УДК 681.3.06: 519.248.681

ПРОТИВОРЕЧИВЫЕ ПОДСТАНОВКИ В АЛГОРИТМЕ ГОСТ 28147-89

И.В. Кононова

Обсуждаются свойства и характеристики множества таблиц, составленных из m противоречивых подстановок степени n при их использовании в качестве долговременных ключей в алгоритме ГОСТ 28147-89. Обосновывается уточненная формула для подсчета числа нормализованных латинских прямоугольников.

Числовые конструкции типа подстановок нашли широкое применение в криптографии [1,2] и, в частности, они использованы в стандарте блочного симметричного шифрования по ГОСТ 28147-89 [3], разрешенному к применению в Украине.

Основным фактором, сдерживающим использование этого алгоритма, можно считать отсутствие в публикациях сведений о принципах построения ключей подстановок, которые засекречены разработчиками. Поэтому наблюдаются попытки использования в качестве долговременных ключей в ГОСТе (**S** - блоков) подстановок, сгенерированных случайным образом [4]. Получение ключей подстановок это актуальная задача и в других схемах симметричного шифрования, построенных на основе чередования слоёв замен и подстановок [5]. Определённого интереса в этом отношении заслуживают противоречивые подстановки [6].

Настоящая работа посвящается обсуждению свойств и характеристик множества таблиц, составленных из таких подстановок. Таблицу $\mathbf{S}_{\mathbf{m,n}}$ (систему) из $\mathbf{m}$ различных подстановок $\mathbf{n}$ - ной степени, содержащую $\mathbf{m}$ строк будем представлять в виде расширения стандартной записи подстановки [8] путем дописывания новых строк, т. е. в виде

$$\mathbf{S}_{\mathbf{m,n}} = \begin{pmatrix} \mathbf{1} & \mathbf{2} & \mathbf{3} & \dots & \mathbf{n} \\ \mathbf{i}_{11} & \mathbf{i}_{12} & \mathbf{i}_{13} & \dots & \mathbf{i}_{1n} \\ \mathbf{i}_{21} & \mathbf{i}_{22} & \mathbf{i}_{23} & \dots & \mathbf{i}_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \mathbf{i}_{m1} & \mathbf{i}_{m2} & \mathbf{i}_{m3} & \dots & \mathbf{i}_{mn} \end{pmatrix} \qquad (1)$$

Верхнюю строку будем называть нулевой, а остальные нумеровать от **1**-ой до **m**-ой. Применительно к ГОСТу $\mathbf{n = 16}$, $\mathbf{m = 8}$ (для долговременного ключа используются ненулевые строки таблицы (1)).

Противоречивые подстановки удовлетворяют требованию **D**: ни в одном столбце таблицы подстановок (1) не должно быть совпадающих (одинаковых) чисел, и может быть интерпретировано в виде двух требований:

Требование **d1**. Запрещенными являются ключи (подстановки), содержащие тождественные переходы.

Требование **d2**. При построении таблицы подстановок в их число не должны включаться подстановки, содержащие одинаковые (совпадающие) переходы (фрагменты).

Ставится задача оценить число таблиц подстановок $\mathbf{W_{m,n}}$, удовлетворяющих требованием **d1** и **d2**.

Прежде всего заметим, что в изложенной постановке задачи таблица (1) представляет собой нормализованный латинский прямоугольник размера $\mathbf{(m+1)\cdot n}$, $\mathbf{m + 1 < n}$ [6], и задача сводится, таким образом, к определению числа нормализованных латинских прямоугольников $\mathbf{W_{m,n} = K(m+1,n)}$. Здесь **K(l, n)** - обозначение числа нормализованных латинских прямоугольников в [8]. По сведениям, представленным в [8], известно решение задачи для двух частных случаев, когда латинский прямоугольник содержит две и когда он содержит три строки ($\mathbf{m = 1}$ и $\mathbf{m = 2}$). Первый из этих случаев мы зафиксируем в виде утверждения.

