УДК 519.246

С.Н. САРАНЧА

МЕТОДИКА ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ АППАРАТНОГО ГЕНЕРАТОРА СЛУЧАЙНЫХ ЧИСЕЛ, РЕАЛИЗОВАННОГО В ПЛИС АРХИТЕКТУРЫ FPGA

Рассматриваются общие подходы к построению аппаратных генераторов случайных чисел. Предлагается методика оценивания качества случайной последовательности с использованием критерия «хи-квадрат» и результатов прохождения пакета статистических тестов NIST STS. Исследуется влияние параметров генератора случайных чисел на качество случайной последовательности.

1. Введение

Генераторы случайных чисел (ГСЧ, TRNG – True Random Numbers Generators) имеют достаточно широкий спектр применения в современных вычислительных системах. К областям применения ГСЧ можно отнести криптографию, моделирование, компьютерные игры и т.д. В настоящий момент широко используются три основных подхода к получению последовательностей случайных чисел: выборка колебаний генератора, хаотические составляющие в цепях или специальное усиление шумов резисторов или диодов. При этом вопросы влияния параметров генератора (вид генератора, количество цепей, количество инверторов в цепях) на качество формируемой случайной последовательности рассмотрены слабо. В данной работе предлагается методика определения оптимальных параметров ГСЧ в FPGA общего назначения.

Аппаратный (физический) генератор случайных чисел [1] – устройство, которое генерирует последовательности случайных чисел на основе измеряемых параметров протекающего физического процесса. Работа таких устройств часто основана на процессах, таких как:

– тепловой шум,
– фотоэлектрический эффект,
– другие квантовые явления,
– неравномерность в задержках логических элементов.

Эти процессы, в теории, абсолютно непредсказуемы. Аппаратные генераторы случайных чисел, основанные на квантовых процессах, обычно состоят из специального усилителя и преобразователя. Усилитель усиливает очень слабые сигналы, получаемые в результате проходящих физических явлений, до приемлемых размеров, которые преобразуются преобразователем к цифровому виду (рис.1).

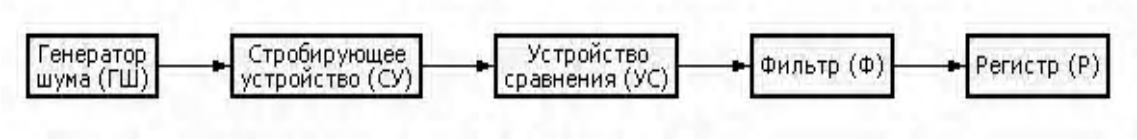

Рис. 1. Ааппаратный метод генерации случайных чисел

Аппаратные генераторы случайных чисел относительно медленны и могут производить смещенные последовательности (когда определенная последовательность чисел повторяется чаще). Использование подобных генераторов зависит от потребностей конкретной предметной области и от устройства самого генератора.

Целью данной работы является описание методики определения параметров генератора случайных чисел, реализованного в программируемых логических интегральных микросхемах архитектуры FPGA.

2. Структура апаратного генератора случайных чисел

В данный момент все большую популярность завоевывают цифровые генераторы случайных чисел. Источниками энтропии в цифровых генераторах чаще всего являются:

– задержки в логических элементах;
– нестабильность внешнего генератора синхросигнала.

В общем случае (рис.2) источник случайного шума генерирует аналоговый сигнал n(t), который является результатом некоторого недетерминированного физического явления. Сигнал аналогового шума оцифровывается (например, с помощью компаратора), получается так называемый оцифрованный аналоговый сигнал s[i]. Недетерминированный источник и преобразователь в цифровую форму вместе формируют преобразованный в цифровую форму шумовой источник.

Оцифрованный цифровой шум передается в модуль постпроцессинга, который затем выдает последовательность m-битных случайных слов r[i] , так называемых внутренних случайных чисел. Прежде всего, пост-процессор должен регулировать распределение вероятностей «сырых» случайных бит s[i], компенсируя таким образом внутренние несовершенства источника энтропии или цифрового преобразователя (например, ответвление компаратора напряжения). Распределение вероятностей внутреннего случайного слова r[i] намного более близко к общепринятому, чем s[i]. Подробнее методы постпроцессинга рассмотрены ниже.

