УДК 681.3.06:519.248.681

И.Д. ГОРБЕНКО, д-р техн. наук, И.В. ЛИСИЦКАЯ

КРИТЕРИИ ОТБОРА СЛУЧАЙНЫХ ТАБЛИЦ ПОДСТАНОВОК ДЛЯ АЛГОРИТМА ШИФРОВАНИЯ ПО ГОСТ 28147 – 89

Числовые конструкции типа подстановок широко применяются в системах криптографических преобразований, в частности в алгоритмах ГОСТ 28147 – 89, DES [1]. В DES основу такой конструкции составляют так называемые S-блоки, принципы построения которых засекречены, но сами таблицы подстановок фиксированны и известны. Принципы построения таблиц подстановок, использованных в ГОСТе, также не разглашаются, но эти таблицы могут изменяться и являются дополнительным секретным параметром алгоритма (долговременным ключом).

Поскольку алгоритм шифрования по ГОСТ 28147 – 89 нашел официальное признание и в Украине [2], стал актуальным вопрос о возможности применения в алгоритме ГОСТа таблиц подстановок, сгенерированных случайным образом [3].

По-видимому, речь должна идти не о выборе в качестве долговременного ключа взятого в чистом виде случайного набора подстановок, а об использовании таблицы подстановок, удовлетворяющих некоторым критериям случайности. Ведь очевидно, что использование подстановок, взятых наугад, не исключает применения "плохого" ключа, если под ним понимать вырожденные конфигурации (повторяющиеся подстановки, подстановки тождественного и подобных им типов, их циклические сдвиги, инверсии тождественных подстановок и их сегментов и целый ряд других подстановок, близких к упорядоченному виду). Во всяком случае, практика применения методов шифрования свидетельствует о том, что всегда предусматриваются дополнительные затраты на исключение может быть и очень маловероятных, но явно "опасных" режимов, снижающих характеристики стойкости системы защиты. В связи с отмеченным рассмотрим некоторые результаты разработки и применения процедур отбора случайных подстановок.

Прежде всего сформулируем кратко основные идеи предлагаемого подхода и сущность разработанных критериев и процедур отбраковки подстановок при формировании множества долговременных ключей (допустимого или разрешенного множества таблиц подстано-

вок). При изложении материала будем опираться на общепринятые понятия и определения теории подстановок [4; 5 и др.], специально оговорив лишь обозначения для таблицы подстановок.

Запишем систему (набор) из m различных подстановок n-й степени S_{mn} (таблицу подстановок) в виде расширения традиционного представления подстановки [4] за счет добавления новых строк:

$$S_{mn} = \begin{bmatrix} 1 & 2 & 3 & \dots & n \\ i_{11} & i_{12} & i_{13} & \dots & i_{1n} \\ i_{21} & i_{22} & i_{23} & \dots & i_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ i_{m1} & i_{m2} & i_{m3} & \dots & i_{mn} \end{bmatrix}. \qquad (1)$$

Верхнюю строку назовем нулевой, а остальным присвоим номера от 1 до m. Ставится задача построить некоторое заданное число случайных таблиц подстановок, т.е. таблиц вида (1), удовлетворяющих определенным критериям случайности.

Предлагается использовать показатели и критерии случайности трех уровней. На первом уровне проверки оценивается соответствие характеристик отдельно взятой подстановки свойствам случайной равновероятной подстановки. Подстановки, прошедшие первый уровень проверки, считаются уже подстановками случайного типа.

На втором уровне определяется соответствие характеристик случайности системы подстановок, попавших в таблицу, свойствам среднестатистической таблицы случайных подстановок. Таблицы подстановок, прошедшие первые два уровня проверки, считаются случайными таблицами подстановок.

На третьем уровне осуществляется оценка характеристик случайности множества таблиц подстановок, из которых отбираются таблицы, удовлетворяющие требованиям к долговременным ключам.

Соответствующие критерии отбора подстановок, таблиц подстановок и множеств таблиц подстановок названы критериями отбора первого, второго и третьего уровня.

