УДК 681.3.06: 519.248.681

В.И. ДОЛГОВ, И.В. ЛИСИЦКАЯ, С.А. ГОЛОВАШИЧ, А.С. БОНДАРЕНКО

ОБЕСПЕЧЕНИЕ СТОЙКОСТИ DES - ПОДОБНЫХ АЛГОРИТМОВ ШИФРОВАНИЯ К АТАКАМ ЛИНЕЙНОГО КРИПТОАНАЛИЗА ПРИ ИСПОЛЬЗОВАНИИ ТАБЛИЦ ПОДСТАНОВОК СЛУЧАЙНОГО ТИПА

В представленных ранее наших работах [1-3] рассматриваются пути и возможности обеспечения защищенности алгоритма шифрования DES от одной из опаснейших криптоаналитических атак – дифференциального криптоанализа. Этой работой мы продолжаем изучение вопросов безопасности шифра DES. Речь будет идти о другой криптоаналитической атаке, с помощью которой удалось поколебать уверенность в надежности американского стандарта – линейном криптоанализе. То, что таблицы стандарта не оптимизированы в отношении линейного криптоанализа, отмечается в ряде публикаций [3,4 и др.]. В [4] удалось найти краткое упоминание о работах по преодолению этого недостатка группы Кванджио Ким. Приводится даже пример построения таблиц S блоков, защищенных от атак линейного и дифференциального криптоанализа, однако не отмечается, насколько группе Кванджио Ким удалось продвинуться в этом направлении, как и не излагается сама методика отбора таблиц S блоков, а приведена лишь ссылка на критическое отношение Эли Бихама к результатам исследований отдельных этапов. Наши исследования показывают, что в отношении дифференциального криптоанализа приведенные в [4] таблицы нельзя считать надежными. Учитывая, что сама методика выполнения линейного криптоанализа, также как и возможности защиты от атак этого типа, остаются все еще мало изученными в Украине, в этой работе мы приводим краткое изложение принципов выполнения такой атаки для шифра DES и предлагаем свою версию решения задачи построения S блоков стандарта, защищенных от атак линейного криптоанализа.

Линейный криптоанализ – сравнительно новый тип криптонападения, предложенный Мацуи [3] в 1993 г. Этот метод использует линейную аппроксимацию для описания процедуры нападения на DES. Она заключается в нахождении ситуаций, когда сумма по модулю 2 некоторых битов открытого текста и некоторых битов соответствующего ему зашифрованного текста равна сумме по модулю 2 некоторых битов ключа. Если такая ситуация выполняется с некоторой вероятностью $p \neq 1/2$, то имеется возможность использовать собранные открытые тексты и соответствующие им шифрованные тексты для определения битов ключа.

Как и при дифференциальном криптоанализе алгоритма DES, сложность линейного криптоанализа определяется правилами построения S блоков [3], для описания свойств которых строятся специальные таблицы.

При построении таких таблиц просматриваются все возможные 4-битные выходы S блока, которые получаются при различных 6-битных значениях его входов. При этом вычисляются поразрядные произведения по модулю 2 входов S блока (6-битное число) и некоторого фиксированного 6-битного числа ("маски" по строкам) и соответствующие поразрядные произведения по модулю 2 выходов S блока и второго фиксированного теперь уже 4-битного числа ("маски" по столбцам). Эти фиксированные числа являются индексами входов в ячейку таблицы размера 64×16. Сама таблица, названная Мацуи линейной аппроксимационной таблицей, получается заполнением каждой из ячеек числом, соответствующим количеству линейных соотношений, выполняющихся для входных битов, прошедших маску по столбцам, и выходных битов, прошедших маску по строкам для этой ячейки, при вариации по всему множеству входов S блока. Сущность линейных соотношений заключается в равенстве нулю суммы по модулю 2 всех входных и выходных бит, прошедших обе маски. Математически отмеченные действия можно описать следующим образом.

Пусть, вектор $\bar{x}_i = \left(x_{i1}, x_{i2}, x_{i3}, x_{i4}, x_{i5}, x_{i6}\right)$ представляет собой 6-битное число, обозначающее один из $i = \overline{1,64}$ возможных входов S блока, а вектор $\bar{a}_l = \left(a_{l1}, a_{l2}, a_{l3}, a_{l4}, a_{l5}, a_{l6}\right)$, $l = \overline{1,64}$ – 6-битную входную «маску» (индекс таблицы по строкам).

