ПОЛИНОМИАЛЬНЫЙ КОНГРУЭНТНЫЙ МЕТОД С ПЕРЕМЕННЫМИ КОЭФФИЦИЕНТАМИ И ЕГО НЕЛИНЕЙНЫЕ РАСШИРЕНИЯ

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

Исследования показывают, что линейный (смешанный) конгруэнтный метод, как один из широко известных методов генерации псевдослучайных чисел, и представляемая им алгебраическая база, еще далеко не полно изучены и содержат в себе еще много нерешенных проблем. На это прямо указывает существование нелинейных расширений полиномиального (линейного, квадратичного, кубичного и более высоких порядков) конгруэнтного метода по двоичному модулю 2^n и возможность использования в его составе переменных коэффициентов. Причем, при соблюдении элементарных условий синхронизации коэффициентов, в дополнение и вопреки общепринятым на сегодня положениям теории чисел, период формируемых таких образом двоичных последовательностей максимален, и равен 2^n, а сами последовательности в пределах периода бесповторны.

Один из практикуемых методов генерации последовательностей псевдослучайных чисел, широко известный под названием ***линейный смешанный конгруэнтный метод*** [1,2], задается рекурсивным уравнением:

$$x_i = a \cdot x_{i-1} + b \mod m. \qquad (1)$$

В дальнейшем, ограничимся наиболее простым для реализации уравнением, с двоичным модулем $m = 2^n$, а именно:

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

с n-битовой двоичной x переменной, и постоянными коэффициентами – множителем $a \equiv 1 (mod\ 4)$ и нечетном приращении b.

Формируемые таким образом двоичные (конгруэнтные) последовательности имеют максимальный период $T_n = T_{\max} = 2^n$, и в пределах его бесповторны (апериодичны). Известно [3,4], что конгруэнтные последовательности данного типа, даже усеченные до одного старшего бита, не являются криптографически стойкими.

Линейный конгруэнтный метод (2), в связи с ограничением, накладываемым на второй бит множителя a, ***носит неполный характер***, и допускает простое развитие [5]:

$$x'_i = a' \cdot (x'_{i-1} \oplus b') \mod 2^n, \qquad (3)$$

с побитовой операцией сложения по модулю 2 – $\oplus$, при $a' \equiv 3 (mod\ 4)$ и нечетном приращении b' (собственно, $b' = b$).

В отличие от предшествующего ему линейного уравнения (2), данное ***уравнение*** (3), ***носит явно выраженный нелинейный характер*** [2]. В силу этого, в отличие линейного конгруэнтного метода (2), криптографический анализ таких последовательностей (3), в особенности сильно усеченных, на 64 и более младших значащих бит, уже затруднен [3,4].

Для сравнительного анализа статистических свойств самых старших разрядов конгруэнтных последовательностей (2) и (3), положим $b' = b = 1$,

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

Статистический анализ на основе пакета тестов DIEHARD [6], формируемых на основе уравнений (2) и (3) конгруэнтных последовательностей, при значениях входящих в их состав коэффициентов, установленных исходя из соотношений (4), показывает статистическую надежность выборки, полученной по самому старшему биту, начиная с платформ, разрядностью $n = 40$ и $n = 36$ бит, соответственно.

Линейный конгруэнтный метод с постоянными коэффициентами (2) и его расширение (3), допускают обобщение на случаи, когда коэффициенты $\{a,b\}$, входящие в состав указан-

ных уравнений, являются переменными. Не вдаваясь в детали, остановимся на вопросах, которые могут быть весьма поучительными и полезными для практики и теории.

Линейный конгруэнтный метод с <u>переменными коэффициентами</u> задается уравнением:

$$x_i = A_{i-1} \cdot x_{i-1} + B_{i-1} \ \ mod\, 2^n, \tag{5}$$

с n-битовой двоичной x переменной, при $A_{i-1} \equiv 1(mod\ 4)$, для всех i, а его ***нелинейное расширение***, аналогичным уравнением с множителем $A'_{i-1} \equiv 3(mod\ 4)$, вида:

$$x'_i = A'_{i-1} \cdot (x'_{i-1} \oplus B_{i-1}) \ mod\, 2^n, \tag{6}$$

с приращением, устанавливаемым исходя из побитовых операций $\vee$ – OR и $\wedge$ – AND:

$$B_{i-1} = (B^*_{i-1} \vee 3) \oplus b_o, \quad \text{при} \quad b_o = const = \{0,2\}. \tag{7}$$