При формировании подходов и критериев отбора подстановок на первом уровне (по индивидуальным характеристикам случайности) за основу взяты результаты хорошо разработанной теории случайных подстановок, в частности известные теоремы об асимптотической нормальности законов распределения для числа инверсий, числа циклов и числа возрастаний случайных равновероятных подстановок [4]. Обоснована правомерность использования асимптотических показа-

телей случайности для оценки характеристик случайности подстановок конечной степени n. Это позволило сформулировать критерии отбора подстановок первого уровня в виде требований 1–3.

***Требование* 1.** Число инверсий η_n в подстановке степени n должно удовлетворять условиям

$$|\eta_n - n(n-1)/4| \le a\sigma_\eta; \quad \sigma_\eta = n^{3/2}/6. \qquad (2)$$

***Требование* 2.** Число циклов ξ_n в подстановке степени n должно удовлетворять условиям

$$|\xi_n - \ln n| \le a\sigma_\xi; \qquad \sigma_\xi = \sqrt{\ln n}. \qquad (3)$$

***Требование* 3.** Число возрастаний θ_n в подстановке степени n должно удовлетворять условиям

$$|\theta_n - n/2| \le a\sigma_\theta; \qquad \sigma_\theta = \sqrt{n/12}. \qquad (4)$$

При n = 16, a = 1 условия отбора подходящих подстановок (2) принимают вид

$$|\eta_n - 60| \le 10; \quad |\xi_n - 3| \le 2; \quad |\theta_n - 8| \le 1.$$

При формировании критериев отбора подстановок на втором уровне изучены подходы к описанию свойств (характеристик) таблиц, составленных из подстановок, прошедших первый уровень проверки. Обоснована методика, строящаяся на использовании понятия противоречивости подстановок числа несовпадений элементов [6].

Практически для каждой таблицы из m подстановок n-й степени формируется двумерный метрический портрет, т.е. таблица определяется двумя "векторами". В первом случае определяется конфигурация $(t_0, t_1, \ldots, t_n)$ совпадений элементов в $N_k = m(m-1)/2$ попарных декомпозициях строк этой таблицы подстановок.

Элемент конфигурации t_i, $i = 1, 2, \ldots, 16$ представляет собой число пар строк из общего их числа N_k с i совпадениями элементов,

так что $\sum_{i=0}^{n} t_i = N_k$. Затем на множестве возможных исходов $\{t_0, t_1, \ldots, t_n\}$ определяется закон распределения вероятностей $P(\gamma = i)$ для числа совпадений $\gamma = i$ элементов, i = 0, 1, 2, ..., n в двух (в паре) равновероятных подстановках n-й степени.

Во втором случае определяется конфигурация $(\varsigma_0, \varsigma_1, \varsigma_2, \ldots, \varsigma_{[m/2]})$ совпадений элементов по столбцам таблицы. Здесь элемент конфигурации ς_k, $k = 0, 1, 2, \ldots, [m/2]$ представляет собой число столбцов с k повторениями (в том числе многократными) элементов столбца, при этом $\sum_{k=0}^{[m/2]} \varsigma_k = n$. Затем на множестве возможных исходов $\{\varsigma_0, \varsigma_1, \ldots, \varsigma_{[m/2]}\}$ определяется закон распределения вероятностей $P(\gamma = k)$ для числа повторений $\gamma = k$ различных элементов, $k = 0, 1, 2, \ldots, [m/2]$ в столбце таблицы подстановок. Здесь $[x]$ обозначает наибольшее целое число x, не превосходящее x.

В итоге требования к отбору таблиц подстановок для алгоритма ГОСТ на втором уровне проверки формулируются в следующем виде.

***Требование* 4.** В таблицу подстановок должны входить подстановки, не имеющие совпадений с нулевой строкой (не имеющие циклов нулевой длины).

***Требование* 5.** Подстановки, вошедшие в таблицу подстановок, должны по конфигурации $(t_0, t_1, \ldots, t_{16})$ совпадений элементов в $N_k = 28$ попарных декомпозициях таблиц подстановок по строкам удовлетворять критерию χ^2 согласия выборки наблюдаемых данных $(t_0, t_1, \ldots, t_{16})$ с биномиальным (с параметрами $m_t = np_0 = 1$, $\sigma_t^2 = np_0(1 - np_0) = 0{,}995$) законом распределения вероятностей для числа совпадений элементов в двух случайных равновероятных подстановках при заданном уровне значимости α_0.

