УДК 681.3.06:519.248.681

*В.И.ДОЛГОВ, д-р техн. наук, И.В. ЛИСИЦКАЯ, канд. техн. наук,*
*Р.В.ОЛЕЙНИКОВ, А.И.ШУМОВ*

# «СЛАБЫЕ» КЛЮЧИ В АЛГОРИТМЕ ШИФРОВАНИЯ ГОСТ 28147-89

Одной из наиболее универсальных и мощных криптоаналитических атак на симметричные системы шифрования в настоящее время является дифференциальный криптоанализ. Он был первым успешным криптонападением на американский стандарт DES, который до этого более 15 лет считался неуязвимым. Поэтому эта атака обязательно учитывается при оценке стойкости любой современной симметричной системы шифрования.

Напомним основные положения дифференциального криптоанализа [1]. Атакующий имеет возможность управлять разностями пар открытых (незашифрованных) блоков на входе шифратора и имеет доступ к его выходу. Для уязвимых алгоритмов существуют разности между парами открытых текстов, которые проходит через все циклы алгоритма шифрования с вероятностью выше пороговой. Далее, зная входные и выходные значения открытых и зашифрованных текстов, криптоаналитик имеет возможность получить наиболее вероятные значения ключа шифрования. Успех атаки зависит от вероятности нахождения пары открытых текстов, разность которых приводит к специфической разности шифртекстов.

Для DES-подобных шифров (к числу которых относится и ГОСТ 28147-89) устойчивость к дифференциальному криптоанализу в значительной мере определяется свойствами таблиц подстановок (так называемых S-блоков). Именно на основе анализа свойств S-блоков была предложена методика определения ключей для нескольких DES-подобных шифров со сложностью, меньшей чем прямой перебор.

Отечественный стандарт ГОСТ 28147-89 введен в действие гораздо позже DES. Несмотря на то, что и в ГОСТе единственным нелинейным преобразованием, как и в DES, является подстановка, тем не менее в открытой литературе практически нет публикаций, посвященных изучению его стойкости к различным атакам. Предполагается, что за счет использования вдвое большего числа циклов, чем DES, ГОСТ обладает более высокой защищенностью от многих известных криптоаналитических атак. В нашей работе сделана попытка применить к ГОСТ 28147-89 элементы дифференциального криптоанализа и доказать существование и в этом алгоритме определённых слабостей.

Хотелось бы обратить внимание на некоторую аналогию между процессом поиска пар со специфическими различиями в дифференциальном криптоанализе и проверкой статистической безопасности симметричного шифра. И в дифференциальном криптоанализе, и при проверке статистической безопасности интересуются изменениями шифрованных текстов при небольших изменениях в соответствующих им открытых текстах. При оценке статистической безопасности изучается так называемый лавинный эффект, для которого требуется, чтобы изменение хотя бы одного бита открытого текста или ключа изменяло бы в шифртексте примерно половину битов. Результаты статистических испытаний и исследование лавинного эффекта в алгоритмах DES и ГОСТ 28147-89, приведенные в [2] для случайных открытых текстов, показывают, что оба шифра реализуют хорошие показатели статистической безопасности. В таблице 1 представлены данные из этой работы, характеризующие лавинный эффект на различных циклах шифрования для шифра ГОСТ. Использованы обозначения: $m_w$ – математическое ожидание числа $W(\Delta_k)$ единичных бит в булевой побитной сумме $\Delta_k$ (сумме по модулю 2) для пары полученных на $k$-ом цикле шифртекстов, открытые тексты которых отличаются одним битом, $\sigma^2_w$ – дисперсия числа единичных бит для этой же побитовой суммы.

Отсюда следует, что устойчивый лавинный эффект достигается уже на восьмом-девятом цикле шифрования. Поскольку стандарт использует 32 цикла шифрования, можно сделать предположение о хорошей статистической безопасности алгоритма.

Исследование лавинного эффекта позволяет оценить лишь среднестатистические характеристики процедуры шифрования и не исключает наличия отдельных пар открытых текстов, отличающихся малым числом битов, результат шифрования которых также отличается малым числом битов. Именно на этих особенностях и строится дифференциальный криптоанализ.