Двоичные конгруэнтные последовательности (5) и (6) 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:

$$A_{i-1} = A^*_{i-1} \vee a_o, \qquad A'_{i-1} = A^*_{i-1} \vee a'_o, \tag{8}$$

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

В общем, криптографический анализ таких последовательностей, если исходить из сложности образующих коэффициенты конгруэнтных последовательностей, входящих в состав выражений (7) и (8), весьма проблематичен, а сильно усеченных, соизмерим с технически нереализуемым перебором.

Для наглядного примера, подтверждающего истинность слов, в Приложении приведены алгоритмы реализации типовых двоичных конгруэнтных генераторов с переменными коэффициентами, построенных в соответствии с уравнениями (5) и (6), с образующей последовательностью B^* смещенной на два бит. Общая длина ключа составляет около $5 \cdot n$ бит.

Статистический анализ на основе пакета тестов DIEHARD, показывает статистическую надежность выборки, полученной по самому старшему биту (5) и (6), начиная с платформ, разрядностью $n = 37$ и $n = 26$ бит, соответственно.

При использовании образующей последовательностью B^* смещенной на один бит, отмечается улучшение статистических показателей. Статистическая надежность выборки, полученной по самому старшему биту, достигается, начиная с платформ, разрядностью $n = 33$ и $n = 26$ бит, соответственно.

Из линейного конгруэнтного метода с переменными коэффициентами (5) и его нелинейного расширения (6) следует квадратичный конгруэнтный метод [1,3]. Рассмотрим одни из его наиболее легко реализуемых на практике характерных примеров.

Квадратичный конгруэнтный метод задается уравнением:

$$x_i = (a + 4 \cdot x_{i-1}) \cdot (x_{i-1} + b) \ \ mod\, 2^n, \tag{9}$$

с n-битовой двоичной x переменной, при $a \equiv 1(mod\ 4)$, а его ***нелинейное расширение***, аналогичным уравнением с множителем $a' \equiv 3(mod\ 4)$, вида:

$$x'_i = (a' \oplus 4 \cdot x'_{i-1}) \cdot (x'_{i-1} \oplus b) \ \ mod\, 2^n, \tag{10}$$

с нечетным приращением b.

Квадратичные конгруэнтные последовательности (9) и (10) имеют максимальный период $T_n = 2^n$ при любых начальных условиях $\{x_o, x'_o\}$. Общая длина ключа составляет $3(n-1)$ бит. Каждому ключу генерации соответствует уникальная квадратичная конгруэнтная последовательность.

Квадратичные конгруэнтные последовательности, формируемые в соответствии с уравнениями (9) и (10), не являются криптографически стойкими. К этому, если судить по материалам исследований [3,4], криптографический анализ усеченных последовательностей (9) и (10) возможен, но затруднен.

Статистический анализ на основе пакета тестов DIEHARD, показывает статистическую надежность выборки, полученной по самому старшему биту (9) и (10), причем, ***независимо от составляющих ключа генерации*** $\{x_o, a$ - множителя, b - приращения$\}$ или $\{x'_o, a', b\}$, начиная с платформ, разрядностью $n=34$ и $n=36$ бит, соответственно.

В отличие от квадратичного конгруэнтного метода (9), его нелинейное расширение (10) допускает ***существенное усиление по статистическим показателям***, для платформ, с разрядностью $n=27$ бит:

$$x'_i = (a' \oplus 4\cdot x'_{i-1})\cdot(x'_{i-1} \oplus b \oplus 2\cdot(x'_{i-1} \vee 1)) \ mod\ 2^n. \qquad (11)$$

Здесь будет уместно привести для сравнения один из простейших вариантов реализации квадратичного конгруэнтного метода, предложенный Р. Ковэю [1]:

$$x_o \equiv 2(mod\ 4), \qquad x_i = x_{i-1}\cdot(x_{i-1}+1) \ mod\ 2^n. \qquad (12)$$

Период формируемой таким образом конгруэнтной последовательности равен 2^{n-2}, при этом, последовательность $x_i/4$ в пределах периода бесповторна. Точно такими же характеристиками обладает и нелинейное расширение:

$$x'_o \equiv 2(mod\ 4), \qquad x'_i = x'_{i-1}\cdot(x'_{i-1} \oplus 1) \ mod\ 2^n. \qquad (13)$$

Анализ показывает статистическую надежность выборки, полученной по самому старшему биту (12) и (13), начиная с платформ, разрядностью $n=36$ бит.