Пусть, вектор $\bar{y}_p = \left(y_{p1}, y_{p2}, y_{p3}, y_{p4}\right)$ обозначает один $p = \overline{1,16}$ 4-битных выходов S блока, $\bar{b}_m = \left(b_{m1}, b_{m2}, b_{m3}, b_{m4}\right)$, $m = \overline{1,16}$ – 4-битную выходную «маску» (индекс таблицы по столбцам).

Обозначим $\bar{a}_l \cdot \bar{x}_i$ – двоичное скалярное произведение векторов $\bar{a}_l$ и $\bar{x}_i$, т.е. $\bar{a}_l \cdot \bar{x}_i = \bigoplus_{k=1}^{6} a_{lk} \cdot x_{ik}$.

Аналогично пусть $\bar{b}_m \cdot \bar{y}_p$ – двоичное скалярное произведение векторов $\bar{b}_m$ и $\bar{y}_p$, и, следовательно, $\bar{b}_m \cdot \bar{y}_p = \bigoplus_{k=1}^{4} b_{mk} \cdot y_{pk}$. В представленных выражениях символом $\oplus$ обозначена операция суммирования по модулю 2 (XOR).

Значения ячеек аппроксимационной таблицы можно представить в следующем аналитическом виде:

$$\theta_{lm} = \underset{\{\bar{x}_i\}}{\mu}\left(\bar{x}_i \cdot \bar{a}_l = \bar{y}_p \cdot \bar{b}_m\right) - 2^5 \Big|_{\,y_p = S(x_i);\ i=\overline{1,64},\ p=\overline{1,16}},$$

где функция $\mu(\cdot)$ обозначает количество случаев выполнения равенства в скобках при вариации по всем возможным значениям аргумента $\bar{x}_i$, $i = \overline{1, 64}$. В приведенной выше формуле использование различных индексов для входных и выходных значений S блока связано с тем, что размер множества входных значений $\{x_i\}$ в 4 раза больше размера множества выходных значений $\{y_p\}$. Слагаемое -2^5 используется для нормирования величин θ_{lm} относительно половинного (среднего) значения. В результате вероятность того, что аппроксимация в S блоке является правильной (соответствует значению линейной аппроксимационной таблицы θ_{lm}), дается выражением $p'_{lm} = (32 - \theta_{lm})/64 = 1/2 - \theta_{lm}/64$, или если ввести обозначение $p_{lm} = \theta_{lm}/64$, то $p'_{lm} = 1/2 - p_{lm}$. Вход со значением $p_{lm} = 0$, или что то же $\theta_{lm} = 0$ имеет вероятность $p' = 1/2$. Такой вход бесполезен для атаки на криптосистему. Любое ненулевое значение (положительное или отрицательное) может быть использовано для атаки.

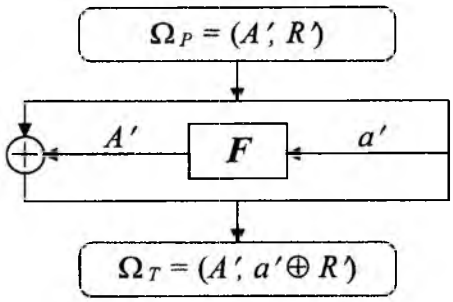
Здесь мы не будем сосредотачивать внимание на самой технике добывания ключей (основная атака Мацуи ориентирована на получение одного (единственного) бита ключа, а определение других битов требует уже дополнительных ухищрений [3]). Нас будет интересовать только то, что связано с оценкой стойкости алгоритма DES к рассматриваемой атаке. Это касается, прежде всего, построения аппроксимационных характеристик, под которыми понимаются системы взаимосвязанных линейных соотношений, распространенных на несколько циклов или S блоков процедуры шифрования, и оценки вероятностей одновременного выполнения всех линейных соотношений, попавших в цепочку (вероятностей аппроксимационных характеристик).

