ЛИНЕЙНЫЕ КОНГРУЭНТНЫЕ И РАНДОМИЗАЦИОННЫЕ ГЕНЕРАТОРЫ

Кулаков Игорь Анатольевич
http://random-art.ru/

Рассматриваются способы получения надежных в статистическом, а при необходимости и в криптографическом отношении двоичных последовательностей псевдослучайных чисел, для которых в качестве исходных используются последовательности, формируемые на примере линейного конгруэнтного метода с переменными коэффициентами и его расширений. За счет охвата обратными связями старших разрядов с младшими, формируемые последовательности даже на крайне малых платформах, начиная с 8-9 бит, отличаются высокими статистическими показателями, приобретают существенно выраженные нелинейные свойства, а с ними и приемлемые для прикладных приложений криптографические показатели.

Цель данной статьи показать на конкретных примерах пригодных для эффективной программной реализации, как по последовательностям, формируемым посредством линейного (смешанного) конгруэнтного метода по двоичному модулю и его расширений [1], получить двоичные последовательности псевдослучайных чисел, надежные в статистическом, а при необходимости и в криптографическом отношении.

Напомним, ***линейный смешанный конгруэнтный метод <u>по двоичному модулю</u>***, задается рекурсивным уравнением [2,3]:

$$x_i = a \cdot x_{i-1} + b \mod 2^n, \quad (1)$$

а его ***нелинейное расширение***, уравнением [1]:

$$x_i = a' \cdot (x_{i-1} \oplus b) \mod 2^n, \quad (2)$$

с побитовой операцией сложения по модулю 2 – $\oplus$, n-битовой двоичной образующей x переменной и постоянными коэффициентами – нечетном приращении b, множителем $a \equiv 1 (mod\ 4)$ и $a' \equiv 3 (mod\ 4)$, соответственно. Формируемые таким образом двоичные конгруэнтные последовательности (1) и (2) имеют максимальный период $T_n = 2^n$, и в пределах его бесповторны.

Существенным недостатком двоичных конгруэнтных последовательностей данного типа является, как показывает элементарный анализ, частотно регулярный, с периодом 2^k, характер изменения $k = \overline{1,n}$ битов, составляющих разряды образующей x переменной.

Указанный недостаток может быть устранен за счет использования функций усложнения $r = R(x)$ образующей x переменной. Различают биективные, т.е. однозначно обратимые, и сюръективные, однозначно необратимые функции. При использовании функций усложнения биективного типа, в отличие от функций усложнения сюръективного типа, результирующая последовательность сохраняет присущие исходной последовательности бесповторные свойства. В силу особой специфики данной проблемы, главным образом связанной с генерацией паролей, идентификаторов и прочих бесповторных ключей и потоков, рассмотрение функций усложнения биективного типа выходят за рамки данной работы.

Среди множества функций сюръективного типа будем опираться на функции, биты результирующих переменных которых носят равночастотный характер, с законом распределения неотличимым по данным статистических тестов от идеального равномерного. Далее такие функции будем именовать рандомизационными функциями.

Для статистического анализа будем использовать пакет тестов DIEHARD [4], как наиболее сильный.

В качестве примеров ***реализации функций усложнения***, воспользуемся следующими, полученными по результатам исследований [5], рандомизационными функциями:

$$r_i = g_{i-1} \oplus x_{i-1}, \quad g_i = rotL(x_{i-1},\hbar_n) \oplus rot(g_{i-1},\eta_n), \tag{3}$$

$$r_i = q_{i-1}, \quad q_i = g_{i-1} \oplus rotL(x_{i-1},\hbar_n), \quad g_i = q_{i-1} \oplus rot(g_{i-1},\eta_n), \tag{4}$$

с n-битовой результирующей r, образующей x и промежуточными $\{g,q\}$ переменными, операциями циклического сдвига влево – $rotL$ и $rot = \{rotL, rotR\}$ – влево или вправо, подбираемой эмпирически по результатам статистики, на $\hbar_n$ и η_n значащих бит, соответственно.

Биты результирующей r переменной, формируемой на основе указанных функций, носят упомянутый выше равночастотный характер, при смещении $\hbar_n$, равным ближайшему простому к n/e, e = 2,718 281… – число Эрмита, не кратному n, и смещении η_n, равным максимальному, меньшему $n/2$, взаимно простому с n. При равенстве смещений, характеристическое смещение η_n понижается до ближайшего, взаимно простого с n.