По аналогии с квадратичным конгруэнтным методом (9) и его нелинейным расширением (10), с тем же успехом, но больше в академическом, нежели чем в практическом плане в связи с многократным падением производительности обработки, можно построить кубичные и более высоких порядков – ***полиномиальные конгруэнтные методы и их нелинейные расширения***:

$$x_i = (a+4\cdot x_{i-1})(x_{i-1}+b)\prod_{j=1}^{m-2}(a_j + 2^{j+2}\cdot x_{i-1}) \ mod\ 2^n, \qquad (3 \le m \le n-1) \qquad (14)$$

$$x'_i = (a' \oplus 4\cdot x'_{i-1})(x'_{i-1} \oplus b \oplus 2\cdot(x'_{i-1} \vee 1))\prod_{j=1}^{m-2}(a_j \oplus 2^{j+2}\cdot x'_{i-1}) \ mod\ 2^n, \qquad (15)$$

с нечетными постоянными коэффициентами $a_j = (a_{oj}\cdot 2^{j+2}+1) \ mod\ 2^n$, при $(0 \le a_{oj} \le 2^n - 1)$.

Анализ показывает, что с увеличением степени m образующих многочленов (14) и (15), статистическая надежность выборок, полученных по самому старшему биту, фактически не изменятся и, как в случаях (9) и (10), составляет $n=$34-35 и $n=$26-27 бит, соответственно. Общая длина ключа составляет $(m+1)(2\cdot n-m)/2$, а при $m=n-1$, $n(n+1)/2$ бит.

В общем случае, руководствуясь положениями линейного конгруэнтного метода и его расширения (5) и (6), коэффициенты, входящие в состав уравнений (14) и (15), могут быть переменными.

Криптографический анализ усеченных последовательностей формируемых на основе указанных уравнений (14) и (15) с постоянными коэффициентами проблематичен, а с введением переменных коэффициентов соизмерим с технически нереализуемым перебором.

Из линейного конгруэнтного метода с переменными коэффициентами (5) и его нелинейного расширения (6) также следует, так называемый, регулярный рандомизационный метод, с коэффициентами, зависящими от предыдущих элементов последовательности. Рассмотрим одни из его наиболее легко реализуемых на практике характерных примеров.

Регулярный рандомизационный метод задается уравнением:

$$x_i = A_{i-1}\cdot x_{i-1} + B_{i-1} \ mod\ 2^n, \qquad A_{i-1} = 4\cdot x_{i-3} + a, \qquad B_{i-1} = 4\cdot x_{i-2} + b, \qquad (16)$$

при $a \equiv 1(mod\ 4)$, а его ***нелинейное расширение*** уравнением:

$$x'_i = A'_{i-1}\cdot(x'_{i-1} \oplus B'_{i-1}) \ mod\ 2^n, \qquad A_{i-1} = 4\cdot x_{i-3} \oplus a', \qquad B_{i-1} = 2\cdot(x_{i-2} \vee 1) \oplus b, \qquad (17)$$

при $a' \equiv 3(mod\ 4)$ и нечетном коэффициенте b.

Характерной особенностью регулярного рандомизационного метода является наличие зависящего от начальных условий переходного нелинейного участка, длиной $L_T \le n$, после прохождения которого последовательности формируемые на основе уравнений (16) и (17), благодаря феноменальной саморегуляции, достигают максимального периода 2^n и далее всюду, в границах каждого из последующих периодов ведет себя стационарно и бесповторно.

Аналогичная саморегуляция с $L_T = 1$ характерна и для квадратичного конгруэнтного метода Р. Ковэю (12) и его нелинейного расширения (13) при нечетных начальных условиях x_o.

Общая длина ключа генерации составляет $3\,(n-1) + 2{\cdot}n = 5{\cdot}n - 3$ бит. Саморегуляция не проходит бесследно. Число уникальных последовательностей составляет всего $2^{3(n-1)}$.