Приведем в связи с этим правила построения таких характеристик. Воспользуемся здесь обозначениями и определениями, предложенными в работе [3].

Определение 1. Одноцикловая характеристика есть форма $(\Omega_P, \Omega_T, \Omega_K, 1/2 + p)$, в которой $(\Omega_P)_L = (\Omega_T)_L = A$, $(\Omega_P)_R \oplus (\Omega_T)_R = a$, и для которой $1/2 + p$ есть вероятность того, что случайный входной блок P и его одноцикловое шифрование C с применением случайного подключа K удовлетворяет условию $P \cdot \Omega_P \oplus C \cdot \Omega_T \oplus K \cdot \Omega_K = 0$, где $(\cdot)$ обозначает двоичное скалярное произведение двух двоичных векторов, Ω_P – «маска», определяющая подмножество бит данных перед циклом, Ω_T – «маска», определяющая подмножество бит данных после цикла, и Ω_K – «маска», определяющая подмножество бит ключа, четность которых аппроксимируется; индексы L и R обозначают соответственно левую и правую половины блока данных.

Равенство $P \cdot \Omega_P \oplus C \cdot \Omega_T \oplus K \cdot \Omega_K = 0$ и есть линейная аппроксимация.

Примеры построения одноцикловых характеристик приведены на рис.1 и рис. 2 (здесь и далее мы пользуемся обозначениями и манерой изображения иллюстративного материала, предложенными в работе [3]).

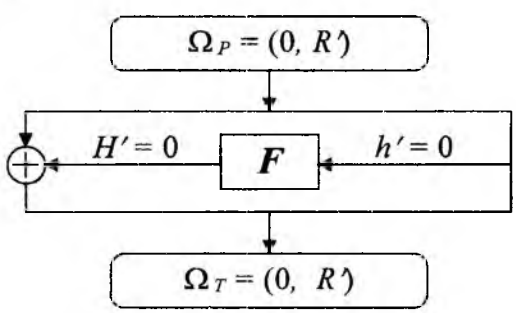


с вероятностью $1/2 - p_a$ задействован один ключевой бит

Рис. 1

Как и при дифференциальном криптоанализе аппроксимационные характеристики могут использовать более чем один S блок и распространяться на большее число циклов шифрования.

Определение 2. n – цикловая характеристика $\Omega^1 = \left(\Omega^1_P, \Omega^1_T, \Omega^1_K, 1/2 + p_1\right)$ может быть объединена с m – цикловой характеристикой $\Omega^2 = \left(\Omega^2_P, \Omega^2_T, \Omega^2_K, 1/2 + p_2\right)$, если Ω^1_T равно переставленному значению двух половинок Ω^2_P, т.е. $\left(\Omega^2_P\right)_L = \left(\Omega^1_O\right)_R$, $\left(\Omega^2_P\right)_R = \left(\Omega^1_O\right)_L$. Конкатенация характеристик Ω^1 и Ω^2 (если они могут быть объединены) есть $(n+m)$ – цикловая характеристика $\Omega = \left(\Omega^1_P, \Omega^2_T, \Omega^1_K \oplus \Omega^2_K, 1/2 + 2 \cdot p_1 \cdot p_2\right)$.



с вероятностью 1 нет задействованных ключевых бит

Рис. 2

Заметим здесь, что вероятность аппроксимации с двумя активными S блоками или вероятность двухцикловой аппроксимации есть $p'_1 p'_2 + \left(1 - p'_1\right)\left(1 - p'_2\right) = 1/2 + 2p_1 p_2$, так как сумма линейных соотношений будет линейным соотношением тогда, когда оба линейных соотношения равны нулю, и тогда, когда оба линейных соотношения равны единице (напомним, что $p'_i = 1/2 + p_i$).

Когда объединяется l характеристик с вероятностью p_i каждая (если это может быть выполнено), то вероятность результирующей характеристики определяется выражением

$$1/2 + p = 1/2 + 2^{l-1} \prod_{i=1}^{l} p_i \,. \qquad (1)$$

Заметим, наконец, что если линейная аппроксимация с вероятностью $1/2 + p$ известна, то атака на полный 16-цикловый DES требует около p^{-2} известных пар открытый – шифрованный текст, которые могут быть выбраны случайно [5].

