УДК 681.3.06: 519.248.681

В. И. ДОЛГОВ, д-р техн. наук, И. В. ЛИСИЦКАЯ, канд. техн. наук,
Р. В. ОЛЕЙНИКОВ, С. А. ГОЛОВАШИЧ, А. С. КОРЯК.

ДОПОЛНИТЕЛЬНЫЕ ТРЕБОВАНИЯ К ОТБОРУ ТАБЛИЦ ПОДСТАНОВОК ДЛЯ ГОСТ 28147-89

В существующих работах обосновывается методика отбора случайных подстановок и случайных таблиц подстановок в интересах построения долговременных ключей для алгоритма ГОСТ 28147-89 (в дальнейшем просто ГОСТ)[1-5]. В последующих публикациях [6-7] делается, однако, вывод, что хотя используемые в шифрах DES и ГОСТ таблицы подстановок (*S* блоков) и удовлетворяют критериям случайности [5], однако выполнение одних только показателей случайности не дает уверенности в безопасности шифра. Необходимо еще выполнить проверку защищенности шифра от известных криптоаналитических атак. Для шифров DES и ГОСТ речь, прежде всего, идет об атаках дифференциального и линейного криптоанализа [8,9]. В работе [8] показано, что случайные *S*-блоки делают шифр DES более слабым к этим атакам по сравнению с *S*-блоками разработчиков стандарта. А в работе [10] прямо указывается на то, что для окончательной уверенности в пригодности случайных таблиц подстановок в качестве долговременных ключей для шифра ГОСТ также необходима проверка устойчивости этого шифра и к отмеченным криптоаналитическим атакам.

Следует сразу подчеркнуть, что анализу стойкости шифра DES уделено очень большое внимание в печати, и с момента его появления он находится под постоянным и пристальным вниманием специалистов. Именно для этого шифра и были впервые разработаны эффективные принципы дифференциального, а в дальнейшем и линейного криптоанализа. Наибольшее различие между шифрами DES и ГОСТ автор работы [11] видит в том, что в цикловой функции ГОСТа используется циклический сдвиг вместо перестановки. Он отмечает, что в DES перестановка приводит к существенному увеличению лавинного эффекта. Если в шифре DES для того, чтобы изменение одного входного бита повлияло на каждый выходной бит, достаточно 5 циклов, то в ГОСТе для этого необходимо уже 8-9 циклов. Однако разработчикам ГОСТа удалось добиться некоторого повышения его надежности или повышения трудности его криптоанализа, о чем по всей вероятности и свидетельствует весьма ограниченное число посвященных этому шифру публикаций. Этому, конечно, в значительной степени способствует также то, что таблица подстановок в шифре ГОСТ является конфиденциальным параметром и может меняться по желанию пользователя, что создает дополнительные трудности для криптоанализа.

В этой работе мы все же хотим высказать некоторые соображения, связанные с анализом устойчивости шифра ГОСТ к атакам дифференциального и линейного криптоанализа, которые позволяют обосновать дополнительные ограничения к отбору *S*-блоков и придать методике отбора случайных таблиц подстановок для шифра ГОСТ 28147-89 более завершенный вид.

Как известно, для шифра DES операцией, позволяющей исключить при вычислении дифференциалов влияние ключевых бит, является операция суммирования по модулю 2 (XOR). Именно с помощью операции XOR в этом алгоритме в цикловую функцию вводятся биты подключа. В ГОСТе ключевые биты вводятся в цикловую функцию путем суммирования по модулю 2^{32}, что при вычислении разностей вызывает определенные проблемы, связанные с необходимостью учета переносов разрядов. Поэтому первые наши предложения по построению атак дифференциального криптоанализа ГОСТ [12] были предприняты при использовании разностей, вычисляемых опять-таки с помощью операции XOR.

Переходя к построению дифференциальной характеристики, полезно напомнить о выводе работы [12] о том, что при использовании для формирования разностей операции суммирования по модулю 2 для ГОСТа существуют "слабые" подстановки, которые позволяют построить характеристики для этого шифра, имеющие высокие вероятности. Основу этих "слабых" *S*-блоков составляют специфические подстановки. Они обладают тем свойством, что для них вероятность перехода фиксированной входной разности в фиксированную выходную разность равна 1.

