УДК 519.115

С.В. АГИЕВИЧ, О.В. СОЛОВЕЙ

О ПЕРИОДИЧЕСКИХ СВОЙСТВАХ САМОПРОРЕЖИВАЮЩИХ ГЕНЕРАТОРОВ ПСЕВДОСЛУЧАЙНЫХ ЧИСЕЛ

We consider the self-decimated generator of pseudorandom-numbers and investigate the pre-period λ and the period μ of its state sequence. We obtain the expectations and variances of λ and μ for the case when decimation steps are chosen randomly, independendly and uniformly from the set {1,2}.

1. Результаты. Пусть N — множество всех натуральных чисел, $\mathrm{N}_0 = \mathrm{N} \cup \{0\}$, $A \subset \mathrm{N}_0$ — конечный алфавит. Через A^* будем обозначать множество всех слов конечной длины в алфавите A (включая пустое слово ε), а через A^ω — множество всех односторонне бесконечных слов (подробнее см. [1, 2]). Если $a \in A^*$, то $l(a)$ — длина a, а $w(a)$ — сумма символов a. Далее c^n — слово, составленное из n символов $c \in A$, и ab — конкатенация слов a и b.

По заданным слову $s = s_0 s_1 \ldots \in A^\omega$ и натуральному T определим числа $\lambda = \lambda(s,T) \in \mathrm{N}_0$ и $\mu = \mu(s,T) \in \mathrm{N}$ такие, что

$$T \mid (s_\lambda + s_{\lambda+1} + \ldots + s_{\lambda+\mu-1}) \tag{1}$$

и T не делит никакую из сумм $s_t + s_{t+1} + \ldots + s_\tau$ с $0 \le t < \tau < \lambda + \mu - 1$. Например, если $s = 2212221\ldots = 2^2 12^3 1\ldots$, то

$$\lambda(s,1) = 0,\ \mu(s,1) = 1,\quad \lambda(s,2) = 0,\ \mu(s,2) = 1,\ldots\quad \lambda(s,8) = 2,\ \mu(s,8) = 5,\ldots.$$

Введенные характеристики λ и μ описывают периодические свойства самопрореживающих генераторов псевдослучайных чисел (см. [3, 4]). Пусть имеется генератор G со множеством внутренних состояний $\mathcal{S}$, $|\mathcal{S}| = T$, и функцией $\varphi : \mathcal{S} \to \mathcal{S}$ перехода между состояниями. Пусть φ является полноцикловой подстановкой, т. е. для $S \in \mathcal{S}$ образы $\varphi(S), \varphi^2(\varphi(S)), \ldots, \varphi^T(S)$ пробегают все $\mathcal{S}$.

При обычной работе генератора G на основании выбранного начального состояния $S_0 \in \mathcal{S}$ определяется последовательность

$$S_{t+1} = \varphi(S_t),\quad t = 0,1,\ldots \tag{2}$$

и текущее внутреннее состояние S_t используется для определения текущего выходного псевдослучайного числа.

Самопрореживание (внутренних состояний G) состоит во введении дополнительной функции $d : \mathcal{S} \to A$ и замене правила (2) на правило

$$S_{t+1} = \varphi^{d(S_t)}(S_t),\quad t = 0,1,\ldots \tag{3}$$

Теперь, если символами слова s являются числа $s_t = d(S_t)$, то $\lambda(s,T)$ и $\mu(s,T)$ есть соответственно предпериод и период последовательности внутренних состояни (3).

Далее будем считать, что символы слова s выбираются из алфавита A случайно, независимо и равновероятно. В этом случае характеристики λ и μ являются случайными величинами, свойства которых представляют интерес в связи с оценкой периода и предпериода самопрореживающих генераторов.

Один из классических результатов теории случайных отображений можно сформулировать так (см. [5]): если $A = \{0,1,\ldots,T-1\}$, то при $T \to \infty$ средние значения

$$\mathrm{E}\lambda(s,T), \mathrm{E}\mu(s,T) = \sqrt{\frac{\pi T}{8}} + O(1).$$

Получение подобных асимптотических (или точных) выражений для произвольного алфавита A является, по-видимому, значительно более сложной задачей.