Утверждение 1. Среди полного числа перестановок **n**-й степени число перестановок, в которых элемент i не может занимать **i**-ю позицию (т.е. число перестановок, не имеющих совпадений с нулевой строкой в (1)), выражается формулой

$$\mathbf{D_n = n!\cdot\left(1-\frac{1}{1!}+\frac{1}{2!}-\ldots+(-1)^n\frac{1}{n!}\right)}.$$

Число $\mathbf{D_n}$ определено в [8] как число беспорядков. Очевидно, что

$$\mathbf{D_n = K(2,n) = W_{1,n}}\ .$$

Известна также формула для числа нормализованных латинских прямоугольников **K(3,n)**, имеющих три строки, но ввиду ее сложности мы её здесь приводить не будем.

Задача перечисления латинских прямоугольников с числом строк более трех считается не решенной [8]. Известна лишь следующая оценка:

$$K(l, n) \geq (n - 1)! \, (n - 2)! \cdots (n - l + 1)!. \qquad (2)$$

Для $n = 16$, $l = 9$ расчеты дают $K(9,16) \geq 7 \cdot 10^{68}$, т. е. это результат уже можно рассматривать как очень хороший.

Здесь будет предложен свой приближенный способ подсчета допустимого числа перестановок $K(l, n)$, задаваемых таблицей (1), позволяющий уточнить этот результат.

Идея предлагаемого подхода основывается на интерпретации получения каждой их возможных перестановок (подстановок) как результата сложного случайного эксперимента с n независимыми элементарными исходами, а ранее представленные соотношения мы используем для подтверждения правомерности развиваемого подхода.

Продемонстрируем идею этого подхода для подсчета числа перестановок первой строки в (1), т.е. опять займемся определением $W_{1,n}$.

Пусть имеется некоторая исходная перестановка из n элементов (в нашем случае нулевая строка, т.е. перестановка $1, 2, ..., n$). Необходимо построить перестановку, которая должна быть произвольным набором из n элементов, но без повторений верхних элементов (элементов нулевой строки).

Возьмем произвольный столбец. Если бы не было никаких ограничений, то нижним могло бы быть любое из n чисел. С другой стороны, если считать появление любого числа из возможного их набора n на любой из позиций нижней строки независимыми и равновероятными событиями, то совпадение нижнего элемента с верхним на любой из позиций возможно в $1/n$ случаев. Для вероятности выпадения элемента на любой позиции, не совпадающего с верхним, получим оценку в виде $1 - \frac{1}{n} = \frac{n-1}{n}$.

В результате для определения вероятности случайного выбора перестановки для первой строки в (1), не имеющей совпадений элементов с исходной (для вероятности несовпадения на n позициях), можно записать выражение $P_{1,n} = \left(\frac{n-1}{n}\right)^n$.

Можно доказать справедливость утверждения.

Утверждение 2. Вероятность случайного выбора подстановки из $\mathbf{n}$ элементов, не имеющей совпадений с некоторой фиксированной подстановкой из $\mathbf{n}$ элементов, ограничена снизу величиной $\left(\frac{\mathbf{n-1}}{\mathbf{n}}\right)^{\mathbf{n}}$.

Если теперь считать, как уже было отмечено выше, что события: - строка образует первую строчку в (1), и каждый столбец содержит разные элементы; -независимыми, то для приближенного значения $\tilde{\mathbf{W}}_{\mathbf{1,n}}$ можно записать выражение

$$\tilde{\mathbf{W}}_{\mathbf{1,n}} = \mathbf{n!}\left(\frac{\mathbf{n-1}}{\mathbf{n}}\right)^{\mathbf{n}}. \qquad (3)$$

