УДК 681.3.06

К.Е. ЛИСИЦКИЙ, Е.Д. МЕЛЬНИЧУК

УТОЧНЁННАЯ МАТЕМАТИЧЕСКАЯ МОДЕЛЬ СЛУЧАЙНОЙ ПОДСТАНОВКИ

Осуждаются подходы к отбору случайных подстановок, основанные на применении системы критериев, построенных с использованием оценок близости законов распределения XOR таблиц и смещений таблиц линейных аппроксимаций подстановок теоретическим законам. Отмечается их неконструктивность. Излагается сущность методики определения законов распределения максимумов для больших по объёму выборок независимых одинаково распределённых случайных величин. Методика применяется для определения распределений максимумов XOR таблиц и смещений таблиц линейных аппроксимаций выборки из байтовых подстановок. На основе полученных результатов предлагается уточнённое определение случайной подстановки, строящееся на свойствах выборки случайных подстановок.

1. Введение

Самые современные традиционные ключевые криптосистемы базируются на идее произведения (product) шифров, которые представляют класс криптосистем, многократно повторяющих сложную операцию, отображающую плайнтекст в шифртекст. Каждое такое повторение (итерация) известно как цикл шифра. Сложная (составная) операция, выполняющаяся в каждом цикле, является обычно комбинацией из набора примитивных операций, таких как сдвиг, линейное преобразование, модульное сложение и подстановка. В частности, комбинация перестановочных и подстановочных операций может привести к криптографически сильному нелинейному преобразованию, если оно применяется достаточное количество раз. Подстановочные операции во многих шифрах выступают при этом как основной нелинейный элемент циклового преобразования (нелинейный элемент замены). Поэтому значительные усилия исследователей направлены на изучение подходов к построению подстановок с высокими криптографическими показателями.

Сегодня наиболее разработанным и наиболее популярным математическим аппаратом оценки криптографических свойств нелинейных элементов замены (S-блоков) стал аппарат линейной алгебры и в частности аппарат булевых функций. Его развитию и применению посвящено много публикаций. Предложено и используется множество критериев и показателей оценки свойств как самих булевых (компонентных) функций S-блоков, так и критериев и показателей криптографических свойств S-блоков в целом. В их числе такие как: сбалансированность булевой функции, нелинейность N_f, корреляционный иммунитет, критерий распространения (строгий лавинный критерий) *KP*(k), алгебраическая степень булевой функции $\deg(f)$, а также соответствующие характеристики S-блоков: критерий битовой независимости (*BIC*), критерий нелинейности, максимальный порядок строгого лавинного критерия (*MOSAC*), максимальное значение линейной аппроксимационной таблицы-*LAT*, δ-гладкость (равномерность) XOR-таблицы S-блока и многие другие.

Следует отметить также предложенный в свое время подход к отбору подстановок [1-3], строящийся на основе оценки показателей их случайности (значений числа циклов, возрастаний и инверсий), дополненных ограничениями на максимально допустимые значения таблиц дифференциальных разностей и линейных аппроксимаций.

Этот подход нашел продолжение в работах [4,5], выполненных, в том числе, и с участием авторов данной работы. Основное внимание в этих публикациях сосредоточилось на разработке дополнительных критериев отбора случайных подстановок, построенных на использовании законов распределения переходов XOR таблиц и смещений таблиц линейных аппроксимаций случайных подстановок. Было предложено два дополнительных (к комбинаторным) критерия отбора, основанных на оценке близости дифференциальных и линейных законов распределения вероятностей подстановок к теоретически полученным законам [5]. Напомним здесь кратко их сущность.

Критерий 1. *Подстановка удовлетворяет критерию случайности* 4, *если закон распределения однотипных переходов* $\Pr(\Lambda_\pi(\Delta X, \Delta Y) = 2k),\ k = 0, 1, \ldots, k^*$ *её таблицы XOR разностей для входов, приписываемых к ненулевым характеристикам, соответствует по критерию согласия Колмогорова теоретическому закону распределения переходов случайной подстановки, т.е. наибольшее значение модуля разности теоретического и эмпирического законов распределения вероятностей удовлетворяет условию* $F_T(x_k) - F(x_k)S \le b$.

