УДК 004.056

МЕТОДЫ ГЕНЕРАЦИИ ПСЕВДОСЛУЧАНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ И ОЦЕНКА ИХ СВОЙСТВ

А.А.ЗАМУЛА, Д.А. СЕМЧЕНКО

Рассмотрены некоторые методы генерации псевдослучайных последовательностей (ПСП), приведены оценки их свойств на случайность с использованием Diehard тестов. На основе математических моделей методов генерации ПСП и результатов исследований статистических свойств последовательностей посредством программной реализации, описаны преимущества и недостатки каждого из рассматриваемых методов. Сделаны выводы о необходимости использования перспективных методов генерации ПСП, обеспечивающих реализацию требований к последовательностям с точки зрения криптографических свойств формируемых последовательностей символов.

Ключевые слова: генератор, псевдослучайная последовательность, метод, конгруэнтный, Xorshift, Фибоначчи, Diehard, сид.

ВВЕДЕНИЕ

На сегодняшний день разработано большое количество различных видов генераторов псевдослучайных последовательностей (ПСП). Одной из основных проблем при генерации ПСП, в частности, для криптографических приложений, является поддержание определенных криптографических свойств. При генерации последовательности чисел $X = x_1, x_2, ...$ необходимо убедиться, что случайная величина X обладает равномерным законом распределения, её реализации случайны и независимы [1]. Для проверки гипотезы о законе распределения используются статистические критерии χ^2 Пирсона, Колмогорова-Смирнова, Мизеса ω^2 [2] и др.

Математическая модель некоторых генераторов ПСП, использует: множество целых чисел Z, представленных 32-х разрядными машинными словами; сид (под сидом понимается начальное значение генератора) z; а также функциию f, определенную на множестве Z [3]. При случайном выборе сида z из Z, последовательность псевдослучайных чисел может быть получена путем многократного применения функции f:

$$f(z), f^2(z), f^3(z)...,$$

где $f^2(z)$ означает $f(f(z))$, а $f^3(z)$ означает $f(f^2(z))$ и так далее.

Однако, для генераторов ПСП, которые отвечают более строгим требованиям (большой период, m_i-ичное основание алфавита, структурная скрытность), множество Z представляется множеством всех m-кортежей $(x_1, x_2, ..., x_m)$ 32-х разрядных целых чисел, а функция f преобразует один из таких m-кортежей в другой.

Если f является биективной функцией (функция $f: X \to Y$ является биективной тогда и только тогда, когда существует обратная функция $f^{-1}: Y \to X$ такая, что $\forall x \in X$, $f^{-1}(f(x)) = x$ и $\forall y \in Y$, $f^{-1}(f(y)) = y$) определенной над множеством Z, то для любого сида z выбранного равномерно из Z, случайная переменная $f(z)$, полученная многократным применением функции f: $f(z), f^2(z), f^3(z)...$ также будет равномерно распределена по Z.

1. КОНГРУЭНТНЫЙ МЕТОД ГЕНЕРАЦИИ ПСП

Наиболее распространенным методом генерации ПСП является конгруэнтный метод, который для генерации ПСП использует правило:

$$x_n = (ax_{n-1} + k) \bmod m,$$

где m – модуль, a – множитель, k – аддитивная константа и x_0 - инициализирующий случайный сид.

Если a является примитивным корнем простого числа p, а x_0 является случайным сидом из множества $Z = \{1, 2, 3, ..., p-1\}$, то последовательность, порожденная $x_n = (ax_{n-1} + k) \bmod m$, будет строго периодическая, с периодом $p-1$, и каждый элемент этой последовательности будет равномерной случайной величиной на множестве Z, однако при этом элементы такой последовательности не будут независимыми между собой.

Если взять четыре последовательных числа x, y, z и w, сгенерированных линейным конгруэнтным генератором ПСП с множителем a, то [4-5]:

– точка (x, y, z) попадает на решетку точек, сгенерированных линейных комбинаций точек $(1, a, a^2)$, $(0, m, 0)$, $(0, 0, m)$ с целочисленными коэффициентами;

– аналогично, любая точка (x, y) будет попадать на решетку сгенерированную всеми линейными комбинациями точек $(1, a)$ и $(0, m)$ с целочисленными коэффициентами;

– точка (x, y, z, w) в четырехмерном пространстве будет попадать на решетку целочисленных комбинаций четырех точек $(1, a, a^2, a^3)$, $(0, m, 0, 0)$, $(0, 0, m, 0)$, $(0, 0, 0m)$.

