УДК 004.382:519.2

Л.О. Кириченко, Р.И. Цехмистро, О.Я. Круг, А.В. Стороженко

Харьковский национальный университет радиоэлектроники, Харьков

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ГЕНЕРАЦИИ ПСЕВДОСЛУЧАЙНЫХ ЧИСЕЛ В СОВРЕМЕННЫХ ТЕХНОЛОГИЯХ БЕСПРОВОДНОЙ ПЕРЕДАЧИ ДАННЫХ

В работе рассмотрена генерация псевдослучайных числовых последовательностей, построенных с помощью рекуррентных алгоритмов и хаотических отображений. Представлены результаты практической реализации данных алгоритмов на современной микропроцессорной технике. Проведен сравнительный анализ генерируемых последовательностей на предмет их соответствия критериям независимости и равномерности.

Ключевые слова: *алгоритм Лемера, хаотическое отображение, аппаратная реализация алгоритмов, технология Wi-Fi.*

Введение

При проведении различных исследований зачастую необходимо формирование последовательности чисел, обладающих свойствами чисто случайных. Такие последовательности могут применяться при решении задач криптографии, навигации, радиотехники, при проведении локации удаленных и быстро движущихся объектов, при стохастических вычислениях и для многих других целей. В настоящее время широкое распространение получила технология Wi-Fi (Wireless Fidelity), что дает возможность осуществлять беспроводную передачу данных с высокой точностью. Спецификация данного способа описана в известном стандарте IEEE 802.11x, который регламентирует осуществление обмена данных по радиоканалам в беспроводных локальных сетях (WLAN).

Стандарт IEEE 802.11 предусматривает применение расширенного спектра сигнала двумя методами передачи: метод скачкообразного переключения частот FHSS (Frequency Hopping Spread Spectrum) и метод прямой последовательности DSSS (Direct Sequence Spread Spectrum). При применении метода FHSS весь диапазон (2,4ГГц) делится на 79 каналов шириной 1МГц каждый. Приемник и передатчик последовательно по псевдослучайному закону переключаются на несущие частоты каналов. Для модуляции несущих используется двухуровневое гауссово переключение частот. При использовании широкополосных (псевдошумовых) сигналов расширение спектра сигнала происходит за счет добавления псевдослучайной последовательности битов, называемых чипами, к каждому информационному биту передаваемого сигнала. Устройства Bluetooth работают на той же частоте 2,4 ГГц, что и устройства стандарта Wi-Fi/IEEE. 802.11b. Но для того, чтобы избежать интерференции с устройствами Wi-Fi, в Bluetooth используется метод передачи сигналов, получивший название расширенного спектра скачкообразной перестройки частоты [1].

Данные обстоятельства подчеркивают актуальность разработки и аппаратной реализации эффективных алгоритмов генерации псевдослучайных чисел. Среди всевозможных типов распределений самым важным является равномерное, так как с его помощью можно получить любое другое распределение. Поэтому генерация случайного процесса фактически сводится к получению множества равномерно распределенных случайных чисел.

Существует много аналитических способов получения равномерных псевдослучайных чисел на ЭВМ, как, например, выбор «середины произведения», использование вычетов, способ перемешивания и т.д. Все они представляют собой некоторое рекуррентное соотношение, в котором каждое последующее значение мы получаем из предыдущего или предыдущих [2]. В последнее время актуальным является поиск альтернативных способов получения псевдослучайных чисел, например с помощью детерминированных хаотических отображений [3].

Задачей данной работы является практическая генерация псевдослучайных чисел, полученных различными алгоритмами, с помощью современных 8-ми разрядных микроконтроллеров с гарвардской регистр-аккумуляторной архитектурой и исследование свойств полученных псевдослучайных последовательностей.

Актуальность данных исследований связана с тем, что даже наиболее эффективный алгоритм будет неприемлем для практической реализации из-за необходимости использования большого объема памяти программ и данных. Он также может требовать достаточно большое время для генерации заданного количества выборок.