Статистический анализ показывает надежность выборки, полученной по самому старшему биту (16) и (17), начиная с платформ, разрядностью $n = 36$ и $n = 26$ бит, соответственно. К этому, криптографический анализ усеченных последовательностей формируемых на основе указанных уравнений (16) и (17) проблематичен и по всему соизмерим с технически нереализуемым перебором.

Для получения криптографически стойких и статистически надежных псевдослучайных последовательностей ***не только <u>по ее самым старшим значащим битам</u>***, но и по каждому отдельному биту платформы и в целом, необходимо использовать функции усложнения. Более подробную информацию по этому поводу можно найти в статье [7].

И в заключение, для получения общей картины о неполноте и незавершенности теории чисел, дадим еще не раскрытые, лежащие в потенциале конгруэнтного метода, многомерные (сетевые) случаи, отмеченные в заявках на изобретения [5]. Согласно с ними, так называемый, ***дуальный рандомизационный метод*** задается уравнениями:

$$u_i = (a_u + 4{\cdot}v_{i-1}) \cdot (u_{i-1} + b_u + (2{\cdot}v_{i-1} \vee 2)) \ \ mod\ 2^n, \qquad (18)$$
$$v_i = (a_v + 4{\cdot}u_{i-1}) \cdot (v_{i-1} + b_v) \ \ mod\ 2^n,$$

с множителями $a_{u,v} \equiv 1 (mod\ 4)$, а уравнения его нелинейного расширения имеют вид:

$$u'_i = (a'_u \oplus 4{\cdot}v'_{i-1}) \cdot (u'_{i-1} \oplus b_u \oplus (2{\cdot}v'_{i-1} \vee 2)) \ \ mod\ 2^n, \qquad (19)$$
$$v'_i = (a'_v \oplus 4{\cdot}u'_{i-1}) \cdot (v'_{i-1} \oplus b_v) \ \ mod\ 2^n,$$

с множителями $a'_{u,v} \equiv 3 (mod\ 4)$ и нечетными приращениями $b_{u,v}$. Формируемые таким образом последовательности имеют период T_{max} при любых начальных условиях. Общая длина ключа составляет $6(n-1)$ бит.

Анализ показывает статистическую надежность выборки, полученной по самому старшему биту образующей u переменной (18) и (19), начиная с платформ, разрядностью $n = 32$ и $n = 26$ бит, соответственно.

При использовании образующей u переменной в качестве результирующей, выходной переменной, криптографический анализ со стороны старших бит формируемых на ее основе последовательностей проблематичен, за счет формального усечения дополняющей ее v переменной.

<u>Общие выводы</u>.

1. Двоичные конгруэнтные последовательности, формируемые в соответствии с уравнениями (2)-(3), (5)-(6), (9)-(11) и (14)-(19), имеют максимальный период, равный 2^n и в пределах периода бесповторны. Анализ, выходящий за рамки данной статьи, показывает, что число всевозможных образующих двоичных конгруэнтных последовательностей огромно, порядка 2^{2^n}, что говорит о еще не раскрытом, неизмеримо высоком потенциале конгруэнтного метода.

2. Для типовых вариантов реализации линейного конгруэнтного метода с переменными коэффициентами (5) и его нелинейного расширения (6), длина ключа генерации составляет порядка $5{\cdot}n$ бит. Каждому ключу генерации, из множества ключей порядка 2^{5n}, соответствует уникальная конгруэнтная последовательность. Для квадратичного конгруэнтного метода общая длина ключа генерации составляет $3\,(n-1)$ бит, а для дуального $6(n-1)$ бит.

3. Статистический анализ линейного конгруэнтного метода с переменными коэффициентами, квадратичного и полиномиального метода и представляемых расширений на основе пакета тестов DIEHARD, показывает возможность достижения высоких статистических показателей старших значащих бит на малых платформах, от среднестатистических 37 до предельных предусмотренных пакетом 26 бит.

4. Криптографический анализ конгруэнтных последовательностей со стороны старших бит существенно затруднен, проблематичен или технически невозможен. К этому, согласно с приведенными расчетами, нелинейные расширения (6), (11), (15), (17) и (19) по статистическим свойствам и криптографическим показателям заметно превосходят свои линейные аналоги (5), (9), (14), (16), (18) и вполне пригодны для практических приложений.

