УДК 004.415.53:004.421.5

Э.В. Фауре, А.И. Щерба, А.А. Лавданский

Черкасский государственный технологический университет, Черкассы

АНАЛИЗ КОРРЕЛЯЦИОННЫХ СВОЙСТВ ПОСЛЕДОВАТЕЛЬНОСТЕЙ (ПСЕВДО) СЛУЧАЙНЫХ ЧИСЕЛ

В работе выполнен сравнительный анализ автокорреляционных свойств последовательностей, порожденных комбинационным генератором с различными его параметрами, генератором типа «Вихрь Мерсенна», и последовательностей «истинно случайных» чисел. Результаты, полученные в работе, подтверждают однородность оценок автокорреляционных функций последовательностей псевдослучайных чисел комбинационного генератора и генератора типа «Вихрь Мерсенна» с оценкой автокорреляционной функции случайной последовательности чисел, а также позволяют использовать комбинационный генератор в задачах, требующих некоррелированных последовательностей чисел.

Ключевые слова: *последовательность (псевдо) случайных чисел, комбинационный генератор, корреляция, автокорреляционная функция, статистический критерий.*

Введение

Методы и средства синтеза псевдослучайных чисел (ПСЧ) широко используются для решения большого круга практических задач. При этом результат решения задачи значительно зависит от свойств используемого генератора. Качественный генератор ПСЧ (ГПСЧ) должен формировать последовательность чисел, статистические свойства которой близки к статистическим свойствам последовательности случайных чисел, порождаемой естественным источником дискретного белого шума (генератором случайных чисел – ГСЧ). Дискретным белым шумом является стационарный дискретный случайный процесс, отсчеты которого некоррелированы друг с другом [1, с. 235], т.е. белый шум является дельта-коррелированным стационарным случайным процессом [2, с. 170].

Общеизвестно, что детерминированные методы и алгоритмы синтеза ПСЧ не способны формировать «истинно случайные» последовательности чисел, они лишь аппроксимируют их свойства. Такое обстоятельство, с одной стороны, требует разработки новых подходов и методов синтеза ПСЧ и стимулирует исследования путей улучшения статистических свойств существующих ГПСЧ, а с другой стороны – ставит задачу разработки и исследования новых методов оценивания качества статистических свойств последовательностей на выходе ГПСЧ с целью их максимального соответствия свойствам последовательностей на выходе ГСЧ.

Постановка проблемы

Синтезированные с помощью ГПСЧ последовательности не могут быть применены для решения теоретических и прикладных задач без детального исследования их статистических свойств и сравнительных количественных оценок.

Одной из важных характеристик дискретных последовательностей случайных чисел является их коэффициенты автокорреляции, составляющие автокорреляционную функцию (АКФ), которая позволяет оценить уровень внутренней корреляции исследуемой последовательности в различные моменты времени. Поэтому важной и актуальной проблемой является объективный анализ полученной оценки АКФ псевдослучайной последовательности чисел и определение степени ее соответствия АКФ случайной последовательности чисел, которая математически описывается дельта-функцией Дирака [3, с. 84]. Например, общепринятый в эконометрике [4, с. 302-303] анализ оценки АКФ проводится таким образом:

1) если наиболее высоким и значимым является коэффициент автокорреляции первого порядка, исследуемая последовательность содержит линейную тенденцию;

2) если наиболее высоким и значимым является коэффициент автокорреляции порядка τ, исследуемая последовательность содержит циклические колебания с периодичностью в τ моментов времени;

3) если ни один из коэффициентов автокорреляции не является значимым, можно сделать одно из предположений относительно структуры исследуемой последовательности:

- последовательность не содержит линейной тенденции и циклических колебаний, а включает только случайную компоненту;
- последовательность содержит сильную нелинейную тенденцию.

Для нормально распределенной совокупности исходных данных значимость коэффициентов автокорреляции может быть определена в соответствии с *t*-критерием Стьюдента [5, с. 26].

Кроме того, актуальным вопросом является выявление и исследование новых статистических

свойств автокорреляционных функций последовательностей случайных и псевдослучайных чисел и их сравнение на основе статистических критериев. Особый интерес представляет распределение коэффициентов корреляции ненулевого порядка, а также распределение знаков этих коэффициентов.

Постановка задачи

Задачей данной работы является:

– исследование и анализ корреляционных связей случайных и псевдослучайных последовательностей чисел с помощью оценок коэффициентов автокорреляции и автокорреляционной функции. Определение значимости оценок коэффициентов автокорреляции. В качестве источников ПСЧ в данном исследовании используются генератор типа «Вихрь Мерсенна» [6] и комбинационный генератор с комбинирующей функцией суммирования в некотором конечном поле [7];

– исследование и анализ распределения боковых лепестков оценки автокорреляционной функции случайной последовательности чисел и статистическая проверка гипотез соответствия этому распределению боковых лепестков оценок автокорреляционных функций псевдослучайных последовательностей чисел. Боковыми лепестками АКФ будем называть точки на графике АКФ (коррелограмме), соответствующие коэффициентам корреляции ненулевого порядка;

– исследование распределения знаков боковых лепестков оценок АКФ случайных и псевдослучайных последовательностей чисел с помощью статистических критериев, а также исследование распределения k-грамм для знаков оценок АКФ.

Решение задачи

Вычисление АКФ является трудоемким алгоритмом и напрямую зависит от размера исследуемой выборки. Для уменьшения времени вычисления АКФ целесообразно рассматривать небольшие случайно выделенные участки выборки для общей оценки АКФ. Размеры таких участков зависят от расстояния, на котором нужно выявить корреляцию символов в исследуемой последовательности.