И в наших исследованиях мы остановились на использовании подстановок, обладающих отмеченным свойством. Однако мы здесь приведем пример построения дифференциальной характеристики, отличающейся от работы [12], которая примечательна тем, что имеет более высокую вероятность. Она включает в себя две шестицикловые характеристики, представленные на рис. 1 и 2.

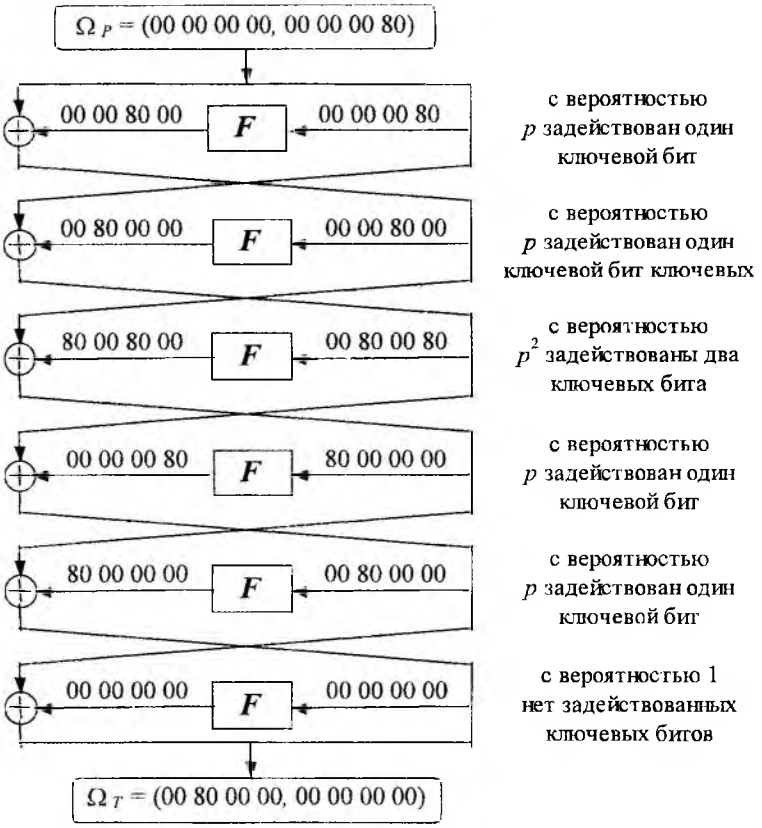
Ω_P = (00 00 00 00, 00 00 00 80)
00 00 80 00
F
00 00 00 80
с вероятностью p задействован один ключевой бит
00 80 00 00
F
00 00 80 00
с вероятностью p задействован один ключевой бит ключевых
80 00 80 00
F
00 80 00 80
с вероятностью p^2 задействованы два ключевых бита
00 00 00 80
F
80 00 00 00
с вероятностью p задействован один ключевой бит
80 00 00 00
F
00 80 00 00
с вероятностью p задействован один ключевой бит
00 00 00 00
F
00 00 00 00
с вероятностью 1 нет задействованных ключевых битов
Ω_T = (00 80 00 00, 00 00 00 00)