Требование 6. Подстановки, вошедшие в таблицу подстановок, должны по конфигурации $(\varsigma_0, \varsigma_1, \varsigma_2, \varsigma_3, \varsigma_4)$ совпадений элементов в столбцах таблицы подстановок удовлетворять критерию χ^2 согласия выборки наблюдаемых данных $(\varsigma_0, \varsigma_1, \varsigma_2, \varsigma_3, \varsigma_4)$ с нормальным (с параметрами $m_\varsigma = 1{,}424$, $\sigma_\varsigma^2 = 0{,}674$) законом распределения вероятностей для числа повторений различных элементов в столбце таблицы при заданном уровне значимости α_0.

При формировании критериев отбора таблиц подстановок на третьем уровне за основу взята идея наложения таблиц подстановок и подсчета совпадающих элементов [5].

В результате требование к отбору таблиц постановок на третьем уровне проверки сформулировано следующим образом.

Требование 7. Множество таблиц подстановок, используемых в качестве долговременных ключей, должно при всех попарных наложениях таблиц давать число совпадающих элементов q, удовлетворяющее условиям $|q - m| \le \sqrt{m}$.

Можно запретить и совпадение строк, стоящих на различных позициях таблиц, если оговорить в требовании 7, что процедура наложения для каждой пары таблиц выполняется со всеми циклическими перестановками строк одной из них.

Приведем теперь аргументы в пользу изложенной методики построения случайных таблиц подстановок. Легко убедиться в том, что предлагаемые процедура и последовательность проверок позволяют исключить из допустимого множества подстановок, образующих таблицы долговременных ключей, большинство, если не все, подстановки вырожденного и близкого к вырожденному видов. К последним, как отмечалось выше, отнесены: тождественная подстановка и подстановки – циклические сдвиги нижней строки тождественной подстановки; инверсная (в отношении порядка элементов) тождественной и ее циклические сдвиги; подстановки, содержащие отрезки циклических сдвигов, и другие близкие к ним конструкции.

Действительно, проверка по числу возрастаний обеспечивает отбраковку подстановок, содержащих сегменты последовательности символов нулевой строки и их перестановки в обратном порядке; отсеиваются, по крайней мере, упорядоченные и близкие к ним наборы символов, содержащие более восьми чисел.

По инверсиям запрещаются группирования (комбинации) на отдельных позициях подстановки значительных групп близких по значениям символов (более шести — восьми символов), т.е. проверяется "однородность" распределения больших и малых чисел в блоке перестановки.

При проверке по числу циклов контролируются уже связи между строками подстановки, т.е. оценивается мера ее детерминизма в виде числа столбцов с совпадающими элементами в общепринятом представлении подстановки с помощью матрицы. Ограничение на число циклов практически определяет допустимое в подстановке число коротких циклов, в частности число циклов единичной длины (тождественных переходов), которое не может быть больше заданного (для $n = 16$ и $\left| \xi_n - \ln n \right| \le \ln n$ это число равно четырем). Действительно, запрещаются подстановки с числом циклов не менее пяти , а в подстановке с k циклами возможно только k — 1 тождественных переходов. Поскольку число циклов однозначно связано с декрементом подстановки, определяющим минимальное число транспозиций, с помощью которых подстановка может быть переведена в единичную, то число циклов практически характеризует "удаленность" данной подстановки от единичной, т.е. затраты, требуемые для приведения подстановку к вырожденному виду.

В результате первый уровень проверки обеспечивает получение подстановок со сбалансированным перемешиванием элементов исходного множества $X = \{0, 1, \ldots, 15\}$, сбалансированным в смысле близости параметров подстановки к среднестатистическим показателям случайной равновероятной подстановки.

На втором уровне проверки исключаются подстановки, содержащие циклы единичной длины (требуется отсутствие совпадений с нулевой строкой), т.е. запрещаются тождественные переходы. Очевидность этого требования следует из того, что при тождественных переходах подстановка как бы не участвует в процессе шифрования. Более детальный анализ показывает, что именно благодаря выполнению этого требования (с учетом механизмов временного сдвига и перекрестного суммирования самого алгоритма ГОСТ 28147 – 89) обеспечивается реализация одного из важных свойств шифра – быстрый рост с увеличением числа циклов зависимости все большего числа выходных бит от значения каждого входного бита. Заметим, что ГОСТ требует для обеспечения зависимости всех выходных бит от значения каждого входного бита 8 циклов (а DES – 5 циклов) [1]. Отсутствие тождественных переходов приводит к влиянию каждого входа S-блока на данном цикле на входы двух S-блоков (с учетом сдвига на 11 тактов) на следующем цикле, в то время как тождественный переход при изме-

нении одного бита на входе S-блока приводит к влиянию только на один S-блок в следующем цикле.