Предположим, что α, β, τ – какие-либо три точки на плоскости с координатами, которые являются последовательными результатами генератора ПСП. Тогда определитель матрицы 2×2 с рядами $(\beta - \alpha)$ и $(\tau - a)$ будет давать объем

параллелепипеда, определенного этими тремя точками и объем параллелепипеда должен быть кратен m. Таким образом, НОД пяти или шести таких определителей будет равен m и, в этом случае, можно определить аддитивную константу k и множитель a.

Гистограмма распределения элементов последовательности, сгенерированной конгруэнтным генератором ПСП, представлена на рис. 1.

Рис. 1. Гистограмма распределения элементов последовательности конгруэнтного генератора ПСП

На рис. 1 по оси X - EBCDIC (*Extended Binary Coded Decimal Interchange Code*) коды символов 00h,01h,02h,...0Ah,0Fh,10h, 11h,..., FEh,FFh, а по оси Y – частота появления символа в последовательности, выраженная в процентах.

Проведенное моделирование конгруэнтного генератора ПСП, показало, что такой генератор не отвечает требованиям определенным тестом DIEHARD [6], в частности, тестам: BINARY RANK TEST для 6x8 матриц, BITSTREAM TEST (OVERLAPPING 20-кортежей), OPSO, OQSO и DNA, COUNT-THE-1's TEST.

2. XORSHIFT МЕТОД ГЕНЕРАЦИИ ПСП

Рассматривая 32-х (или 64-х) битное целое число как элемент векторного пространства с компонентами в поле $\bmod 2$, сложение двух векторов может быть реализовано как исключающее или (XOR) операция, что в сочетании с операцией сдвига, может быть использовано для создания некоторых линейных преобразований над этим векторным пространством. Множество сидов Z является множеством всех ненулевых двоичных векторов 1×32, а f является линейной трансформацией на множестве Z, представленной двоичной матрицей T 32×32, где T – невырожденная.

Для случайного сида $y\in Z$ последовательность $yT, yT^2, yT^3...$ будет иметь период $2^{32}-1$ тогда и только тогда, когда порядок матрицы T равен $2^{32}-1$ в группе 32×32 невырожденных двоичных матриц. Если $T=(I+L^a)\cdot(I+R^b)\cdot(I+L^c)$, где L – матрица, которая дает сдвиг влево на 1 (в C, $y^{\wedge}=(y<<1)$, таким образом yL^a в C – есть $y^{\wedge}=(y<<a)$, то можно получить простой и быстрый способ для формирования матричного произведения [7]. Матрица R, являющаяся транспонированной L, дает сдвиг вправо на единицу. Таким образом, для $T=(I+L^a)\cdot(I+R^b)\cdot(I+L^c)$, для случайного 32-битного сида y из Z, каждый новый y в последовательности $yT, yT^2, yT^3...$ может быть получен посредством последовательного применения следующих трех инструкций: $y^{\wedge}=y<<13$, $y^{\wedge}=y>>17$, $y^{\wedge}=y<<5$.

Для 32-х (или 64-х) битных двоичных векторов отсутствуют двухсдвиговые матрицы $T=(I+L^a)\cdot(I+R^b)$, которые имеют полный период и нет односдвиговых, поэтому необходима 3-х сдвиговая матрица T. Существует точно 81 тройка [7] $[a,b,c]$, $a<c$, для которых 32×32 двоичная матрица $T=(I+L^a)\cdot(I+R^b)\cdot(I+L^c)$ имеет порядок: $2^{32}-1$.

Если матрица $T=(I+L^a)\cdot(I+R^b)\cdot(I+L^c)$ имеет полный период, то и $(I+L^c)\cdot(I+R^b)\cdot(I+L^a)$ и $(I+L^a)\cdot(I+L^c)\cdot(I+R^b)$ также имеют полный период, что дает возможность получить 4×81 матриц T с порядком $2^{32}-1$. Но тогда и транспонированная матрица каждой из них имеет полный период. Таким образом, могут быть получены 648 матриц.