Алгоритмы генерации псевдослучайных последовательностей и их практическая реализация

На практике разработка и применение устройств телекоммуникаций с реализацией псевдослучайных сигналов невозможна без использования современной микропроцессорной техники. Данное обстоятельство накладывает дополнительные условия

для отбора эффективных алгоритмов создания генераторов случайных чисел, поскольку заказчики зачастую предъявляют требования к малогабаритности и быстродействию. Последние обстоятельства приводят к необходимости учета ограниченного объема памяти программ и данных. Следовательно, количество случайных чисел будет ограничено, что подчеркивает необходимость проверки существующих алгоритмов по указанным выше критериям.

Одним из таких алгоритмов является алгоритм, предложенный Лемером, который известен как метод линейного конгруента, так как он в полной мере соответствует вышеизложенным требованиям [4, 5]. Этот алгоритм параметризируется четырьмя числами следующим образом: m – модуль (основание системы), $m > 0$; a – множитель, $0 \le a < m$; c – приращение, $0 \le c < m$; X_0 – начальное значение, или зерно, $0 \le X_0 < m$.

Последовательность случайных чисел $\{X_n\}$ получается с помощью следующего итерационного равенства:

$$X_{n+1} = (aX_n + c) \bmod m .$$

Если значения m, a и c являются целыми, то создается последовательность целых чисел в диапазоне $0 \le X_n < m$. Такая последовательность и будет нужна нам для выбора очередной частоты из используемого диапазона. Время между итерациями может быть выбрано с помощью внутреннего таймера микроконтроллера.

Выбор значений для параметров m, a и c является критичным для разработки хорошего генератора случайных чисел.

Существует три критерия, используемые при выборе генератора случайных чисел:

– функция должна создавать полный период, т.е. все числа между 0 и m до того, как создаваемые числа начнут повторяться;

– создаваемая последовательность должна появляться случайно. Последовательность не является случайной, так как она создается детерминировано, но различные статистические тесты, которые могут применяться, должны показывать, что последовательность случайна;

– функция должна эффективно реализовываться на процессорах либо микроконтроллерах.

Значения параметров m, a и c должны быть выбраны таким образом, чтобы эти три критерия выполнялись. В соответствии с первым критерием можно показать, что если число m является простым и $c = 0$, то при определенном значении a период, создаваемый функцией, будет равен $m-1$. Для 32-битной арифметики соответствующее простое значение $m = 2^{31} - 1$.

Очевидно, что m должно быть очень большим, чтобы была возможность создать много случайных чисел. Считается, что m должно быть приблизительно равно максимальному положительному целому числу для данного процессора либо микроконтроллера. Таким образом, обычно m близко или равно 2^{31} для 32-разрядных процессоров, либо 2^{15} для 16-разрядных микроконтроллеров.

Только небольшое число значений параметра a удовлетворяет всем трем критериям. Одно из таких значений есть $a = 7^5 = 16807$, которое использовалось в семействе компьютеров IBM 360. Этот генератор широко применяется и прошел более тысячи тестов – больше, чем все другие генераторы псевдослучайных чисел.

Особенность алгоритма линейного конгруента заключается в том, что если сомножитель и модуль (основание) соответствующим образом подобраны, то результирующая последовательность чисел будет статистически неотличима от последовательности, являющейся случайной из набора $1, 2, ..., m-1$. Но не может быть случайности в последовательности, полученной с использованием алгоритма, независимо от выбора начального значения X_0. Если значение выбрано, то оставшиеся числа в последовательности будут предопределены. Это всегда учитывается при криптоанализе. Для нас это является важным аспектом, так как мы должны иметь строго определённый алгоритм шифрования для создания адекватного генератора случайных сигналов. Характерной особенностью исследования эффективности алгоритмов генерации псевдослучайной последовательности на микропроцессорной технике является возможность оценки времени генерации (быстродействия), объема памяти программ и данных, необходимых для устойчивого функционирования какого-либо алгоритма.