Оценка $\tilde{\mathbf{W}}_{\mathbf{1,n}}$ приближается к $\mathbf{W}_{\mathbf{1,n}}$ с большой точностью. Так, $\tilde{\mathbf{W}}_{\mathbf{1,5}} = \mathbf{120 \cdot 0{,}328 = 39{,}322}$, что составляет 87,4% от истинного значения, но уже для $\mathbf{n = 6}$ $\tilde{\mathbf{W}}_{\mathbf{1,6}} / \mathbf{W}_{\mathbf{1,6}} = \mathbf{0.91}$. Это отношение при увеличении $\mathbf{n}$ быстро сходится к $\mathbf{1}$, поскольку,

$$\lim_{\mathbf{n}\to\infty}\left(\mathbf{1}-\frac{\mathbf{1}}{\mathbf{n}}\right)^{\mathbf{n}} = \lim_{\mathbf{n}\to\infty}\frac{\mathbf{1}}{\left(\mathbf{1}+\frac{\mathbf{1}}{\mathbf{n-1}}\right)^{\mathbf{n}}} = \frac{\mathbf{1}}{\mathbf{e}},$$

и можно заключить, что при увеличении $\mathbf{n}$ (точнее, при $\mathbf{n}\to\infty$) $\tilde{\mathbf{W}}_{\mathbf{1,n}} \to \mathbf{n!\,e^{-1}}$.

С другой стороны, $\mathbf{e^{-x} = 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \ldots}$, а так как

$$\tilde{\mathbf{W}}_{\mathbf{1,n}} = \mathbf{n!}\left(\mathbf{1}-\frac{\mathbf{1}}{\mathbf{1!}}+\frac{\mathbf{1}}{\mathbf{2!}}\cdot\frac{\mathbf{n-1}}{\mathbf{n}}-\frac{\mathbf{1}}{\mathbf{3!}}\cdot\frac{\mathbf{(n-1)(n-2)}}{\mathbf{n^2}}+\ldots+\mathbf{(-1)^n}\frac{\mathbf{1}}{\mathbf{n!}}\cdot\frac{\mathbf{n!}}{\mathbf{n^n}}\right),$$

то при $\mathbf{n}\to\infty$ имеем $\mathbf{W_{1,n} \to n!\,e^{-1}}$.

Покажем, что для конечных значений $\mathbf{n}$ выполняется соотношение

$$\mathbf{D_n = W_{1,n} > \tilde{W}_{1,n}}, \qquad (4)$$

т.е. $\tilde{\mathbf{W}}_{\mathbf{1,n}}$ является оценкой снизу, причем как будет показано далее более точной, чем оценка (2).

Используем теперь этот же подход для подсчета числа допустимых перестановок для **i** - той строки в (1). Рассуждения здесь можно построить следующим образом.

Пусть построен набор из **i** первых допустимых строк в (1). Требуется построить **(i + 1)** - ю строку - произвольный набор из **n** чисел, но таких, чтобы в каждом столбце было поставлено число, не равное ни одному из стоящих выше в этом столбце чисел (также разных). Строка **i** может рассматриваться как исходная (фиксированная) в первоначальном рассуждении, только каждый из ее элементов теперь выбран из **n + 1 - i** возможных. В **i + 1** -ой строке каждый элемент столбца выбирается из **n - i** оставшихся (не выбранных в предыдущих строчках), так что вероятность всей **(i + 1)** -ой строке быть допустимой равна

$$\mathbf{P_{i+1,n} \approx \left(\frac{n-i}{n-i+1} \right)^n}, \ \mathbf{i = 1, 2, ..., m\text{-}1}.$$

Тогда в прежнем предположении о независимости выбора каждого из элементов строки для приближенного числа допустимых **(i +1)** строк можно записать выражение $\mathbf{\widetilde{W}_{i+1,n} = \widetilde{W}_{i,n} \left(\frac{n-i}{n-i+1} \right)^n}$.

Чтобы проиллюстрировать точность оценки $\mathbf{\widetilde{W}}$ по отношению к **W** в табл.1 приведены значения $\mathbf{\widetilde{W}_{i,n}}$ и $\mathbf{W_{i,n}}$ для **i = 0, 1, 2, 3, 4** и **n = 5-9**.