Для сравнительного анализа статистических свойств самых старших разрядов конгруэнтных последовательностей (1) и (2), положим $b' = b = 1$,

$$a = m_o \oplus 2, \quad a' = m_o \oplus 4, \quad \text{при} \quad m_o = (2^n - 1)/2^{n/2}. \tag{5}$$

Статистический анализ на основе пакета тестов DIEHARD [4], формируемых на основе уравнений (1) и (2) конгруэнтных последовательностей, при значениях входящих в их состав коэффициентов [1], установленных исходя из соотношений (5), показывает статистическую надежность выборки, полученной по самому старшему биту, начиная с платформ, разрядностью $n = 40$ и $n = 36$ бит, соответственно.

При введении функций, задаваемых соотношениями (3) / (4), в состав уравнений (1) и (2), результирующие r переменные носят по каждому своему биту статистически надежный характер, начиная с платформ, разрядностью $n = 25/17$(с гарантией 18) и $n = 23/17(18)$ бит, соответственно.

Линейный конгруэнтный метод с постоянными коэффициентами (2) и его расширение (3), допускают обобщение на случаи, когда коэффициенты $\{a,b\}$, входящие в состав указанных уравнений, являются переменными [1].

Линейный конгруэнтный метод с переменными коэффициентами задается уравнением:

$$x_i = A_{i-1} \cdot x_{i-1} + B_{i-1} \ mod\ 2^n, \tag{6}$$

а его ***нелинейное расширение***, уравнением:

$$x_i = A'_{i-1} \cdot (x_{i-1} \oplus B_{i-1}) \ mod\ 2^n, \tag{7}$$

с n-битовой образующей x переменной, при $A_{i-1} \equiv 1(mod\ 4)$ и $A'_{i-1} \equiv 3(mod\ 4)$, для всех i, с приращением, устанавливаемым исходя из побитовых операций $\vee$ – OR и $\wedge$ – AND:

$$B_{i-1} = (B^*_{i-1} \vee 3) \oplus b_o, \quad \text{при} \quad b_o = const = \{0,2\}. \tag{8}$$

Двоичные конгруэнтные последовательности (6) и (7) имеют максимальный период $T_n = 2^n$, если образующая последовательность B^*, есть смещенная двоичная конгруэнтная последовательность не менее чем на один бит в сторону старших разрядов, а последовательности $\{A,A'\}$ формируются исходя из условий:

$$A_{i-1} = A^*_{i-1} \vee a_o, \qquad A'_{i-1} = A^*_{i-1} \vee a'_o, \tag{9}$$

при константах $a_o = 1$ и $a'_o = 3$, с образующей последовательностью A^*, являющейся смещенной двоичной конгруэнтной последовательностью, не менее чем на два бит.

Для сравнительного анализа статистических свойств конгруэнтных последовательностей (6) и (7), к примеру, как это было сделано ранее в [1], зададим образующие последовательности $\{A^*,B^*\}$ изменения коэффициентов инкрементными счетчиками, вида:

$$A^*_i = A^*_{i-1} + H_{A,} \qquad B^*_i = B^*_{i-1} + H_{B,} \tag{10}$$

с нечетными приращениями $\{H_A,H_B\}$, соответственно.

Статистический анализ показывает статистическую надежность выборок (6) и (7), по-

лученных по самому старшему биту, при использовании образующей последовательности B^* смещенной на два или на один бит, начиная с платформ, разрядностью $n = 37$ и $n = 26$, $n =$ 32(с гарантией 33) и $n = 26$ бит, соответственно.

При введении функций, задаваемых соотношениями (3) / (4), в состав уравнений (6) и (7), результирующие r переменные носят в целом и по каждому отдельно взятому биту, как и в случаях (1) и (2), статистически надежный характер, начиная с платформ, разрядностью $n = 25/17$(с гарантией 18) и $n = 23/17(18)$ бит, соответственно.

Отдельно проведенные исследования показывают, что предположительный период формируемых таким образом последовательностей составляет порядка $e^n \approx 2^{1.443 \cdot n}$.

Общая длина ключа генерации при использовании функций усложнения (3) и (4) составляет $5 \cdot n - 9$ и $6 \cdot n - 9$ бит, соответственно.

Характерной особенностью двоичных конгруэнтных последовательностей, формируемых в соответствии с уравнениями (1), (2) и (:6), (7), является строго однонаправленный в сторону старших разрядов, регулярный характер изменения битов образующей x переменной. В соответствии с этим, представляемый метод относится к ***регулярному рандомизационному способу*** генерации псевдослучайных двоичных последовательностей [5]. К тому же необходимо отметить, что для регулярного метода характерно неравнозначность вкладов битов ключа генерации, а это может негативно сказаться на общей стойкости построенного криптографического алгоритма.