Нами рассмотрен часто используемый при построении генераторов выбор $A = \{1,2\}$ и получены следующие результаты.

Теорема. *Если символы слова $s \in \{1,2\}^{\omega}$ выбираются случайно, независимо и равновероятно, то средние значения*

$$\mathrm{E}\lambda(s,T) = \frac{4}{9} - \frac{1}{2^{T+2}}\left(T + 2 - (-1)^T \frac{3T+2}{9}\right), \tag{4}$$

$$\mathrm{E}\mu(s,T) = \frac{2T}{3} + \frac{T}{2^{T+2}}\left(1 - (-1)^T \frac{5}{3}\right), \tag{5}$$

а при $T \to \infty$ дисперсии

$$\mathrm{D}\lambda(s,T) = \frac{44}{81} + O\left(\frac{T^2}{2^T}\right), \quad \mathrm{D}\mu(s,T) = \frac{2T}{27} + O\left(\frac{T^2}{2^T}\right). \tag{6}$$

Отметим, что при умножении символов слова s на взаимно простое с T число q характеристики λ и μ не изменяются. Поэтому оценки (4)–(6) остаются справедливыми и для случайного слова $s \in \{q,2q\}^{\omega}$.

Отметим также, что Р. Рюппель в работе [4] рассмотрел случай, когда слово s формируется на основании линейной рекуррентной последовательности $\sigma_0,\sigma_1,\ldots$ над полем из двух элементов по следующему правилу:

$$s_t = \begin{cases} 1, & \sigma_t = 0, \\ 2, & \sigma_t = 1, \end{cases} \quad t = 0,1,\ldots.$$

Для случая, когда характеристический многочлен последовательности $\sigma_0,\sigma_1,\ldots$ является примитивным многочленом степени k получена оценка

$$\mu(s,T)=\left\lfloor\frac{2T}{3}\right\rfloor,\quad T=2^k-1,$$

где $\lfloor z\rfloor$ есть максимальное целое $\le z$. Как видим, данная оценка согласуется со средним значением характеристики $\mu(s,T)$ случайного слова s.

2. Доказательство. По заданным s и T определим слова $s_0\dots s_{\lambda-1}$ и $s_\lambda\dots s_{\lambda+\mu-1}$, которые назовем префиксом и циклической частью s соответственно. Обратно, всякое непустое слово $a\in A^*$ является циклической частью некоторого слова s. По a можно определить возможные значения $T=T(a)$ и множество $B(a)$ допустимых префиксов s, используя следующие ограничения:

P1) $w(a)\equiv 0 \pmod T$;

P2) вычеты

$$0,\quad a_0,\quad a_0+a_1,\dots,\quad a_0+a_1+\dots+a_{l(a)-2}\pmod T \tag{7}$$

попарно различны;

P3) если $b\in B(a)$ и $b\ne\varepsilon$, то каждый из вычетов

$$-b_{l(b)-1},\quad -b_{l(b)-1}-b_{l(b)-2},\dots,\quad -b_{l(b)-1}-b_{l(b)-2}-\dots-b_0\pmod T$$

отличается от вычетов (7).

Действительно, ограничение P1 следует непосредственно из определения циклической части. Если нарушено P2, то для некоторых t и τ, $0\le t<\tau<l(a)-1$ выполняется сравнение

$$a_t+\dots+a_\tau\equiv 0 \pmod T$$

и слово a не может являться циклической частью. Наконец, при нарушении P3 для некоторых $t\le l(b)-1$ и $\tau<l(a)-1$ выполняется одно из сравнений

$$b_t+\dots+b_{l(b)-1}\equiv 0 \pmod T,\quad b_t+\dots+b_{l(b)-1}+a_0+\dots+a_\tau\equiv 0 \pmod T$$

и слово a снова не может являться циклической частью.

Для $\Omega\subseteq A^*$ введем в рассмотрение производящую функцию

$$G_\Omega(x,y,z)=\sum_{n,t\ge 1,\,m\ge 0} g(n,m,t)x^n y^m z^t, \tag{8}$$

где $g(n,m,t)$ — число слов s с циклической частью $a\in\Omega$ и префиксом $b\in B(a)$ такими, что $l(a)=n$, $l(b)=m$ и $T(a)=t$.