5. Варианты реализации линейного и полиномиального конгруэнтного метода могут характеризоваться наличием короткого переходного нелинейного участка (предпериода), после прохождения которого, благодаря феноменальной саморегуляции, достигается максимальный период 2^n и апериодичное (бесповторное) поведение.

6. С представленными вариантами генерации псевдослучайных последовательностей снимаются проблемы выбора начальных условий и коэффициентов, входящих в состав образующих их уравнений. Иными словами статистические показатели генерируемых последовательностей фактически не зависят от изменения составляющих ключа генерации, т. е. все ключи равнозначны.

7. Полиномиальный метод и его разновидности, как показано в [5], допускают не только двумерное – дуальное представление типа (18) и (19), но и многомерное сетевое обобщение любой структурной сложности.

Как видим, потенциальные возможности линейного конгруэнтного метода и его разновидностей далеко не исчерпаны, и если судить по доступной научной литературе, ***математические доказательства*** построения последовательностей максимального периода в соответствии с представленными выше уравнениями (3)-(6), (10)-(11), (13) и (15)-(19) – ***неизвестны***, за исключением достаточно хорошо исследованных и изученных [1,4] элементарных случаев (2), (12) и случаев (9), (14) в некотором роде пересекающимся с аналогами, представляемыми квадратичными конгруэнтными методами [1].

Для того чтобы указанные доказательства стали возможными, похоже, по результатам проведенных исследований http://t.random-art.ru/ , так или иначе, придется с новых позиций взглянуть на имеющиеся и вскрыть еще не известные нам математические структуры в алгебре и теории чисел, а также понять свойства и определить место этих структур в математике, физике и естествознании в целом.

Несмотря на явные выявленные по ходу изложения материала признаки неполноты и незавершенности теории чисел, а также, несмотря на отсутствие математических доказательств (к которым даже и неизвестно еще и как подойти), позволяющих вместо обычной операции сложения использовать побитовую операцию сложения по модулю 2 – все утверждения подтверждены результатами натурного моделирования и расчетами.

Все это, дополненное существованием переходных нелинейных участков, наблюдаемой феноменологической саморегуляцией и многомерными сетевыми решениями прямо указывает на необходимость проведения дальнейших прикладных и фундаментальных исследований.

Между тем, линейный и полиномиальные конгруэнтные методы с переменными коэффициентами и его разновидности не так просто и эффективно устроены, как это кажется. В первую очередь это связано с использованием высоко затратной операции умножения и даже, относительно затратной операции сложения, на которые приходится большая доля вычислительных, аппаратных и энергетических затрат.

Более того, следует помнить, понимать и учитывать при оценке сложности реализации и аппаратных затрат, что арифметические операции не допускают распараллеливание вычислений по каждому из отдельно взятых разрядов, что крайне негативно сказывается на конечной производительности конгруэнтных генераторов и в силу чего, выводит их не в практическую, а в академическую 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, нелинейный, квадратичный, кубичный, полиномиальный, конгруэнтный, метод, дихотомический, генератор, псевдослучайных, число, линейная, нелинейная, квадратичная, кубичная, полиномиальная, конгруэнтная, дихотомическая, последовательность, симметричная, криптография.

Литература

1. Кнут Дональд Э. Искусство программирования.
   Третье издание, Том 2, М.: Издательский дом “Вильямс”, 2002.
2. Шнайер Брюс. Прикладная криптография.
   Изд. ТРИУМФ, Москва, 2002.
3. Brickell et al. A Surkey of Recent Results.
   Proc. of the IEEE, Vol. 76, no. 5, May 1988.
4. Поточные шифры. Результаты открытой зарубежной криптологии.
   Москва,1997, http://www.ssl.stu.neva.ru/psw/crypto/potok/str_ciph.htm
5. Кулаков И.А. Способ придания реальному объекту рандомизационных свойств
   и рандомизационная система.
   Международная заявка PCT/RU03/00141 от 7 апреля 2003.
   Заявка на Евразийский патент №200500946 от 11 июля 2005.
6. 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.
7. Кулаков И.А. Линейные конгруэнтные и рандомизационные генераторы.
   Рукопись статьи, Москва, 2012, http://random-art.ru/?download=LCG_RNG_ru.pdf