Типовой алгоритм и тестовый пример генерации псевдослучайных последовательностей на основе регулярного рандомизационного способа приведен в Приложении, на Рис.1, для 5-ти разрядной платформы.

В целях придания результирующим последовательностям существенно выраженных лавинных и нелинейных свойств, а также для достижения равнозначных вкладов битов ключа генерации, старшие биты охватываются обратными связями с младшими битами. Подобные методы относятся к нерегулярному рандомизационному способу генерации псевдослучайных двоичных последовательностей [5].

Нерегулярный рандомизационный способ может быть представлен следующими уравнениями:

$$r_i = g_{i-1} \oplus x_{i-1}, \quad g_i = rotL(x_{i-1}, \hbar_n) \oplus rot(g_{i-1}, \eta_n), \tag{11}$$
$$x_i = A_{i-1} \cdot g_{i-1} \,\Delta\, B_{i-1} \ \ mod\ 2^n,$$
$$r_i = q_{i-1}, \quad q_i = g_{i-1} \oplus rotL(x_{i-1}, \hbar_n), \quad g_i = q_{i-1} \oplus rot(g_{i-1}, \eta_n), \tag{12}$$
$$x_i = A_{i-1} \cdot g_{i-1} \,\Delta\, B_{i-1} \ \ mod\ 2^n,$$

с одной из операций сложения $\Delta = \{+, \oplus\}$, нечетными коэффициентами $\{A, B\}$, для всех i, и двоичными образующими конгруэнтными последовательностями $\{A^*, B^*\}$, смещенными не менее чем на один бит в сторону старших разрядов. Формируемые на основе уравнений (11) и (12) последовательности не вырождаются, в силу периодического изменения коэффициентов $\{A, B\}$. Предположительный период формируемых таким образом последовательностей составляет порядка $e^{2 \cdot n} \approx 2^{2.885 \cdot n}$.

Для сравнительного анализа статистических свойств рандомизационных последовательностей (11) и (12), зададим образующие последовательности $\{A^*, B^*\}$изменения коэффициентов, как это было сделано выше, инкрементными счетчиками (10), с нечетными приращениями $\{H_A, H_B\}$. Общая длина ключа генерации при использовании функций усложнения (3) и (4) составляет $6 \cdot n - 6$ и $7 \cdot n - 6$ бит, соответственно.

Анализ показывает статистическую надежность результирующих r переменных, формируемых в соответствии с уравнениями (11) и (12), в целом и по каждому отдельно взятому биту, начиная с платформ, разрядностью $n = 10$ и $n = 8(9)$ бит, соответственно.

Типовой алгоритм и тестовый пример генерации псевдослучайных последовательностей на основе нерегулярного рандомизационного способа приведен в Приложении, на Рис.2, для 5-ти разрядной платформы.

И в заключение, для подробного исследования и анализа приведем вариант реализации нерегулярного рандомизационного способа, построенного на основе ***дуального рандомизационного метода***, приведенного в статье [1]:

$$r_i = u_{i-1} \oplus v_{i-1}, \qquad u = rotL(u_{i-1}, \hbar_n), \qquad v = rotL(v_{i-1}, \eta_n)\ , \tag{13}$$
$$u_i = (a_{\mathrm{u}}\ \Delta\ 4{\cdot}v)\cdot(u\ \Delta\ b_{\mathrm{u}})\ \ mod\ 2^{\,n}, \qquad v_i = (a_{\mathrm{v}}\ \Delta\ 4{\cdot}u)\cdot(v\ \Delta\ b_{\mathrm{v}})\ \ mod\ 2^{\,n},$$

с нечетными коэффициентами $\{a_{\mathrm{u}}, b_{\mathrm{u}}, a_{\mathrm{v}}, b_{\mathrm{v}}\}$.

Предположительный период формируемых таким образом последовательностей составляет порядка $n{\cdot}e^n \approx n{\cdot}2^{1.443{\cdot}n}$. Общая длина ключа генерации составляет $6{\cdot}n-4$ бит.

Анализ показывает статистическую надежность результирующей r переменной, в целом и по каждому отдельно взятому биту, начиная с платформ, разрядностью $n=14$ бит.

Общие выводы.

1. Линейный конгруэнтный метод и его расширения позволяет посредством элементарных функций, исполняемых параллельно, получить надежные в статистическом отношении результаты, в целом и по каждому отдельно взятому биту результирующей переменной.

