УДК 004.052

С.Ю. Мельников

Статистические свойства неавтономных обобщенных двоичных регистров сдвига

Исследуются двоичные обобщенные регистры сдвига, включающие в себя обычные регистры сдвига. Получено выражение для вероятностной функции, описывающей предел относительной частоты единиц в выходной последовательности регистра при случайном бернуллиевском входе. Показано, что обобщенные регистры, вообще говоря, не обладают свойством чезарово-наследственности, которым обладают обычные регистры сдвига.
Ключевые слова: регистр сдвига, граф де Брейна, вероятностная функция.
doi: 10.21293/1818-0442-2017-20-1-93-95

При построении генераторов случайных последовательностей [1] широко используются как линейные, так и нелинейные регистры сдвига, с теми или иными элементами усложнения или обратной связи [2]. Это во многом обусловлено совокупностью «хороших» комбинаторных и структурных свойств графов де Брейна, описывающих преобразования информации в таких регистрах.

В 80-е гг. прошлого века для конструирования сетевых архитектур с минимальными временными задержками были предложены [3] так называемые обобщенные графы де Брейна, которые с тех пор активно изучаются, в частности, при организации пиринговых сетей [4], в криптографических приложениях [5], при разработке методов сборки генома [6] и других областях.

В [7] определены обобщенные регистры сдвига (ОРС), графами переходов которых являются обобщенные графы де Брейна. Двоичный ОРС порядка m, $m=1,2,...$ – это автомат Мура $A_f^{(m)}=(X,Y,Q,h,f)$, где входной и выходной алфавиты есть $X=Y=\{0,1\}$, множество состояний $Q=\{0,1,...,m-1\}$, функция переходов определена правилом $h(q,\varepsilon)=(2q+\varepsilon)\bmod m$, $q\in Q$, $\varepsilon=0,1$, функция выходов есть некоторое отображение $f:Q\to\{0,1\}$. При $m=2^t$ двоичный ОРС является обычным двоичным проходным регистром сдвига с накопителем размера t.

В работе сравниваются значковые статистические свойства обобщенных и обычных регистров сдвига.

Вероятностная функция обобщенного регистра сдвига

Граф переходов ОРС обозначим G_m. Граф G_m является ориентированным графом на m вершинах с дугами, ведущими из вершины из i в вершину $2i+\varepsilon\pmod m$, $0\le i\le m-1$, $0\le\varepsilon\le 1$. Такой граф имеет $2m$ дуг, является сильносвязным и регулярным степени 2. При $m=2^t$, $t=1,2,...$ рассматриваемый граф является классическим двоичным графом де Брейна степени t.

Рассмотрим случайное блуждание на графе G_m, при котором начальная вершина выбирается случайно из множества вершин графа, а шаг блуждания проходит по исходящим из нее дугам. Предположим, что переходы из вершины в вершину независимы и вероятность шага блуждания из вершины i в вершину $2i+1\pmod m$ равна p, в вершину $2i\pmod m$ равна $1-p$, $0<p<1$. Задачи, связанные с изучением случайных блужданий на классических графах де Брейна, рассматривались в [8].

Пусть $m=s2^k$, s – нечетно, $k\ge 0$. Для $0\le q\le m-1$ обозначим $b(q)$ – количество единиц в двоичной записи числа $q\bmod 2^k$.

Утверждение 1 ([9]). Стационарное распределение описанного случайного блуждания на вершинах графа G_m имеет вид

$$P(q)=\frac{1}{s}p^{b(q)}(1-p)^{t-b(q)},\ 0\le q\le m-1. \qquad (1)$$

Следствие. Если m – нечетно, то стационарное распределение на вершинах графа является равномерным и не зависит от p:

$$P(q)=\frac{1}{m},\ 0\le q\le m-1.$$

Пусть на вход автомата $A_f^{(m)}$, который находился в начальном состоянии q_0, поступает бернуллиевская последовательность случайных величин с параметром p, $0<p<1$. Как нетрудно видеть, последовательность состояний автомата образует эргодическую цепь Маркова. Согласно закону больших чисел для цепей Маркова существует предел относительной частоты встречаемости знака «1» в растущих начальных отрезках выходной последовательности автомата. Этот предел описывается вероятностной функцией автомата [10], которую мы обозначим $P_A(p)$. Используя формулу полной вероятности и соотношение (1), нетрудно получить следующее

Утверждение 2. $P_A(p)=\frac{1}{s}\sum_{i=0}^{k}\|f/S_i\|\,p^i(1-p)^{k-i}$, где $S_i=\{q\,|\,b(q)=i\}$, $0\le i\le k$.