Здесь граничный параметр b подлежит уточнению по результатам экспериментов.

Критерий 2. *Подстановка удовлетворяет критерию случайности* 5, *если закон распределения однотипных переходов* $\Pr(\lambda^*(\alpha,\beta) = 2k),\ k = 0, 1, \ldots, k^*$ *её таблицы линейных аппроксимаций соответствует по критерию согласия Колмогорова теоретическому закону распределения переходов случайной подстановки, т.е. наибольшее значение модуля разности теоретического и эмпирического законов распределения вероятностей удовлетворяет условию* $F_T(x_k) - F(x_k)S \le c$.

Здесь параметр с также подлежит уточнению по результатам экспериментов.

В работе [6] были рассмотрены вопросы установления границ при использовании критерия Колмогорова для оценки близости законов распределения переходов дифференциальных и линейных таблиц подстановок теоретическим (мы их назвали Iэталонными), на основе результатов которых принимается решение – можно ли отнести проверяемую подстановку к случайной или нет.

Ожидалось, что подстановки, отобранные по предлагаемой системе критериев, окажутся более предпочтительными, чем известные конструкции. Однако, с одной стороны, формируемые в этом случае подстановки, как показал анализ, не имеют заметных преимуществ по сравнению с известными, а с другой - применение представленных выше критериев для практического отбора случайных подстановок встретило определённые затруднения, так как не ясной стала сама стратегия применения этих критериев. Вроде бы мы порождаем случайные подстановки, а потом начинаем их фильтровать. Не ясно, какие же показатели отбора являются предпочтительными.

В этой работе мы хотим изменить позицию к определению показателей случайности. Мы хотим ответить на вопросы, а какими свойствами будет обладать выборка случайно порождаемых подстановок? С какими подстановками в этом случае реально мы имеем дело? Как они соотносятся с приведенными критериями отбора?

Здесь будут изучаться показатели последовательности случайных подстановок, порождаемых случайным генератором. Итогом их изучения станет усовершенствованная модель случайной подстановки, отличающаяся от известных использованием свойств выборки случайных подстановок, что позволило существенно упростить правила отбора случайных подстановок (а практически использовать подстановки, порождаемые генератором случайных подстановок без каких-либо ограничений).

Задача практически сводится к определению законов распределения выборки, составленной из максимумов таблиц XOR разностей и максимумов смещений таблиц линейных аппроксимаций случайных подстановок.

Математические аспекты этой задачи рассмотрены в приложении работы [7]. В ней изучаются случаи, когда все значения выборки имеют одно и то же распределение и их плотности уменьшаются с ростом переменной *x* экспоненциально. Но это как раз и есть наши случаи.

2. Сущность методики определения законов распределения максимумов для больших по объёму выборок независимых одинаково распределённых случайных величин

Нас интересуют два случая.

1-й случай, когда выборка состоит из случайных значений переходов XOR таблицы случайной подстановки. Как известно [8], в этом случае закон распределения вероятностей переходов подчиняется Пуассоновскому закону:

$$\Pr(\Lambda_\pi(\Delta X, \Delta Y) = 2k) = e^{-1/2} \cdot \frac{1}{2^k \cdot k!} \cdot \qquad (1)$$

Здесь $\Pr(\Lambda_\pi(\Delta X, \Delta Y) = 2k)$ – вероятность, что значение дифференциальной таблицы случайно взятой подстановки p порядка 2^n для перехода входной разности ΔX в соответствующую выходную разность ΔY будет равно 2k.

2-й случай, когда выборка состоит из случайных значений, являющихся смещениями таблиц линейных аппроксимаций случайных подстановок, подчиняющихся нормальному закону распределения. Как показано в [7], в этом случае справедливо.

Утверждение. *Для случайной* n-*битовой подстановки, с* $n \geq 5$ *дисбаланс* Imb(v,u) *аппроксимации является случайным значением с распределением, которое может быть аппроксимировано в виде*

$$\Pr(\mathrm{Imb}(v, u) = z) \approx 2Z\left(\frac{z}{2^{(n-2)/2}}\right) \tag{2}$$

для z четного и ноль для z нечетного.

В наших обозначениях дисбаланс Imb(v, u) = z при z = 2k как раз соответствует значению смещения таблицы линейных аппроксимаций.