Различают два типа АКФ по способу ее построения: битовая и символьная. Для построения битовой АКФ используется двоичное представление исследуемой последовательности, для символьной – представление в системе счисления, равной мощности алфавита исследуемого генератора. Следовательно, битовая АКФ отображает межбитовую корреляцию, символьная – межсимвольную. Для построения битовой АКФ дополнительно производится нормирование двоичной последовательности.

В данной работе оценку корреляционных свойств последовательности символов – (псевдо) случайных чисел – будем проводить по некоторой фиксированной выборке объемом n. Обозначим элемент последовательности случайных чисел в дискретный момент времени «t» в виде x_t, $t \in [0, n-1]$. Оценку нормированной АКФ для последовательности случайных чисел будем рассчитывать таким образом [8, с. 460], [9, с. 402]:

$$r_x(\tau) = \frac{\sum_{i=0}^{n-1-\tau} \left[\left(x_i - \bar{x} \right) \cdot \left(x_{i+\tau} - \bar{x} \right) \right]}{\sqrt{\sum_{i=0}^{n-1-\tau} \left(x_i - \bar{x} \right)^2 \cdot \sum_{i=0}^{n-1-\tau} \left(x_{i+\tau} - \bar{x} \right)^2}}, \quad (1)$$

где $\bar{x} = \frac{1}{n} \sum_{i=0}^{n-1} x_i$ – статистическая оценка математического ожидания.

Очевидно, что $r_x(0) = 1$, а при $\bar{x} = 0$

$$r_x(\tau) = \sum_{i=0}^{n-1-\tau} \left(x_i \cdot x_{i+\tau} \right) \Bigg/ \sqrt{\sum_{i=0}^{n-1-\tau} x_i^2 \cdot \sum_{i=0}^{n-1-\tau} x_{i+\tau}^2} \, .$$

Выполним построение графиков оценок нормированных АКФ для различных последовательностей:

– случайной последовательности чисел (оцифрованных радиошумов [10] и квантового ГСЧ [11]);

– последовательности чисел на выходе генератора типа «Вихрь Мерсенна» [6] MT19937;

– последовательности на выходе комбинационного генератора [7] с различными его параметрами.

Последовательности псевдослучайных чисел комбинационного генератора [7] формируются путем применения комбинирующей функции суммы по модулю M к нескольким последовательностям, порожденным группой независимых первичных генераторов, где M – требуемая мощность алфавита. Первичные генераторы представляют собой циклические сдвиговые регистры, в которые записаны перестановки на множествах с мощностями алфавитов M_i. В данной работе рассматривается комбинационный генератор с различным количеством исходных циклических сдвиговых регистров (от 4 до 8) и различным их заполнением: с помощью линейного конгруэнтного метода [12, с. 275-277], аддитивного генератора [13] и генератора случайных чисел (в данной работе используется квантовый ГСЧ [11]).

Для всех генераторов объем выборки составляет $n = 2^{16}$ значений, а мощность алфавита M выбрана равной 256, что позволяет без дополнительных преобразований формировать бинарный файл на выходе генератора.

Графики оценок нормированных АКФ по (1) для $\tau \in [1, n/4]$ для некоторых из указанных последовательностей представлены на рис. 1.

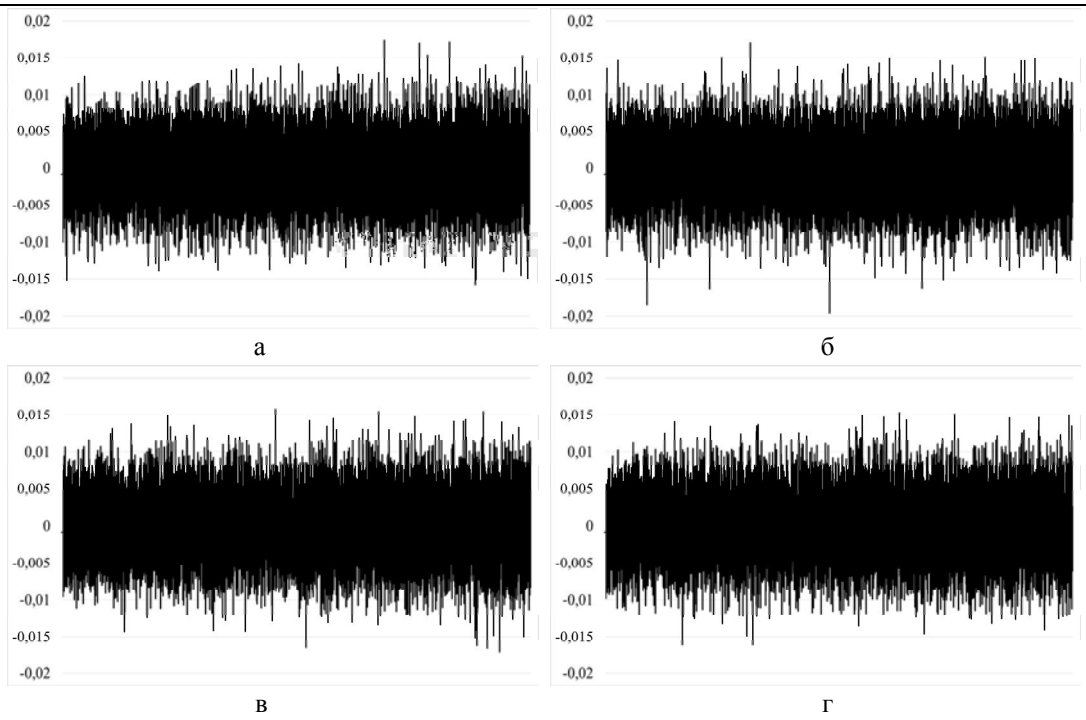

Рис. 1. Оценки нормированных АКФ различных последовательностей:
а – оцифрованных радиошумов; б – вихря Мерсенна; в – комбинационного генератора с использованием четырех циклических регистров сдвига и их заполнением с помощью линейного конгруэнтного метода; г – комбинационного генератора с использованием восьми циклических регистров сдвига и их заполнением с помощью квантового генератора случайных чисел