2. Криптографический анализ рандомизационных последовательностей, в зависимости от сложности изменения коэффициентов входящих в состав уравнений (6) и (7) существенно затруднен, проблематичен или технически невозможен. К этому, согласно с приведенным анализом, нелинейные расширения по статистическим свойствам и криптографическим показателям заметно превосходят свои линейные аналоги.

3. Посредством введения обратных связей, последовательности, формируемые в соответствии с уравнениями (11)-(13), характеризуются существенно выраженными лавинными и нелинейными свойствами и высокими статистическими показателями.

4. В отличие от последовательностей, формируемых посредством уравнения (13), формируемые на основе уравнений (11) и (12) последовательности не вырождаются, в силу периодического изменения коэффициентов $\{A,B\}$. Оценка вероятности вырождения в случае (13), неизвестна, но если судить по аналогии с блочными шифром ГОСТ 28147-89 используемым в режиме обратной связи по выходу, крайне низка.

5. Для увеличения периода, равно как и для предупреждения вырождений, следует использовать k-разрядные регистры сдвига с линейной обратной связью (LFSR). При комплексировании с ними, длина периода увеличивается в (2^k-1) раз.

Как видим, ***потенциальные возможности линейного конгруэнтного метода и его разновидностей далеко не исчерпаны, а представляемые ими алгебраические системы еще подобающим образом не изучены***.

К сожалению, к этому можно добавить, что ***линейный конгруэнтный метод и его разновидности не так просто и эффективно устроены, как это кажется, особенно для аппаратной реализации***. В первую очередь это связано с использованием в уравнениях генерации псевдослучайных последовательностей высоко затратной операции умножения и даже, относительно затратной операции сложения, на которые приходится большая доля вычислительных, аппаратных и энергетических затрат.

Ко всему, следует помнить, что арифметические операции не допускают распараллеливание вычислений по каждому из отдельно взятых разрядов, что крайне негативно сказывается на конечной производительности рандомизационных генераторов, особенно при аппаратной реализации, и в силу чего, при крайне ограниченном ресурсе, выводит их больше в академическую, нежели чем в практическую плоскость исследований.

Разрешить указанные проблемы и довести формирование двоичных последовательностей, аналогичным представленным двоичным конгруэнтным методом, до математической прозрачности, должного прикладного и эффективного практического результата, ***оказалось возможным благодаря введению неполной арифметики (предарифметики) и развитию представляемых ими стохастических технологий***: http://t.random-art.ru/recommendation/

Ключевые слова: линейный, квадратичный, кубичный, полиномиальный, нелинейный, конгруэнтный, метод, генератор, генерация, случайных, псевдослучайных, чисел, линейная, квадратичная, кубичная, полиномиальная, нелинейная, конгруэнтная, дихотомическая, последовательность, регулярный, нерегулярный, рандомизационный, стохастический, способ, функция, симметричная, криптография.

Литература

1. Кулаков И.А. Полиномиальный конгруэнтный метод с переменными коэффициентами и его нелинейные расширения.
   Рукопись статьи, Москва, 2012, http://random-art.ru/?download=LCM_ru.pdf
2. Кнут Дональд Э. Искусство программирования.
   Третье издание, Том 2, М.: Издательский дом “Вильямс”, 2002.
3. Шнайер Брюс. Прикладная криптография.
   Изд. ТРИУМФ, Москва, 2002.
4. Marsaglia G. DIEHARD Tests, 1997, http://en.wikipedia.org/wiki/diehard_tests/
   A Statistical Test Suite for the Validation of Pseudorandom Number Generators.
   NIST Special Publication 800-22, (FIPS PUB 140-1,2). NIST, 2001.
5. Кулаков И.А. Рандомизационный способ.
   Москва, 2011, http://random-art.ru/randomizacionnii-sposob/