В работе [7] отмечается, что распределение максимумов больших по объёму выборок независимых одинаково распределённых случайных величин хорошо изучено в теории вероятностей и описывается распределением экстремальных значений Фишера-Типпета или log-Вейбула в виде:

$$D_{\max}(X) \approx e^{-e^{\frac{a-X}{b}}}.$$

Это распределение имеет математическое ожидание $\mu(X) = a + b\gamma$ с $\gamma \approx 0{,}58$ и среднеквадратическое отклонение $\frac{\pi}{\sqrt{6}} b \approx 1{,}3b$. Параметр *a* является решением уравнения

$$\ln(2)Y = f(x), \tag{3}$$

a b – единицей, делённой на производную функции $f(x)$ в точке a (здесь используется линейная аппроксимация функции $f(x)$ в точке a).

В работе [7] также показано, что решение уравнения (3) для выборки из 2^Y случайных значений, распределённых по Пуассоновскому закону, имеет вид

$$i = \frac{\ln(2)y - \frac{1}{2}\ln(2\pi i) - \lambda}{\ln\left(\frac{i}{\lambda}\right) - 1}. \tag{4}$$

Это уравнение может быть решено итеративно. Производная $f(i)$ вычисляется по формуле:

$$\ln\left(\frac{i}{\lambda}\right) + \frac{1}{2i}. \tag{5}$$

Определяя a и используя условие $a >> \lambda$, имеем: $b = \frac{1}{\ln\left(\frac{a}{\lambda}\right)}$.

Для нормального распределения (2) параметр, a_S (подстрочный индекс *s* для стандарта) является решением уравнения

$$a_S = \sqrt{2\ln(2)y - \ln(2\pi) - 2\ln(a_S)}, \tag{6}$$

которое может быть найдено итеративным путём, не обращая внимания на правый член в первой итерации. Производная $f(x)$ определяется по формуле:

$$x + \frac{1}{x} \tag{7}$$

и, следовательно,
$$b_s = \frac{a_s}{a_s^2 + 1} \approx \frac{1}{a_s}. \tag{8}$$

Грубо говоря, максимум имеет распределение со средним значением $1{,}17\sqrt{y}$ и стандартным отклонением $1.11/\sqrt{y}$. Авторы работы [7] отмечают, что можно найти значения a и b для любого нормального распределения со средним значением $\mu(X)$ и стандартным отклонением σ, заменив x на $\frac{X - \mu(X)}{\sigma}$. Это дает:

$$a = \sigma a_s + \mu(X),\ b = \sigma b_s. \tag{9}$$

3. Распределение максимумов XOR таблиц и смещений таблиц линейных аппроксимаций выборки из байтовых подстановок

Будем рассматривать выборку размера 2^n, n = 8. Для n = 8 из (4) имеем (табл.1).

Таблица 1

i	$\frac{\ln(2)\cdot 16 - \frac{1}{2}\ln(2\pi i) - \frac{1}{2}}{\ln(2i) - 1}$
5	6,8
5,5	6,3
5,9	5,98
6	5.9
7	5,3

И, следовательно, решением уравнения (4) является значение i = a, близкое к числу 6. Соответственно $b = \frac{1}{\ln(12)} = 0{,}4$. Но заметим, что формула (4), по которой мы определяли значение a, работает с половинным значением перехода дифференциальной таблицы. Поэтому при подсчёте действительного среднего значения мы должны полученный результат удвоить.

И тогда $\mu(X) = 2 \cdot 6 + 2 \cdot 0{,}4 \cdot 0{,}58 = 12{,}4$.

Если ориентироваться на результаты реального эксперимента, то среднее значение максимума должно быть близким к 11,55. Поэтому скорректируем наше значение до a = 5.

Это значение хорошо согласуется с результатами расчётов и экспериментов, представленных в работе [9].

Выше отмечалось, что поскольку распределение максимумов дискретное, то малая величина стандартного отклонения $b = \frac{1}{\ln(12)} = 0{,}4$ приводит к тому, что распределение сосредоточено в двух целочисленных значениях вблизи $\mu(X) \approx 2a$. В наших экспериментах с байтовыми подстановками это два значения: 10 и 12.