Все графики на рис. 1 не имеют ярко выраженных корреляционных всплесков, что может являться грубой оценкой отсутствия линейной корреляции внутри исследуемых последовательностей.

Несмотря на то, что автокорреляционная функция относится к классу графических тестов [14], анализ ее численных значений представляет собой значительный интерес.

Значимость оценок коэффициентов автокорреляции

В основе классического анализа значимости оценок коэффициентов автокорреляции $\left(r_x(\tau)=0\right)$ в соответствии с *t*-критерием Стьюдента [5, с. 26] лежит предположение о принадлежности исследуемых случайных последовательностей многомерному нормальному закону распределения. Однако, как показано в [15, 16], распределения статистик критериев, используемых при проверке гипотезы о равенстве нулю для парных, частных и множественных коэффициентов корреляции (в том числе и *t*-критерия Стьюдента), устойчивы к отклонениям наблюдаемого многомерного закона от нормального. Эмпирические распределения данных статистик хорошо описываются предельными законами, полученными в предположении о нормальности наблюдаемых величин.

Таким образом, несмотря на то, что случайные вектора $\left(x_i, x_{i+\tau}\right)$ не принадлежат двумерному нормальному распределению, проверка значимости полученных оценок коэффициентов автокорреляции в соответствии с *t*-критерием Стьюдента является правомерной. В силу того, что анализу подвергаются 2^{14} значений оценок нормированных коэффициентов автокорреляции для каждого источника, результатом проверки будут являться частости попадания статистики критерия в критическую область при уровне значимости 0,05. Полученные данные сведены в табл. 1.

В соответствии с теоремой Бернулли [8, с. 149], уклонение частости попадания статистики *t*-критерия Стьюдента в критическую область от вероятности (заданного уровня значимости) стремится к нулю при возрастании количества измерений. Для доверительной вероятности, принятой на уровне 0.95, анализируемых 2^{14} значений эмпирических нормированных коэффициентов автокорреляции и теоретической вероятности 0,05 доверительным интервалом [8, с. 153] для частости попадания статистики *t*-критерия Стьюдента в критическую область является интервал (0,04666;0,05334), который покрывает все полученные значения частости.

Оценим, кроме того, максимальные абсолютные значения эмпирических нормированных коэффициентов автокорреляции $\max\left|r_x(\tau)\right|$, а также максимальные абсолютные значения нормированных статистик $r_x^*(\tau)=\left[r_x(\tau)-M_r(\tau)\right]/\sigma_r(\tau)$ для всех оценок АКФ.

Таблица 1

Расчетные значения статистик при анализе значимости оценок коэффициентов автокорреляции

		Заполнение первичных циклических сдвиговых регистров	Количество первичных циклических сдвиговых регистров	Частость попадания статистики критерия Стьюдента в критическую область	$\max\left\|r_x(\tau)\right\|$	$\max\left\|r_x^*(\tau)\right\|$	Частость $\left\|r_x^*(\tau)\right\| > 3$
Источник ПСЧ	Комбинационный генератор	Линейный конгруэнтный генератор	4	0,05145	0,01704	3,93863	0,00262
			6	0,04700	0,01665	3,76514	0,00183
			8	0,05109	0,01662	3,98361	0,00269
		Аддитивный генератор	4	0,05103	0,01740	4,22484	0,00275
			6	0,05066	0,01613	3,88696	0,00238
			8	0,04901	0,01601	3,95014	0,00250
		Квантовый ГСЧ	4	0,04700	0,01555	3,66197	0,00214
			6	0,04974	0,01709	3,97397	0,00232
			8	0,05255	0,01612	4,03652	0,00220
	Вихрь Мерсенна			0,05005	0,01960	4,70412	0,00342
Квантовый ГСЧ				0,05194	0,01599	3,88178	0,00262
Оцифрованные радиошумы				0,04803	0,01733	4,03954	0,00287

Применение методики, изложенной в [17], позволяет показать, что $M_r(\tau) = -(n-1)^{-1}$, а в соответствии с [9, с. 412] $\sigma_r(\tau) \approx \sqrt{1/(n-\tau)}$ для больших значений *n*. Результаты анализа представлены также в таблице 1, где кроме непосредственно максимальных значений $\max\left|r_x^*(\tau)\right|$, представлена частость превышения величиной $\left|r_x^*(\tau)\right|$ значения 3 (с целью проверки выполнения «правила трех сигм»).