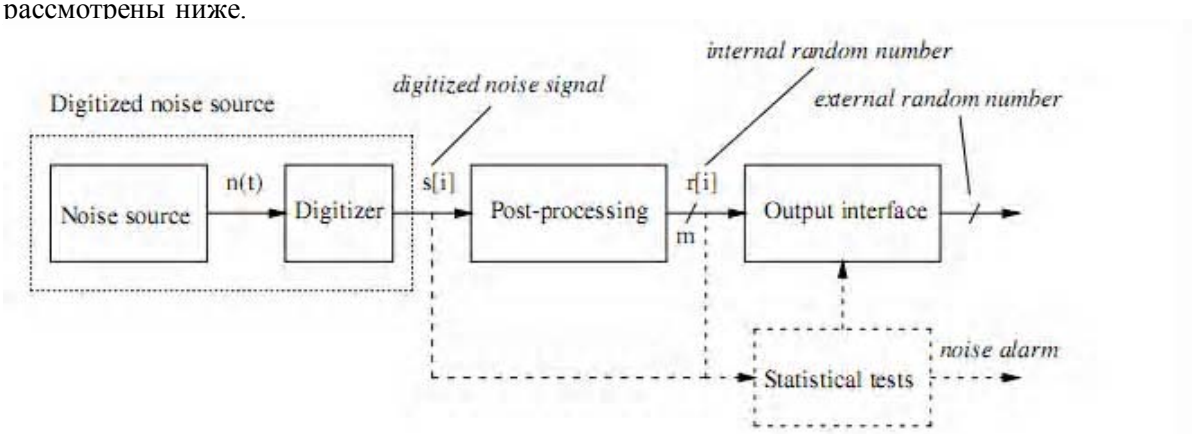

Рис. 2. Структура аппаратного ГСЧ

Шаг постпроцессинга используется, чтобы увеличить энтропию с помощью внутреннего случайного слова r[i], применяя функцию компрессии к входному потоку s[i] , что приводит к получению потока с меньшей скоростью и большей случайностью. Это становится особенно важно, если используется источник шума с низкими характеристиками энтропии на бит. Компрессия также обеспечивает устойчивость к различного рода воздействиям.

Метод, использующий неравномерность в задержках логических элементов, идет исключительно по цифровому пути - использует кольцевые генераторы в качестве источника шума (рис.3). Если в кольцевую структуру объединить нечетное количество инверторов, выходное значение каждого инвертора будет колебаться от логического ноля к логической единице и обратно из-за неустойчивой природы цепи. В любой момент времени в цепи наблюдается периодическая квадратная волна.

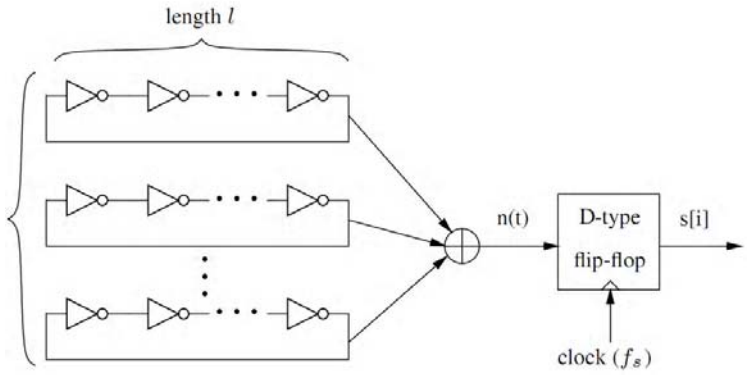

Рис. 3. Структура генератора

В идеале период волны линейно зависит от числа инверторов (т.е. от длины цепи) и задержки одного инвертора. На практике существует некоторая случайность в момент подачи сигнала. Это явление обычно называют вибрацией. Цель цифрового TRNG – обработать полученную энтропию, выбрав неопределенные зоны и недетерминированные части волны.

В общем, существуют две технологии для извлечения случайности из вибраций:

– выборка результатов кольца генераторов с помощью другого генератора (связные генераторы);

– комбинация сигналов ряда генераторов.

Структура, представленная на рис. 3, основана на втором подходе. Сигналы от k ëîêàëüíûõ äâóõâõîäîâûõ ëîãè÷åñêèõ $l_1 \ldots l_k$ объединяются исключающим или и используются в качестве сигнала n(t).