В дальнейшем речь будет идти об атаках, использующих одноблочные характеристики. Введем понятие минимальной итеративной характеристики, под которой будем понимать характеристику, содержащую минимальное число циклов с задействованными ключевыми битами, среди которых имеется хотя бы один цикл тождественного типа (в котором нет задействованных ключевых битов), допускающую циклическое продолжение.

Для итеративной характеристики выполняется условие: выход характеристики является перестановкой левой и правой половинок входа, т.е. при входе $\Omega_P = (A', R')$ имеем $\Omega_T = (R', A')$, что и обеспечивает в соответствии с правилами построения характеристик (определения 1 и 2) ее циклическое продолжение. Заметим здесь сразу, что одной из особенностей линейного криптоанализа по сравнению с дифференциальным является то, что в линейной характеристике «свободной» является правая половина используемого множества входных бит, в то время как в дифференциальном криптоанализе это левая половина. Но тогда, если свободную часть взять равной нулю, то есть $R' = 0$, то тождественное одноцикловое преобразование, представленное на рис. 2, позволяет сразу ориентироваться на формирование тождественного нетривиального многоциклового преобразования $\Omega_P = (A', 0) \to \Omega_T = (A', 0)$, поскольку оно с помощью дополнительного тождественного (тривиального) преобразования приводится к требуемому виду: $\Omega_P = (A', 0) \to \Omega_T = (0, A')$. При этом удается еще один раз воспользоваться тождественным одноцикловым преобразованием, выполняющимся с вероятностью единица.

Заметим далее, что тождественное преобразование (нетривиального типа) можно реализовать только для нечетного числа циклов (в соответствии с правилом обмена левых и правых частей цепи Фейстеля). Наконец, можно сразу отметить, что нас интересуют возможности циклического продолжения характеристик, т.е. нас интересуют многоцикловые итеративные характеристики.

Теперь наша ближайшая задача построить такую характеристику для шифра DES. Будем исходить из принципа симметрии, который необходимо реализовать для входа и выхода итеративной характеристики с нечетным числом циклов. В качестве «центра симметрии» очевидно и должно выступить тождественное тривиальное преобразование (поскольку изначально рассматривается задача, когда такое преобразование имеется в единственном числе). В этой ситуации представляется естественным использование характеристики, представленной на рис. 3.

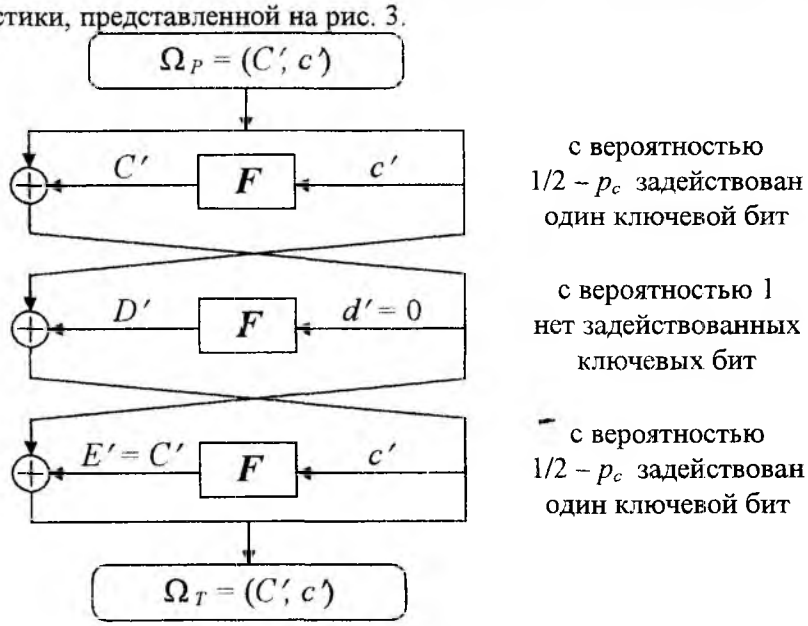


Рис.3

На этом рисунке представлена трехцикловая характеристика тождественного типа с ненулевыми значениями входов $\Omega_P = (C', c')$, причем $C' = F(c')$. Поскольку $E' = C'$, то, очевидно, что $e' = c'$. Очевидно также, что для этой характеристики симметричного типа c' и соответственно C' не могут быть равными нулю. Теперь нужно к этой характеристике подобрать начальную и конечную части, обеспечивающие получение итеративно продолжающейся характеристики $\Omega_P = (A', 0) \to \Omega_T = (A', 0)$. Это удается сделать, используя в обоих случаях двухцикловые характеристики (преобразования), представленные на рис. 4 и рис. 5.