Гистограмма распределения элементов последовательности, сгенерированной Xorshift генератором ПСП, представлена на рис. 2.

Рис. 2. Гистограмма распределения элементов последовательности Xorshift генератора ПСП

Проведенное моделирование Xorshift генератора ПСП, показало, что такой генератор не отвечает требованиям определенным тестом DIEHARD, в частности таким, как: BINARY RANK TEST для 6x8 матриц, OPSO, OQSO, COUNT-THE-1's TEST.

3. МЕТОД ГЕНЕРАЦИИ ПСП ФИБОНАЧЧИ С ЗАПАЗДЫВАНИЯМИ

Базовое рекуррентное соотношение для генераторов ПСП Фибоначчи [1] с запаздываниями представляется в виде $x_n=x_{n-r}\bullet x_{n-s}$ для r и s при $r>s$ [7]. Операция $\bullet$ определяется как бинарные отношения для пар элементов в некотором множестве χ, и множество сидов Z представляет собой множество r-кортежей $(x_1,x_2,...,x_r)$, где $x\in\chi$. Обычно χ является множеством 32-х битных целых чисел, а операция $\bullet$ является сложением или вычитанием по $\bmod 2^{32}$. Правило для $\bullet$ может быть основано на выражении элементов x, y из χ в форме $x=\pm 3^a, y=\pm 3^b \mod 2^{32}$ так, что $x\bullet y=\pm 3^{(a+b)\bmod 2^{30}}$ и верно рекуррентное правило для сложения $\bmod 2^{30}$. Для генераторов ПСП Фибоначчи с запаздываниями используется обозначение $F(r,s,\bullet)$. Для правильного выбора запаздываний r, s период $F(r,s,\pm\bmod 2^{32})$ должен быть 2^{32+r}, в то время как $F(r,s,\oplus)$ будет 2^r не зависимо от размера слова. Для правильного выбора $r>s$ период генератора $F(r,s,*$ нечетных по модулю $2^{32})$ будет 2^{30}.

В отличии от конгруэнтного и Xorshift методов генерации ПСП, метод генерации ПСП Фибоначчи с запаздываниями требует множество

сидов Z и функцию f, определенную над Z. Под Z в данном случае подразумевается множество кортежей $\{[x_1, x_2, ...x_r]\}$, где x из множества χ, на котором определено бинарное отношение и функция f :. $f([x_1, x_2, ..., x_r]) = [x_2, ..., x_r, x_1 \bullet x_{r-s+1}]$..

Реализация последовательностей Фибоначчи, с запаздываниями $r > s$ требует хранения таблицы последних r значений. Одно из самых полезных применений генератора ПСП Фибоначчи с запаздываниями является непосредственная генерация равномерных случайных чисел с плавающей запятой на интервале [0,1) .

Предположим, необходимо сгенерировать 64-х битные равномерные [0,1) случайные переменные, используя IEEE 754 стандарт (стандарт формата представления чисел с плавающей точкой, используемый как в программных реализациях арифметических действий, так и аппаратных реализациях): 1 знаковый бит, 11 битов экспоненты, 52 бита мантиссы с подразумеваемой ведущей единицей. Для указанного бинарного отношения $x \bullet y$ используют правило: если $x \geq y$, тогда $x - y$, в противном случае – $x - y + 1$. Если x и y – представление рациональных чисел $a / 2^{53}$ и $b / 2^{53}$ с плавающей запятой, то $x \bullet y$ будет давать в результате точное значение $c / 2^{53}$ с плавающей запятой, где $c = (x - y) \bmod 2^{53}$.

Гистограмма распределения элементов последовательности, сгенерированной генератором ПСП Фибоначчи с запаздываниями, представлена на рис. 3.

Рис. 3. Гистограмма распределения элементов последовательности сгенерированной генератором ПСП Фибоначчи с запаздываниями

Проведенное моделирование генератора ПСП Фибоначчи с запаздываниями, показало, что такой генератор не отвечает требованиям определенным тестом DIEHARD, в частности требованиям: OPSO, OQSO, DNA.