В наших исследованиях использовался микроконтроллер AVR-atmega 128 фирмы АТМЕЛ, который имеет:

– объем FLASH-памяти программ128 кБт;

– объем статической оперативной памяти 4кБт;

– объем памяти данных на основе электрически стираемого ПЗУ (EEPROM) оперативной памяти 4кБт.

Входящие в состав лабораторного макета матричная клавиатура и жидкокристаллический индикатор (ЖКИ) Toshiba T6963C позволяют пронаблюдать выборки произвольного и фиксированного количества чисел для разных алгоритмов генерации. Входящие в состав микроконтроллера 8 и 16 разрядные таймеры-счетчики и ЖК-дисплей позволяют оценивать время генерации фиксированного количества случайных чисел для конкретных алгоритмов, используя процедуры прерывания для фиксации и последующего отображения времени генерации периода последовательности. Устройство, которое практически реализует метод линейного конгруента, представлено на рис. 1.

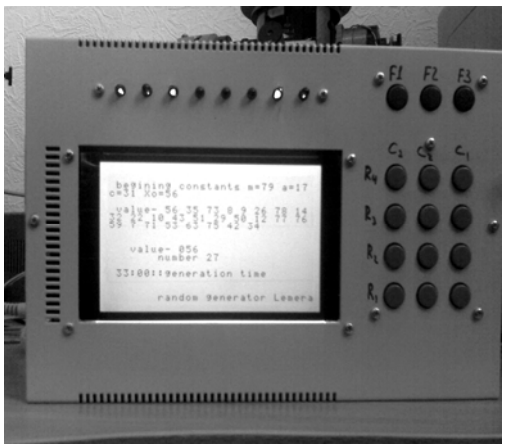


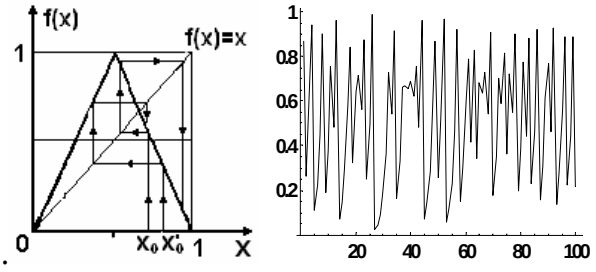
Рис. 1. Лабораторный макет
на базе микроконтроллера AVR-Атмега 128

Данный макет позволяет демонстрировать возможности влияния исходных параметров (m , a , c , X_0) на эффективность метода. Он отображает результат, время генерации случайных сигналов, значение суммы случайных чисел в байтах на экране жидкокристаллического индикатора.

С помощью вышеописанного лабораторного устройства был реализован альтернативный метод генерации случайных чисел. В основе предлагаемого генератора лежит одномерное итерационное отображение, часто называемое треугольным, которое описывается формулой

$$X_{n+1} = r\left(1-2\left|0,5-X_n\right|\right),$$

где r –параметр отображения, а последовательность $\{X_n\}$ изменяется на промежутке (0,1). При значении $r > \frac{1}{2}$ в процессе итерирования изначально близкие точки отдаляются друг от друга (рис. 1, а) и треугольное отображение порождает хаотическую последовательность чисел $\{X_n\}$, которые можно рассматривать как случайные (рис. 1, б).



а б

Рис. 2. Треугольное отображение и хаотическая последовательность чисел

Инвариантная мера $\rho(x)$ задает плотность итераций отображения $X_{n+1} = f(X_n)$, $X_n \in [0,1]$. Инвариантную меру $\rho(x)$ можно рассматривать как аналог плотности распределения случайной величины. Для треугольного отображения при $r = 1$ $\rho(x) = 1$ [3]. Это означает, что последовательность итераций $X_0, f(X_0), f(f(X_0)),...$ равномерно покрывает интервал (0,1). Таким образом, мы можем считать, что треугольное отображение является генератором равномерно распределенных случайных чисел.

Тестирование псевдослучайных чисел на независимость и равномерность

Очевидно, что свойства чисто случайной последовательности должны быть следующими: отсутствие корреляции между числами (независимость или случайность выборочных данных) и соответствие заданному закону распределения, в данном случае равномерному.

