СТАТИСТИЧЕСКИЕ ХАРАКТЕРИСТИКИ ПСЕВДОСЛУЧАЙНЫХ СИГНАЛОВ, ФОРМИРУЕМЫХ ДИСКРЕТНЫМИ АЛГОРИТМАМИ С ЗАПАЗДЫВАНИЕМ

Беляев Р.В., Воронцов Г.М., Кислов В.Я., Колесов В.В., Попов А.М., Рябенков В.И.

Институт радиотехники и электроники РАН, Москва

Аннотация. Численным моделированием показано, что статистические свойства формируемых дискретных последовательностей, близкие к случайному процессу, обеспечивают такие порождающие кодирующие алгоритмы, у которых как одномерное распределение вероятностей, так и распределения условных вероятностей генерируемых чисел близки к равномерному.

Наиболее полную информацию о статистических свойствах дискретных последовательностей дает анализ распределений попарных условных вероятностей членов последовательности $p(i+j, x_n|i, x_k)$, j=1, 2, 3… k, l=1, 2, 3,…M, т.е. вероятности генерации числа x_n на (i+j)- том шаге алгоритма, если на i- том шаге было получено число x_k. При этом областью определения дискретного алгоритма является произвольный замкнутый целочисленный интервал $[M_1, M_2]$, $M=M_2-M_1+1$, $x_n \in [M_1, M_2]$.

Если распределение условных вероятностей при любом j практически совпадает с равномерным распределением, то отсюда следует, что все вероятности перехода $p(i+j, x_n|i, x_k) \approx 1/M$, j=1, 2, 3… при произвольном выборе i. В то же время, если распределение вероятностей генерируемых чисел $p(x_n)$ близко к равномерному, то вероятность значения x_n практически также равна 1/M. Тем самым вероятности перехода в состояние x_n на j-том шаге совпадают с вероятностью этого значения на этом шаге независимо от значений последовательности на предыдущих шагах алгоритма, что характерно для случайных последовательностей при независимых испытаниях. Более того, формируемая таким алгоритмом псевдослучайная последовательность по своим вероятностным характеристикам будет близка к последовательности независимых равновероятных чисел из интервала $[M_1, M_2]$ [1]. В последнем случае можно ожидать, что данная последовательность будет обладать наилучшими статистическими свойствами. Установление подобного факта подчеркивает важность исследования распределений условных вероятностей для априорного суждения о качестве формируемых псевдослучайных последовательностей.

Для характеристики условных распределений $p(x_{i+j}|x_i)$ большое значение имеет вид расположения точек (x_{i+j}, x_i) на плоскости для отображения $x_{i+j}=\text{func}(x_i)$, задаваемого дискретным алгоритмом, при соответствующих значениях j=1, 2, 3,…и i=1, 2, 3,…N [2, 3]. Получение разброса точек (x_{i+j}, x_i) и визуализация на экране не требует большого машинного времени по сравнению с непосредственным вычислением условных вероятностей, и хотя характер этого разброса не дает непосредственно формы распределения условных вероятностей, тем не менее визуализация разброса свидетельствует о степени регулярности этих распределений, наличии функциональных связей, существовании запретных переходов, а то и целых запретных зон, что неизбежно сказывается на корреляционных и других статистических свойствах последовательности.

Рассмотрим простейшую формулу алгоритма с запаздыванием: $x_n=x_{n-1}+x_{n-Nz}$, где Nz- параметр запаздывания, дополненную операцией возврата в интервал $[M_1, M_2]$ в случае выброса вновь полученного значения x_n за его пределы. В данном алгоритме перемешивание, хаотизация формируемого процесса осуществляется добавлением вообще говоря случайного значения запаздывающего члена последовательности и, операцией клипирования получаемой суммы чисел на границе области определения M_2.

Для численного моделирования были приняты следующие значения параметров алгоритма: $M_1=1$, $M_2=255$, Nz=16. Анализируемый алгоритм при выбранных значениях параметров и заданных начальных условиях, представляющих собой набор из 16-ти случайных чисел (вектор запаздывания), формируют псевдослучайную последовательность с распределением вероятностей, близким к равномерному распределению p(x)=1/M. Отличие от равномерного распределения характеризуется суммарным расхождением наблюдаемых в численном эксперименте частот появления в генерируемой последовательности чисел x_m и величины 1/M: $\Delta p_{\text{Ó}} = \sum_{m=1}^{M} \left| n(x_m)/N - 1/M \right|$, где $n(x_m)$- количество чисел x_m в последовательности из N членов. Это суммарное расхождение численно совпадает с относительным средним отличием от равномерного закона $\Delta p_{\text{ср.отн.}}$. Наибольшее относительное отличие частоты выпадения наблюдаемых чисел от значения 1/M: $\Delta p_{\text{макс.отн}} = [1/(1/M)] \cdot \left| n(x_m)/N - 1/M \right|_{\text{макс}}$.