МОСКВА, март 2012

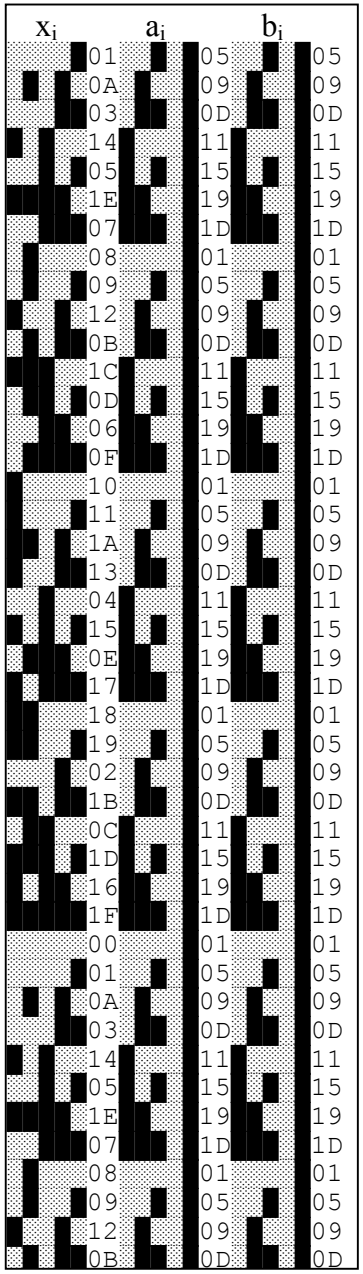

Рис.1

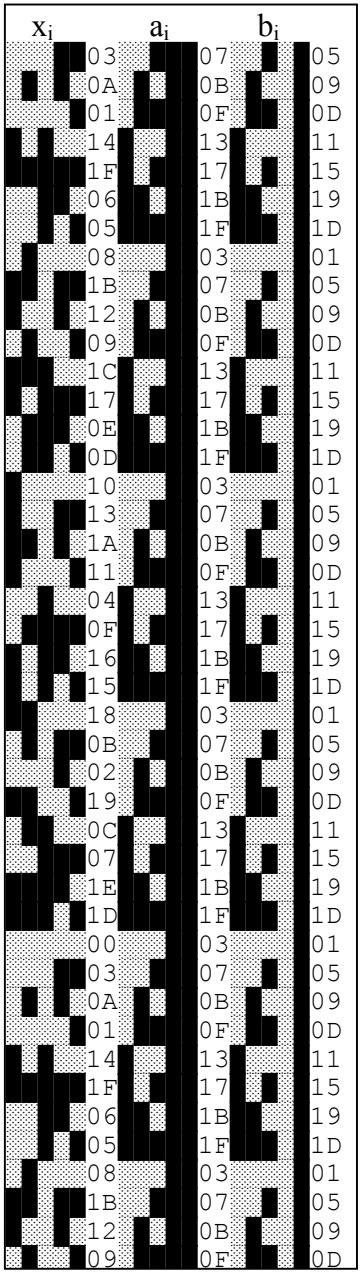

Рис.2

Исходные данные:
n – длина платформы генерации, бит.
Rx, Ra, Rb – задающие параметры генерации, длина = n + (2·n – 3) бит.
RHa, RHb – компоненты ключа генерации, длина = 2·n – (5÷6) бит.
Синхронизация параметров генератора:
x = Rx; a = Ra | 3; b = Rb | 1; Ha = (RHa | 7) ^ 3; Hb = (RHb | 7) ^ 3;
x – смещение элементов последовательности,
Ha, Hb – приращения коэффициентов генерации.
Общая длина ключа генерации:
k = n + (2·n – 3) + (2·n – (5÷6)) = 5·n – (8÷9) бит.
Каждому ключу генерации соответствует уникальная двоичная конгруэнтная последовательность, с периодом 2^n. В пределах периода последовательность бесповторна. Число уникальных последовательностей велико и равно $2^k \approx 2^{5n}$.

Типовой пример реализации
линейного конгруэнтного метода с переменными коэффициентами

начальная синхронизация множителя: a ^= 2;

формирование очередного элемента конгруэнтной последовательности:
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 = 0; a = 1; b = 1; Ha = 4; Hb = 4;

Типовой пример реализации нелинейного расширения
линейного конгруэнтного метода с переменными коэффициентами

формирование очередного элемента конгруэнтной последовательности:
x = (a*(x ^ 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 = 0; a = 3; b = 1; Ha = 4; Hb = 4;