*Таблица 1*

| Номер цикла | Изменен 1-й бит сообщения | | Изменен 31-й бит сообщения | | Изменен 63-й бит сообщения | |
|---|---|---|---|---|---|---|
| $k$ | $m_w$ | $\sigma^2_w$ | $m_w$ | $\sigma^2_w$ | $m_w$ | $\sigma^2_w$ |
| 1 | 3.33 | 0.67 | 2.33 | 0.55 | 1.00 | 0.00 |
| 2 | 7.30 | 4.32 | 5.39 | 3.15 | 3.00 | 0.78 |
| 3 | 12.33 | 15.88 | 10.02 | 9.81 | 5.99 | 3.21 |
| 4 | 18.14 | 23.50 | 15.50 | 19.20 | 9.98 | 10.03 |
| 5 | 24.28 | 29.44 | 21.24 | 28.03 | 15.38 | 20.46 |
| 6 | 28.97 | 23.86 | 26.45 | 25.36 | 21.60 | 26.91 |
| 7 | 31.32 | 18.75 | 30.41 | 20.93 | 26.23 | 32.13 |
| 8 | 31.80 | 16.14 | 31.65 | 16.26 | 29.64 | 24.04 |
| 9 | 32.17 | 16.07 | 32.10 | 16.16 | 31.61 | 16.51 |

Задачей этой работы и является изучение возможностей прохождения через циклы шифрования ГОСТ 28147-89 таких специфических различий. Для решения этой задачи воспользуемся элементами дифференциального криптоанализа.

Нас будет интересовать вероятность прохождения некоторой фиксированной разности через нелинейное преобразование (подстановку). Под разностью будем понимать, как и в классической атаке [1], сложение по модулю 2 пар открытых и зашифрованных текстов, а также промежуточных значений. Для вычисления интересующей нас вероятности используются так называемые таблицы распределения разностей $S$ блоков, которые строятся следующим образом. Перебираются все возможные комбинации пар входов (все возможные пары из чисел от 0 до 15), а затем для каждой из них определяются значения результатов подстановки. Составляется таблица (размер которой 16x16), в которой индексом (входом) ячейки по строкам будет побитовая сумма по модулю 2 входных значений, а индексом (входом) по столбцам – сумма по модулю 2 выходных значений. Сами ячейки заполняются числами, соответствующими количеству попаданий в каждую из них при заданных входах. Для получения вероятностей нужно разделить эти числа на 16 (количество возможных пар входов).

Пример случайной подстановки и её таблицы распределения разностей приведен в табл.2 и 3. Отметим, что индексы в таблицах записаны в шестнадцатеричной системе счисления.

*Таблица 2*

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 7 | 4 | 2 | 13 | 9 | 3 | 1 | 8 | 0 | 6 | 14 | 10 | 15 | 5 | 11 | 12 |

*Таблица 3*

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 2 | 2 | 0 | 2 | 2 | 0 | 2 | 4 | 0 | 0 | 0 | 0 | 2 |
| 2 | 0 | 0 | 0 | 0 | 2 | 2 | 0 | 0 | 2 | 4 | 0 | 2 | 2 | 0 | 2 | 0 |
| 3 | 0 | 2 | 2 | 2 | 0 | 0 | 2 | 0 | 2 | 0 | 4 | 0 | 0 | 0 | 2 | 0 |
| 4 | 0 | 0 | 0 | 4 | 0 | 4 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 |
| 5 | 0 | 2 | 2 | 0 | 2 | 2 | 0 | 0 | 0 | 2 | 2 | 0 | 2 | 2 | 0 | 0 |
| 6 | 0 | 2 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 2 | 4 | 2 | 0 | 2 | 2 |
| | | | | | | | | | | | | | | | | |

*Окончание табл. 3*

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 7 | 0 | 2 | 0 | 0 | 2 | 4 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 0 | 2 |
| 8 | 0 | 0 | 2 | 0 | 2 | 0 | 4 | 4 | 0 | 0 | 2 | 0 | 2 | 0 | 0 | 0 |
| 9 | 0 | 2 | 0 | 4 | 2 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 4 | 2 | 0 | 0 |
| A | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 2 | 0 | 2 | 4 | 2 |
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
| B | 0 | 0 | 0 | 0 | 4 | 2 | 0 | 2 | 2 | 0 | 2 | 0 | 0 | 4 | 0 | 0 |
| C | 0 | 4 | 2 | 0 | 0 | 2 | 0 | 0 | 2 | 4 | 0 | 0 | 0 | 0 | 0 | 2 |
| D | 0 | 0 | 2 | 2 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 2 | 2 |
| E | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 2 | 4 | 2 | 0 | 0 | 2 | 2 | 2 | 0 |
| F | 0 | 0 | 2 | 2 | 0 | 0 | 0 | 4 | 2 | 0 | 0 | 2 | 0 | 2 | 0 | 2 |

Из табл.3 следует, что с вероятностью $\frac{2}{16}$ входная разность 01h[1] (пары чисел 0 и 1, 2 и 3, 4 и 5 и т.д. – всего 16 комбинаций) переходит в выходную 03h. Аналогично можно определить вероятность перехода между произвольными заранее заданными входными и выходными разностями.