Очевидно также, что проверки второго уровня исключают попадание в таблицу одинаковых подстановок и (что еще более важно для эффективного перемешивания исходного текста) приводят к наибольшему в статистическом смысле различию подстановок, их среднестатистической "непохожести" друг на друга. Символы текста, проходя многократно через всю систему подстановок, приобретают необходимую непредсказуемость значений именно благодаря многообразию воздействий на них.

Наконец, третий уровень проверки обеспечивает заданную степень непохожести уже самих долговременных ключей.

Описанные выше подходы к формированию методов и критериев отбора случайных подстановок были положены в основу разработки программного комплекса генерации и сертификации долговременных ключей для алгоритма ГОСТ 28147 – 89. В табл. 1–3 представлены примеры долговременных ключей, построенных с помощью предлагаемой методики, а в табл. 4 – конфигурации совпадений в столбцах и строках для этих ключей. В правой части последней строки табл. 1 —3 указано количество совпадений в столбцах.

Таблица 1

№ строки	Число			Ненулевые строки таблицы подстановок S_{mn}
	инверсий	возрастаний	циклов	
1	67	9	3	7 10 6 8 11 4 14 15 13 1 0 3 5 12 2 9
2	57	7	3	3 11 6 4 2 7 15 12 13 0 14 8 9 1 10 5
3	69	8	1	8 13 10 2 15 14 1 9 4 3 5 7 11 12 0 6
4	68	9	4	12 2 1 10 15 7 11 8 13 14 3 5 6 0 4 9
5	58	8	4	9 3 5 1 15 14 10 2 13 4 6 8 11 0 7 12
6	65	7	3	11 4 9 5 2 8 12 14 13 6 15 7 0 3 1 10
7	50	8	2	3 15 11 6 0 4 8 1 2 13 7 10 14 9 5 12
8	62	8	4	14 3 7 10 1 0 11 12 15 8 4 6 9 5 13 12
Среднее	62	8	3	1 1 1 1 2 3 1 1 1 0 0 2 2 2 0 2

Таблица 2

№ строки	Число			Ненулевые строки таблицы подстановок S_{mn}
	инверсий	возрастаний	циклов	
1	65	9	1	8 10 6 5 7 12 15 9 1 2 4 13 11 14 0 3
2	54	8	4	7 0 5 6 14 2 11 13 4 15 8 3 9 12 10 1
3	50	8	2	11 2 7 5 0 8 15 6 9 10 1 4 14 3 13 12
4	65	7	1	1 15 8 10 7 9 13 0 6 2 12 4 14 11 5 3
5	58	9	2	11 4 7 14 6 1 2 12 13 0 5 10 8 9 15 3
6	66	9	2	4 8 1 7 12 13 15 11 14 6 0 9 3 10 2 5
7	59	7	3	1 13 4 5 3 11 9 12 10 15 8 7 6 0 2 14
8	61	8	1	10 8 0 12 6 15 1 5 9 14 3 7 11 2 13 4
Среднее	60	8	2	2 1 1 1 2 0 1 1 1 2 1 2 2 0 2 1

Таблица 3

№ строки	Число			Ненулевые строки таблицы подстановок S_{mn}
	инверсий	возрастаний	циклов	
1	56	9	2	2 5 15 4 8 12 7 11 0 14 3 10 6 1 9 13
2	58	8	2	6 3 13 4 8 11 15 5 2 7 12 1 0 14 9 10
3	61	7	3	7 4 12 2 9 10 8 14 13 1 3 15 5 0 11 6
4	55	9	1	1 6 11 0 10 14 15 4 12 5 2 13 3 8 7 9
5	69	8	1	3 14 13 11 12 9 1 10 5 0 2 4 6 15 7 8
6	60	7	2	11 6 0 7 8 14 10 2 1 15 9 12 13 4 5 3
7	59	8	3	7 6 8 15 2 0 1 3 11 1 9 4 10 1 5 13
8	57	8	3	2 5 0 9 14 8 7 15 11 13 3 6 10 4 12 1
Среднее	59	8	2	2 2 2 1 1 1 2 0 1 0 3 1 2 2 3 1

