе случайных перестановок увеличение количества запросов непрерывно ведет к максимизации вероятности различения.

Из рассмотренных алгоритмов различения цепи Фейстеля на основе случайных функций наиболее эффективным является алгоритм №1.

Дополнительные аппроксимационные соотношения позволяют значительно упростить расчет вероятностей различения и преимущества алгоритмов-различителей с высокой точностью.

Список литературы: 1. *ГОСТ 28147-89*: Системы обработки информации. Защита криптографическая. Алгоритм криптограф. преобразования / http://protect.gost.ru/document.aspx?control=7&id=139177. 2. *FIPS 46-3*: Data Encryption Standard (DES) / http://csrc.nist.gov/publications/fips/fips46-3/fips46-3.pdf. 3. *RFC 4312*: The Camellia Cipher Algorithm and Its Use With IPsec / http://datatracker.ietf.org/doc/rfc4312/. 4. *Vaudenay, S.* Decorrelation: a theory for block cipher security. Journal of Cryptology, 16(4):249{286, 2003. 5. *Luby, M.* How to construct pseudorandom permutations from pseudorandom functions / M. Luby and C. Rackoff / SIAM J. Comput, Vol. 17, No. 2, April 1988. 6. *Maurer, M.* A simplified and Generalized Treatment of Luby – Rackoff Pseudorandom Permutations Generator, Advances in Cryptology EuroCrypt '92. 7. *Patarin, J.* Generic Attacks on Feistel Schemes, Asiacrypt '01 (Lecture Notes in Computer Science 2248), pp. 222-238, Springer-Verlag. 8. *Patarin, J.* Security of random Feistel schemes with 5 or more rounds. / In Matthew K.Franklin, editor, CRYPTO, volume 3152 of Lecture Notes in Computer Science, pages 106-122. Springer, 2004. 9. *Patarin, J.* New results on pseudorandom permutation generators based on the DES scheme. Crypto '91, pp. 301-312, Springer-Verlag. 10. *Patarin, J.* About Feistel Schemes with 6 (or More) Rounds. Fast Software Encryption 1998, pp. 103-121. 11. *Naor, M., Reingold, O.* On the Construction of pseudorandom permutations: Luby-Rackoff revisited. Journal of Cryptology, vol. 12, 1999, pp. 29-66. Extended abstract was published in Proc. 29th Ann. ACM Symp. on Theory of Computing, 1997, pp. 189-199. 12. *Gilbert, H., Minier, M.* New Results on the Pseudorandomness of Some Blockcipher Constructions // Mitsuru Matsui, editor, Fast Software Encryption – FSE 2001, volume 2355 of Lecture Notes in Computer Science, pages 248–266. Springer, 2001. 13. *Maurer, M., Oswald, Y., Pietrzak, K., Sjödin, J.* Luby-Rackoff Ciphers from Weak Round Functions?, pp. 391 – 408, EUROCRYPT 2006. 14. *Piret, G.* Luby-Rackoff Revisited: On the Use of Permutations as Inner Functions of a Feistel Scheme, Designs, Codes and Cryptography, Volume 39(2), pp. 233 – 245, Springer, 2006. 15. *Moriai, S., Vaudenay, S.* On the Pseudorandomness of Top-Level Schemes of Block Ciphers, pp. 289 – 302, ASIACRYPT 2000. 16. *Бондаренко, М. Ф.* Анализ основных атак на трёхраундовую цепь Фейстеля / М. Ф. Бондаренко, А. Б. Небывайлов // Прикладная радиоэлектроника. – ХНУРЭ, 2009. – Т.8. – С.278-281. 17. *Олейников, Р.В.* Эффективность различения случайной перестановки и цепи Фейстеля на основе биективной раундовой функции / Р.В. Олейников, Д.С. Кайдалов // VI Междунар. науч.-практ. конф. «Наука и социальные проблемы общества: информатизация и информационные технологии». – Харьков: ХНУРЭ, 2011.

Харьковский национальный университет радиоэлектроники

Поступила в редколлегию 03.12.2011