Отметим, что именно свойства таблиц распределения разностей в значительной мере определяют устойчивость шифра к дифференциальной атаке. Приведенная в качестве примера случайная таблица подстановки не имеет переходов с вероятностью, равной единице, и поэтому обладает неплохими свойствами устойчивости к дифференциальному криптоанализу. Однако всего существует $N_p = 16! \approx 2{,}09 \cdot 10^{13}$ подстановок типа «4 бита в 4», и некоторые из них гораздо менее устойчивы к дифференциальному криптоанализу, чем рассмотренная. Пример такой «слабой» подстановки и соответствующая таблица распределения разностей приведены в табл.4 и 5.

*Таблица 4*

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 6 | 5 | 2 | 9 | 11 | 13 | 15 | 10 | 3 | 14 | 7 | 0 | 12 | 4 | 8 |

*Таблица 5*

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 4 | 0 | 4 | 0 | 0 | 0 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 4 | 0 | 4 | 0 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 0 | 4 | 0 | 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 12 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 8 | 0 | 0 | 0 | 4 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 8 | 0 | 4 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 8 | 0 | 0 | 0 | 4 | 0 | 4 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 8 | 0 | 4 | 0 | 4 | 0 |
| A | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 4 | 0 | 0 | 0 | 8 |
| B | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 4 | 0 | 8 | 0 | 0 |
| C | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 4 | 0 | 0 | 0 | 8 | 0 |
| D | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 4 | 0 | 8 | 0 | 0 | 0 |
| E | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 8 | 0 | 0 | 0 | 4 | 0 | 4 |
| F | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 8 | 0 | 4 | 0 | 4 |

Как видно из табл.5, входная разность 01h всегда (с вероятностью $p = \frac{16}{16} = 1$) переходит в выходную разность 02h. Отметим, что при случайной генерации перестановок заданному свойству удовле-

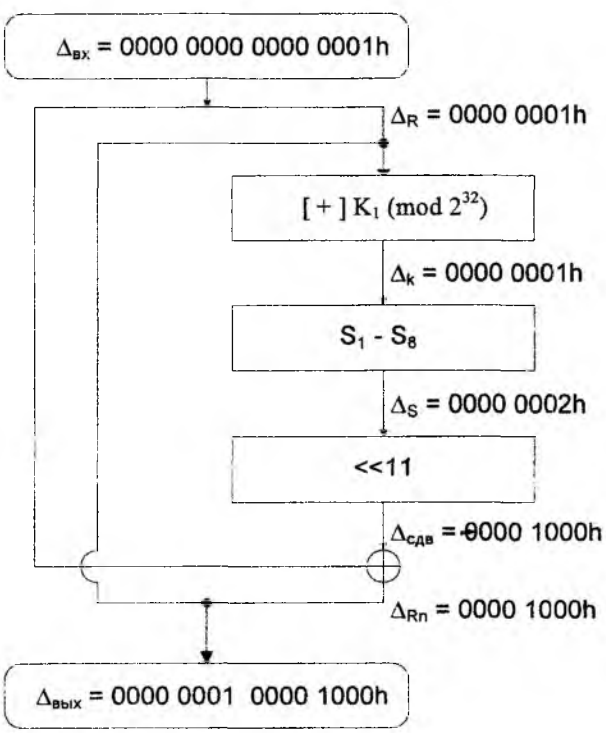
[1] Символ h после числа здесь и далее определяет запись числа в шестнадцатеричной системе счисления

творяет, в среднем, одна из 100 000, поэтому количество «слабых» подстановок можно оценить, приблизительно как $N_w = 10^{-5} \cdot N_p \approx 2{,}09 \cdot 10^8$.

Покажем, что при использовании таких «слабых» подстановок в ГОСТе для большого количества сеансовых ключей отмеченные выше показатели статистической безопасности не обеспечиваются.

Рассмотрим долговременный ключ, составленный из подстановок, обладающих заданным переходом, и будем проводить шифрование на одном и том же сеансовом ключе двух 64-битовых блоков, отличающихся младшим битом в правом полублоке. Входная разность перед первым циклом шифрования будет 0000 0000 0000 0001h. Левая половина (все нули в разности) будет гаммироваться с выходным значением цикловой функции, а правая без изменений поступит на следующий цикл и на вход цикловой функции в текущем цикле.