Анализ максимальных абсолютных значений эмпирических нормированных коэффициентов автокорреляции свидетельствует об отсутствии значимых корреляционных всплесков для всех оценок АКФ. Практически все оценки частости превышения абсолютных значений нормированных статистик $r_x^*(\tau)$ значения 3 лежат в пределах доверительного интервала (0,00191;0,00349) с доверительной вероятностью 0,95. Следует отметить только, что данный показатель для последовательности комбинационного генератора с шестью первичными циклическими сдвиговыми регистрами и их заполнением с помощью линейного конгруэнтного метода лежит ниже нижнего предела доверительного интервала (что может свидетельствовать о меньшей выборочной дисперсии данных оценок коэффициентов автокорреляции), а для вихря Мерсенна практически равняется верхнему пределу доверительной вероятности.

Проведенное исследование распределения максимальных абсолютных значений нормированных статистик $r_x^*(\tau)$ для 1000 выборок для каждого генератора показывает их однородность для всех рассматриваемых источников, поэтому отклонения представленных в табл. 1 статистик (в частности, для вихря Мерсенна и комбинационного генератора с шестью первичными циклическими сдвиговыми регистрами и их заполнением с помощью линейного конгруэнтного метода) не следует считать значимым и закономерным.

Распределение боковых лепестков АКФ

Напомним, что боковыми лепестками АКФ будем называть точки на коррелограмме при $\tau \neq 0$.

Построим гистограмму распределения нормированных значений $r_x^*(\tau)$ боковых лепестков представленной на рис. 1 оценки нормированной АКФ случайной последовательности чисел. Аналогичные гистограммы построим и для других оценок. В силу того, что $\max\left|r_x^*(\tau)\right| < 4.8$ (см. таблицу 1), разделим диапазон значений $(-4.8, 4.8)$ на поддиапазоны с шагом $\Delta = 0,2$. Определим для каждой оценки АКФ статистику распределения амплитуды нормированных боковых лепестков по поддиапазонам. Гистограммы распределения нормированных значений боковых лепестков некоторых оценок нормированных АКФ показаны на рис. 2.

Визуальный анализ гистограмм распределения нормированных значений боковых лепестков полученных оценок нормированных АКФ показывает, что распределения схожи между собой и имеют форму нормального закона распределения.

В табл. 2 отражены результаты проверки согласия эмпирического распределения нормированных значений боковых лепестков оценки нормированной

АКФ оцифрованных радиошумов с теоретическим предельным (стандартным нормальным) по критериям Колмогорова [5, с. 80-82], Смирнова [5, с. 80-82], ω^2 Крамера-Мизеса-Смирнова [5, с.83], [8, с.277-281], Ω^2 Андерсона-Дарлинга [20], [5, с.83] и χ^2 Пирсона [8, с. 267-275].

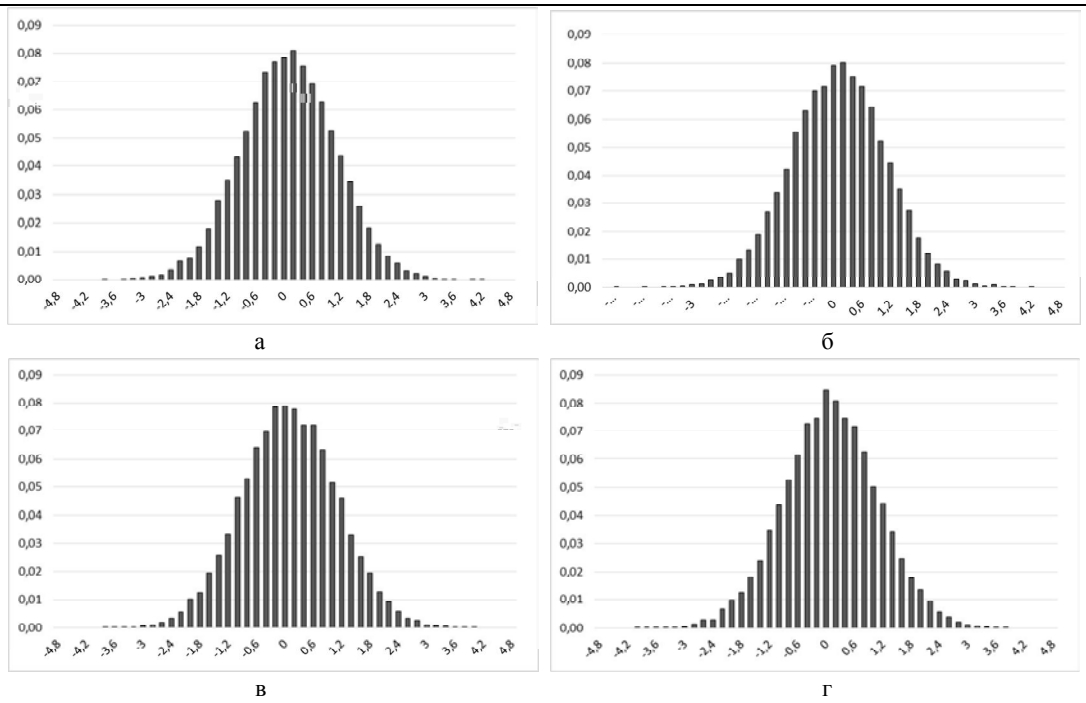

Рис. 2. Гистограммы распределения нормированных значений боковых лепестков оценок нормированных АКФ для: а – оцифрованных радиошумов; б – вихря Мерсенна; в – комбинационного генератора с использованием четырех циклических регистров сдвига и их заполнением с помощью линейного конгруэнтного метода; г – комбинационного генератора с использованием восьми циклических регистров сдвига и их заполнением с помощью квантового генератора случайных чисел