Из этого сигнала производится выборка с регулярной частотой f(s), результат заносится в D-триггер защелку. Так получается битовый поток s[i].

Альтернативное решение реализации генератора случайных чисел на FPGA работает, сэмплируя высокоточный высокочастотный синхросигнал Fh нестабильным низкочастотным сигналом Ft.

Для этого используется триггер-защелка D-типа, на вход синхронизации которого поступает сигнал Ft, на вход данных – сигнал Fh. Выдача данных ведется с частотой F1 (рис.4).

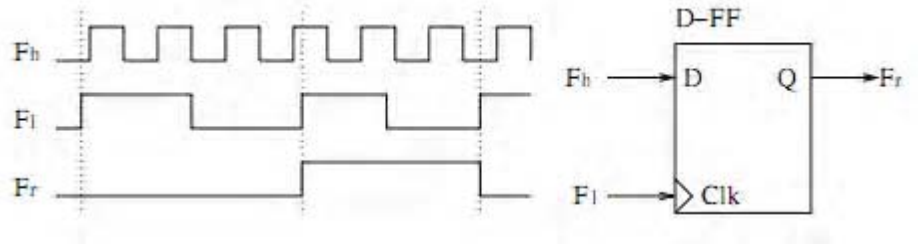

Рис. 4. Использование триггера-защелки для ГСЧ на основе нестабильности внешнего источника тактовой частоты

Для данного генератора существует несколько факторов, влияющих на случайность получаемых последовательностей.

Во-первых, скважность сигнала Fh не обязательно 50%, а следовательно, Ft может случайным образом принимать значения «1» и «0». Для выравнивания неравномерного распределения используется N-битный фильтр частоты.

Во-вторых, на качество получаемых случайных последовательностей влияет выбор частоты синхросигнала. Если изменения периода F1 недостаточны, корреляция между битами позволит в определенной степени «предсказать» следующий бит на основе предыдущих.

Для улучшения характеристик случайных последовательностей (особенно в случае применения генератора с низкой энтропией) может использоваться постпроцессинг. Компрессия, неизбежная в ходе постпроцессинга, обеспечивает устойчивость к различного рода воздействиям.

Существует два популярных метода постпроцессинга:

– метод фон Ньюмона,

– XOR-коррекция.

XOR-коррекция предполагает обработку входных бит – из двух входных путем суммирования по модулю два получается один выходной. Таким образом, поток сжимается в два раза.

Коррекция фон Ньюмона также рассматривает пары битов, но использует первый из них в случае, если они одинаковы, в противном случае отбрасывает. Результирующий поток будет иметь переменную разрядность, но в среднем фактор сжатия составит 4.

Наряду с простыми коррекциями применяются и более сложные, такие как псевдогенератор BBS или эластичная функция, схема аппаратной реализации которой приведена на рис. 5.

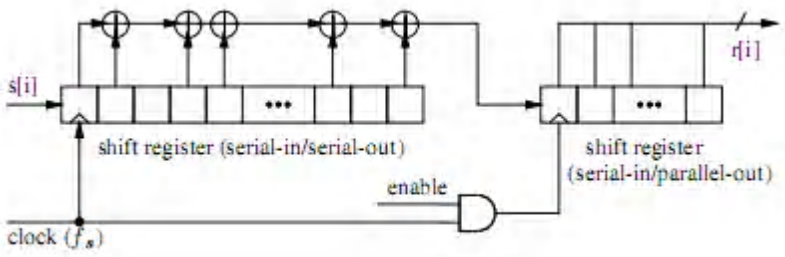

Рис. 5. Реализация алгоритма постпроцессинга, основанная на циклических кодах

Также для постпроцессинга применяется псевдогенератор (Pseudo random number generator - PRNG) BBS (Blum Blum Shub), названный так по именам трех своих создателей.

BBS использует уравнение квадратичного вычета (заметим, что это псевдослучайный генератор бит вместо генератора псевдослучайных чисел; он генерирует последовательность битов (0 или 1)).

Ниже приведены шаги генерации:

– Найдем два простых числа p и q в форме 4k+3, где k – целое число (p и q являются конгруэнтными).

– Выберем модуль n=p*q.

– Выберем случайное число r, взаимно-простое с n.

– Вычислим начальное число как $x_0=r^2 \bmod n$.