МОСКВА, март 2012

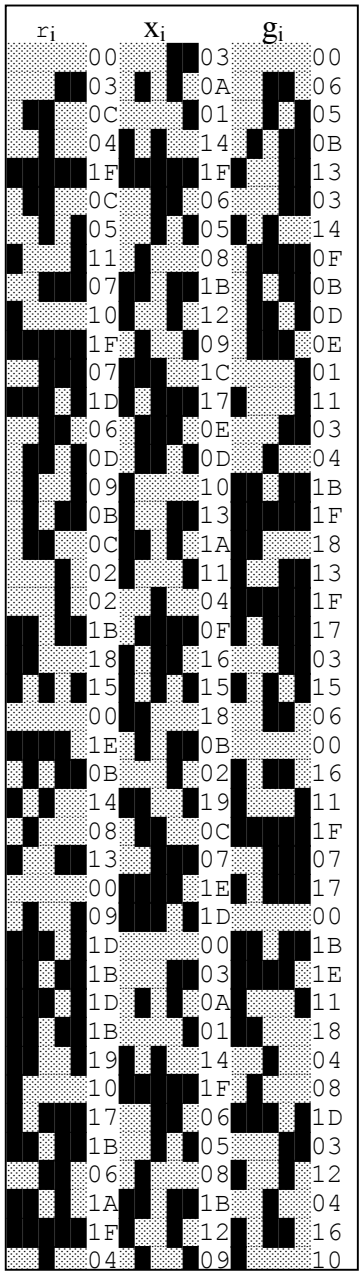

Рис.1

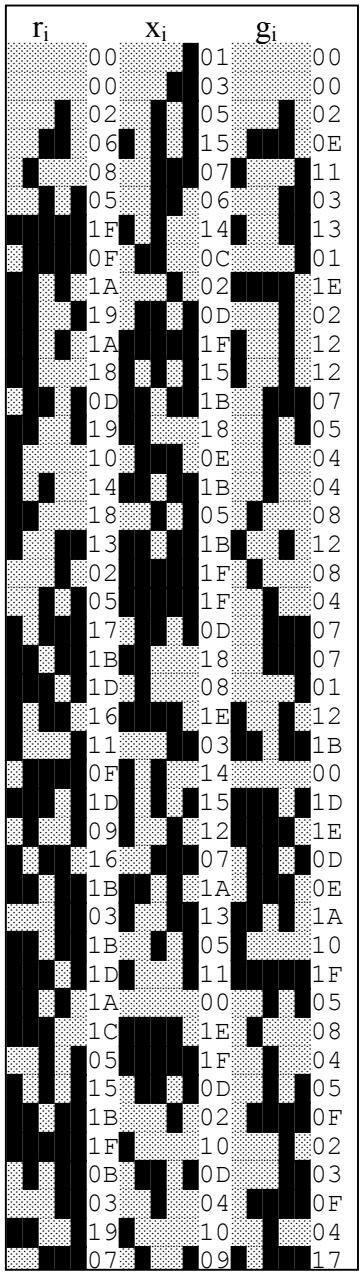

Рис.2

Исходные данные:
n – длина платформы генерации, бит.
Rx, Ra, Rb, Rg, Rq – задающие параметры генерации.
RHa, RHb – компоненты ключа генерации.
Ha, Hb – приращения коэффициентов генерации.
r, x – результирующая и образующая двоичные переменные генератора.
g, q – промежуточные двоичные переменные генератора.
a, b – множитель и приращение.

Типовой пример реализации
регулярного рандомизационного способа (период $\approx 2^{1.443 \cdot n}$)

Синхронизация параметров генератора (длина ключа = 6·n – 9, бит):
x = Rx; g = Rg; a = Ra | 3; b = Rb | 1; Ha = (RHa | 7) & 4; Hb = (RHb | 7) & 4;

Формирование очередного элемента рандомизационной последовательности:
r = g ^ x; g = RotL(x, $\hbar$) ^ RotR(g, η);
x = (a*(x ^ b)) & MOD_N; (MOD_N = $2^n - 1$)
a += Ha; b += Hb;

Тестовый пример генерации (см. эпюру на Рис.1), при n = 5.
Начальные условия: x = g = 0; a = 3; b = 1; Ha = Hb = 4; $\hbar$ = 1; η = 2;

Типовой пример реализации
нерегулярного рандомизационного способа (период $\approx 2^{2.885 \cdot n}$)

Синхронизация параметров генератора (длина ключа = 7·n – 6, бит):
x = Rx; g = Rg; q = Rq; a = Ra | 1; b = Rb | 1; Ha = (RHa | 3) & 2; Hb = (RHb | 3) & 2;

Формирование очередного элемента рандомизационной последовательности:
z = g; r = q; q = z ^ RotL(x, $\hbar$); g = r ^ RotL(z, η);
x = (a*z + b) & MOD_N; (MOD_N = $2^n - 1$)
a += Ha; b += Hb;

Тестовый пример генерации (см. эпюру на Рис.2), при n = 5.
Начальные условия: x = g = q = 0; a = b = 1; Ha = Hb = 2; $\hbar$ = 1; η = 2;