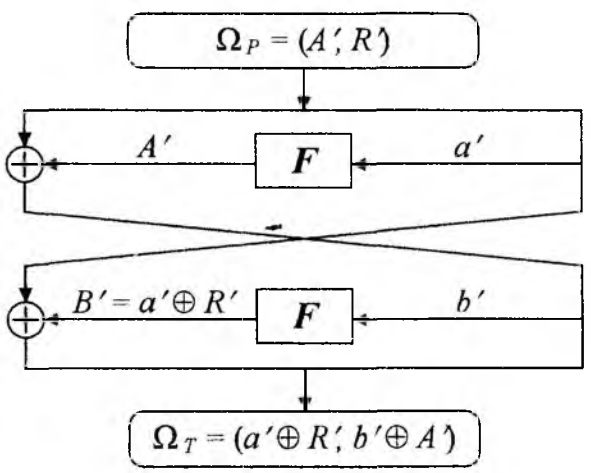
Условием сшивки характеристик, представленных на рис. 3, рис. 4 и рис. 5 выступают соотношения $C' = a' \oplus R'$, $\tilde{n}' = b' \oplus A'$.

Наша задача получить на выходе в качестве результата $\Omega_T = (A', R')$, что совпадает с исходным множеством битов $\Omega_P = (A', R')$ с точностью до порядка следования левой и правой половинок.

Для построенной семицикловой характеристики при $f' = a'$ и $g' = b'$ получаем $\Omega_T = (R', A')$. Остается сделать обмен левой и правой половинок множеств битов, участвующих в аппроксимации. Но как следует из рис. 3 такой обмен можно выполнить, если воспользоваться дополнительным тождественным циклом, который в свою очередь требует выполнения условия $R' = 0$.

В результате мы приходим к итеративной восьмицикловой характеристике, представленной на рис. 6. Именно характеристика такого типа использована в атаке Бихама [3]. Важно здесь подчеркнуть, что нельзя построить характеристики с большим числом тождественных преобразований, приходящихся на 8 циклов.

Любая другая характеристика с уменьшенным числом тождественных преобразований будет иметь результирующую вероятность не выше, чем вероятность характеристики минимального типа.



с вероятностью $1/2 - p_a$ задействован один ключевой бит

с вероятностью $1/2 - p_b$ задействован один ключевой бит

Рис.4

Для вероятности этой восьмицикловой характеристики в соответствии с (1) имеем:

$$P_8 = 2 \cdot \left(4 \cdot p_a p_b p_c\right)^2 = 2^5 \cdot \left(p_a p_b p_c\right)^2.$$

Для 16 цикловой характеристики, получающейся при итеративном продолжении минимальной 8-цикловой характеристики, соответственно получим:

$$P_{16} = 2 \cdot P_8^2 = 2^{11} \cdot \left(p_a p_b p_c\right)^4.$$

Следовательно, все показатели стойкости S блоков к атакам, использующим минимальные характеристики, определяются возможными значениями произведения трех вероятностей $p_a p_b p_c$, две из которых относятся к одному и тому же входу одного и того же S блока, а третья относится к другому S блоку, участвующему в формировании аппроксимационной характеристики. Действительно, из правил построения минимальной характеристики следует

