УДК 004.421.5:519.2

# Оценка качества криптографических генераторов на основе цепей Маркова высокого порядка

*В. Ю. Палуха*; *Ю. С. Харин,* д-р физ.-мат. наук
НИИ прикладных проблем математики и информатики БГУ, Минск, Беларусь

*Рассмотрен подход к оценке качества криптографических генераторов на основе аппроксимации их выходных последовательностей малопараметрическими моделями цепей Маркова. Представлены результаты применения этого подхода к регистру сдвига с нелинейной обратной связью. Приведены результаты компьютерных экспериментов.*

*Ключевые слова*: криптографические генераторы, регистр сдвига с нелинейной обратной связью, малопараметрические модели цепей Маркова, цепь Маркова с частичными связями.

В средствах криптографической защиты информации широко используются генераторы псевдослучайных чисел. Для того чтобы системы защиты могли успешно противодействовать атакам, генераторы должны соответствовать определенным требованиям. В частности, если генератор аппроксимируется некоторой вероятностной моделью, то он может быть подвержен статистическим атакам, позволяющим оценить параметры этой модели.

Большинство криптографических генераторов является рекуррентной функцией некоторого порядка. Вероятностной моделью таких функций является цепь Маркова высокого порядка. Успешное оценивание параметров цепи Маркова, аппроксимирующей генератор, будет свидетельствовать о недостаточном качестве генератора. Продемонстрирована аппроксимация регистра сдвига с нелинейной обратной связью цепью Маркова с частичными связями.

## Математическая модель

Пусть на вероятностном пространстве $\{\Omega, F, P\}$ определена эргодическая цепь Маркова $s$-го порядка (ЦМ($s$)) $x_t \in V = \{0,1\}, t \in \boldsymbol{N}$ [1]:

$$P\{x_{t+1} = i_{t+1} \mid x_t = i_t, \ldots, x_1 = i_1\} =$$
$$= P\{x_{t+1} = i_{t+1} \mid x_t = i_t, \ldots, x_{t-s} = i_{t-s}\}, t > s.$$

Обозначим условное распределение вероятностей одношаговых переходов

$$p_{i_1,\ldots,i_{s+1}} = P\{x_{t+s} = i_{s+1} \mid x_t = i_1, \ldots, x_{t+s-1} = i_s\},$$
$$\sum_{i_{s+1}} p_{i_1,\ldots,i_{s+1}} = 1, \quad i_1, \ldots, i_{s+1} \in V,$$

которое в силу однородности цепи Маркова не зависит от $t \in \boldsymbol{N}$.

Содержательный смысл Марковского условия заключается в том, что распределение вероятностей элемента последовательности $x_{t+1}$ зависит не от всей предыстории $x_1, \ldots, x_t$, а только от $s$ предыдущих элементов $x_{t-s+1}, \ldots, x_t$. По такому принципу работает большинство криптографических генераторов.

На практике модель полносвязной цепи Маркова высокого порядка используется редко из-за экспоненциального роста числа параметров модели ($2^s$) с ростом порядка $s$. Вместо них используются малопараметрические модели [1]. Одной из таких моделей является цепь Маркова $s$-го порядка с $r$ частичными связями (ЦМ($s$, $r$)), предложенная в 2004 г. [2]:

$$p_{i_1,\ldots,i_{s+1}} = q_{i_{m_1},\ldots,i_{m_r},i_{s+1}}, \quad i_1,\ldots,i_{s+1} \in V,$$
$$M = \{m_1, \ldots, m_r\}, \quad I = \{1,\ldots,s\}, \quad M \subseteq I.$$

Вероятности одношаговых переходов $q_{i_{m_1},\ldots,i_{m_r},i_{s+1}}$, $i_1, \ldots, i_{s+1} \in V$ образуют матрицу вероятностей одношаговых переходов $Q$ размерности $2^r \times 2$. Суть модели (ЦМ($s$, $r$)) состоит в том, что следующий элемент последовательности $x_{t+1}$ зависит не от всех $s$ предыдущих элементов $x_{t-s+1}$, …, $x_t$, а от некоторых $r \le s$ элементов $x_{t-s+m_1}, \ldots, x_{t-s+m_r}$. Если $r = s$, то получаем полносвязную цепь Маркова, если $r < s$, то число параметров модели сокращается до $2^r$. В этом заключается преимущество малопараметрических моделей цепей Маркова.

Одними из самых простых в реализации и удобных в использовании генераторов являются

**Палуха Владимир Юрьевич**, аспирант.
E-mail: palukha@bsu.by
**Харин Юрий Семенович**, директор, профессор.
E-mail: Kharin@bsu.by

*Статья поступила в редакцию 14 июля 2014 г.*

регистры сдвига с функциональной обратной связью. Эти генераторы характеризуются состоянием ($y_1$, $y_2$, …, $y_s$) и функцией обратной связи $f(y_1, y_2, \ldots, y_s)$, $y_i \in V = \{0,1\}$, $i \in \{1,\ldots,s\}$. Приведем алгоритм выработки последовательности регистром сдвига с обратной связью.