Таблица 2

Результаты проверки гипотезы нормальности распределения нормированных значений боковых лепестков оценки нормированной АКФ оцифрованных радиошумов

Критерий	Расчетное значение статистики S*	Критическая область	Достигнутый уровень значимости P(S > S*)
Колмогорова	0,7082	$S_K > 1,3581$	0,6976
Смирнова	2,0059	$S_m > 5,9915$	0,3668
ω^2 Крамера-Мизеса-Смирнова	0,0820	$n\omega^2 > 0,4614$	0,6807
Ω^2 Андерсона-Дарлинга	0,5694	$n\Omega^2 > 2,4924$	0,6773
χ^2 Пирсона	21,14	$\chi^2 > 47,40$	0,9451

По каждому из критериев приведены расчетное значение статистики, критическая область при уровне значимости $\alpha = 0,05$, а также достигнутый уровень значимости $P\left(S > S^*\right) = 1 - G\left(S|H_0\right)$, где $G\left(S|H_0\right)$ – предельное распределение статистики S соответствующего критерия согласия при справедливости проверяемой гипотезы H_0, S^* – расчетное значение статистики критерия.

Достигнутые уровни значимости статистик критериев свидетельствуют о согласии эмпирического распределения нормированных значений боковых лепестков оценки нормированной АКФ оцифрованных радиошумов со стандартным нормальным законом.

В соответствии с критерием χ^2 [8, с. 275-276] выполним проверку гипотезы однородности распределения нормированных значений боковых лепестков каждой из оценок нормированных АКФ рассматриваемых последовательностей ПСЧ и распределения нормированных значений боковых лепестков оценки нормированной АКФ последовательности случайных чисел оцифрованных радиошумов. Расчетные значения статистик критерия, их критические области для уровня значимости $\alpha = 0,05$, а также достигнутые уровни значимости приведены в табл. 3.

Таблица 3

Результаты проверки гипотезы однородности распределения нормированных значений боковых лепестков оценок нормированных АКФ

		Заполнение первичных циклических сдвиговых регистров	Кол-во первичных циклических сдвиговых регистров	Расчетное значение статистики χ^2	Достигнутый уровень значимости P(S > S*)
Источник ПСЧ	Комбинационный генератор	Линейный конгруэнтный генератор	4	28,20	0,71
			6	34,36	0,40
			8	25,49	0,82
		Аддитивный генератор	4	22,44	0,92
			6	29,21	0,66
			8	18,70	0,98
		Квантовый ГСЧ	4	27,94	0,72
			6	19,63	0,97
			8	29,82	0,63
	Вихрь Мерсенна			25,90	0,81
Квантовый ГСЧ				22,83	0,91
Критическая область статистики критерия при $\alpha = 0{,}05$				$\chi^2 > 47{,}40$	

Приведенные результаты анализа (табл. 3) показывают, что расчетные значения статистик не попадают в критические области, достигнутые уровни значимости существенно превышают $\alpha = 0{,}05$, а гипотезу однородности распределений нормированных значений боковых лепестков оценок нормированных АКФ следует принять.

Также с целью проверки статистической гипотезы об однородности распределений боковых лепестков оценок нормированных АКФ последовательностей ПСЧ распределению боковых лепестков оценки нормированной АКФ последовательности случайных чисел оцифрованных радиошумов воспользуемся статистическими критериями знаков [8, с. 254-260], [5, с. 89-91] и серий [5, с. 91-93], а также ранговыми критериями Смирнова [5, с. 81], Вилкоксона [5, с. 93-95] и типа омега-квадрат Лемана-Розенблатта [5, с. 86]. Учтем, что в соответствии с рекомендациями, изложенными в [18, 19], при проверке гипотезы абсолютной однородности двух независимых выборок состоятельными критериями являются только критерии Смирнова и Лемана-Розенблатта. Расчетные значения статистик всех рассматриваемых критериев, а также их критические области для уровня значимости $\alpha = 0{,}05$ приведены в табл. 4.

Расчетные значения статистик (табл. 4) не попадают в критические области, поэтому гипотезу однородности распределений боковых лепестков оценок нормированных АКФ последовательностей ПСЧ распределению боковых лепестков оценки нормированной АКФ последовательности случайных чисел оцифрованных радиошумов следует принять.

Распределение знаков боковых лепестков АКФ

Проведем анализ распределения знаков в полученных последовательностях с помощью статистических критериев знаков [8, с. 254-260], [5, с. 89-91] и серий [5, с. 91-93]. Результаты анализа сведем в табл. 5. Результаты проверки подтверждают гипотезу о равновероятности положительного и отрицательного знака оценок АКФ для критерия знаков и гипотезу о случайности расположения положительных и отрицательных знаков оценок АКФ для критерия серий.

Распределение k-грамм для знаков боковых лепестков АКФ

С целью проверки статистической гипотезы о равномерном распределении знаков боковых лепестков оценок АКФ в k-мерном пространстве в данной работе используется критерий χ^2 [8, с. 267-275]. Проведем замену всех положительных значений оценок АКФ единицей, отрицательных – нулем. С помощью критерия χ^2 оценим распределение нулей и единиц в последовательности, а также распределение k-грамм для $k \geq 2$. Из [21, определение Q1] известно, что *b*-ичная последовательность длины *V* случайна, если она k-распределена для всех положительных целых чисел $k \leq \log_b V$. Таким образом, достаточно провести оценку распределения k-грамм знаков боковых лепестков для $k \leq \log_2 T$. Но при таком условии ожидаемые частоты появления k-грамм для максимальных значений k будут менее единицы, что неприемлемо для критерия хи-квадрат. Для использования критерия требуются, чтобы в 80% случаев ожидаемые частоты были более 5. Поэтому в данной работе верхний предел значения k выбирается таким образом, чтобы выполнялось неравенство $[T/k]/2^k \geq 5$. Таким образом, для объема выборки $n = 2^{16}$ $T = 2^{14}$, а $1 \leq k \leq 8$.