Таблица 1 - Сравнение значений оценок $\mathbf{\widetilde{W}}$ и $\mathbf{\widetilde{\widetilde{W}}}$ с точным значением **W** для подстановок степени **n = 5 - 9**

i	n = 5					n = 6				
	$\widetilde{W}$	$\widetilde{W}/W$	W	$\widetilde{\widetilde{W}}$	$\widetilde{\widetilde{W}}/W$	$\widetilde{W}$	$\widetilde{W}/W$	W	$\widetilde{\widetilde{W}}$	$\widetilde{\widetilde{W}}/W$
0			120					720		
1	39,3	0,89	44	39,3	0,89	241,1	0,91	264	265	0,91
2	9,32	0,78	12	12,85	1,07	63,2	0,79	80	80,76	1,01
3	1,23	0,10	2	4,2	2,1	11,2	0,56	20	27	1,35
4	0,04	0,04	1	1,37	1,37	1	0,25	4	9	2,25

i	n = 7					n =8				
	$\tilde{W}$	$\tilde{W}/W$	W	$\tilde{\tilde{W}}$	$\tilde{\tilde{W}}/W$	$\tilde{W}$	$\tilde{W}/W$	W	$\tilde{\tilde{W}}$	$\tilde{\tilde{W}}/W$
0			5040					40320		
1	1713,6	0,92	1854	1713	0,92	13854	0,934	14834	13854	0,934
2	478,2	0,82	580	582	1,003	4036	0,849	4752	4765,8	1,003
3	100,3	0,67	144	198	1,375	938,7	0,715	1313	1639	1,248
4	13,4	0,50	27	67	2,48	157,4	0,579	272	564	2,07

i	n = 9				
	$\tilde{W}$	$\tilde{W}/W$	W	$\tilde{\tilde{W}}$	$\tilde{\tilde{W}}/W$
0			362880		
1	125716	0,94	133496	125716	0.94
2	37400,5	0,86	43424	43497.7	1.0016
3	9296,7	0,77	12102	15050.2	1.243
4	1782	0,66	2706	5207	1.924

Из таблицы видно, что с увеличением степени подстановок **n** величина $\tilde{W}_{i,n}$ приближается к своему асимптотическому значению $W_{i,n}$. Точность приближения зависит и от положения строки в таблице подстановок (1), однако из представленных результатов следует, что окончательная погрешность будет вполне приемлемой для практических расчетов.

Вспомним теперь, что в данной задаче $n = 16$, а $m = 8$. Расчеты соответствующих значений $\tilde{W}_{i,n}$ в этом случае приводят к результатам:

$$W_{0,16} = 16! \cong 2{,}09 \cdot 10^{13};$$

$$W_{1,16} > \tilde{W}_{1,16} \cong 16!\left(\frac{15}{16}\right)^{16} \approx 0{,}74 \cdot 10^{13};$$

$$W_{2,16} > \tilde{W}_{2,16} \cong 0{,}74 \cdot 10^{13} \cdot \left(\frac{14}{15}\right)^{16} \approx 0{,}24 \cdot 10^{13};$$

$$W_{3,16} > \tilde{W}_{3,16} \cong 0{,}73 \cdot 10^{12};$$

$$W_{4,16} > \tilde{W}_{4,16} \cong 0{,}2 \cdot 10^{12};$$

$$W_{5,16} > \tilde{W}_{5,16} \cong 0{,}5 \cdot 10^{11};$$

$$W_{6,16} > \tilde{W}_{6,16} \cong 0{,}11 \cdot 10^{11};$$

$$W_{7,16} > \tilde{W}_{7,16} \cong 0{,}2 \cdot 10^{10};$$

$$W_{8,16} > \tilde{W}_{8,16} \cong 0{,}3 \cdot 10^{9}.$$

Таким образом, первая строка системы выбирается из более, чем $0{,}74 \cdot 10^{13}$ перестановок и т.д., 8-я строка выбирается из более, чем $3 \cdot 10^{8}$ перестановок, а система $S_{8,16}$ из 8×16 перестановок может быть выбрана более, чем

$$K(9,16) \geq \prod_{i=1}^{8} \tilde{W}_{i,16} \approx 0.86 \cdot 10^{87} \qquad (5)$$

способами. Это вполне достаточное пространство долговременных ключей.