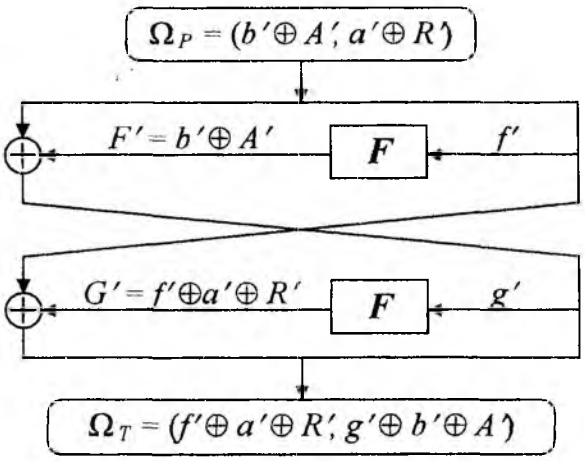
$$A' = F(a'),\ C' = F(a') \text{ и при этом } B' = F(b'), \text{ где } b' = A' \oplus C'. \qquad (2)$$

Но тогда, чтобы защититься от атаки, использующей минимальную характеристику, достаточно выбрать таблицы S блоков, исходя из условия, что максимально возможное значение произведения этих трех вероятностей удовлетворяет требованию

$$\left(P_{16}\right)^2 \le 2^{-55} \rightarrow p_a p_b p_c \le \sqrt[8]{2^{-77}} = 2^{-10}. \qquad (3)$$

Заметим здесь, что $2^{-10} = \dfrac{256}{64^3}$.

Теперь можно перейти к формированию критериев отбора S блоков устойчивых к атакам линейного криптоанализа. Его можно сформулировать как критерий для проверки линейных аппроксимационных таблиц для 3-цикловых аппроксимационных характеристик (являющихся первыми тремя циклами 8-цикловой минимальной характеристики).



с вероятностью $1/2 - p_f$ задействован один ключевой бит

с вероятностью $1/2 - p_g$ задействован один ключевой бит

Рис.5

Основная идея построения такой 3-цикловой аппроксимационной характеристики, как следует из соотношений (2), состоит в использовании двух наиболее вероятных одноблочных характеристик для одних и тех же однобитных входов (одного и того же S блока), разнесенных на один цикл. При этом сумма по модулю два выходов этих S блоков должна образовывать однобитный вход другого S блока на промежуточном цикле, также участвующего в формировании аппроксимационной характеристики. Выход этого S блока в свою очередь должен совпадать с входами разнесенных S блоков.

Тогда критерии для отбора таблиц S блоков, позволяющих защититься от атак, построенных на использовании таких 3-цикловых характеристик, можно сформулировать следующим образом.

Требование. Для обеспечения устойчивости шифра DES к известным атакам линейного криптоанализа необходимо и достаточно, чтобы максимальное значение произведения вероятностей $p_a p_b p_c$ одноцикловых характеристик, соответствующих формам $\left(\Omega_P^A, \Omega_T^A, \Omega_K^A, 1/2 + p_a\right)$, $\left(\Omega_P^B, \Omega_T^B, \Omega_K^B, 1/2 + p_b\right)$ и $\left(\Omega_P^C, \Omega_T^C, \Omega_K^C, 1/2 + p_c\right)$, где $\Omega_P^A = \left(A', 0\right) \rightarrow \Omega_T^A = \left(A', a'\right)$, $\Omega_P^B = \left(a', A'\right) \rightarrow \Omega_T^B = \left(a', A' \oplus b'\right)$, $\Omega_P^C = \left(A' \oplus b', a'\right) \rightarrow \Omega_T^C = \left(A' \oplus b', 0\right)$ (т.е. выполняются ограничения (2)), было меньше порогового значения 2^{-10}.

Фактически это трехцикловая характеристика вида $\left(\Omega_P^1, \Omega_T^3, \Omega_K^A \oplus \Omega_K^B \oplus \Omega_K^C, 1/2 + 4 p_a p_b p_c\right)$, для которой выполняется условие: при $\Omega_P^1 = \Omega_P^A = \left(A', 0\right)$ имеем $\Omega_T^3 = \Omega_T^C = \left(A' \oplus b', 0\right)$.

Приведем несколько замечаний относительно самой методики выполнения проверок.