Рис.1

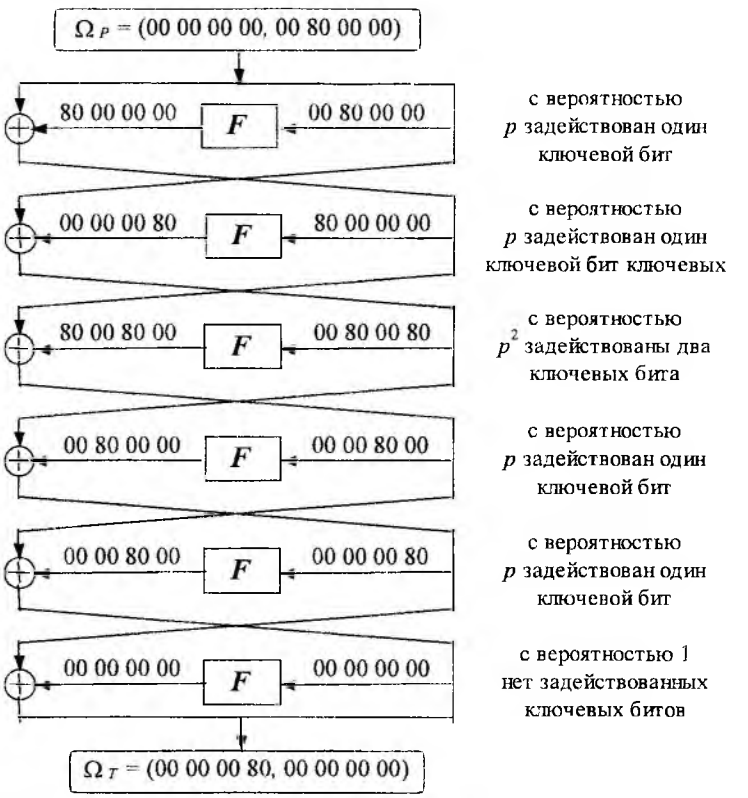
Ω_P = (00 00 00 00, 00 80 00 00)
80 00 00 00
F
00 80 00 00
с вероятностью p задействован один ключевой бит
00 00 00 80
F
80 00 00 00
с вероятностью p задействован один ключевой бит ключевых
80 00 80 00
F
00 80 00 80
с вероятностью p^2 задействованы два ключевых бита
00 80 00 00
F
00 00 80 00
с вероятностью p задействован один ключевой бит
00 00 80 00
F
00 00 00 80
с вероятностью p задействован один ключевой бит
00 00 00 00
F
00 00 00 00
с вероятностью 1 нет задействованных ключевых битов
Ω_T = (00 00 00 80, 00 00 00 00)

Рис.2

При построении этой дифференциальной характеристики использованы таблицы подстановок с переходами фиксированной разности 8h (1000_2) в 1h (0001_2). Всего в 32-цикловой характеристике, получающейся при итеративном продолжении шестицикловых характеристик (рис.1 и 2), используется 31 активный *S*-блок. Важной особенностью этой характеристики следует считать то, что из восьми *S*-блоков в ее построении участвуют только 4, а именно 2-й, 4-й, 6-й и 8-й (четные *S*-блоки), т.е. отмеченное выше ограничение на переходные характеристики касается только четырех из восьми *S*-блоков. На нечетные *S* блоки никаких ограничений не накладывается. Однако при этом возможно восстановить лишь биты ключа, соответствующие активным подстановкам на последних циклах, а не весь 256-битовый сеансовый ключ.

Приведем соображения по оценке возможностей реализации атак с применением указанной характеристики. Нас будет интересовать вероятность получения на входе *S*-блока разностей, не зависящих от битов ключа.

Рассмотрим тетрады полублоков, формирующие фиксированные входные разности $\Delta = 8\text{h}$. Всего возможно 16 значений таких разностей: 0-8, 1-9, 2-10, 3-11, 4-12, 5-13, 6-14, 7-15, 8-0, 9-1, 10-2, 11-3, 12-4, 13-5, 14-6, 15-7 (8 значений в прямом и 8 этих же значений в обратном сочетании). При суммировании тетрад, формирующих фиксированные разности, с соответствующими ключевыми битами возможны три ситуации:

1) переносов разрядов нет ни для одного из слагаемых;
2) имеется перенос только для одного из слагаемых;
3) оба четырехбитных слагаемых имеют переносы разрядов в следующую тетраду.

Заметим, что здесь, как и ранее, идентичные переносы разрядов в каждом из четырехбитных слагаемых не мешают построению интересующей нас характеристики, ей будет препятствовать только ситуация 2. Для того, чтобы оценить вероятность прохождения разности через сумматор по модулю 2^{32} без искажения, рассмотрим ситуации, которые могут возникнуть для различных пар входов. Они представлены вместе с показателями прохода через сумматор в табл. 1.