– Генерируем случайную последовательность $x_{i+1}=x_i^2 \bmod n$.

– Возьмем самый младший бит сгенерированного случайного целого числа (LSB - Least Significant Bit) как случайный бит.

Схема данного метода приведена на рис. 6.

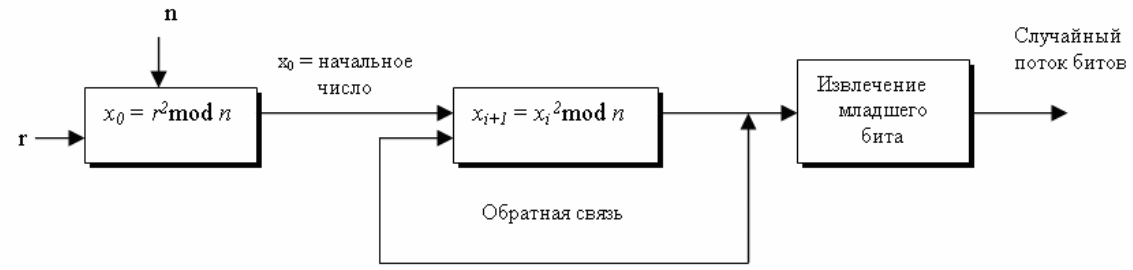

Рис.6. BBS-метод генерации случайной последовательности

В общем случае, постпроцессинг используют, когда имеется возможность пожертвовать скоростью генерации при необходимости повысить энтропию.

3. Методы оценки качества случайной последовательности

Для оценки качества случайной последовательности применяются два вида критериев:

– эмпирические (статистические) критерии, при использовании которых компьютер манипулирует группами чисел последовательности и вычисляет определенные статистики;

– теоретические (формальные) критерии, для которых характеристики последовательности определяются с помощью теоретико-числовых методов, основанных на рекуррентных правилах, которые используются для образования последовательности.

Идеальный ГСЧ должен выдавать близкие к следующим значения статистических параметров, характерных для равномерного случайного закона:

$$m_r = (\sum_{i=1}^{n} r_i)/n,$$

где m_r – математическое ожидание, стремится к значению 0.5; r_i – количество попаданий в i-й интервал; n – число интервалов;

$$D_r = \frac{\sum_{i=1}^{n}(r_i - m_r)^2}{n};$$

здесь D_r – дисперсия, стремится к значению 0.0833; σ_r – среднеквадратическое отклонение $\sigma_r = \sqrt{D_r}$, которое стремится к 0,2886.

Для доказательства гипотезы о равномерном распределении наиболее часто используются рассмотренные ниже критерии.

Одним из стандартных наборов статистических тестов является NIST STS[2], состоящий из 15 различных тестов. Отличительная особенность NIST STS – открытость алгоритмов и однозначность интерпретации результатов анализа.

На основе открытых алгоритмов NIST STS было разработано программное обеспечение, позволяющее анализировать свойства потока случайных чисел, поступающих от аппаратного ГСЧ через COM-порт в реальном масштабе времени. Целью разработки программного обеспечения является получение зависимости интегральной оценки качества случайной последовательности от параметров конфигурации аппаратного ГСЧ.

Если критерии T_1-T_{n-1} подтверждают, что последовательность ведет себя случайным образом, это еще не означает, вообще говоря, что проверка с помощью критерия T_n будет успешной. Однако каждая успешная проверка дает все больше и больше уверенности в случайности последовательности. Обычно к последовательности применяется около полудюжины критериев, и если она удовлетворяет им, то считается случайной.

4. Методика проведения экспериментальных исследований свойств аппаратного генератора случайных чисел

Для реализаций аппаратного генератора с различными параметрами с помощью разработанного программного обеспечения для анализа случайных последовательностей была проведена серия экспериментов, целью которых являлось:

– сбор информации о корреляции между качеством работы генератора и его характеристиками;

– построение аналитической модели этой зависимости;

– выявление факторов, наиболее влияющих на эту зависимость.

Каждый эксперимент проводился по следующей схеме:

На первом шаге с помощью разработанного программного обеспечения, генерирующего VHDL генератора, получали код генератора с необходимыми параметрами. Данное приложение позволяет получить код генератора с заданным количеством петель и их длин, а также задать произвольную их длину.