Первой операцией цикловой функции является сложение с ключом по модулю $2^{32}$, при которой могут возникнуть переносы из младших разрядов в старшие. Поскольку для вычисления разности мы используем операцию сложения по модулю 2, то из-за влияния переносов существует некоторая вероятность, что после сложения с ключом разность изменится. Для исключения возникновения переноса в младшем разряде будем рассматривать ключи, у которых соответствующий бит равен нулю. Тогда разность 0000 0001h в младшей тетраде без изменения перейдет на вход подстановки и после неё трансформируется в значение 0000 0002h. После сдвига влево на 11 разрядов разность принимает вид 0000 1000h. Гаммирование с левой половиной двоичной разности на входе цикла даёт следующее значение разности на выходе первого цикла: 0000 0001 0000 1000h. Рассмотренные преобразования представлены на рисунке:




Описанные преобразования выполняются всегда для любых входных блоков, отличающихся младшим битом и чётным подключом. Вероятность прохождения разности через ключевое преобразование (сложение с ключом по модулю $2^{32}$) зависит от значений битов ключа. В дальнейшем под «абсолютно слабым» ключом будем понимать такой, на котором вся 32-битная разность проходит через ключевой сумматор без изменений. Назовём ключ «слабым», если 32-битная разность проходит через ключевой сумматор без изменений с некоторой вероятностью, отличной от нуля.

Можно убедиться, что необходимым условием для прохождения разности через ключевое преобразование без изменения на «абсолютно слабом» ключе будет отсутствие переносов в младших 29 разрядах. Оно выполняется при любых входных значениях, если применяется подключ, у которого 29 младших битов нулевые. Если все подключи удовлетворяют этому условию, то прохождение разностей можно продлить на два цикла и более – вплоть до последнего цикла. Этапы этого преобразования в ходе шифрования иллюстрирует табл.6 (все значения разностей даны в шестнадцатеричной системе счисления).

*Таблица 6*

| Номер | Разность (hex) | |
|---|---|---|
| цикла | левая половина | правая половина |
| 0 | 0000 0000 | 0000 0001 |
| 1 | 0000 0001 | 0000 1000 |
| 2 | 0000 1000 | 0100 0001 |
| 3 | 0100 0001 | 0000 0010 |
| 4 | 0000 0010 | 0101 0001 |
| 5 | 0101 0001 | 1000 1000 |
| 6 | 1000 1000 | 0001 0101 |
| 7 | 0001 0101 | 0010 0000 |
| 8 | 0010 0000 | 0001 0100 |
| 9 | 0001 0100 | 1000 0000 |
| 10 | 1000 0000 | 0001 0000 |
| 11 | 0001 0000 | 0000 0000 |
| 12 | 0000 0000 | 0001 0000 |
| 13 | 0001 0000 | 1000 0000 |
| 14 | 1000 0000 | 0001 0100 |
| 15 | 0001 0100 | 0010 0000 |
| 16 | 0010 0000 | 0001 0101 |
| 17 | 0001 0101 | 1000 1000 |
| 18 | 1000 1000 | 0101 0001 |
| 19 | 0101 0001 | 0000 0010 |
| 20 | 0000 0010 | 0100 0001 |
| 21 | 0100 0001 | 0000 1000 |
| 22 | 0000 1000 | 0000 0001 |
| 23 | 0000 0001 | 0000 0000 |
| 24 | 0000 0000 | 0000 0001 |
| 25 | 0000 0001 | 0000 1000 |
| 26 | 0000 1000 | 0100 0001 |
| 27 | 0100 0001 | 0000 0010 |
| 28 | 0000 0010 | 0101 0001 |
| 29 | 0101 0001 | 1000 1000 |
| 30 | 1000 1000 | 0001 0101 |
| 31 | 0001 0101 | 0010 0000 |
| 32 | 0010 0000 | 0001 0100 |

«Абсолютно слабые» сеансовые ключи имеют нулевые младшие 29 битов (варьировать мы можем лишь 32 - 29 = 3 бита) в каждом из 8 подключей. Поэтому существует всего $\left(2^3\right)^8 = 2^{24}$ «абсолютно слабых» ключей, которые всегда дают заданное преобразование.

