ЦЕПИ МАРКОВА УСЛОВНОГО ПОРЯДКА И ИХ ПРИМЕНЕНИЕ ДЛЯ МОДЕЛИРОВАНИЯ КРИПТОГРАФИЧЕСКИХ ГЕНЕРАТОРОВ

М. В. Мальцев

Факультет прикладной математики и информатики БГУ
Минск, Беларусь
E-mail: maltsew@mail.ru

Рассматривается цепь Маркова условного порядка, используемая для моделирования криптографических генераторов, представлены обобщения данной модели. Приведены статистические оценки параметров. Представлены результаты компьютерных экспериментов.

Ключевые слова: цепь Маркова условного порядка, криптографический генератор, статистическое тестирование.

Введение

Методы статистического анализа временных рядов широко применяются для моделирования элементов систем криптографической защиты информации (СКЗИ) [1, 2]. Генератор, используемый в СКЗИ, должен порождать выходную последовательность, неотличимую от равномерно распределенной случайной последовательности (РРСП) [1]. Для обнаружения отклонения от модели РРСП используются статистические тесты, которые объединяются в так называемые батареи. Наиболее известными батареями тестов являются батарея NIST [3], батарея DIEHARD [4], батарея Дональда Кнута [5]. Но для более глубокого анализа, для выявления зависимостей высокого порядка требуются дополнительные исследования. Математической моделью, в которой распределение вероятностей будущего состояния зависит от s прошлых, является цепь Маркова s-го порядка [6]. К сожалению, число независимых параметров полносвязной цепи Маркова s-го порядка возрастает экспоненциально с увеличением s, поэтому использовать ее напрямую для обнаружения зависимостей большой глубины в выходных последовательностях криптографических генераторов практически невозможно. Для решения этой проблемы разрабатываются так называемые «малопараметрические» модели цепи Маркова s-го порядка. Примерами таких моделей являются цепь Маркова с частичными связями [7], модель Рафтери [8]. В настоящей статье рассматривается модель, предложенная в [9, 10] – цепь Маркова условного порядка.

Цепь Маркова условного порядка и ее обобщения

Приведем математическое описание цепи Маркова условного порядка. Примем обозначения: N – множество натуральных чисел; $A = \{0, 1, ..., N - 1\}$ – пространство состояний мощности $N \in \mathrm{N}$, $2 \le N < \infty$; $J_n^m = (j_n, j_{n+1}, \ldots, j_m) \in A^{m-n+1}$, $m, n \in \mathrm{N}$, $m \ge n$ – мультииндекс (цепочка $m - n + 1$ индексов); $< J_n^m > = \sum_{k=n}^{m} N^{k-n} j_k \in \{0, 1, \ldots, N^{m-n+1} - 1\}$ – числовое представление мультииндекса J_n^m; $G_k^l J_n^m$ – конкатенация мультииндек-

сов G_k^l и J_n^m; I{B} – индикаторная функция события B; $\{x_t \in A : t \in \mathrm{N}\}$ – однородная цепь Маркова s-го порядка ($2 \le s < \infty$) с (s + 1)-мерной матрицей вероятностей одношаговых переходов $P = (p_{J_1^{s+1}})$, $p_{J_1^{s+1}} = \mathrm{P}\{x_{t+s} = j_{s+1} \mid x_{t+s-1} = j_s, \ldots, x_t = j_1\}$, $J_1^{s+1} \in A^{s+1}$, $t \in \mathrm{N}$; $L \in \{1, \ldots, s-1\}$, $K = N^{L-1}$ – натуральные числа; $Q^{(1)}, \ldots, Q^{(M)}$ – семейство M ($1 \le M \le K+1$) различных квадратных стохастических матриц порядка N: $Q^{(m)} = (q_{ij}^{(m)})$, $0 \le q_{ij}^{(m)} \le 1$, $\sum_{j \in A} q_{ij}^{(m)} = 1$, $i, j \in A$, $1 \le m \le M$; $\nu_{l,y}(J_1^l) = \sum_{t=1}^{n-s} \mathrm{I}\{x_{t+s-l-y+1} = j_1, X_{t+s-l+2}^{t+s} = J_2^l\}$, $l \ge 2$, $y \ge 0$, $l + y \le s + 1$ – частота состояния $J_1^l \in A^l$ с пропуском в y символов между j_1 и J_2^l.