На втором шаге с помощью программного пакета Quartus генерировали файл прошивки.

На третьем шаге с помощью утилиты Programmer из пакета Quartus ПЛИС программировали для работы в качестве генератора случайных чисел

На четвертом шаге с помощью разработанного ПО для оценки статистических свойств случайных последовательностей проводили анализ – по 100 последовательностей по 1024 4-байтных чисел каждая для каждой модификации генератора.

Результаты всех проведенных экспериментов заносили в сводную таблицу, которую анализировали с помощью статистических пакетов. Под результатом следует понимать сводный рейтинг, вычисленный на основании прохождения последовательностью формальных и статистических тестов, подробно описанных в разделе 2, максимальное значение которого составляет 1; за прохождение каждого из шести статистических тестов рейтинг последовательности увеличивался на 0,13, оставшаяся величина определялась по результатам вычисленной доверительной вероятности для формального критерия.

Для удобства анализа результаты всех измерений были занесены в таблицу с тремя независимыми – количество петель, максимальное и минимальное количество инверторов в одной петле – столбцами, и одним зависимым – вычисленный рейтинг для последовательности, порожденной генератором с заданными параметрами. Всего было проведено 128 серий испытаний по 100 последовательностей каждая.

Следует отметить, что существует множество методов аппроксимации (приближение функций многочленами, формула Тейлора, приближение функций тригонометрическими многочленами).

Поскольку необходимая функция имеет три аргумента, применение таких методов потребует большое количество вычислительных ресурсов и времени. Для решения такой задачи наиболее рациональным будет применение нейронных сетей [3], как наиболее перспективное, простое в реализации и наглядное средство для анализа сложных закономерностей.

Для анализа статистических данных использовались программные пакеты Deductor Academic и STATISTICA 6.

Для проверки адекватности обучения сети все данные были разбиты на обучающее (95%) и тестовое (5%) множества.

В качестве активационной функции была выбрана одна из самых распространенных – сигмоида.

В качестве метода обучения нейронной сети был выбран алгоритм Back Propagation.

После выбора параметров было проведено обучение сети и проверка качества обучения на тестовом наборе. Предсказанные таким образом значения с высокой точностью совпадали с экспериментально полученными данными тестового множества, что дает возможность утверждать правильность выбора структуры сети, активационной функции и подбора параметров сети на этапе обучения.

5. Выводы

Оценка полученных результатов проводилась по двум направлениям – выявление общих свойств и закономерностей и построение аналитической модели.

Анализируя полученные результаты, можно сделать следующие выводы:

– случайность в большей мере зависит от длины цепочек, чем от их количества;

– случайность возрастает при увеличении длины цепочек и в меньшей мере – при увеличении их числа;

– целесообразно использовать цепочки одинаковой длины, так как разброс в длинах цепочек снижает случайность.

Экспериментальные исследования проводились с применением ПЛИС архитектуры FPGA двух ведущих фирм-производителей – Altera ACEX1K и Xilinx Spartan 3E. Результаты экспериментальных исследований показали практическую идентичность параметров генераторов случайных чисел с одинаковыми параметрами. Это дает возможность утверждать, что данные зависимости являются общими и платформенно-независимыми.

Список литературы: 1. *Рябко Б.Я., Фионов А.Н.* Криптографические методы защиты информации. М.: Горяч.Линия-Телеком, 2005. 229 с. **2.** *Харин Ю.С., Ярмола А.Н., Петлицкий А.И.* Методы и алгоритмы статистического тестирования генераторов случайных и псевдослучайных последовательностей в системах информационной безопасности // Штучний інтелект. 2006. Вып.3. С. 793-803. **3.** *Руденко О.Г., Бодянский Е.В.* Основы теории искусственных нейронных сетей. Харьков: ТЕЛЕТЕХ, 2002. 317 с.

Поступила в редколлегию 12.10.2011

Саранча Сергей Николаевич, канд. техн. наук, доцент кафедры ЭВМ ХНУРЭ. Научные интересы: проектирование и моделирование цифровых систем. Увлечения и хобби: интернет-программирование, музыка. Адрес: Украина, 61000, Харьков, пр.Ленина, 14, корпус «з», ауд. 35. Контактный телефон (+38057)702-13-54, e-mail: softpro@kture.kharkov.ua