Вход: начальное состояние ($x_1$, $x_2$, …, $x_s$), длина последовательности $T$.

Шаг 0: ($y_1$, $y_2$, …, $y_s$) := ($x_1$, $x_2$, …, $x_s$).

Шаг $i$, $i$ = 1, …, $T$:

а) $x_i := y_1$; $c := f(y_1, y_2, \ldots, y_s)$.

б) Для $j$ = 1, …, $s$ – 1: $y_j = y_{j+1}$.

в) $y_s = c$.

Выход: $x_1$, …, $x_T$.

Если функция $f$ является линейной, то в этом случае генератор называется регистром сдвига с линейной обратной связью (РСЛОС). РСЛОС является хорошо изученным с точки зрения криптоанализа генератором. Функция обратной связи восстанавливается при помощи алгоритма Берлекэмпа—Месси. Практический интерес представляют генераторы с нелинейной функцией — регистры сдвига с нелинейной обратной связью. Их свойства до сих пор являются недостаточно изученными [3]. Рассматриваются регистры с функциями, алгебраическая степень которых равна двум.

## Статистическое оценивание параметров цепи Маркова и регистра сдвига

В [4] предложен алгоритм статистического оценивания параметров модели ЦМ($s$, $r$). С помощью этого алгоритма строится оценка порядка цепи Маркова $s \in [s_-, s_+]$, числа частичных связей $r \in [r_-, r_+]$, шаблона $M$ и матрицы $Q$. Вычислительная сложность алгоритма оценивания $M$ и $Q$ при известных $r$ и $s$ и длине последовательности $T$ равна $O(2^{r+1}s^{r-1} + Ts^r)$.

Оценка матрицы вероятностей одношаговых переходов $Q$, построенная в результате работы алгоритма оценивания параметров ЦМ($s$, $r$), может быть использована для построения таблицы истинности булевой функции обратной связи. Она строится следующим образом. Матрица $Q$ имеет размерность $2^r \times 2$. Каждая строка матрицы $(q_{i_{m_1},\ldots,i_{m_r},0}\ q_{i_{m_1},\ldots,i_{m_r},1})$ содержит вероятности генерации 0 или 1 в зависимости от предыстории $i_{m_1},\ldots,i_{m_r}$. Если вероятность генерации 0 больше, чем вероятность генерации 1, то булева функция принимает значение 0, в противном случае значение функции будет равным 1:

$$f(i_{m_1},\ldots,i_{m_r}) = \begin{cases} 0, & q_{i_{m_1},\ldots,i_{m_r},0} > q_{i_{m_1},\ldots,i_{m_r},1}; \\ 1, & q_{i_{m_1},\ldots,i_{m_r},0} \le q_{i_{m_1},\ldots,i_{m_r},1}. \end{cases}$$

Если генератор без искажений, то строки матрицы будут иметь вид (0 1) либо (1 0). Это значит, что в качестве вектора значений булевой функции будет выступать второй столбец матрицы $Q$.

Система Wolfram Mathematica позволяет по таблице истинности построить алгебраическую нормальную форму булевой функции. Написана программа, которая преобразует оценку матрицы $Q$, полученную в результате работы алгоритма оценивания параметров ЦМ($s$, $r$), в таблицу истинности в формате Wolfram Mathematica. Также в системе Mathematica была написана программа, которая восстанавливает булеву функцию по таблице истинности.

## Компьютерные эксперименты

Реализованы шесть регистров сдвига с нелинейной обратной связью [3]. С помощью каждого из регистров получены по 5 последовательностей длины $T$ = 1000, которые отличались начальным заполнением регистра. В таблице представлены значения функций обратной связи каждого из регистров, истинные значения порядка и числа частичных связей, значения интервалов для работы алгоритма оценивания параметров.

На рисунке проиллюстрировано восстановление функции $f_2$ по таблице истинности, построенной по оценке матрицы, полученной в результате работы алгоритма оценивания параметров ЦМ($s$, $r$). Во всех 30 случаях исходная функция была восстановлена верно.