Цепь Маркова s-го порядка ($2 \le s < \infty$) $\{x_t \in A : t \in \mathrm{N}\}$ называется цепью Маркова условного порядка [9, 10], если ее вероятности одношаговых переходов имеют вид:

$$p_{J_1^{s+1}} = \sum_{k=0}^{K} \mathrm{I}\{< J_{s-L+1}^s >= k\} q_{j_{b_k} j_{s+1}}^{(m_k)}, \; J_1^{s+1} \in A^{s+1}, \tag{1}$$

где $1 \le m_k \le M \le K + 1$, $1 \le b_k \le s - L$, $0 \le k \le K$, $\min\limits_{0 \le k \le K} b_k = 1$; при этом в последовательности $m_0, \ldots, m_K$ встречаются все элементы множества $\{1, 2, \ldots, M\}$. Последовательность L элементов J_{s-L+1}^s, определяющая условие в формуле (1), называется базовым фрагментом памяти (БФП), L – длина БФП; величина $s_k = s - b_k + 1$ называется условным порядком. Таким образом, распределение вероятностей состояния процесса в момент времени t зависит не от всех s предыдущих состояний, как это было бы для полносвязной цепи Маркова порядка s, а от $L + 1$ состояний (j_{b_k}, J_{s-L+1}^s). Отметим, что при $L = s - 1$, $s_0 = \ldots = s_K = s$, получаем полносвязную цепь Маркова порядка s.

Если последовательность, порождаемая моделируемым криптографическим генератором, представляет собой эргодическую цепь Маркова, то с течением времени текущее распределение вероятностей стремится к стационарному распределению вероятностей. Поэтому для исследования надежности таких генераторов важно определить условия, при которых стационарное распределение вероятностей является равномерным. Следующая теорема устанавливает такие условия для цепи Маркова условного порядка.

Теорема. [10] Если цепь Маркова условного порядка (1) – эргодическая, то ее стационарное распределение вероятностей является равномерным тогда и только тогда, когда

$$\begin{cases} Q^{(m_k)} = N^{-1} \cdot 1_{N \times N}, \text{ если } s_k \in \{L+1, \ldots, s-1\}, \\ Q^{(m_k)} - \text{бистохастическая матрица, т. е. } \sum\limits_{i \in A} Q_{ij}^{(m_k)} = 1, \forall\, j \in A, \text{ если } s_k = s, \end{cases}$$

где $1_{N \times N}$ – ($N \times N$) – матрица, все элементы которой равны 1, k = 0,1, … , K.

В дальнейшем будем рассматривать случай, когда каждому значению БФП соответствует своя матрица вероятностей переходов, т. е. $m_k = k + 1$, $k = 0, 1, \ldots, K$.

В [9] построены оценки максимального правдоподобия вероятностей переходов, условных порядков:

$$\hat{q}^{(k+1)}_{j_0 j_{L+1}} = \begin{cases} \sum\limits_{J_1^L \in A^L} \mathrm{I}\{< J_1^L >= k\} \dfrac{\nu_{L+2, s_k - L-1}(J_0^{L+1})}{\nu_{L+1, s_k - L-1}(J_0^L)}, \text{ если } \nu_{L+1, s_k - L-1}(J_0^L) > 0, \\ 1/N, \text{ если } \nu_{L+1, s_k - L-1}(J_0^L) = 0, \end{cases} \tag{2}$$

$$\hat{s}_k = \arg\max_{L+1 \le y \le s} \sum_{J_1^L \in A^L} \mathrm{I}\{< J_1^L >= k\} \sum_{j_0, j_{L+1} \in A} \nu_{L+2, y-L-1}(J_0^{L+1}) \ln \hat{q}^{(k+1)}_{j_0 j_{L+1}}, \; k = 0, \ldots, K.$$

В [9] на основе асимптотических свойств оценок (2) построен статистический тест для проверки гипотез о значении матриц вероятностей одношаговых переходов $Q^{(k)}$, $k = 1, \ldots, K+1$: $H_0 = \{Q^{(1)} = Q_0^{(1)}, \ldots, Q^{(K+1)} = Q_0^{(K+1)}\}$, $H_1 = \overline{H}_0$, где $Q_0^{(1)}, \ldots, Q_0^{(K+1)}$ – некоторые фиксированные матрицы вероятностей переходов. Если $N = 2$, $A = \{0, 1\}$, все элементы матриц $Q_0^{(1)}, \ldots, Q_0^{(K+1)}$ равны 1/2, то тест позволяет обнаруживать отклонения от РРСП и имеет вид (порядок цепи Маркова, длина БФП и условные порядки предполагаются фиксированными):

$$\text{принимается } \begin{cases} H_0 (PPCП): \; \rho \le \Delta, \\ H_1: \; \rho > \Delta, \end{cases} \tag{3}$$

где

$$\rho = \sum_{J_0^{L+1} \in A^{L+1}} \sum_{k=0}^{K} I\{< J_1^L >= k\} 2 \nu_{L+1, s_k - L-1}(J_0^L) \left(\hat{q}^{(k+1)}_{j_0 j_{L+1}} - 1/2 \right)^2, \tag{4}$$

$\Delta = G_z^{-1}(1 - \alpha)$ – квантиль уровня $1 - \alpha$ ($\alpha \in (0, 1)$) стандартного χ^2-распределения с $z = 2^L + 1$ степенями свободы.