Заметим здесь, что если бы каждая подстановка выбиралась без всяких ограничений, то систему $S_{8,16}$ можно было бы построить $(n!)^8 = (16!)^8 = (2{,}09 \cdot 10^{13})^8$ способами, что примерно равно $364 \cdot 10^{104} \approx 0{,}36 \cdot 10^{107}$. Допустимыми являются примерно 2 системы из 10^{20}, так что случайно их порождать безнадежно.

Отметим также, что результат (3), определяющий нижнюю границу числа латинских прямоугольников, для $n=16$ оказался существенно (более, чем на 18 порядков) точнее известного результата (2).

Наконец отметим, что расчеты, приведенные выше выполнены для схемы зависимых исходов. Каждый последующий результат зависит от предыдущего. Если расчеты выполнить по схеме независимых исходов, т.е. вместо (3) воспользоваться соотношением

$$\tilde{\tilde{W}}_{i+1,n} = n!\left(P_{2,n}\right)^{i+1}, \; i=1, 2, ..., m-1, \qquad (6)$$

то в этом случае получим $\prod_{k=1}^{8} \tilde{\tilde{W}}_{i,16} \approx 2{,}57 \cdot 10^{90}$, что в $3 \cdot 10^{3}$ раз превышает оценку (4).

Этот результат подтверждает возможность использования при изучении свойств пространства подстановок рассматриваемого порядка отмеченного выше способа интерпретации каждой из случайно сформированных подстановок как исхода сложного эксперимента с независимыми элементарными событиями.

Для точного значения допустимого множества подстановок каждой из строчек таблицы (1) вместо (4) для $\mathbf{i > 2}$ можно записать ограничения в виде $\tilde{\tilde{W}}_{i,n} > W_{i,n} > \tilde{W}_{i,n}$.

В табл.1 приведены также оценочные значения и для значений $\tilde{\tilde{W}}_{i,n}$, рассчитанные в соответствии с формулой (6).

Важным является то, что для порождения таблиц, составленных из противоречивых подстановок, могут быть построены достаточно простые практически реализуемые алгоритмы, которые будут обсуждены в отдельной работе.

ЛИТЕРАТУРА

1. Барсуков В.С., Дворянкин С.В., Шеремет И.А. Безопасность связи в каналах телекоммуникаций. - М.: Россия, 1993. - Т.20, - 123 с.

2. Мафтик С. Механизмы защиты в сетях ЭВМ. - М.: МИР, 1993.- 216 с.

3. ГОСТ 28147-89. Системы обработки информации. Защита криптографическая. Алгоритмы криптографического преобразования. Введ. 01.01.89. - М.: Изд-во стандартов, 1989. - 78 с.

4. Фаль А.М. Алгоритм шифрования по ГОСТу 28147-89 и способы применения блочных шифров // Безопасность информации. - 1995. – №3. – С. 8-11

5. Месси Дж. Введение в современную криптологию // ТИИЭР. - 1988. - Т. 76, №5. - С.24-42.

6. Сачков В.Н. Комбинаторные методы дискретной математики. - М., Наука, 1977. - 319 с.

7. Завало С.Т., Костарчук В.Н., Хацет Б.И. Алгебра и теория чисел. - ч.2. - К.: Вища школа , 1980. - 407 с.

8. Математическая энциклопедия: в 5 т./ гл. ред. И.М. Виноградов - Советская энциклопедия. - М., 1979. - Т.2: Д-КОО. - 1103 с.