Тесты на независимость (случайность) выборочных данных были проведены по критериям серий, инверсий и поворотных точек. В каждом случае мы выдвигаем гипотезу H_0 о том, что независимыми исходами одной и той же случайной величины, и задаемся уровнем значимости α. Кратко опишем используемые критерии [2, 6].

Критерий серий. Рассмотрим последовательность из N наблюденных значений случайной величины, причем каждое наблюдение отнесем к одному из двух взаимно исключающих классов. Серией называется последовательность однотипных наблюдений, перед и после которой следуют наблюдения противоположного типа. Число серий, появившееся в последовательности наблюдений, позволяет выяснить, являются ли отдельные результаты независимыми наблюдениями. Всегда можно провести разбиение на классы, так чтобы $N_1=N_2=N/2$, например, путем сравнения наблюдений с медианой выборки.

Если последовательность из N наблюдений состоит из независимых исходов одной и той же случайной величины, то число серий в последовательности является случайной величиной ξ с математическим ожиданием $m_\xi = N/2+1$ и дисперсией $\sigma_\xi^2 = \frac{N(N-2)}{4(N-1)}$. Распределение вероятностей числа серий ξ является затабулированным.

В качестве нулевой гипотезы принимается, что наблюдения являются независимыми. Для проверки гипотезы с требуемым уровнем значимости α надо сравнить наблюденное значение числа серий с границами области принятия гипотезы. Если число серий окажется вне этой области, то гипотеза отвергается. В противоположном случае гипотезу можно принять с уровнем значимости α.

Критерий инверсий. Рассмотрим, сколько раз в последовательности из N наблюдений имеют место неравенства $x_i > x_j$ при $i < j$ (инверсии). Рассчитаем общее число инверсий. Если последовательность из N наблюдений состоит из независимых исходов

одной и той же случайной величины, то число инверсий является случайной величиной ξ с математическим ожиданием $m_\xi = \frac{N(N-1)}{4}$ и дисперсией $\sigma_\xi^2 = \frac{2N^3+3N^2-5N}{7^2}$. Число инверсий, появившееся в последовательности наблюдений, будет иметь выборочное распределение, которое протабулировано. Для проверки гипотезы о независимости данных с требуемым уровнем значимости α надо сравнить наблюденное значение числа инверсий с границами области принятия гипотезы. Если число серий окажется вне этой области, то гипотеза отвергается. В противоположном случае гипотезу можно принять с уровнем значимости α.

Критерий поворотных точек. В последовательности из N наблюдений подсчитаем количество пиков и впадин, т.е. сколько раз в последовательности имеют место неравенства $x_i < x_{i+1} > x_{i+2}$ или $x_i > x_{i+1} < x_{i+2}$. Каждое такое неравенство определяет поворотную точку. Определим общее число поворотных точек.

Если последовательность из N наблюдений состоит из независимых исходов, то число поворотных точек является случайной величиной ξ с математическим ожиданием $m_\xi = \frac{2}{3}(n-2)$ и дисперсией $\sigma_\xi^2 = \frac{16n-29}{90}$. Число поворотных точек, появившееся в последовательности наблюдений, стремится к нормальному распределению $N(m_\xi, \sigma_\xi)$. Для проверки гипотезы о независимости данных с требуемым уровнем значимости α надо сравнить наблюденное значение числа поворотных точек с границами области принятия гипотезы $[m_\xi - t*\sigma_\xi, m_\xi + t*\sigma_\xi]$, где $2\Phi(t) = 1-\alpha$, Φ – интеграл Лапласа. Если число серий окажется вне этой области, то гипотеза отвергается. В противоположном случае гипотезу можно принять с уровнем значимости α. Для исследования равномерности данных была осуществлена проверка статистической гипотезы о законе распределения с помощью критериев согласия. В работе были использованы три критерия: критерий Пирсона, Колмогорова и Мизеса. В каждом случае выдвигалась гипотезу H_0 о том, что полученная выборка имеет равномерный закон распределения, и задавался уровень значимости α. Кратко опишем используемые критерии [2, 7].