У «слабого» подключа в нуль установлены разряды, которым в двоичной разности соответствуют «1». Для описанного преобразования существует $2^{224}$ «слабых» ключей из $2^{256}$ всего возможных (соответственно вероятность генерации «слабого» ключа $p_w = \frac{2^{224}}{2^{256}} = 2^{-32}$). Для этих ключей единичные биты в разности без изменения будут проходить через ключевое преобразование, однако переносы могут с некоторой вероятностью возникнуть в предшествующих разрядах, что исказит разность. Оценим вероятность прохождения заданной разности через все 32 цикла с учетом возникновения переносов. Поскольку мы используем подстановки, всегда дающие нам заданный переход, вероятность будет зависеть только от сеансового ключа. На проявление этого свойства для каждого ключа будут влиять и шифруемые блоки. Назовём разряд, которому соответствует «1» в двоичной разности, актив-

ным. Разность будет искажена, если в шифруемых блоках возникнет перенос в активный разряд (поскольку двоичная разность до активного разряда равна нулю, то эта часть блоков совпадает, и перенос, если появится, то сразу в двух блоках). В активном разряде биты шифруемых блоков противоположны (один из блоков содержит 0, другой – 1). Ключ в разряде, соответствующем активному, содержит нуль. Поэтому в одном из блоков переноса из активного разряда не будет, а в другом блоке перенос обязательно возникнет. Практически всегда это приводит к искажению разности, но не исключено, что и искаженная разность будет преобразована в единичную на выходе сумматора. Однако вероятность этого события достаточно низкая, поэтому его рассматривать не будем. Итак, искажение разности произойдет при возникновении переносов в одном из блоков. Вероятность возникновения переноса в активный разряд – $\frac{1}{2}$. Это справедливо для любого активного разряда, исключая самые младшие. Отсюда вероятность искажения разности на первом цикле равна нулю (см. табл. 6, цикл 0), на втором и третьем – $\frac{1}{2}$ и т.д. Соответственно вероятность прохождения разности без искажений восьми циклов шифрования – $p_8 = 2^{-10}$, шестнадцати – $p_{16} = 2^{-19}$, и всего алгоритма – $p_{32} = 2^{-38}$. Можно дополнительно повысить вероятность прохождения заданной разности на выход алгоритма, подбирая значения левого полублока.

Соответственно для описанного преобразования разностей существует $2^{224}$ слабых ключей (при общем их числе $2^{256}$), на каждом из которых заданная разность транслируется на выход алгоритма с вероятностью не менее $2^{-38}$.

Здесь описан всего лишь один «потенциально опасный» вариант входной разности. В действительности для рассмотренной слабой подстановки существует 255 «потенциально опасных» входных значений разности только для правого полублока (от 0000 0001h до 1111 1111h), причем для восьми из них существует $2^{224}$ «слабых» ключей, для одного – $2^{192}$, остальные располагают некоторым промежуточным количеством.

Можно рассматривать и сеансовые ключи, состоящие из комбинации «слабых» и «абсолютно слабых» подключей. Например, если второй подключ является «абсолютно слабым», а все остальные – «слабыми», то эффективная длина ключа составит $256 - 11 - 29 = 216$ битов, а вероятность распространения разности на выход – $2^{-32}$. Использование сразу двух и более «абсолютно слабых» подключей ещё больше увеличивает вероятность успешной атаки.

Остались нерассмотренными и ряд других типов «слабых» долговременных ключей, для которых существуют «слабые» сеансовые ключи (существует много вариантов долговременных ключей, для которых в таблицах распределения разностей существует единичный переход).

Однако из уже представленных результатов следует, что для алгоритма ГОСТ 28147-89 существуют «слабые» подстановки и соответствующие им «слабые» сеансовые ключи, на которых не выполняются требования статистической безопасности. Более того, на «абсолютно слабых» сеансовых ключах заданная входная разность всегда преобразуется в определённую выходную и не зависит от шифруемых текстов.

Это ставит под сомнение вопрос о надёжности и безопасности алгоритма ГОСТ 28147-89 при использовании «слабых» долговременных ключей. На наш взгляд, представленные результаты свидетельствуют о том, что существующие методики генерации долговременных ключей, основанные на критериях проверки случайности[3],необходимо дополнить требованиями фильтрации «слабых» подстановок.

**Список литературы**: **1.** *E. Biham, A. Shamir.* Differential Cryptanalysis of DES-like cryptosystems. The Weizmann Institute of Science. Department of Applied Mathematics. Technical Report CS90-16. 1990. **2.** *Лисицкая И.В., Бондаренко А.С., Цепурит Т.В.* Сравнительный анализ механизмов образования лавинного эффекта в алгоритмах DES и ГОСТ 28147-89. 5-я Международная конференция «Теория и техника передачи, приёма и обработки информации». ХТУРЭ: Харьков, 1999. **3.** *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*