Таблица 4

№ таблицы	Число совпадений : число		Значение χ_q^2 для	
	столбцов	строк	столбцов	строк
1	0:3; 1:7; 2:5; 3:1	0:8; 1:11; 2:7; 3:2	1,43	1,36
2	0:2; 1:8; 2:6	0:9; 1:13; 2:5; 3:1	1,35	1,26
3	0:2; 1:6; 2:6; 3:2	0:10; 1:10; 2:5; 3:2; 4:1	0,98	1,20

Остается лишь показать, что используемые на всех уровнях проверки ограничения не приводят к сколько-нибудь существенному уменьшению результирующего пространства (множества) долговременных ключей, т.е. к снижению стойкости системы в этом смысле.

Действительно, если ориентироваться на приведенные здесь значения параметров отбора подстановок, то, пройдя последовательно по всем уровням отбраковки подстановок и таблиц подстановок, можно получить следующие результаты.

Из общего числа $n! = 2 \cdot 10^{13}$ возможных подстановок степени n (при $n = 16$) первый уровень проверки проходит около 50 % всех подстановок. На втором уровне сначала исключается примерно 37 % из оставшегося количества из-за наличия совпадения с нулевой строкой в (1), поэтому в формировании таблиц подстановок участвует $0{,}5(1 - 0{,}37)\,n! > 10^{12}$ подстановок. Из 10^{12} подстановок можно составить $C^8_{10^{12}}$ таблиц подстановок, что дает число $O(10^{84})$ возможных таблиц. Как показывают расчеты и данные экспериментов, из этого числа примерно 70 % будет отброшено из-за совпадений в столбцах и около 95 % — из-за совпадений в парах строк ($\chi_q^2 = 1{,}5$). В результате на третий уровень проверки пройдет примерно $O(10^{82})$ таблиц.

Самым жестким в отношении усечения пространства итоговых вариантов оказывается третий уровень проверки. Если ориентироваться на использование на этом уровне указанных выше граничных значений числа совпадений элементов в паре наложенных таблиц подстановок, то для $m = 8$ и $n = 16$, как следует из результатов статистического эксперимента, вероятность попадания в указанный интервал произвольной пары таблиц составляет около 0,78.

Простые расчеты показывают, что если для получения 10 долговременных ключей (случайных таблиц подстановок, прошедших третий уровень проверки) нужно проверить в среднем 55 таблиц, то для получения 50 ключей надо проверить более 10^6 пар таблиц.

Очевидно, что при необходимости можно изменить границы отбора таблиц подстановок, но эта задача требует отдельного рассмотрения.

Таким образом, представленные подходы и результаты позволяют сформулировать конкретные правила отбора случайных таблиц подстановок при построении долговременных ключей (генерации S-блоков), для которых уже можно решать задачи обоснования криптографической стойкости алгоритма шифрования по ГОСТ 28147 – 89 при использовании подстановок случайного типа.

Список литературы: 1. *Scheier B.* Applied Cryptography Second Edition: protocols, algoritms and soirce code in Communication. New York: John Wiley a. Sons Incorp., 1996. 675 p. 2. *Интервью* начальника Главного управления правительственной связи СБУ Лазарева Г.П. журналу «Безопасность информации» // Безопасность информ. 1995. № 3. С. 2 – 3. 3. *Фаль А.М.* Алгоритм шифрования по ГОСТ 28147 – 89 и способы применения блочных шифров // Там же. С. 8 – 11. 4. *Скачков В.Н.* Введение в комбинаторные методы дискретной математики. М.: Наука, 1982. 384 с. 5. *Скачков В.Н.* Комбинаторные методы дискретной математики. М.: Наука, 1977. 319 с. 6. *Математическая* энциклопедия: В 5 т. / Гл. ред. И.М. Виноградов. М.: Сов. энцикл., 1979. Т. 2. 278 с.

Поступила в редколлегию 31.07.97