ВЫВОДЫ

Результаты проведенных исследований различных генераторов ПСП показали, что метод генерации Xorshift является более предпочтительным, чем конгруэнтный метод. Это обосновывается более простой технической реализацией и лучшей производительностью [8], а также свойствами случайности генерируемых последовательностей (показателями прохождения статических тестов DIEHARD). При использовании конгруэнтного метода достаточно легко определить аддитивную константу k и множитель a, поэтому этот метод генерации нельзя считать криптостойким, однако благодаря его простоте, он вполне может использоваться в открытых системах.

Анализ методов генерации ПСП показывает, что для обеспечения криптографических свойств ПСП, необходимо, чтобы сид был выбран независимо и базировался на случайных явлениях, либо процессах. Методы, обеспечивающие реализацию криптографических свойств ПСП, будут рассмотрены в следующих статьях.

Литература

[1] *Б. Шнаер*. Прикладная криптография / *Б. Шнаер* // 2-е издание.

[2] *Вентцель Е.С.*, Овчаров Л.А. “Теория вероятностей и ее инженерные приложения” 2-е изд. М.: Высшая школа, 2000.— 480 с.

[3] *Marsaglia, G* (2003). Seeds for random number generators.

[4] *Marsaglia, G* (1970). Regularities in congruential random number generators. Numerische Mathematic 16, 8-10.

[5] *Marsaglia, G* (1972). The structure of linear congruential sequences. In Applications of Number Theory to Numerical analysis, Z.K. Zaremba, ed. Academic Press, 249-285;

[6] The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness // [Электронныйресурс]: http://www.stat.fsu.edu/pub/diehard/.

[7] *Marsaglia, G*, 2003, Random number generators, Journal of Modern Applied Statistical Methods, No. 2.

[8] *Дональд Кнут*. Искусство программирования / Дональд Кнут // Получисленные алгоритмы. The Art of Computer Programming. – Vol.2. Seminumerical Algorithms. – 3-е изд. – М.: «Вильямс»/ – 2007. – С. 832.

Поступила в редколлегию 23.04.2012

Замула Александр Андреевич, профессор кафедры БИТ ХНУРЭ, кандидат технических наук, доцент. Область научных интересов: технологии защиты информации в информационно-телекоммуникационных системах.

Семченко Денис Александрович, аспирант кафедры БИТ ХНУРЭ. Область научных интересов: разработка и применение методов генерации псевдослучайных последовательностей, тестирование программного обеспечения.

УДК 004.056

Методи генерації псевдовипадкових послідовностей та оцінка їх властивостей / О. А. Замула, Д.О. Семченко // Прикладна радіоелектроніка: наук.-техн. журнал. – 2012. – Том 11. № 2. – С. 191–194.

Розглянуті деякі методи генерації псевдовипадкових послідовностей (ПСП), наведено оцінки їх властивостей на випадковість з використанням Diehard тестів. На основі математичних моделей методів генерації ПСП та результатів досліджень статистичних власти-

востей послідовностей за допомогою програмної реалізації, описані переваги та недоліки кожного з розглянутих методів. Зроблені висновки про необхідність використання перспективних методів генерації ПСП, що забезпечують реалізацію вимог до послідовностей з точки зору криптографічних властивостей формованих послідовностей символів.

Ключові слова: генератор, псевдовипадкова послідовність, метод, конгруентний, алгоритм Xorshift, коди Фібоначчі, Diehard тести, сід.

Іл. 3. Бібліогр.: 8 найм.

UDC 004.056

Methods of generating pseudorandom sequences and evaluation of their properties / A. A. Zamula, D.A. Semchenko // Applied Radio Electronics: Sci. Journ. – 2012. Vol. 11. № 2. – P. 191–194.

The paper considers some methods for generating pseudo-random sequences (PRS), evaluates their properties with the Diehard tests. On the basis of mathematical models of PRS methods for generating and research results of the statistical properties of sequences through the program implementation the advantages and disadvantages of each of these methods are described. Conclusions about the need for promising methods of generating PRSs to ensure the satisfaction of the requirements for the sequences in terms of cryptographic properties of the generated sequences of symbols are made.

Keywords: generator, pseudorandom sequence, method, congruent, Xorshift algorithm, Fibonacci codes, Diehard test, seed.

Fig. 3. Ref.: 8 items.