С учетом того, что значения ключевых тетрад равновероятны, для среднего значения числа входных разностей, которые проходят через сумматор по модулю 2^{32} без искажения, можно получить результат

$$N_{пр} = 1 + \frac{1}{16}(2 + 4 + 6 + 8 + 10 + 12 + 14) = 7.$$

В этой формуле дробь 1/16 перед скобкой есть вероятность выбора одного из равновероятных значений ключевого слова. Поэтому результирующая вероятность прохода через сумматор входной разности $\Delta = 8h$ без изменений равна $p_{пр} = 1/2$.

Заметим, что переносы, возникающие при прохождении через сумматор по модулю 2^{32} в предшествующей паре четырехбитных слов (а они могут быть только двойными), легко учитываются увеличением значения четырехбитного ключа для текущей пары слов на единицу. Это значит, что полученный выше результат не меняется. Кроме того, для полублоков, отличающихся 32-м битом, единица переноса уходит за границы 32-битного блока, и, следовательно, вероятности прохода соответствующих разностей ($\Delta = 8h$) через сумматор по модулю 2^{32} без переносов равны 1.

Таблица 1

Значение ключевых бит	Благоприятные сочетания пар выходов (от → до)	Число благоприятных исходов
0	0 - 8 → 7 - 15	8 (переносов нет)
1	0 - 8 → 6 - 14	7 (переносов нет)
2	0 - 8 → 5 - 13	6 (переносов нет)
3	0 - 8 →4 - 12	5 (переносов нет)
4	0 - 8 → 3 - 11	4 (переносов нет)
5	0 - 8 → 2 - 10	3 (переносов нет)
6	0 - 8 → 1 - 9	2 (переносов нет)
7	0 - 8 → 0 - 8	1 (переносов нет)

Продолжение табл. 1

Значение ключевых бит	Благоприятные сочетания пар выходов (от → до)	Число благоприятных исходов
8	0	0
9	7 - 15 → 7 - 15	1 (двойной перенос)
10	6 - 13 → 7 - 15	2 (двойной перенос)
11	5 - 12 → 7 - 15	3 (двойной перенос)
12	4 - 11 → 7 - 15	4 (двойной перенос)
13	3 - 10 → 7 - 15	5 (двойной перенос)
14	2 - 9 → 7 - 15	6 (двойной перенос)
15	1 - 8 → 7 - 15	7 (двойной перенос)

Все представленные выше результаты получены при использовании случайного набора пар шифруемых текстов и полного перебора по всему множеству возможных ключей. Для вероятности 32-цикловой характеристики, использующей переходы рис. 1 и 2, с учетом того, что из 32-х активных *S*-блоков, задействованных в ней, шесть имеют вероятности прохода входной разности через сумматор по модулю 2^{32} без изменения, равные 1, а 26 – равные $p_{пр} = 1/2$, приходим к результату:

$$P_{32} = \left(p_{пр}\right)^{26} \cdot p^{32} = 2^{-26} \cdot p^{32}. \qquad (1)$$

Заметим, что при получении этого и предыдущих результатов использовано предположение о взаимной независимости ключей на различных циклах шифрования, в то время как в шифре ГОСТ используется восемь сеансовых 32-битных подключей. При шифровании они циклически повторяются, причем, на последних восьми циклах порядок использования подключей меняется на обратный. Распределение подключей по циклам шифрования в привязке к дифференциальной характеристике, строящейся с использованием переходов рис. 1 и 2, представлено в табл. 2. В ячейках таблицы указаны ненулевые биты разностей пар правых полублоков, участвующих в построении этой характеристики.

