

ПРИМЕНЕНИЕ ДВУЛИКИХ ПРОЦЕССОВ К ГЕНЕРИРОВАНИЮ ПСЕВДОСЛУЧАЙНЫХ ЧИСЕЛ¹

Б. Я. Рябко

Описываются случайные процессы, у которых энтропия может быть сколь угодно близка к нулю, но при этом, как для полностью случайных последовательностей, частота встречаемости любого двоичного слова u стремится к $2^{-|u|}$, где $|u|$ — длина u . Это позволяет строить генераторы псевдослучайных чисел с доказанными свойствами, что представляет большой интерес для криптографических систем защиты информации.

Ключевые слова: случайные числа, псевдослучайные числа, энтропия Шеннона.

Генераторы случайных и псевдослучайных чисел (ГСЧ и ГПСЧ) находят широкое применение в системах защиты информации, причём используемые в таких системах генераторы должны удовлетворять целому ряду требований, одно из которых — статистическая неотличимость порождаемых генератором последовательностей от бернуллиевских с $p(0) = p(1) = 1/2$ [1]. С другой стороны, ГПСЧ по своему построению существенно отличаются от бернуллиевских процессов: энтропия (на символ) у выходной последовательности значительно меньше одного бита, тогда как у полностью случайной — один (см. описание схемы ГПСЧ, например, в [2, 3]). Напомним определение энтропии стационарного процесса μ : условная энтропия порядка m , $m = 1, 2, \dots$, и (предельная) энтропия даются равенствами $h_m = - \sum_{u \in \{0,1\}^{m-1}} \mu(u) \sum_{v \in \{0,1\}} \mu(v/u) \log \mu(v/u)$, $h_\infty = \lim_{m \rightarrow \infty} h_m$ [4].

В работе описываются процессы, для которых с вероятностью 1 в порождаемой последовательности $x_1 x_2 \dots$ для каждого двоичного слова u выполняется равенство

$$\lim_{t \rightarrow \infty} \nu_t(u)/(t - |u|) = 2^{-|u|},$$

где $\nu_t(u)$ — число встреч слов u в последовательности $x_1 \dots x_{|u|}$, $x_{|u|+1}, \dots$, $x_{t-|u|+1} \dots x_t$, что должно выполняться для полностью случайных последовательностей. Однако энтропия процесса может быть много меньше единицы, что обычно выполняется для ГПСЧ.

Сначала определим два семейства процессов $T_{k,\pi}$ и $\bar{T}_{k,\pi}$, где $k = 1, 2, \dots$ и $\pi \in (0, 1)$ — параметры. Оба процесса — марковские цепи связности, или памяти k , которые генерируют буквы из алфавита $\{0, 1\}$. Определим их матрицы переходов по индукции: матрица для $T_{1,\pi}$ определяется равенствами $P_{T_{1,\pi}}(0/0) = \pi$, $P_{T_{1,\pi}}(0/1) = 1 - \pi$ (очевидно, $P_{T_{1,\pi}}(1/0) = 1 - \pi$, $P_{T_{1,\pi}}(1/1) = \pi$). Процесс $\bar{T}_{1,\pi}$ определяется равенствами $P_{\bar{T}_{1,\pi}}(0/0) = 1 - \pi$, $P_{\bar{T}_{1,\pi}}(0/1) = \pi$. Предположим теперь, что $T_{k,\pi}$ и $\bar{T}_{k,\pi}$ определены. Тогда $T_{k+1,\pi}$ и $\bar{T}_{k+1,\pi}$ определяются так:

$$\begin{aligned} P_{T_{k+1,\pi}}(0/0u) &= P_{T_{k,\pi}}(0/u), \quad P_{T_{k+1,\pi}}(1/0u) = P_{T_{k,\pi}}(1/u), \\ P_{T_{k+1,\pi}}(0/1u) &= P_{\bar{T}_{k,\pi}}(0/u), \quad P_{T_{k+1,\pi}}(1/1u) = P_{\bar{T}_{k,\pi}}(1/u), \end{aligned}$$

¹Работа поддержана грантом РФФИ, проект № 15-29-07932.

и наоборот,

$$\begin{aligned} P_{\bar{T}(k+1,\pi)}(0/0u) &= P_{\bar{T}(k,\pi)}(0/u), & P_{\bar{T}(k+1,\pi)}(1/0u) &= P_{\bar{T}(k,\pi)}(1/u), \\ P_{\bar{T}(k+1,\pi)}(0/1u) &= P_{T(k,\pi)}(0/u), & P_{\bar{T}(k+1,\pi)}(1/1u) &= P_{T(k,\pi)}(1/u) \end{aligned}$$

для каждого $u \in \{0, 1\}^k$ (здесь vu — конкатенация слов v и u). Например,

$$P_{T(2,\pi)}(0/00) = \pi, \quad P_{T(2,\pi)}(0/01) = 1 - \pi, \quad P_{T(2,\pi)}(0/10) = 1 - \pi, \quad P_{T(2,\pi)}(0/11) = \pi.$$

Будем считать, что начальное распределение равномерное на $\{0, 1\}^k$, т. е. $P\{x_1 \dots x_k = u\} = 2^{-k}$ для $u \in \{0, 1\}^k$.