Результаты исследования приведены в табл. 6.

Таблица 4

Расчетные и критические значения статистик
критериев знаков, серий, Смирнова, Вилкоксона и Лемана-Розенблатта при проверке гипотезы
об однородности распределений боковых лепестков оценок нормированных АКФ

		Заполнение первичных циклических сдвиговых регистров	Количество первичных циклических сдвиговых регистров	Расчетные значения статистик				
				$F_{в1}$, $F_{в2}$ критерия знаков	$s_в$ критерия серий	$D_{m,n}$ критерия Смирнова	$z_в$ критерия Вилкоксона	t критерия Лемана-Розенблатта
Источник ПСЧ	Комбинационный генератор	Линейный конгруэнтный генератор	4	$F_{в1}$=1,0093; $F_{в2}$=0,9905	0,5262	0,0054	-0,1175	0,0457
			6	$F_{в1}$=1,0098; $F_{в2}$=0,9900	-1,9267	0,0052	-0,0313	0,0451
			8	$F_{в1}$=1,0113, ; $F_{в2}$=0,9886	1,9651	0,0080	-1,0645	0,1349
		Аддитивный генератор	4	$F_{в1}$=1,0066; $F_{в2}$=0,9932	-1,1627	0,0079	-0,3760	0,0739
			6	$F_{в1}$=0,9956; $F_{в2}$=1,0042	0,0241	0,0099	-0,0865	0,1779
			8	$F_{в1}$=1,0034; $F_{в2}$=0,9963	0,3051	0,0070	-0,3205	0,0547
		Квантовый ГСЧ	4	$F_{в1}$=1,0088; $F_{в2}$=0,9910	0,9635	0,0078	-0,5695	0,0807
			6	$F_{в1}$=1,0150; $F_{в2}$=0,9850	-0,0632	0,0063	-0,3159	0,0338
			8	$F_{в1}$=1,0103; $F_{в2}$=0,9896	-2,5046	0,0075	-0,3828	0,0512
	Вихрь Мерсенна			$F_{в1}$=0,9912; $F_{в2}$=1,0086	0,7447	0,0064	-0,1831	0,0759
Квантовый ГСЧ				$F_{в1}$=1,0093; $F_{в2}$=0,9905	0,2762	0,0048	-0,2634	0,0262
Критическая область статистики критерия				$F_в$>1,0311	$\|s_в\|$>2,576	$D_{m,n} \geq 0,015$	$\|z_в\|$>1,96	t > 0,461

Таблица 5

Расчетные и критические значения статистик
критериев знаков и серий для знаков боковых лепестков оценок АКФ

		Заполнение первичных циклических сдвиговых регистров	Количество первичных циклических сдвиговых регистров	Расчетные значения статистики $F_{в1}$, $F_{в2}$ критерия знаков	Расчетное значение статистики $s_в$ критерия серий
Источник ПСЧ	Комбинационный генератор	Линейный конгруэнтный генератор	4	$F_{в1}$=1,0201; $F_{в2}$=0,9801	-1,0576
			6	$F_{в1}$=1,0228; $F_{в2}$=0,9774	-1,5228
			8	$F_{в1}$=0,9950; $F_{в2}$=1,0048	2,5555
		Аддитивный генератор	4	$F_{в1}$=1,0136; $F_{в2}$=0,9863	-1,3301
			6	$F_{в1}$=1,0058; $F_{в2}$=0,9940	1,8059
			8	$F_{в1}$=0,9962; $F_{в2}$=1,0035	0,5551
		Квантовый ГСЧ	4	$F_{в1}$=0,9902; $F_{в2}$=1,0097	0,4953
			6	$F_{в1}$=1,0176; $F_{в2}$=0,9825	-2,2482
			8	$F_{в1}$=1,0161; $F_{в2}$=0,9839	-1,2652
	Вихрь Мерсенна			$F_{в1}$=0,9996; $F_{в2}$=1,0001	0,7422
Квантовый ГСЧ				$F_{в1}$=1,0174; $F_{в2}$=0,9827	0,9394
Оцифрованные радиошумы				$F_{в1}$=1,0139; $F_{в2}$=0,9861	0,0921
Критическая область статистики критерия				$F_в$ > 1,0311	$\|s_в\|$ > 2,576

Таблица 6

Результаты исследования распределения k-грамм для знаков боковых лепестков оценок нормированных АКФ