Полученные оценки распределений вероятностей и распределений условных вероятностей для последовательностей, формируемых анализируемыми в данной работе алгоритмами в сопоставлении со стандартным генератором RND, сгруппированы в таблице:

	p(x)		$P(X_{l+J} \mid X_l)$		
	N=210 000		N=52 000 000, k=6		
алгоритм	$\Delta p_{ср.отн.}$	$\Delta p_{макс.отн}$	j	$\Delta p_{ср.отн.}$	$\Delta p_{макс.отн}$
1	0.03	0.10	1	0.03	0.12
			16	0.031	0.12
2	0.03	0.10	1	0.997	1.12
			2	0.50	1.04
			16	0.03	0.10
RND	0.03	0.15	–	–	–

Двумерное распределение пар точек (x_{i+j}, x_i), где i=1,2,…N представляет собой равномерно заполненную область хаотически разбросанных точек, при любом j=1, 2, 3,…16,…32. Такой характер равномерного, полного и случайного заполнения точками (x_{i+j}, x_i) численного интервала [M_1, M_2] свидетельствует о хаотизации исследуемого процесса.

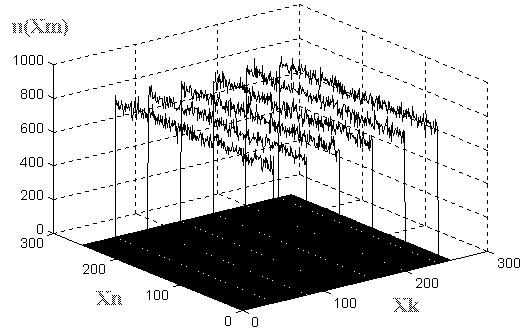

Рис. 1.

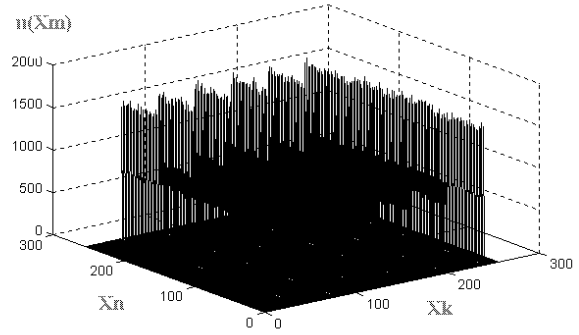

Рис.2.

Более точные сведения о вероятностях перехода к каждому из значений x_n дают построения распределений условных вероятностей. Для наглядного представления на рис. 1 построены гистограммы частот переходов только для 6-ти значений генерируемых чисел x_k =(k-1)·50+1, k=1, 2,…6 с шагом j=1. Гистограммы частот рассчитаны по реализации последовательности из 5,2 10^7 членов. Значения $\Delta p_{ср,отн}$ и $\Delta p_{макс,отн}$, характеризующие отличия от равномерного распределения одной из гистограмм (x_k=251) для j=1 и 16, приведены в таблице. Из полученных данных, следует, что для алгоритма №1 распределения условных вероятностей при любом j практически совпадают с равномерным распределением.

После соответствующего изменения алгоритма, при котором одномерное распределение остается равномерным, а функции условных распределений вероятностей $p(x_{i+1} \mid x_i)$=0 (j=1) через одно значение (рис.2) в зависимости от четности числа x_i на предыдущем шаге. Картина распределения точек на плоскости в этом случае вместо случайного носит регулярный характер. Это подтверждается зафиксированными количественными отличиями $\Delta p_{ср.отн}$ и $\Delta p_{макс.отн}$ от равномерного распределения (см. таблицу).

Оценка корреляционных характеристик последовательностей, сформированных алгоритмами № 1 и 2, проводилась на основе анализа клипированных и неклипированных 100 сегментов размером N=128 и 1024 символа, последовательно генерируемых алгоритмами без какого-либо отбора. Несмотря на существенные различия в форме распределений $p(x_{i+j} \mid x_i)$ величина боковых выбросов АКФ и ВКФ относительно уровня $1/\sqrt{N}$ для обоих алгоритмов примерно одинакова и составляет: (1.3÷3.8) для АКФ N=128, (1.5÷4.3) для ВКФ N=128, (2.2÷4.8) для АКФ и ВКФ N=1024. Эти данные свидетельствует о том, что рассмотренные выше особенности алгоритмов мало сказались на корреляционных свойствах как самих последовательностей, так и результатов их клипирования. В то же время численный эксперимент по блочной структуры на длине клипированной последовательности из 2,7 10^5 членов показал, что если алгоритм №1 генерирует последовательности с блоковой структурой, близкой к закону $p(k)=1/2^k$ вплоть до блоков размером k=17-18, то блоковая структура последовательностей, порождаемых алгоритмом №2, существенно

отклоняется от этого закона (рис.3), что непосредственно связано с неравномерностью распределений вероятностей переходов даже на одном шаге j=1.