Базовый метод, разработанный Мацуи и развитый Бихамом, использует одноблочные характеристики. Это значит, что вход b' является однобитным, так как в соответствии с (2) имеем $b' = F(a') \oplus F(c')$ – это сумма по модулю 2 выходных битов одного и того же S блока ($a' = c'$). А в соответствии с правилом завершающей цикловую функцию P подстановки выходы идентичных S блоков могут сформировать вход в один из S блоков только в случае, когда сумма $F(a') \oplus F(c')$ есть один единственный бит. Очевидно также, что должны быть однобитными и входы a' и соответственно $c' = a'$. В результате анализу подлежат все однобитные входные «маски» линейных аппроксимационных таблиц S блоков (индексы по строкам): 1_x, 2_x, 4_x, 8_x, 10_x и 20_x.

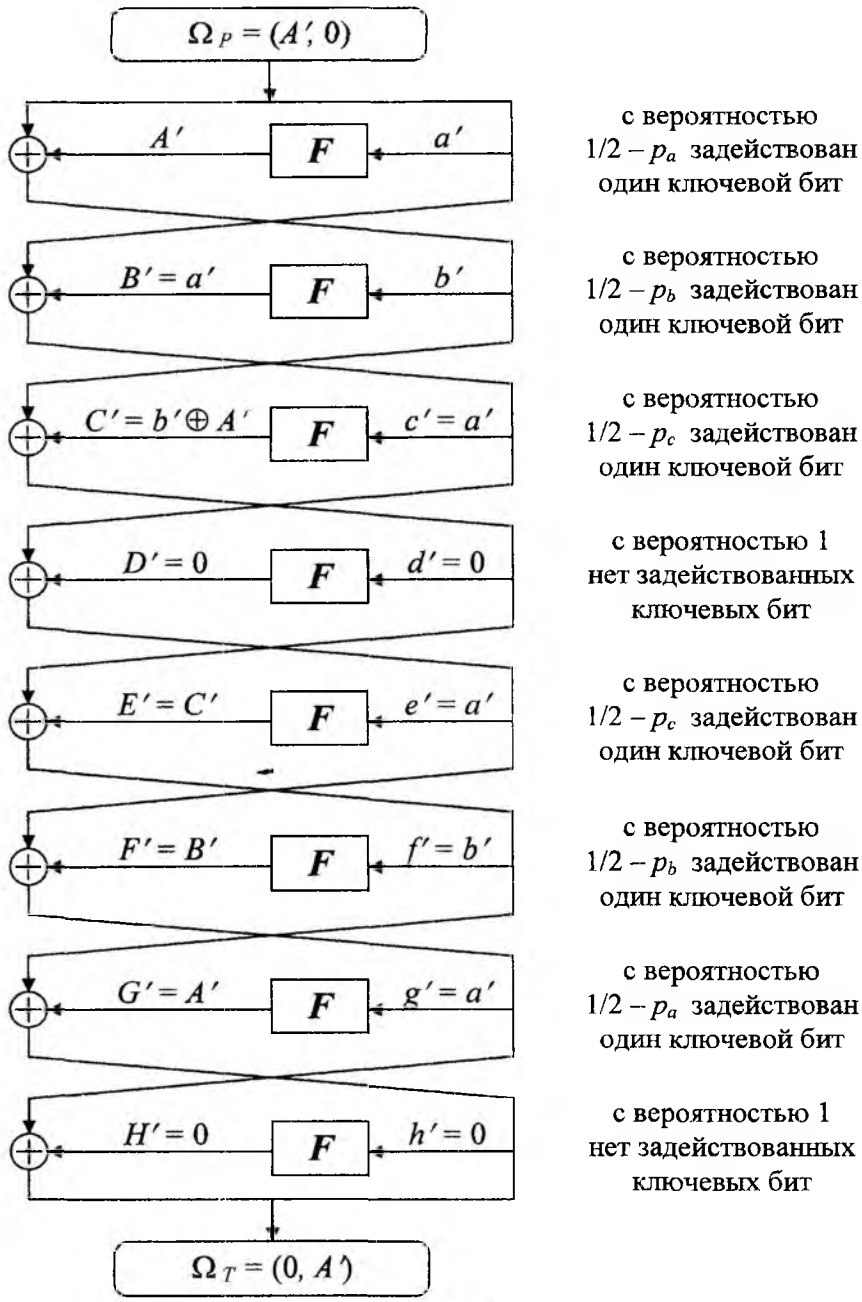


Рис. 6

Заметим, что в соответствии с принципами построения S блоков DES крайние пары битов 6-битных входов каждого из S блоков являются одновременно и входными битами соседних S блоков. Представленные варианты значений входов 1_x, 2_x, 4_x, 8_x, 10_x и 20_x (000001, 000010, 000100, 001000, 010000, 100000), тем не менее, всегда активизируют лишь один из S блоков, так как для входов (входных «масок») 1_x и 20_x линейных аппроксимационных таблиц S блоков все значения выходов (линейных аппроксимационных таблиц) являются нулевыми (для любой маски выходных значений одна и та же подстановка дает сбалансированный результат для числа одинаковых выходов), т.к. биты, соответствующие этим двум «маскам»,осуществляют выбор одной из четырёх перестановок, которые составляют S блок (уравновешенный результат для всех множеств битов, высекаемых из всех элементов перестановки).

При осуществлении проверки следует учитывать только те пары ячеек таблицы, для которых при одинаковой входной «маске» сумма по модулю два выходных «масок» даёт один единственный бит. Это пары входов по столбцам 1_x и 3_x, 1_x и 5_x, 1_x и 9_x, 2_x и 3_x, 2_x и 6_x, 2_x и A_x, 3_x и 7_x, 3_x и B_x, 4_x и 5_x, 4_x и 6_x, 4_x и C_x, 5_x и 7_x, 5_x и D_x, 6_x и 7_x, 6_x и E_x, 7_x и F_x, 8_x и 9_x, 8_x и A_x, 8_x и C_x, 9_x и