Из таблицы следует, что преобразование выполняется с вполне определенным набором цикловых побитных разностей. Для определения вероятности благоприятного прохода соответствующих разностей через k активных *S*-блоков с переносами разрядов через циклы с одним и тем же ключом можно записать выражение:

$$P_k = \frac{1}{16} \cdot \left(1 + 2 \cdot \left(\left(\frac{14}{16}\right)^k + \left(\frac{12}{16}\right)^k + \left(\frac{10}{16}\right)^k + \left(\frac{8}{16}\right)^k + \left(\frac{6}{16}\right)^k + \left(\frac{4}{16}\right)^k + \left(\frac{2}{16}\right)^k\right)\right).$$

Отметим здесь, что переносы из 32-го бита "уходят" за пределы полублока и их можно игнорировать.

В итоге, с учетом распределения активных *S*-блоков с переносами по подключам, представленного в табл. 2, для вероятности 32-цикловой дифференциальной характеристики получаем результат

$$p_{пр} = \left(P_4\right)^4 \cdot P_5 \cdot P_3 \cdot \left(P_1\right)^2 = 6{,}28 \cdot 10^{-7} \approx 2^{-20}, \qquad (2)$$

что подтверждается экспериментально полученным $p_{пр} = 2 \cdot 10^{-7}$.

Таблица 2

Подключи K_i	K_1	K_2	K_3	K_4	K_5	K_6	K_7	K_8
Биты 1 - 8 циклов	8	16	8,24	32	24	–	24	32
Биты 9 - 16 циклов	8,24	16	8	–	8	16	8,24	32
Биты 17 - 24 циклов	24	–	24	32	8,24	16	8	–
Биты 32 - 25 циклов	32	24	–	24	32	8,24	16	8
Число активных *S* блоков с переносами	4	3	4	1	4	4	5	1

Как видно из результатов (1) и (2), наличие детерминированной связи в используемых ключах приводит к заметному увеличению вероятности получения приемлемой для атаки дифференциальной характеристики.

Таким образом, и для шифра ГОСТ 28147-89 при выполнении описанных условий открываются возможности для выполнения криптоанализа на основе атаки с выбранными открытыми текстами. Все рассмотренные атаки строятся на том, что существуют подстановки, для которых фиксированная входная разность с большой вероятностью переходит в фиксированную выходную (более того, как мы увидим из дальнейшего, существует значительное число подстановок, для которых этот переход осуществляется с вероятностью 1). Будем, как и ранее [12], такие подстановки называть "слабыми".

Очевидно, что в этих условиях значительный практический интерес приобретают ответы на два принципиальных вопроса:

1) обеспечивают ли предлагаемые в [5] критерии отбора подстановок и таблиц подстановок фильтрацию "слабых" подстановок и таблиц подстановок?

2) если же это не так, то какие нужно ввести дополнительные ограничения на отбор подстановок и таблиц подстановок, чтобы исключить "слабые" ключи?

Для ответа на эти вопросы и были предприняты исследования, направленные на изучение особенностей формирования и оценки характеристик множества подстановок и таблиц подстановок, являющиеся "слабыми" в рассматриваемом смысле.

Отметим в связи с этим результаты еще одной из редких работ [13], посвященных анализу шифра ГОСТ 28147-89. В этой работе при оценке устойчивости шифра ГОСТ к атакам дифференциального криптоанализа авторы обнаружили подстановки, для которых ненулевой входной дифференциал ΔX приводит к ненулевому выходному значению ΔY с вероятностью 1. К примеру, в подстановке $\pi = (15,13,9,11,7,14,10,3,2,12,8,6,0,5,1,4)$, все входы с дифференциалом $\Delta X = 3 = 0011$ приводят к выходному дифференциалу $\Delta Y = 4 = 0100$. Всего, как отмечают авторы [13], было найдено 1096 перестановок с таким свойством.