Хорошо известный результат о виде вероятностной функции обычного регистра сдвига с накопителем размера k (см., например, [10]) $P_A(p)=$ $=\sum_{i=0}^{k}\|f/S_i\|p^i(1-p)^{k-i}$, где $S_i=\{(\varepsilon_1\varepsilon_2...\varepsilon_k),\ \varepsilon_j=0,1|\sum\varepsilon_j=i\}$, $\|f/S_i\|=\sum_{(\varepsilon_1\varepsilon_2...\varepsilon_k)\in S_i} f(\varepsilon_1\varepsilon_2...\varepsilon_k)$, следует из доказанной формулы при $s=1$.

Чезарово-наследственность обобщенного регистра сдвига

Двоичное слово называется [11] чезаровским для бесконечной последовательности, если существует предел относительной частоты его встречаемости в растущих начальных отрезках. Последовательность называется чезаровской, если произвольное двоичное слово является для нее чезаровским. Конечный автомат называется чезарово-наследственным, если из любого начального состояния он чезаровские последовательности перерабатывает в чезаровские.

Оказывается, ОРС, в отличие от обычных регистров, не являются чезарово-наследственными.

Утверждение 3. Если $m=2^t$, $t\geq 0$, то при любой функции выходов f автомат $A_f^{(m)}$ является чезарово-наследственным.

Если $m\neq 2^t$, то найдется функция выходов f, для которой автомат $A_f^{(m)}$ не является чезарово-наследственным.

Доказательство. Пусть $m=2^t$. ОРС является обычным проходным регистром сдвига с накопителем размера $t+1$ и его чезарово-наследственность следует из того, что он является автоматом с конечным запоминанием.

Пусть $m\neq 2^t$. Тогда $m=2^k s$, $k\geq 1$, $s\geq 3$ – нечетно. Докажем, что в графе ОРС найдутся по крайней мере два цикла c_1 и c_2, входная разметка которых состоит только из нулей.

Последовательность состояний регистра

$$q_0 \xrightarrow{0} q_1 \xrightarrow{0} q_2 \xrightarrow{0} \ldots \xrightarrow{0} q_l = q_0$$

является циклом длины l в графе ОРС с входной разметкой, состоящей только из нулей, в том случае, если выполнены соотношения

$$q_i = 2^i q_0 \bmod 2m,\ i=1,...,l-1,$$

$$q_0 = 2^l q_0 \bmod 2m.$$

Отсюда $q_0 = 2^l q_0 \bmod 2^{k+1}s$, и тогда

$$\begin{cases} q_0 = 2^l q_0 \bmod 2^{k+1}, \\ q_0 = 2^l q_0 \bmod s. \end{cases}$$

В качестве q_0, удовлетворяющего приведенным соотношениям, можно взять 2^{k+1}. Таким образом, в качестве цикла c_1 выступает следующая последовательность вершин графа:

$$c_1 = \left(2^{k+1} \bmod 2m, 2^{k+2} \bmod 2m, ..., 2^{k+l} \bmod 2m\right).$$

Как нетрудно убедиться, l здесь равно порядку элемента «2» в мультипликативной группе кольца вычетов по модулю s.

В качестве цикла c_2 можно взять петлю в нулевой вершине, $c_2=(0)$.

Итак, мы указали два различных цикла в графе переходов автомата, входная разметка которых состоит из нулей. Определим функцию f выходов так, чтобы на состояниях цикла c_1 она принимала только нулевые значения, а на состояниях цикла c_2 (т.е. в состоянии 0) – единичное. На остальных состояниях функцию f можно задать произвольно.

Обозначим для удобства состояния цикла c_i через $q_1^{(i)},...,q_{l_i}^{(i)}$, l_i – длина цикла c_i, $i=1,2$. Пусть $\chi^{(i)}=0^{l_i}$ – входная последовательность, под действием которой $A_f^{(m)}$ последовательно проходят состояния цикла c_i, начиная с $q_1^{(i)}$, $i=1,2$. Через ξ_{12} (ξ_{21}) обозначим кратчайшую последовательность, переводящую рассматриваемый ОРС из состояния $q_1^{(1)}$ в состояние $q_1^{(2)}$ (из состояния $q_1^{(2)}$ в состояние $q_1^{(1)}$) (рис. 1).