Критерий Пирсона (критерий χ^2). Использование этого критерия основано на применении такой меры расхождения между теоретическим $F(x)$ и эмпирическим распределением $F_n(x)$, которая приближенно подчиняется закону распределения χ^2. Разбиваем интервал [a,b] на l непересекающихся интервалов (разрядов). Для нахождения общей степени расхождения между $F(x)$ и $F_n(x)$ необходимо подсчитать статистику $\chi^2 = \sum_{i=1}^{l} \frac{n_i - n\cdot p_i}{n\cdot p_i}$, где n_i – наблюдаемая частота попаданий в i-й разряд; n – объем выборки; p_i – теоретическая вероятность попадания в i-й разряд. Если полученное значение статистики $\chi^2 < \chi^2(k,\alpha)$, где $\chi^2(k,\alpha)$ – табличное значение χ^2 с $k = l-3$ степенями свободы и уровнем значимости α, то гипотезу H_0 принимаем, в противном случае – отвергаем.

Критерий Колмогорова. Для применения этого критерия выборку вместо разбиения на разряды представляют в виде вариационного ряда. В качестве меры расхождения между теоретической $F(x)$ и эмпирической $F_n(x)$ функциями распределения непрерывной случайной величины X используется модуль максимальной разности $d_n = \max_{x\in\{x_n\}} |F(x) - F_n(x)|$.

При неограниченном увеличении количества наблюдений n функция распределения случайной величины $d_n\sqrt{n}$ асимптотически приближается к функции распределения $K(\lambda) = P(d_n\sqrt{n} < \lambda) = \sum_{k=-\infty}^{\infty} (-1)^k \exp(-2k^2\lambda^2)$. Т.е. гипотезу H_0 принимаем, если полученное значение статистики $d_n\sqrt{n} < \lambda$. При заданном уровне значимости α число λ выбирается из соотношения $\alpha = 1 - K(\lambda)$. В противном случае гипотезу H_0 отвергаем.

Критерий Мизеса (критерий w^2). В качестве меры различия теоретической функции распределения $F(x)$ и эмпирической $F_n(x)$ по критерию Мизеса выступает средний квадрат отклонений по всем значениям аргумента $x: w_n^2 = \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 dF(x)$. Статистикой критерия является величина $nw_n^2 = \frac{1}{12n} + \sum_{i=1}^{n}\left[F(x_i) - \frac{i-0,5}{n}\right]^2$. При неограниченном увеличении n существует предельное распределение статистики nw_n^2. Выбрав уровень значимости α можно определить критические значения $nw_n^2(\alpha)$. Если фактическое значение nw_n^2 окажется больше критического или равно ему, то согласно критерию Мизеса с уровнем значимости α гипотеза H_0 отвергается, в противном случае – принимается.

Результатыятестированияя

Используемый нами лабораторный макет позволяет реализовывать вышеописанные и другие алгоритмы образования случайных чисел и производить следующий анализ результатов:

– оценивать время генерации одинакового количества случайных чисел (при одинаковом объеме, занимаемой памятью программы);

– проводить оценку объема памяти, которую занимают сгенерированные числа (для одинакового их числа в каждом алгоритме);

– проводить исследование влияния начальных параметров, задающихся в каждом алгоритме;

– оценивать выборочные моментные характеристики последовательностей случайных чисел при фиксированном и разном времени их генерации.

В ходе работы было проведено численное моделирование псевдослучайных чисел на макете с помощью хаотического отображения и алгоритма Лемера. В каждом случае генерировалась выборка из 100 элементов, для нее проверялись гипотеза случайности по критериям серий, инверсий и поворотных точек и гипотеза равномерности по критериям Пирсона, Колмогорова и Мизеса с уровнем значимости $\alpha = 0,05$.