Далее расчёт предлагается вести по распределению

$$D_{max}(X) \approx e^{-e^{\frac{10 - 2\cdot X}{0{,}87}}}, \tag{10}$$

в котором использовано значение a = 5 (формула записана с учётом реального удвоения значений переходов XOR таблицы).

В табл. 2 приводим распределение значений максимумов для 256 битовых подстановок, рассчитанных по выражению (10), и результаты эксперимента.

Таблица 2

k^* (X_1, X_2)	Pr(k^*)	Расчётное значение	Эксперимент
8	0,00004	0,01	0
10 (10,8)	0,368 – 0,00004 = 0,368	94	92
12 (12,10)	0,905 – 0,368 = 0,537	137	147
14 (14, 12)	0,9901 – 0,905 = 0,008	22	14
16 (16,14)	0,9967 – 0,9901 = 0,0066	1,71	3
18 (18,16)	0,9999– 0,9967 = 0,0032	0,819	0

В табл. 3 представлены результаты решения уравнения (5) способом подбора. Для ориентировочного выбора начальных значений, используемых в переборе, вполне можно опираться на результаты расчётов и экспериментов, приведенных в работе [9].

Таблица 3

a_S	$\sqrt{\ln(2)32 - \ln(2\pi) - 2\ln(a_S)}$
4	4,19
5	4.13
6	4,09
8	4

Мы и в этом случае сделали небольшую коррекцию результата, ориентируясь на данные экспериментов. В качестве значения a_s рассматривалось значение $a_S = 4$ и соответственно

$$b_S = \frac{a_S}{a_S^2 + 1} \approx \frac{1}{a_S} = \frac{1}{4} = 0{,}25 \qquad (11)$$

(здесь уже учитываем результаты выполненных экспериментов, представленных в табл. 3). Для подстановок степени 2^8 имеем $\sigma = 2^{\frac{8-4}{2}} = 2^2$ и тогда $a = \sigma a_s + \mu(X) = 4 \cdot 4 + 0 = 16$ и в соответствии с (9) b = 4·0,25 = 1. Приходим к интегральному закону распределения максимумов полных дифференциалов уменьшенной 16-битной модели шифра в виде:

$$D_{max}(X) \approx e^{-e^{\frac{16-X}{1}}}, \qquad (12)$$

или с учётом реального удвоения результатов смещений таблицы линейных аппроксимаций

$$D_{max}(X) \approx e^{-e^{\frac{32-X}{2}}}. \qquad (13)$$

В табл. 4 представлены результаты расчётов по определению распределения значений максимумов линейных корпусов на основе интегрального закона распределения вероятностей (13).

Заметим, что по результатам ранее выполненной теоретической и экспериментальной оценки значения максимума смещения линейной аппроксимационной таблицы случайной подстановки степени 2^8 равно 32 (расчёт) и 34 (эксперимент) [10].

Видно, что и в этом случае результаты экспериментов практически повторяют результаты расчётов.

Самое интересное, как показали эксперименты, именно случайные подстановки, полученные без всяких ограничений, с очень большой вероятностью оказались подходящими с точки зрения криптографических приложений. Они позволили обеспечить динамические показатели выхода шифров с сильным линейным преобразованием к асимптотическим

показателям случайных подстановок, не уступающие считающимся лучшими (отобранными по специальным методикам) S-блокам практически всех современных шифров [11].

Таблица 4

k* (X_1, X_2)	Pr(k*)	Число значений	Эксперимент
< 26	$3.41\ 10^{-7}$	0	0
28 (28,26)	$5,6\ 10^{-4} - 3,4110^{-7} = 5,6\ 10^{-4}$	0,14	0
30 (30,28)	$0,064 - 5,6\ 10^{-4} = 0,0638$	16	10
32 (32,30)	0,368–0,064 = 0,304	78	86
34 (34,32)	0,692 – 0,304 = 0,388	99	98
36 (36,34)	0,874 – 0,692 = 0,181	46	46
38(38,36)	0,9518 – 0,874 = 0,078	19	10
40 (40,38)	0,9821 – 0,9518 = 0,03	8	6
42 (42,40)	0,9933 – 0,9821 = 0,011	3	0
44 (44,42)	0,9975 – 0,9973 = 0,00028	0,07	0

Подводя итог приведенным соображениям, можно теперь в соответствии с представленными результатами ввести и более практичное определение случайной подстановки.