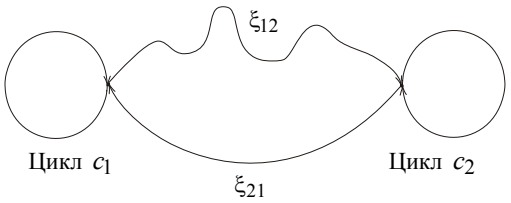

Рис. 1. Структура последовательности χ

Определим бесконечную двоичную последовательность

$$\chi=\left(\chi^{(1)}\right)^{k_1}\wedge\xi_{12}\wedge\left(\chi^{(2)}\right)^{k_2}\wedge\xi_{21}\wedge\left(\chi^{(1)}\right)^{k_3}\wedge\xi_{12}\wedge\left(\chi^{(2)}\right)^{k_4}\ldots,$$

где символ $\wedge$ означает конкатенацию последовательностей. При $k_i=2^{2^i}$ последовательность χ является чезаровской, поскольку, как нетрудно видеть, предел относительной частоты встречаемости в ней любого слова, отличного от серии нулей, существует и равен нулю. Для слов вида 0^j, $j=1,2,...$ такие пределы равны единице. Пусть $Ext_{(A,q_0)}(\chi)$ – выходная последовательность ОРС $A_f^{(m)}$. Рассмотрим ее начальный отрезок длины $k_1+|\xi_{12}|+k_2+|\xi_{21}|+...+k_N$. Поскольку $|\xi_{12}|$ и $|\xi_{21}|$ не превосходят диаметра графа ОРС, то, как несложно показать, относитель-

ная частота встречаемости единиц на этом отрезке равна $0+O(2^{-2^{N-1}})$ при нечетном N и $1+O(2^{-2^{N-1}})$ при четном N. Это означает, что у последовательности относительных частот единиц в растущих начальных отрезках не существует предела и, следовательно, последовательность $Ext_{(A,q_0)}(\chi)$ не является чезаровской, что завершает доказательство.

Заключение

Регистры сдвига широко используются при разработке генераторов случайных последовательностей. В работе исследовано семейство двоичных обобщенных регистров сдвига, которое включает в себя обычные регистры сдвига. Описаны значковые статистические свойства обобщенных регистров при обработке ими случайных бернуллиевских последовательностей. Показано, что обобщенные регистры, вообще говоря, не обладают свойством чезарово-наследственности, которым обладают обычные регистры сдвига.

Литература

1. Вильданов Р.Р. Тесты псевдослучайных последовательностей и реализующее их программное средство / Р.Р. Вильданов, Р.В. Мещеряков, А.А. Шелупанов, С.С. Бондарчук // Доклады ТУСУРа. – 2012. – № 2 (25). – С. 108–111.

2. Golomb S.W. Shift Register Sequences. – USA, Calif., Laguna Hills: Aegean Park Press, 1981. – 247 p.

3. Imase M., Itoh M. Design to minimize diameter on building-block network // IEEE Trans. Comput. – 1981. – Vol. 30. – PP. 439–442.

4. Graph-theoretic analysis of structured peer-to-peer systems: Routing distances and fault resilience / D. Loguinov, A. Kumar, V. Rai, S. Ganesh // ACM SIGCOMM. – 2003. – PP. 395–406.

5. Maurer U.M. Asymptotically-tight bounds on the number of cycles in generalized de Bruijn-Good graphs // Discrete applied mathematics. – 1992. – Vol. 37. – PP. 421–436.

6. Compeau P. How to apply de Bruijn graphs to genome assembly / P. Compeau, P. Pevzner, G. Tesler // Nature Biotechnology. – 2011. – Vol. 29. – PP. 987–991.

7. Максимовский А.Ю., Мельников С.Ю. О числе обобщенных в смысле Imase и Itoh регистров сдвига, устанавливаемых постоянным входом в фиксированное состояние // Обозрение прикладной и промышленной математики. – 2015. – Т. 22, вып. 5. – http://tvp.ru/conferen/vsppm16/chelso144.pdf

8. Мори Т. Случайные блуждания на графах де Брейна // Теория вероятностей и ее применения. – 1992. – Т. 37, вып. 1. – С. 194–197.

9. Melnikov S.Yu. Stationary distribution of random walk on the generalized de Bruijn digraphs // XXXIV International Seminar on Stability Problems for Stochastic Models, Svetlogorsk, June 12–18, 2016. – http://tvp.ru/conferen/bisaimII/kisvd053.pdf

10. Мельников С.Ю. О задаче определения функции выходов автомата со случайным входом по статистике встречаемости слова в выходной последовательности // Доклады ТУСУРа. – 2011. – № 1 (23). – С. 107–123.

11. Мельников С.Ю. О переработке конечными автоматами чезаровских последовательностей // Лесной вестник Моск. гос. ун-т леса. – 2004. – № 1 (32). – С. 169–174.

Мельников Сергей Юрьевич
Канд. физ.-мат. наук, зам. ген. директора
ООО «Лингвистические и информационные технологии», Москва
Тел.: 8 (495) 249-90-53
Эл. почта: melnikov@linfotech.ru

Melnikov S.Yu.

Statistical properties of generalized binary shift registers

The generalized shift registers are the finite state machines with next state function defined as the generalized de Bruijn digraph. The probability function describing the limit of the relative frequency of «1» in the output sequence of the register with Bernoulli input is obtained.

Keywords: shift register, generalized de Bruijn graph, random input.