Выполненные нами теоретические расчеты позволили получить выражение для общего числа подстановок, для которых существует переход фиксированной входной разности в фиксированную выходную разность с вероятностью 1, в виде:

$$N_1 = 15^2 \cdot 8!2^8 = 2{,}32 \cdot 10^9.$$

Для вероятности P_{N_1} случайного выбора подстановки, удовлетворяющей рассматриваемому ограничению, соответственно получаем результат

$$P_{N_1} = \left(\frac{2 \cdot 10^{13}}{2{,}33 \cdot 10^9} \right)^{-1} = (9009)^{-1}.$$

Это означает, что на каждые 9000 случайных подстановок встречается одна подстановка с переходом фиксированной входной разности в фиксированную выходную разность с вероятностью 1. Полученный результат, очевидно, хорошо согласуется с данными статистического эксперимента. Более того, как показывает анализ и эксперименты, существуют подстановки, которые обладают отмеченным свойством одновременно для нескольких вариантов сочетаний входных и выходных разностей.

В результате существует вполне реальная возможность построения таблиц подстановок с одним и тем же фиксированным переходом входной разности в выходную. Так, если считать, что всего существует $N_1' = 8!2^8 = \cdot 10^7$ подстановок с одним и тем же фиксированным переходом входной разности в выходную, то из этих подстановок можно сформировать $\sim \left(10^7\right)^8 = 10^{56}$ различных таблиц подстановок полностью специфического типа.

Заметим также, что общее число подстановок, имеющих переход фиксированной разности $\Delta X = 8h$ в $\Delta Y = 1h$, определяется формулой

$$N_1 = 8! \cdot 2^8 = 10^7.$$

Наконец, отметим дополнительно, что для получения "слабой" таблицы подстановок в рассмотренном примере из восьми подстановок только четыре должны иметь специфический вид (четные *S* блоки).

Наш анализ показал, что из всех случайных подстановок, проходящих критерии, приведенные в [5], примерно половина оказались слабыми.

Изложенные соображения и результаты позволяют сформулировать дополнительные требования к отбору *S*-блоков, выполнение которых, на наш взгляд, позволит гарантировать устойчивость шифра ГОСТ 28147-89 к атакам дифференциального криптоанализа.

Требование 1. Для защиты от атак дифференциального криптоанализа максимально допустимое значение таблиц распределения разностей *S*-блоков для ненулевых входов не должно превышать значения 8, при этом не менее, чем 99% ячеек таблицы дифференциальных разностей *S* блоков не должны превышать значения, равного 4.

При этих ограничениях даже в том случае, если встречается таблица, которая имеет идентичные для всех подстановок переходы фиксированных входных разностей в выходные, равные максимально допустимым (8), для вероятности ранее рассмотренной дифференциальной характеристики, составленной из 31 активного *S* блока, получим оценочное значение:

$$2^{-21}\cdot\left(\frac{8}{16}\right)^{31}=2^{-53}.$$

Более высокий уровень защищенности можно обеспечить, если предъявить дополнительное ограничение на подстановки, из которых составляется таблица, по степени подобия соответствующих им таблиц распределения разностей. Такое ограничение может быть, например, представлено в виде дополнительного требования.

Требование 2. Подстановки, попавшие в таблицу, при выполнении требования 1, не должны иметь сколько-нибудь существенных идентичных переходов (переходов одного типа) входной разности в выходную, или в более конкретном выражении – *S*-блоки не должны иметь более двух максимально возможных значений переходов одного типа.

Тогда, полагая, что два из четырёх активных *S*-блоков, участвующих в формировании рассматриваемой характеристики, дают максимальное значение вероятности прохода *S*-блока с заданными (фиксированными) разностями (например, если считать, что максимальное значение в табл. 3 дают второй и шестой *S*-блоки), получим: $2^{-21}\cdot\left(\frac{8}{16}\right)^{8}\cdot\left(\frac{4}{16}\right)^{23}=2^{-73}$.

Этот результат уже представляется существенно более надежным.

Во всех представленных выше оценках определяются значения вероятностей дифференциальных характеристик путем статистического усреднения по всему возможному множеству значений ключей. Естественно, что эффективность атаки зависит от значения неизвестного ключа. Пределы, в которых меняется успех атаки с характеристикой, составленной из шестицикловых характеристик рис.1, рис.2, определяются границами изменения первого множителя в (1). Он может быть равным 1 при ключевых битах $K_h = 0000$ (с нулевыми значениями позиций) четырехбитных входов четных *S*-блоков и равным 0 при $K_h = 8h = 1000_2$ (K_h – четырехбитный сегмент ключа, взаимодействующий с ненулевой разностью $\Delta = 8h$).