		Заполнение первичных циклических сдвиговых регистров	Кол-во первичных циклических сдвиговых регистров	Расчетное значение статистики χ^2 для k							
				1	2	3	4	5	6	7	8
Источник ПСЧ	Комбинационный генератор	Линейный конгруэнтный генератор	4	1,64	1,69	4,40	15,02	26,33	52,71	88,83	263,00
			6	2,11	3,88	13,83	18,40	30,14	68,89	134,67	238,50
			8	0,10	3,03	15,14	13,72	30,74	90,50	142,76	293,50
		Аддитивный генератор	4	0,77	1,77	10,02	13,46	23,79	59,09	156,55	250,75
			6	0,14	1,30	4,78	14,20	41,06	55,81	122,85	226,00
			8	0,05	0,54	1,85	13,20	26,51	44,69	93,85	229,25
		Квантовый ГСЧ	4	0,39	0,54	21,87	9,97	28,67	80,42	101,41	257,25
			6	1,27	1,50	5,93	15,58	28,42	61,95	146,81	239,25
			8	1,06	2,38	2,55	12,37	56,22	69,17	137,18	245,25
	Вихрь Мерсенна			0,00	10,27	5,95	17,10	24,73	57,96	148,78	246,75
Квантовый ГСЧ				1,23	1,46	16,13	28,91	33,17	74,33	126,79	295,00
Оцифрованные радиошумы				0,73	5,11	21,51	17,70	31,30	58,76	151,19	211,50
Критическая область статистики критерия $\chi^2 > \chi^2_{1-\alpha,2^k-1}$			$\chi^2_{0.95,2^k-1}$ =	3,84	7,81	14,07	25,00	44,99	82,53	154,30	293,25
			$\chi^2_{0.99,2^k-1}$ =	6,63	11,34	18,48	30,58	52,19	92,01	166,99	310,46

В силу наличия случаев попадания статистики в критическую область, в том числе и для оцифрованных радиошумов, полученные результаты не могут свидетельствовать о том, что распределения k-грамм для знаков оценок нормированных АКФ для исследуемых последовательностей чисел являются равномерными.

Рассмотрим частость

$$P^*\left(\chi^2 > \chi^2_{0.95,2^k-1}\right)$$

попадания статистики χ^2 в критическую область для N = 1000 выборок оцифрованных радиошумов (табл. 7).

Таблица 7

Частость попадания статистики в критическую область для $N = 1000$ выборок оцифрованных радиошумов

k	1	2	3	4	5	6	7	8
$P^*\left(\chi^2 > \chi^2_{0.95,2^k-1}\right)$	0,027	0,071	0,080	0,070	0,076	0,078	0,059	0,053

Для доверительной вероятности, принятой на уровне 0.95, и теоретической вероятности 0,05 доверительным интервалом для частости попадания статистики χ^2 в критическую область является интервал $(0.036, 0.064)$.

Большинство полученных частостей превышает верхнюю границу доверительного интервала. Поэтому гипотезу о равномерности распределения знаков боковых лепестков оценок АКФ в k-мерном пространстве при $k \geq 2$, несмотря на его близость к равномерному распределению, следует отклонить.

При $k = 1$ данная гипотеза подтверждается для всех последовательностей.

Несоответствие закона распределения знаков боковых лепестков оценок АКФ в k-мерном пространстве при $k \geq 2$ равномерному закону свидетельствует о корреляции вычисленных по (1) коэффициентов автокорреляции $r_x(\tau_i)$ и $r_x(\tau_j)$.

Выводы

Результаты проведенного в работе исследования позволяют сформулировать следующие утверждения:

– анализ корреляционных связей случайных последовательностей чисел (оцифрованных радиошумов, квантового генератора) и псевдослучайных последовательностей чисел (порожденных комбинационным генератором с комбинирующей функцией суммирования в некотором конечном поле и генератором типа «Вихрь Мерсенна») с помощью оценок их автокорреляционных функций показывает отсутствие в них линейной зависимости;

– анализ распределения оценок нормированных коэффициентов автокорреляции ненулевого порядка (боковых лепестков оценки нормированной автокорреляционной функции) случайной последовательности чисел и статистическая проверка гипотез соответствия этому распределению боковых лепестков оценок автокорреляционных функций исследуемых псевдослучайных последовательностей чисел подтверждают их однородность;

– несмотря на то, что результаты анализа распределения знаков боковых лепестков оценок АКФ случайных и псевдослучайных последовательностей чисел с помощью критериев знаков и серий свидетельствуют о равновероятности и случайности расположения положительного и отрицательного знаков оценок АКФ, результаты анализа распределения k-грамм для знаков оценок нормированных АКФ не могут свидетельствовать об их равномерном распределении при $k \geq 2$ для всех исследуемых последовательностей чисел;

– в результате исследования не выявлено зависимости корреляционных свойств последовательности чисел комбинационного генератора от его структуры (количества первичных циклических сдвиговых регистров (4, 6 или 8) и их заполнения (с помощью линейного конгруэнтного метода, аддитивного или квантового генератора)).

Сформулированные результаты имеют теоретическую и практическую ценность при использовании генераторов случайных и псевдослучайных чисел, в том числе комбинационного генератора, в задачах, требующих высокого качества генерируемых последовательностей.

Список литературы

1. Радиотехнические цепи и сигналы: [учебник для вузов] / [М.Т. Иванов, А.Б. Сергиенко, В.Н. Ушаков]; под ред. В.Н. Ушакова – СПб.: Питер, 2014. – 336 с.

2. Баскаков С.И. Радиотехнические цепи и сигналы / С.И. Баскаков. – М.: Высш. шк., 2000. – 462 с.

3. Дирак П.А.М. Принципы квантовой механики / П.А.М. Дирак. – [2-е издание]. – М.: Наука, 1979. – 481 с.

4. Эконометрика: Учебник / Под ред. И. И. Елисеевой. –М.: Финансы и статистика, 2007. – 576 с.

5. Большев Л.Н. Таблицы математической статистики / Л.Н. Большев, Н.В. Смирнов – [3-е изд.] – М.: Наука, 1983. – 416 с.