Обобщения модели (1) возможны по двум направлениям:

использование обобщенного базового фрагмента памяти;

использование многомерных матриц $Q^{(1)}, \ldots, Q^{(M)}$.

Дадим краткое описание этих обобщений.

Обозначим: $W = \{1, \ldots, s\}$; $W_L = \{\nu_1, \ldots, \nu_L\}$, $\nu_1, \ldots, \nu_L \in W$, $\nu_i \ne \nu_j$ при $i \ne j$; $\overline{W} = W \setminus W_L$.

Цепь Маркова s-го порядка $\{x_t \in A : t \in \mathrm{N}\}$ назовем цепью Маркова условного порядка с обобщенным базовым фрагментом памяти W_L, если ее вероятности одношаговых переходов имеют вид:

$$p_{J_1^{s+1}} = \sum_{k=0}^{K} \mathrm{I}(< S(J_1^s, W_L) >= k) q^{(m_k)}_{j_{b_k} j_{s+1}}, \tag{5}$$

где S – функция-селектор: $S(J_1^s, W_L) = (j_{\nu_1}, \ldots, j_{\nu_L})$. Если $W_L = \{s - L + 1, \ldots, s\}$, то (5) преобразуется в (1) и в этом частном случае мы приходим к введенной ранее модели.

В рамках второго направления обобщения модели цепи Маркова условного порядка соотношение (1) имеет вид:

$$p_{J_1^{s+1}} = \sum_{k=0}^{K} \mathrm{I}(< J_{s-L+1}^s >= k) q^{(m_k)}_{j_{b_k^{(1)}} \ldots j_{b_k^{(r)}} j_{s+1}}, \tag{6}$$

где $Q^{(m_k)} = (q^{(m_k)}_{j_{b_k^{(1)}}, \ldots, j_{b_k^{(r)}}, j_{s+1}})$ – $(r+1)$ – мерная матрица вероятностей одношаговых переходов, $1 \le r \le s - L$, $\{b_k^{(1)}, \ldots, b_k^{(r)}\} \subseteq \{1, \ldots, s - L\}$ – множество r различных элементов. Если $r = 1$, то (6) преобразуется в исходную модель (1).

Результаты компьютерного моделирования

Приведем результаты применения цепей Маркова условного порядка для выявления зависимостей большой глубины в выходных последовательностях самосжимающего генератора [11] и регистра сдвига с переменной обратной связью [12] на основе теста (3). Характеристический многочлен самосжимающего генератора $f(x) = x^{36} + x^{11} + 1$ (период выходной последовательности $T_1 = 2^{18}$), характеристические многочлены регистра сдвига с переменной обратной связью $f_1(x) = x^{23} + x^5 + 1$, $f_2(x) = x^{23} + x^9 + 1$, управляющий многочлен $g(x) = x^{18} + x^{16} + x^{11} + x^4 + 1$ (период выходной последовательности $T_2 \geq 2^{18} - 1$). Все вышеперечисленные многочлены являются примитивными. Генерировалось по $U = 1000$ реализаций выходной последовательности каждого генератора длительности $n = 100000$. Для вычисления тестовой статистики ρ (4) использовались следующие значения длины БФП и порядка: $L = 1, \ldots, 8$; $s = L+1, \ldots, 200$. Для каждой пары (s, L) вычислялась доля решений в пользу H_1: $\hat{\alpha} = \frac{1}{U}\sum_{u=1}^{U} I\{\rho_u > \Delta\}$. Отметим, что в формуле (4) использовались оценки условных порядков, при этом первая половина реализации использовалась для построения оценок по формуле (2), вторая половина – для вычисления ρ. На рис. 1 представлены результаты для самосжимающего генератора при $L = 8$. При таком значении длины БФП достигался наилучший результат – наибольший процент решений в пользу H_1. По горизонтальной оси откладывался порядок s, по вертикальной – значение величины $\hat{\alpha}$. Сплошная линия – уровень значимости $\alpha = 0{,}05$.