Что касается линейного криптоанализа, то, как уже указывалось выше, в ГОСТе биты подключа вводятся на каждом цикле с помощью операции суммирования по модулю 2^{32}. Эта операция выполняется с переносом разрядов, что не позволяет построить линейные соотношения, зависящие только от битов подключей (не удается избавиться от зависимости результирующего линейного соотношения от битов шифруемого текста), т.е. биты ключа не могут быть объединены непосредственно. Эта же мысль проводится в [13].

Отмеченное позволяет утверждать, что атаки линейного криптоанализа (по крайней мере, в том варианте, который предлагается автором линейного криптоанализа М. Мацуи) для шифра ГОСТ неприменимы.

В результате приведенные выше дополнительные ограничения к отбору случайных *S*-блоков позволят, на наш взгляд, сформировать таблицы подстановок (долговременные ключи) для шифра

ГОСТ 28147-89, которые обеспечат стойкость этого алгоритма от рассмотренных криптоаналитических атак.

Остается заметить, что известные нам два варианта таблиц подстановок, построенных разработчиками алгоритма ГОСТ 28147-89, полностью удовлетворяют предложенным выше критериям отбора.

Список литературы: 1. *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. 2. *Лисицкая И.В.* К вопросу построения долговременных ключей для алгоритма ГОСТ 28147-89 // Информационно-управляющие системы на железнодорожном транспорте. 1997. № 3. С. 54–57. 3. *Бильчук В.М., Лисицкая И.В.* К вопросу построения долговременных ключей для алгоритма ГОСТ 28147-89. Критерии отбора второго уровня // Информационно-управляющие системы на железнодорожном транспорте. 1998. № 1. С. 10–17. 4. *Горбенко И.Д., Лисицкая И.В., Коряк А.С.* Анализ стойкости алгоритма ГОСТ 28147-89 при использовании подстановок случайного типа // Радиоэлектроника и информатика. 1998. №1 (02). С. 39–43. 5. *Лисицкая И.В., Коряк А.С.* Уточненные критерии отбора таблиц подстановок с заданными критериями случайности. Радиотехника. 2000. Вып 114. С. 47–56. 6. *Лисицкая И.В., Головашич С.А., Олейников Р.В.* Построение таблиц подстановок для стандарта шифрования данных // Проблемы бионики. 1999. Вып 50. С. 185-194. 7. *Лисицкая, И.В., Олейников Р.В., Головашич С.А.* Анализ стойкости DES - подобных алгоритмов шифрования при использовании таблиц подстановок случайного типа. // Радиоэлектроника и информатика. 1999. № 1. С. 109-115. 8. *Biham E., Shamir A.* Differential Cryptanalysis of the full 16-round DES. Technical Report - Computer Science Department, Technion, Israiel, 1993. 9. *Matsui M.* Linear Cryptanalysis Method for the DES Cipher. // Proc. of Eurocrypt'93, Norway, 1993. 10. *Фаль А.М.* Алгоритм шифрования по ГОСТу 28147-89 и способы применения блочных шифров // Безопасность информации. 1995. №3. С. 8-11. 11. *Schneier B.* Applied Cryptography. Second Edition: protocols, algorithms, and Source code in C. Published by John Wiley & SonS. Inc, New York: ChicheSter BriSbane Toronto Singapore, 1996 – 758 p. 12. *Долгов В.И., Лисицкая И.В., Олейников Р.В.* "Слабые" ключи в алгоритме шифрования ГОСТ 28147-89 // Радиотехника. 2000. Вып 114. С. 63–68. 13. *C. Charnes, L. O'Connor, J. Pieprzyk, R. Safavi-Naini, Y. Zheng.* Further Comments on the Soviet Encryption Algorithm // Wollongong, NSW 2500, AUSTRALIA 1994. pp. 1-10.

Харьковский государственный технический университет радиоэлектроники

Поступила в редколлегию 19.03.2001