6. Matsumoto M. Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator / M. Matsumoto, T. Nishimura // ACM Transactions on Modeling and Computer Simulation. – 1998. – V.8. – pp. 3-30.

7. Лавданский А.А. Комбинационный метод формирования последовательности псевдослучайных чисел / А.А. Лавданский, Э.В. Фауре // Системний аналіз та інформаційні технології: матеріали 16-ї Міжн.НТК SAIT-2014, Київ, 26-30 травня 2014р. / ННК «ІПСА» НТУУ «КПІ». – К.: ННК «ІПСА» НТУУ «КПІ», 2014. – С. 403-404.

8. Смирнов Н.В. Курс теории вероятностей и математической статистики для технических приложений / Н.В. Смирнов, И.В. Дунин-Барковский – М.: Наука. 1969. – 512 с.

9. Kendall, M.G. The Advanced Theory of Statistics / Maurice G. Kendall. – V. II. – London: C. Griffin & Company limited, 1946. – 521 p.

10. True Random Number Service [Электронный ресурс] – Режим доступа: http:// random.org/.

11. QRNG Service [Электронный ресурс] – Режим доступа: http://qrng.physik.hu-berlin.de/.

12. Шнайер Б. Прикладная криптография. Протоколы, алгоритмы, исходные тексты на языке Си; [пер. с англ. под ред. Семьянова П.В.]. – [2-е изд.]. – М.: Триумф, 2002. – 816 с.

13. Random Class [Электронный ресурс]. – Режим доступа: http://msdn.microsoft.com/library/system.random%28v=vs.110%29.aspx.

14. Иванов М.А. Теория, применение и оценка качества генераторов псевдослучайных последовательностей / М.А. Иванов, И.В. Чугунков – М.: КУДИЦ-ОБРАЗ, 2003. – 240 с.

15. Лемешко Б.Ю. Корреляционный анализ наблюдений многомерных случайных величин при нарушении предположений о нормальности / Б.Ю. Лемешко, С.С. Помадин // Сибирский журнал индустриальной математики. – 2002. – Т. V. – №3(11). – С. 115-130.

16. Помадин С.С. Исследование распределений статистик многомерного анализа данных при нарушении предположений о нормальности [Текст] : дис. ... канд. техн. наук : 05.13.17 / С.С. Помадин. – Новосибирск, 2004.

17. Moran, P.A.P. Some Theorems on Time Series. II. The Significance of the Serial Correlation Coefficient / P.A.P. Moran // Biometrica. – 1948. – V. 35. – № 3/4. – P. 255-260.

18. Орлов А.И. О проверке однородности двух независимых выборок / А.И. Орлов // Заводская лаборатория. – 2003. – Т.69. – №1. – С.55-60.

19. Орлов А.И. Состоятельные критерии проверки абсолютной однородности независимых выборок / А.И. Орлов // Заводская лаборатория. Диагностика материалов. – 2012. – Т.78. – №11. – С.66-70.

20. Anderson T.W. Asymptotic Theory of Certain «Goodness of Fit» Criteria Based on Stochastic Processes / T.W. Anderson, D.A. Darling // Ann. Math. Statist. – 1952. – V. 23. – P. 193-212.

21. Кнут Дональд Э. Искусство программирования: В 7 т. Т.2: Получисленные алгоритмы. / Дональд Эрвин Кнут, Станфордский университет ; [пер. с англ. В. Тертышный] . –М.: ООО «И.Д. Вильямс», 2007. – 832 с.

Поступила в редколлегию 22.01.2015

Рецензент: д-р техн. наук проф. В.Н. Рудницкий, Черкасский государственный технологический университет, Черкассы.

АНАЛІЗ КОРЕЛЯЦІЙНИХ ВЛАСТИВОСТЕЙ ПОСЛІДОВНОСТЕЙ (ПСЕВДО) ВИПАДКОВИХ ЧИСЕЛ

Е.В. Фауре, А.І. Щерба, А.О. Лавданський

У роботі виконано порівняльний аналіз автокореляційних властивостей послідовностей, породжених комбінаційним генератором із різними параметрами, генератором типу «Вихор Мерсенна», і послідовностей «істинно випадкових» чисел. Результати, отримані в роботі, підтверджують однорідність оцінок автокореляційних функцій послідовностей псевдовипадкових чисел комбінаційного генератора і генератора типу «Вихор Мерсенна» з оцінкою автокореляційної функції випадкової послідовності чисел, а також дозволяють використовувати комбінаційний генератор у задачах, що потребують некорельованих послідовностей чисел.

Ключові слова: *послідовність (псевдо) випадкових чисел, комбінаційний генератор, кореляція, автокореляційна функція, статистичний критерій.*

ANALYSIS OF CORRELATION PROPERTIES OF (PSEUDO) RANDOM NUMBERS SEQUENCES

E.V. Faure, A.I. Shcherba, A.O. Lavdanskyi

The comparative analysis of autocorrelation properties of sequences generated by combination generator with its various parameters, "Mersenne twister" generator and the sequence of "true random" numbers is done. The results obtained in this work confirm the homogeneity of estimates of autocorrelation functions of pseudorandom numbers from combination generator and "Mersenne twister" generator with an estimate of autocorrelation function of a random sequence of numbers, as well as allow to use the combination generator in tasks requiring uncorrelated sequences of numbers.

Keywords: *(pseudo) random numbers sequence, combination generator, correlation, autocorrelation function, statistical criterion.*