Далее полагаем $A=\{1,2\}$. Разобьем A^* на подмножества Ω_1, Ω_2, Ω_3, Ω_4 (они будут определены ниже), и найдем соответствующие им производящие функции вида (8).

1. $\Omega_1=\{2\}^*1$. Если $a\in\Omega_1$, то $T\mid w(a)$ согласно P1 и $T\ge l(a)>w(a)/2$ согласно P2. Следовательно, $T=w(a)$. Теперь, если $a=2^m1$, то $B(a)=\{2^k:k=0,\dots,m\}$ согласно P3. Искомая производящая функция

$$G_{\Omega_1}(x,y,z)=\sum_{m\ge 0}x^{m+1}z^{2m+1}\sum_{k=0}^{m}y^k=$$

$$=\sum_{m\ge 0}x^{m+1}z^{2m+1}\frac{1-y^{m+1}}{1-y}=$$

$$=\frac{xz}{1-y}\sum_{m\ge 0}\left(xz^2\right)^m-\frac{xyz}{1-y}\sum_{m\ge 0}\left(xyz^2\right)^m=$$

$$=\frac{xz}{(1-y)(1-xz^2)}-\frac{xyz}{(1-y)(1-xyz^2)}.$$

2. $\Omega_2=\{1,2\}^*1\{2\}^*1$. Снова $T=w(a)$ на основании P1, P2. Пусть $a=\alpha 12^m1$, $\alpha\in A^*$. Используя P3, получаем $B(a)=\{2^k:k=0,\dots,m\}$. Поэтому,

$$G_{\Omega_2}(x,y,z)=\sum_{a\in A^*}x^{l(a)}z^{w(a)}\sum_{m\ge 0}x^{m+2}z^{2m+2}\sum_{k=0}^{m}y^k=$$

$$=\sum_{l\ge 0}(xz+xz^2)^l\sum_{m\ge 0}x^{m+2}z^{2m+2}\frac{1-y^{m+1}}{1-y}=$$

$$=\frac{1}{1-xz-xz^2}\left(\frac{x^2z^2}{1-y}\sum_{m\ge 0}\left(xz^2\right)^m-\frac{x^2yz^2}{1-y}\sum_{m\ge 0}\left(xyz^2\right)^m\right)=$$

$$=\frac{1}{1-xz-xz^2}\left(\frac{x^2z^2}{(1-y)(1-xz^2)}-\frac{x^2yz^2}{(1-y)(1-xyz^2)}\right).$$

3. $\Omega_3=\{1,2\}^*12\{2\}^*$. В этом случае $T=w(a)$ и $B(a)=\{\varepsilon\}\cup\{2^k1:k=0,\dots,m\}$ для слова $a=\alpha 122^m$. Поэтому

$$G_{\Omega_3}(x,y,z)=\sum_{a\in A^*}x^{l(a)}z^{w(a)}\sum_{m\ge 0}x^{m+2}z^{2m+3}\sum_{k=0}^{m+1}y^k=$$

$$=\sum_{l\ge 0}(xz+xz^2)^l\sum_{m\ge 0}x^{m+2}z^{2m+3}\frac{1-y^{m+2}}{1-y}=$$

$$=\frac{1}{1-xz-xz^2}\left(\frac{x^2z^3}{(1-y)(1-xz^2)}-\frac{x^2y^2z^3}{(1-y)(1-xyz^2)}\right).$$

4. $\Omega_4=\{2\}^*$. Если $a=2^m$, то $T(a)=2m$ и $B(a)=\{\varepsilon\}\cup\{2^k1:k=0,\dots,m-1\}$. Дополнительно $T(a)=m$ для нечетных m, при этом $B(a)=\{\varepsilon\}$. Производящая функция

$$G_{\Omega_4}(x,y,z)=\sum_{m\ge 0}x^mz^{2m}\sum_{k=0}^{m}y^k+\sum_{m\ge 0}x^{2m+1}z^{2m+1}=$$