В частности, байтовая подстановка является случайной, если:

1) значение максимума её XOR таблицы равняется 10,12;

2) значения максимумов смещений её таблицы линейных аппроксимаций находятся в диапазоне 32-38.

Этими определениями мы уточняем критерии 4 и 5, введенные ранее. Уточнение касается наложения (выполнение) ограничений лишь на максимальные значения переходов XOR таблиц и смещений таблиц линейных аппроксимаций.

Выводы

Результатом выполненных исследований является уточнённоё определение случайной подстановки (уточнённая математическая модель случайной подстановки), строящееся на свойствах выборки из случайных подстановок. Теперь появилось полное понимание того, с какими подстановками мы имеем дело при их случайном формировании. Как оказалось, с очень большой вероятностью мы будем получать подстановки, для которых значения максимумов дифференциальных таблиц и значения максимумов смещений таблиц линейных аппроксимаций принимают существенно ограниченное число возможных значений. Все они концентрируются вокруг теоретических значений максимумов случайных подстановок соответствующей степени.

Как показывают эксперименты [11], случайные подстановки, взятые с выхода генератора случайных подстановок без всяких ограничений, вполне могут конкурировать с лучшими известными конструкциями S-блоков, используемыми в современных шифрах.

Список литературы: 1. *Лисицкая И.В.* К вопросу построения долговременных ключей для алгоритма ГОСТ 28147-89 // Информационно-управляющие системы на железнодорожном транспорте. 1997. № 3. С. 54-57. **2.** *Lysytska I.V., Koriak A.S., Golovashich S.A., Oleshko O.I., Oleinik R.V.* The selection criteria of random substitution tables for symmetric enciphering algorithms // Abstracts of XXVIth General Assembly. Toronto, Ontario Canada, August 13-21, 1999. P. 204. **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. **4.** *Лисицкая И.В.* Оценка числа случайных подстановок с заданным распределением парных разностей XOR таблиц и смещений таблиц линейных аппроксимаций. / И.В. Лисицкая, А.В. Широков, Е.Д. Мельничук, К.Е. Лисицкий // Прикладная радиоэлектроника. Харьков: ХНУРЭ. 2010. Т. 9, № 3. С. 341-345. **5.** *Долгов В.И.* Случайные подстановки в криптографии. / В.И. Долгов, И.В. Лисицкая, К.Е. Лисицкий // Радіоелектронні та комп'ютерні системи. 2010. № 5 (46). С. 79-85. **6.** *Лисицкая И.В.* Экспериментальная проверка работоспособности новых критериев отбора случайных подстановок / И.В. Лисицкая, К.Е. Лисицкий, А.В. Широков, Е.Д. Мельничук // Радіоелектронні та комп'ютерні системи, 2010. № 6 (47). С. 87-93. **7.** *Joan Daemen, Vincent Rijmen.* Probability distributions of Correlation and Differentials in Block Ciphers. / Joan Daemen, Vincent Rijmen // April 13, 2006. P. 1-38. **8.** *Лисицкая И.В.* Свойства законов распределения XOR таблиц и таблиц линейных аппроксимаций случайных подстано-

вок // Вісник Харківського національного університету ім. В.Н. Каразіна. 2011. №960. Вип.16. С. 196-206. **9.** *Олейников Р.В.* Дифференциальные свойства подстановок / Р.В. Олейников, О.И. Олешко, К.Е. Лисицкий, А.Д. Тевяшев. // Прикладная радиоэлектроника. 2010. Т.9, № 3. С. 326-333. **10.** *Долгов В.И.* Свойства таблиц линейных аппроксимаций случайных подстановок. / В.И. Долгов, И.В. Лисицкая, О.И. Олешко. // Прикладная радиоэлектроника. Харьков: ХНУРЭ. 2010. Т. 9, № 3. С. 334-340. **11.** *Долгов В.И.* S-блоки для современных шифров. / В.И. Долгов, Е.В. Мельничук // Радиотехника. 2012. Вып.171. С. 121-133.