Теорема 1. Пусть последовательность $x_1 x_2 \dots$ порождается процессом $T(k, \pi)$ (или $\bar{T}(k, \pi)$), $k \geq 1$ и u — двоичное слово длины k . Тогда

i) имеет место

$$P(x_{j+1} \dots x_{j+k} = u) = 2^{-|u|}; \quad (1)$$

ii) для каждого $\pi \in (0, 1)$ энтропия h_k процессов $T(k, \pi)$ и $\bar{T}(k, \pi)$ равна 1 бит, тогда как предельная энтропия h_∞ равна $-(\pi \log_2 \pi + (1 - \pi) \log_2 (1 - \pi))$.

Определение 1. Назовём случайный процесс *двуликим k -го порядка*, если для него выполняются свойства i) и ii).

Поясним название «двуликие». Если мы рассматриваем слова длины, не превосходящей k , то они все равновероятны (энтропия равна 1), т. е. такие, какие должны быть у полностью случайного процесса. С другой стороны, если длина слова больше k , то энтропия меньше единицы, распределение вероятностей слов отличается от равномерного и процесс явно не «полностью случайный».

Рассмотрим на простом примере свойства этих процессов. Возьмём возможные типичные последовательности для $T(1, \pi)$ и $\bar{T}(1, \pi)$ для $\pi = 1/5$. Пусть последовательности такие: 010101101010100101... и 00001111100011111000... Каждая последовательность содержит примерно половину единиц и нулей и поэтому энтропия первого порядка равна 1, что должно выполняться для полностью случайной последовательности. С другой стороны, последовательности не выглядят как случайные, так как они содержат слишком длинные подпоследовательности вида 101010... или 000...11111... Поэтому энтропия второго и высших порядков меньше 1. Другими словами, данные последовательности имитируют полностью случайные, если учитываются только слова длины 1; при большей длине это не выполняется.

Следующая теорема показывает, что, в некотором смысле, двуликих процессов довольно много.

Теорема 2. Пусть $X = x_1 x_2 \dots$ и $Y = y_1 y_2 \dots$ — случайные процессы и процесс $Z = z_1 z_2 \dots$ задаётся равенствами $z_1 = x_1 \oplus y_1$, $z_2 = x_2 \oplus y_2, \dots$. Тогда если X — двуликий процесс k -го порядка ($k \geq 1$), то Z тоже двуликий k -го порядка.

Двуликие процессы k -го порядка имитируют полностью случайные только при длине слова, не превосходящей k . Оказывается, существуют процессы, для которых это свойство выполняется для слов любой длины.

Теорема 3. Существуют процессы, для которых (1) выполняется для слов любой длины.

В работе показано, как на основе двуликих процессов могут быть построены ГСЧ и ГПСЧ с доказанными статистическими свойствами.

ЛИТЕРАТУРА

1. *Rukhin A., Soto J., Nechvatal J., et al.* A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications. National Institute of Standards and Technology, 2010.
2. *Barker E. and Kelsey D.* Recommendation for Random Bit Generator (RBG) Constructions (DRAFT NIST Special Publication 800-90C). National Institute of Standards and Technology, 2012.
3. *Рябко Б. Я., Фионов А. Н., Шокин Ю. И.* Криптография и стеганография в информационных технологиях. Новосибирск: Наука, 2015.
4. *Cover T. M. and Thomas J. A.* Elements of Information Theory. N.Y., USA: Wiley-Interscience, 2006.

УДК 519.1

DOI 10.17223/2226308X/9/28

О КЛЮЧЕВОМ РАСПИСАНИИ БЛОЧНЫХ ШИФРОВ
БЕЗ СЛАБЫХ КЛЮЧЕЙ

В. М. Фомичев

Исследовано ключевое расписание симметричного r -раундового блочного шифра, при котором все раундовые ключи различны. Ключевое расписание реализуется как последовательное соединение автоматов: автономного автомата A , генерирующего выходную последовательность бинарных векторов с длиной периода не меньше r , и внутренне автономного автомата с постоянной памятью, в которой записан основной ключ блочного шифра. Рассмотрен пример, использующий в качестве автомата A линейный регистр сдвига с максимальной длиной периода.

Ключевые слова: *блочный шифр, раундовый ключ, неповторная последовательность, показатель неповторности последовательности.*

Введение

Используем следующие обозначения:

V_n — множество двоичных n -мерных векторов, $n \in \mathbb{N}$;

$X_{\rightarrow} = \{x_0, x_1, \dots\}$ — последовательность над множеством X ;

$\Gamma(A)$ — граф автомата Мили A ;

$\langle H \rangle$ — линейная оболочка множества векторов H .

Свойства ключевого расписания, характеризующие взаимосвязи основного ключа с раундовыми ключами, являются определяющими при оценке стойкости итеративного блочного шифра (ИБШ) относительно ряда методов криптоанализа: согласования, дифференциального и др. Например, нежелательно ключевое расписание, при котором генерируемая из основного ключа последовательность раундовых ключей содержит определённое число повторяющихся элементов. Так, по отношению к основному ключу при криптографическом анализе DES-алгоритма введено понятие слабого ключа, то есть основного ключа, порождающего 16 одинаковых раундовых ключей. В [1, с. 298] для r -раундового блочного алгоритма это понятие обобщено до μ -слабого ключа, порождающего в наборе раундовых ключей $q_1, \dots, q_r$ ровно μ различных элементов, $1 \leq \mu < r$. Показано, что при определённых условиях использование слабых ключей может привести к негативным последствиям с точки зрения обеспечения конфиденциальности данных. Криптографические свойства ИБШ считаются хорошими, если шифрующие подстановки близки по свойствам к случайным подстановкам, в частности, когда набор раундовых ключей $q_1, \dots, q_r$ есть случайная неповторная