**Параметры регистров сдвига и алгоритма оценивания**

| $f$ | $s$ | $r$ | $s_-$ | $s_+$ | $r_-$ | $r_+$ |
|---|---|---|---|---|---|---|
| $f_1 = x_1 \oplus x_2 \oplus x_8x_{11} \oplus x_{10}x_{16}$ | 17 | 6 | 16 | 20 | 4 | 10 |
| $f_2 = x_1 \oplus x_7 \oplus x_3x_{10} \oplus x_8x_{13}$ | 17 | 6 | 16 | 20 | 4 | 10 |
| $f_3 = x_1 \oplus x_2 \oplus x_4 \oplus x_{10} \oplus x_{13} \oplus x_8x_{14}$ | 17 | 7 | 16 | 20 | 4 | 10 |
| $f_4 = x_1 \oplus x_2 \oplus x_8 \oplus x_{12} \oplus x_{14} \oplus x_7x_{15}$ | 17 | 7 | 16 | 20 | 4 | 10 |
| $f_5 = x_1 \oplus x_4 \oplus x_9 \oplus x_{12} \oplus x_{13} \oplus x_4x_{12}$ | 17 | 6 | 16 | 20 | 4 | 10 |
| $f_6 = x_1 \oplus x_2 \oplus x_9 \oplus x_{10} \oplus x_{16} \oplus x_8x_{19}$ | 24 | 7 | 21 | 25 | 4 | 10 |

```
BooleanFunction[{{False, False, False, False, False, False} → False, {False, False, False, False, False, True} → False, {False, False, False, False, True, False} → False,
  {False, False, False, False, True, True} → False, {False, False, False, True, False, False} → False, {False, False, False, True, False, True} → True,
  {False, False, False, True, True, False} → False, {False, False, False, True, True, True} → True, {False, False, True, False, False, False} → False,
  {False, False, True, False, False, True} → False, {False, False, True, False, True, False} → True, {False, False, True, False, True, True} → False,
  {False, False, True, True, False, False} → False, {False, False, True, True, False, True} → True, {False, False, True, True, True, False} → True,
  {False, False, True, True, True, True} → False, {False, True, False, False, False, False} → True, {False, True, False, False, False, True} → True,
  {False, True, False, False, True, False} → True, {False, True, False, False, True, True} → True, {False, True, False, True, False, False} → True,
  {False, True, False, True, False, True} → False, {False, True, False, True, True, False} → True, {False, True, False, True, True, True} → False,
  {False, True, True, False, False, False} → True, {False, True, True, False, False, True} → True, {False, True, True, False, True, False} → False,
  {False, True, True, False, True, True} → False, {False, True, True, True, False, False} → True, {False, True, True, True, False, True} → False,
  {False, True, True, True, True, False} → False, {False, True, True, True, True, True} → True, {True, False, False, False, False, False} → True,
  {True, False, False, False, False, True} → True, {True, False, False, False, True, False} → True, {True, False, False, False, True, True} → True,
  {True, False, False, True, False, False} → True, {True, False, False, True, False, True} → False, {True, False, False, True, True, False} → True,
  {True, False, False, True, True, True} → False, {True, False, True, False, False, False} → True, {True, False, True, False, False, True} → True,
  {True, False, True, False, True, False} → False, {True, False, True, False, True, True} → False, {True, False, True, True, False, False} → True,
  {True, False, True, True, False, True} → False, {True, False, True, True, True, False} → False, {True, False, True, True, True, True} → True,
  {True, True, False, False, False, False} → False, {True, True, False, False, False, True} → False, {True, True, False, False, True, False} → False,
  {True, True, False, False, True, True} → False, {True, True, False, True, False, False} → False, {True, True, False, True, False, True} → True,
  {True, True, False, True, True, False} → False, {True, True, False, True, True, True} → True, {True, True, True, False, False, False} → False,
  {True, True, True, False, False, True} → False, {True, True, True, False, True, False} → True, {True, True, True, False, True, True} → True,
  {True, True, True, True, False, False} → False, {True, True, True, True, False, True} → True, {True, True, True, True, True, False} → True,
  {True, True, True, True, True, True} → False}, {x1, x2, x8, x10, x11, x16}, "ANF"]

x1 ⩒ x2 ⩒ (x10 && x16) ⩒ (x11 && x8)
```

***Восстановление функции в Wolfram Mathematica***

## Заключение

Проведенные компьютерные эксперименты показали, что регистр сдвига с нелинейной обратной связью хорошо аппроксимируется цепью Маркова с частичными связями. С помощью методов математической статистики строится оценка параметров аппроксимационной модели, на основании которых восстанавливается функция обратной связи исходного генератора.

**Литература**


1. *Харин Ю. С.* Оптимальность и робастность в статистическом прогнозировании / Ю. С. Харин. — Минск: БГУ, 2008. — 263 с.
2. *Харин Ю. С.* Цепи Маркова с *r* частичными связями и их статистическое оценивание / Ю. С. Харин // Доклады НАН Беларуси. 2004. Т. 48. № 1. С. 40—44.
3. *Dubrova E.* A List of Maximum Period NLFSRs / E. Dubrova // Cryptology ePrint Archive, Report 2012/166 (2012). http://eprint.iacr.org/2012/166.
4. *Харин Ю. С.* Цепь Маркова *s*-го порядка с *r* частичными связями и статистические выводы о ее параметрах / Ю. С. Харин, А. И. Петлицкий // Дискретная математика. 2007. Т. 19. № 2. С. 109—130.


# Quality assessment for cryptographic generators based on high-order Markov chains


*V. Yu. Palukha, Yu. S. Kharin*

Research Institute for Applied Problems of Mathematics and Informatics of BSU, Minsk, Belarus



*An approach to quality assessment of cryptographic generators based on the approximation of their output sequences by parsimonious Markov chain models is considered. This approach is illustrated on nonlinear feedback shift register. The results of computer experiments are presented.*



*Keywords*: cryptographic generators, nonlinear feedback shift register, parsimonious Markov chain models, Markov chain with partial connections.

Bibliography — 4 references. *Received July 14, 2014*


* * *