$$=\sum_{m\ge 0}(xz^2)^m\frac{1-y^{m+1}}{1-y}+xz\sum_{m\ge 0}(x^2z^2)^m=$$

$$=\frac{1}{(1-y)(1-xz^2)}-\frac{y}{(1-y)(1-xyz^2)}+\frac{xz}{1-x^2z^2}.$$

Окончательно получаем

$$G_{A^*}(x,y,z) = G_{\Omega_1}(x,y,z) + G_{\Omega_2}(x,y,z) + G_{\Omega_3}(x,y,z) + G_{\Omega_4}(x,y,z) =$$

$$= \frac{1 + x^2yz^3}{(1-xyz^2)(1-xz-xz^2)} + \frac{xz}{1-x^2z^2}.$$

Найдем $\mathrm{E}\lambda(\mathrm{s},\mathrm{T})$. Если $g(n,m)$ — число слов $s \in A^*$ с характеристиками $\lambda(s,T) = m$ и $\mu(s,T) = n$, то по построению

$$\mathrm{E}\lambda(s,T) = \sum_{n\geq 1,\ m\geq 0} g(n,\ m)\frac{m}{2^{n+m}} = \frac{1}{2}[z^T]\frac{\partial G_{A^*}(1/2,y,z)}{\partial y}\bigg|_{y=1/2},$$

где $[z^T]f(z)$ — есть коэффициент f_T в производящей функции $f(z) = f_0 + f_1 z + f_2 z^2 + \ldots$.

Имеем

$$\frac{1}{2}\frac{\partial G_{A^*}(1/2,y,z)}{\partial y}\bigg|_{y=1/2} = \frac{4z^2}{(1-z)(4-z^2)^2} =$$

$$= \frac{1}{3(2+z)^2} - \frac{1}{18(2+z)} - \frac{1}{(2-z)^2} - \frac{1}{2(2-z)} + \frac{4}{9(1-z)} =$$

$$= \frac{1}{12}\sum_{t\geq 0}(t+1)\left(-\frac{z}{2}\right)^t - \frac{1}{36}\sum_{t\geq 0}\left(-\frac{z}{2}\right)^t -$$

$$-\frac{1}{4}\sum_{t\geq 0}(t+1)\left(\frac{z}{2}\right)^t - \frac{1}{4}\sum_{t\geq 0}\left(\frac{z}{2}\right)^t + \frac{4}{9}\sum_{t\geq 0} z^t,$$

откуда и следует (4).

Для определения $\mathrm{E}\mu(s,T)$, $\mathrm{D}\lambda(s,T)$ и $\mathrm{D}\mu(s,T)$ будем действовать аналогичным образом. Из формул

$$\mathrm{E}\mu(s,T) = \frac{1}{2}[z^T]\frac{\partial G_{A^*}(x,1/2,z)}{\partial x}\bigg|_{x=1/2},$$

$$\mathrm{D}\lambda(s,T) = \frac{1}{4}[z^T]\frac{\partial^2 G_{A^*}(1/2,y,z)}{\partial y^2}\bigg|_{y=1/2} + \mathrm{E}\lambda(s,T) - \left(\mathrm{E}\lambda(s,T)\right)^2,$$

$$\mathrm{D}\mu(s,T) = \frac{1}{4}[z^T]\frac{\partial^2 G_{A^*}(x,1/2,z)}{\partial x^2}\bigg|_{x=1/2} + \mathrm{E}\mu(s,T) - \left(\mathrm{E}\mu(s,T)\right)^2,$$

после несложных вычислений получаем (5)–(6). Точные выражения для дисперсий в теореме не приводятся ввиду громоздкости.

1. Гульден Я., Джексон Д. Перечислительная комбинаторика. М.: Наука, 1990.

2. Choffrut Ch., Karhumaki J. Combinatorics of Words, in Handbook of Formal Languages (Rozenberg G., Saloma A. eds.). Berlin, Heidelberg: Springer-Verlag, 1997.

3. Chambers W., Gollmann D. Generators for sequences with nearmaximal linear eguivalence // IEE Proc. E., Vol. 135. 1998. P. 67–69.

4. Rueppel R. When shift registers clock themselves // Advances in Cryptology: Proceedings of Eurocrypt 87, LNCS 304. 1988. P. 53–64.

5. Flajolet P., Odlyzko A. Random mapping statistics // Advances in Cryptology: Proceedings of Eurocrypt 89, LNCS 434. 1990. P. 329–354.