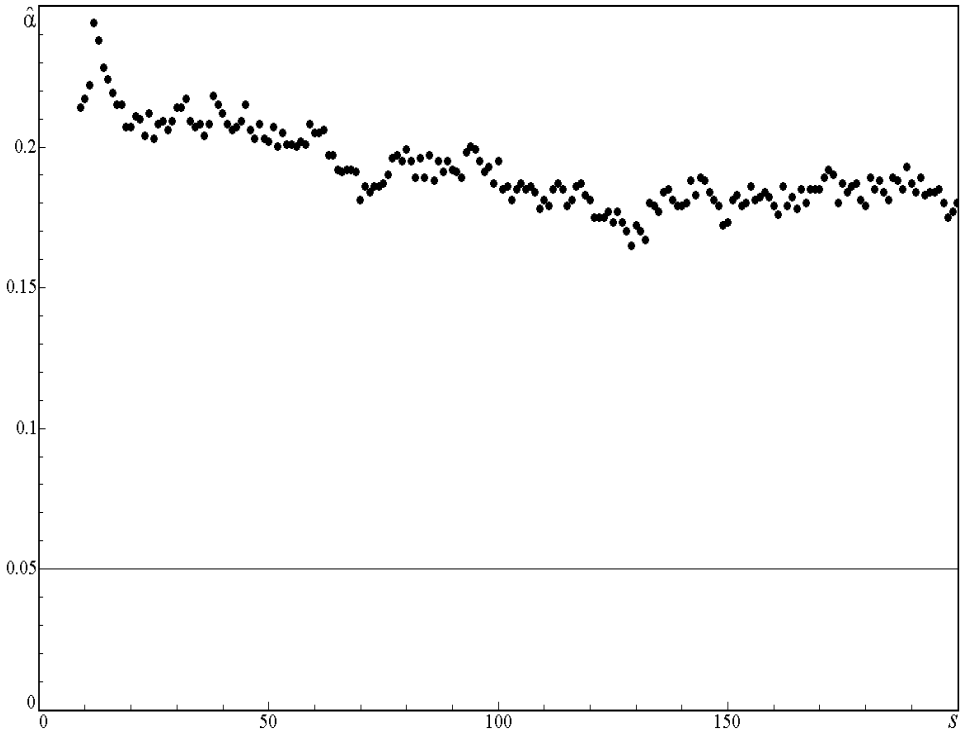

Рис. 1. Результаты моделирования при $L = 8$

Таким образом, результаты компьютерных экспериментов демонстрируют применимость теста (3) при увеличении параметра L. Отметим выявленный в ходе экспериментов нелинейный характер зависимости $\hat{\alpha}$ как от порядка s, так и от длины БФП L. К примеру, для самосжимающего генератора при $s = 3$ число решений в поль-

зу H_1 больше, чем при $s = 4$, а при $L = 6$ число решений в пользу H_1 наибольшее при $L < 50$.

Библиографические ссылки

1. *Харин Ю. С., Берник В. И., Матвеев Г. В., Агиевич С. В.* Математические и компьютерные основы криптологии. Минск : Новое знание, 2003.
2. *Максимов Ю. И.* О цепях Маркова, связанных с двоичными регистрами сдвига со случайными элементами // Труды по дискретной математике. Т. 1. М.: ТВП, 1997. С. 203–220.
3. A statistiical test suite for random and pseudorandom number generators for cryptographic applicatiions: NIST Special Publication 800-22 Rev. 1a. National Institute of Standards and Technology, 2010.
4. *Marsaglia G.* The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness [Electronic resource] // Florida State University, 1995. Mode of access : http://www.stat.fsu.edu/pub/diehard/. Date of access : 01.06.2013.
5. *Кнут Д. Э.* Искусство программирования: в 3 т. Т. 1: Получисленные методы. М.: Вильямс, 2007.
6. *Doob J. L.* Stochastic processes. New York: Wiley, 1953.
7. *Харин Ю. С., Петлицкий А. И.* Цепь Маркова с частичными связями ЦМ (*s*, *r*) и статистические выводы о ее параметрах // Дискретная математика. 2007. Т. 19. № 2. С. 109–130.
8. *Raftery A. E.* A model for high-order Markov chains // J. Royal Statistical Society. 1985. V. B-47. № 3. P. 528–539.
9. *Харин Ю. С., Мальцев М. В.* Статистическая проверка гипотез о параметрах цепи Маркова условного порядка // Весці НАН Беларуси. Серыя. фіз.-мат. навук. 2012. № 3. С. 5–12.
10. *Мальцев М. В.*, *Харин Ю. С.* Об асимптотических свойствах статистических оценок параметров цепи Маркова условного порядка // Весці НАН Беларуси, Сер. фіз.-мат. навук. 2013. № 1. С. 5–12.
11. *Meier W.*, *Staffelbach O.* The self-shrinking generator. // Advances in Cryptology EUROCRYPT 94. Springer-Verlag LNSC 950. 1995. P. 205–214.
12. *Алферов А. П., Зубов А. Ю. , Кузьмин А. С. , Черемушкин А. В*. Основы криптографии. М.: Гелиос АРВ, 2001.