B_x, 9_x и D_x, A_x и B_x, A_x и E_x, B_x и F_x, C_x и D_x, C_x и E_x, D_x и F_x – всего 27 вариантов. Из этого числа сразу исключаются пары, для которых абсолютное значение хотя бы одной из ячеек аппроксимационной таблицы меньше или равно 2 (результирующая вероятность $p_a p_b p_c$ для этих случаев выходит за допустимые границы). Для отобранных пар определяются значения p_a и p_c, вычисляется соответствующее им значение $b' = F(a') \oplus F(c')$ и S блок для этого входа, затем по аппроксимационным таблицам находится значение p_b, при котором $F(b') = a'$.

Найденные значения p_a, p_b, p_c и проверяются на соответствие установленному критерию. При самом пессимистическом подходе всего потребуется выполнить $27 \cdot 4 \cdot 8 = 864$ проверок. На самом деле их будет на много меньше.

Разработанные критерии отбора таблиц подстановок устойчивых к атакам линейного криптоанализа были применены к таблицам S блоков группы Кванджио Ким, приведенным в [4]. Наши комментарии в отношении дифференциального криптоанализа представлены в начале статьи. Что касается линейного криптоанализа, то по результатам нашей проверки они полностью удовлетворяют выдвинутым в работе критериям отбора S боков.

Список литературы: **1.** *Лисицкая И.В., Головашич С.А., Олешко О.И., Олейников Р.В., Коряк А.С.* Построение таблиц подстановок для стандарта шифрования данных // Проблемы бионики. 1999. Вып.50. С. 185–194. **2.** *Лисицкая И.В., Олейников Р.В., Головашич С.А., Коряк А.С., Олешко О.И.* Анализ стойкости DES подобных алгоритмов шифрования при использовании таблиц подстановок случайного типа // Радиоэлектроника и информатика 1999. № 1. Стр 111–114. **3.** *Eli Biham* On Matsyi's Linear Cryptanalysis. Technion – Comput Science Department -Technic Report CS0813 - 1994. P 1-17. **4.** *Schneier B.* Applied Cryptography. *S*econd Edition: protocols, algorithms, and Source code in C. Published by John Wiley & *S*on*S*. Inc, New York: Chiche*S*ter Bri*S*bane Toronto *S*ingapore, 1996 – 758 p. **5.** *Горбенко И.Д., Лисицкая И.В.* Критерии отбора случайных таблиц подстановок для алгоритма шифрования по ГОСТ 28147-89 // Радиотехника. 1997. Вып. 103. С. 121–130.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 15.03.2000