Оценка объема системы сигналов, формируемых алгоритмами, производилась путем отбора из сформированной клипированной последовательности сбалансированных кодов длиной из 128, 256 и 512 символов со следующими корреляционными свойствами: боковые выбросы апериодической автокорреляционной функции не превышают значения $R_{макс}=2.26/\sqrt{N_{код}}$, а выбросы апериодических взаимокорреляционных функций по всему массиву отбираемых кодов меньше или равны $R_{макс}=3.39/\sqrt{N_{код}}$. Корреляционные функции вычислялись по формуле для сбалансированных последовательностей [4]. Как показал численный эксперимент неравномерность условных распределений и существование запретных переходов сказалась на скорости отбора кодов в систему сигналов систему по мере увеличения длительности последовательности N (рис.4). На этом рисунке все кривые, соответствующие алгоритму №2, проходят существенно ниже, чем у алгоритма №1.

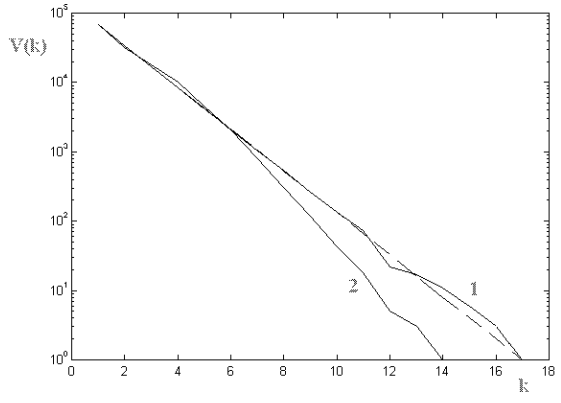

Рис.3.

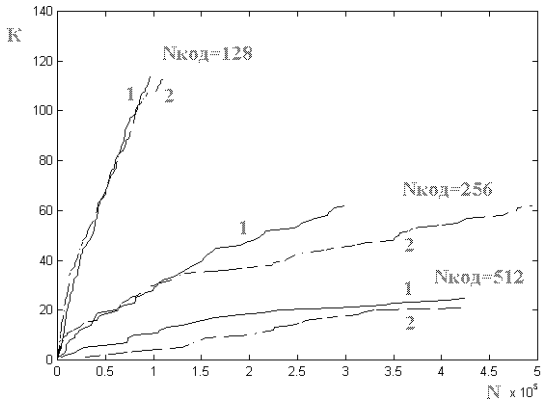

Рис.4.

Работа выполнена при финансовой поддержке РФФИ (проекты 00-0790147, 01-0790349).

Литература

1. Рытов С.М. Введение в статистическую радиофизику. М.: Наука. 1966. 404 С.
2. Жельников В. Криптография от папируса до компьютера- М.: ABF, 1996, 336 с.
3. Каханер Д., Моулер К., Нэш С. Численные методы и программное обеспечение. М.: Мир, 2001, 575с.
4. Варакин Л.Е. // Радиотехника и электроника. 1978. Т.33. №4. С.735.

STATISTICAL CHARACTERISTICS OF PSEUDORANDOM SIGNALS FORMED BY DISCRETE ALGORITHM WITH DELAY

Belyaev R., Vorontsov G., Kislov V., Kolesov V. Popov A., Ryabencov V.

Institute of Radioengineering and Electronics RAS, Moscow

The most complete information about statistical characteristics of discrete sequences are given by analysis of conditional probability distribution for pairs of sequences' terms. The great importance for characterization of the conditional distribution $p(x_{i+j}|x_i)$ is the locations point disposition (x_{i+j}, x_i) on planes for map function $x_{i+j}=\text{func}(x_i)$, assigned by discrete algorithm at corresponding values j=1, 2, 3, and i=1, 2, 3 N. The calculation of the points scattering of (x_{i+j}, x_i) and visualization on the screen does not require the big machine time in contrast with direct calculation of conditional probability. Though nature of this scattering does not give directly forms of the distribution of conditional probability, however visualization of the scattering indicates the degree of regularity these distribution, presence of the functional links, the existence of forbidden transition, that inevitably affects onto correlative and other statistical characteristic of sequences.

These characteristics have been studied for discrete sequences formed by two generalized Fibonaccy types algorithms. It was supposed that these algorithms were defined on arbitrary finite interval M of integer numbers of the natural sequence. These algorithms were expanded by adding a conditions of return back in the defined interval M if it happened that the result of calculations would exit out of these interval. This is a so-called mechanism of "mixing" . There were used also for comparison a standard RND-algorithm.

By computational modeling it have been shown that the statistical characteristics of formed discrete sequences were nearly to random processes. The generic algorithms forming such coding sequences provided that last one's have uniform as well as one-dimensional function of probability distribution and also a function of conditional probability distribution.