Время генерации 100 символов по алгоритму Лемера при тактовой частоте 11,059МГц без учета затрат времени на вывод символов на дисплей ЖКИ составило примерно 160 мкс; аналогичное время генерации с помощью хаотического отображения составило 105 мкс. Однако данная оценка времени генерации существенно зависит от параметров алгоритмов, также как и объем памяти, занимаемый сгенерированными числами, что является предметом дальнейших исследований.

Результаты проверки случайности показали, что данные, полученные с помощью треугольного отображения, также как и псевдослучайные числа встроенных генераторов, удовлетворяют требованиям независимости, корреляция между числами отсутствует.

Результаты проверки равномерности приведены в табл. 1:

Таблица 1

Процент принятия гипотезы равномерности

	Критерий Пирсона	Критерий Колмогорова	Критерий Мизеса
Алгоритм Лемера	89,9	96,9	93,8
Хаотическое отображение	87,3	97,2	94,5

Выводыя

В данной работе предложена аппаратная реализация генераторов псевдослучайных чисел различными методами. Исследование полученных чисел показали их некоррелируемость и равномерность. Тестирование алгоритмов на микропроцессорном макете позволило ввести практическую реализацию данных алгоритмов в учебный процесс.

Списокяитературыя

1. Журавлев В.И. Поиск и синхронизация в широкополосных системах / В.И. Журавлев. – М.: Радио и связь, 1986. – 102 с.

2. Ермаков С.М. Метод Монте-Карло и смежные вопросы / С.М. Ермаков. – М.: Наука, 1975. – 211 с.

3. Шустер Г. Детерминированный хаос: Введение / Г. Шустер. – М.: Мир, 1988. – 240 с.

4. Нечаев В.И. Элементы криптографии / В.И. Нечаев. – М.: Высш. школа ,1999. – 109 с.

5. Сарвате Д.В. Взаимокорреляционные свойства последовательных и родственных последовательностей / Д.В. Сарвате, М.Б. Пресли // ТИИЭТ. – 1980. – Т.68, № 5. – С. 59-88.

6. Кендэл М. Временные ряды / М. Кендэл. – М.: Финансы и статистика,1981. – 198 с.

7. Тюрин Ю.Н. Статистический анализ данных на компьютере / Ю.Н. Тюрин, А.А. Макаров. – М: Инфра, 1997. – 528 с.

Поступила в редколлегию 19.06.2009

Рецензент: д-р техн. наук, проф. В.О. Тимофеев, Харьковский национальный университет радиоэлектроники, Харьков.

я

ПОРІВНсЛЬНИЙяАНАЛІЗяГЕНЕРАЦІЇяПСЕВДОВИПАДКОВИХяЧИСЕЛяя УяСУЧАСНИХяТЕХНОЛОГІсХяБЕЗПРОВІДНОЇяПЕРЕДАЧІяДАНИХя

Л.О. Кіріченко, Р.І. Цехмістро, О.Я. Круг, О.В. Стороженко

У роботі розглянута генерація псевдовипадкових числових послідовностей, побудованих за допомогою рекурентних алгоритмів і хаотичних відображень. Представлені результати практичної реалізації даних алгоритмів на сучасній мікропроцесорній техніці. Проведений порівняльний аналіз послідовностей, що генеруються, на предмет їх відповідності критеріям незалежності і рівномірності.

Ключові слова: *алгоритм Лемера, хаотичне відображення, апаратна реалізація алгоритмів, технологія Wi-Fi.*

COMPARATIVEяANALYSISяOFяTHEяPSEUDORANDOMяNUMBERSяGENERATIONяя ATяCURRENTяTECHNOLOGYяOFяWIRELESSяDATAяTRANSMISSIONя

L.O. Kirichenko, R.I. Tsekhmistro, O.Ya. Krug, A.V. Storozhenko

Is considered the generation of pseudorandom numbers built with the help of recurrent algorithms and chaotic reflections. The results of practical realization of these algorithms are presented on a modern microprocessor technique. The comparative analysis of the generated sequences is conducted for the purpose their accordance the criteria of independence and evenness.

Keywords: *algorithm of Lemer, chaotic reflection, hardware representation of algorithms, technology of Wi-Fi.*