УДК 621.391:004.73

В.В. Швыдкий, Э.В. Фауре, В.В. Веретельник, В.А. Щерба

Черкасский государственный технологический университет, Черкассы

ГЕНЕРАЦИЯ СТОХАСТИЧЕСКОЙ ПОСЛЕДОВАТЕЛЬНОСТИ ГЕНЕРАТОРОМ КОНГРУЭНТНЫХ ЧИСЕЛ

Рассматривается способ генерации стохастической последовательности чисел конгруэнтным генератором. Дается оценка максимального значения периода повторения стохастической последовательности. Показана возможность получения некоррелированной равномерно распределенной случайной последовательности конгруэнтных чисел с максимальным периодом повторения, исчисляемым десятками лет (при производительности генератора случайных чисел до 10^{12} случайных чисел/сек).

Ключевые слова: *генератор случайных чисел, генератор конгруэнтных чисел, стохастическая последовательность, равномерно распределенная последовательность чисел.*

Введение

Стохастический генератор случайных чисел (ГСЧ) – это генератор последовательности случайных равномерно распределенных на интервале [0,(M–1)] некоррелированных чисел, которая обладает свойствами стохастического процесса[1].

В настоящей работе рассматривается стохастический ГСЧ как конечный автомат, сочетающий свойства стохастического конечного автомата и генератора конгруэнтных чисел – чисел, связанных соотношением:

$$S(n) = \left|KS(n-1) + C\right|_M, \qquad (1)$$

где K, C, M – численные параметры генератора, $S(n)$ и $S(n-1)$ – слова, порожденные генератором в текущий – «n» – и в предшествующий – «n–1» – дискретные моменты времени.

Генераторы конгруэнтных чисел широко используются во всех сферах человеческой деятельности, где необходимо использование случайных последовательностей чисел.

Такие генераторы порождают периодически повторяемую псевдослучайную последовательность чисел (ПСП), которая представляет последовательность вычетов по модулю M и принимает конечное множество значений из интервала [0, (M–1)].

Основными недостатками данного метода формирования конгруэнтной последовательности чисел является то, что порождаемая последовательность, как правило, не является равновероятной и всегда коррелирована. Кроме того, правила выбора параметров K, C, M для получения последовательности с нужными свойствами не определены.

В работе [2] показано, что период цикла конгруэнц-генератора с параметрами K, C, M, содержащего более чем одно слово, соответствует периоду кольца вычетов по модулю M с порождающим элементом K. Кроме того, в [2] сформулированы некоторые свойства, выявляющие зависимость периода кольца вычетов по модулю M от порождающего элемента K. Указанные свойства приводят к упрощению процедуры выбора параметров конгруэнц-генератора для формирования циклов конгруэнтной последовательности чисел заданной длины.

Многочисленные попытки модифицировать конгруэнтный метод с целью получения равномерно распределенной случайной величины, например, за счет использования свойств чисел Фибоначчи [3], по существу ничего не дали. Не намного лучше обстоят дела в области теоретического анализа процесса формирования ПСП с нужными параметрами и методов тестирования ГСЧ. В уникальном труде Д.Э. Кнута [4] показана вся сложность этой проблемы, обосновано несколько различных критериев оценки ПСП и приведены общие рекомендации по выбору параметров K, C, M. Проведенные сравнительно недавно исследования [5, 6] показали возможность рандомизации последовательности конгруэнтных чисел, а также определения показателей, связывающих конструктивные свойства ГСЧ с отклонением статистических свойств выходной последовательности ГСЧ от теоретического распределения. Вместе с тем, эти работы не обеспечивают возможности создания генератора некоррелированной равномерно распределенной последовательности чисел, что и стимулирует выполнение исследований в данной области.

Целью настоящей работы является разработка метода генерации равномерно распределенной некоррелированной стохастической последовательности чисел интервала $[0, (2^m-1)]$, где m – количество разрядов в двоичном представлении числа M, который может быть реализован на современных персональных компьютерах, и обеспечивающий период повторения порождаемой стохастической последовательности, исчисляемый десятками лет.

[1] *Стохастический процесс – явление, развивающееся во времени согласно законам теории вероятности [1]. В теории вероятностей итог стохастического процесса не может быть определен по изначальному состоянию системы.*

Постановка задачи. Для достижения поставленной цели необходимо решить научно-техническую задачу определения алгоритма, который обеспечивает обход всех вершин графа состояний генератора строго по одному разу в цикле обхода и случайное изменение порядка обхода графа в каждом стохастическом цикле. Период повторения порождаемой случайной последовательности при некоторой фиксированной производительности ГСЧ (например, равной 10^{12} случайных чисел/сек) должен составлять десятки или сотни лет.

Дополнительной задачей работы является разработка метода расширения области определения случайной величины от значения $[0, (2^m–1)]$ до значения $[0, (2^{qm}–1)]$, где q – коэффициент расширения.

Решение задачи

В настоящей работе будем придерживаться определений, которые сформулированы в [5]:

– цикл генератора конгруэнтных чисел – упорядоченное по времени подмножество слов из множества мощности M;

– нуль-цикл генератора конгруэнтных чисел – цикл, содержащий одно слово из множества мощности M;

– сверхцикл генератора конгруэнтных чисел – объединение циклов генератора конгруэнтных чисел в единый цикл длины M (конкатенация отдельных циклов в единый цикл длины M);

– вектор начальной загрузки (ВНЗ) – одно из слов множества, загружаемое в формулу рекурсии (1) в начале цикла, как число S(0).

Дополнительно определим следующие понятия для предлагаемого метода построения ГСЧ.

Цикл стохастического ГСЧ (стохастический цикл) – упорядоченное по времени множество слов, содержащее ровно по одному слову интервала [0,(M–1)].

Конкатенация – объединение нескольких подстрок в одну строку, объединение подслов (старшей и младшей части) в одно слово, объединение двух последовательностей в одну с порядком следования одна за другой.

Слово – одно число интервала [0,(M–1)], выданное на выход ГСЧ.

Фраза – множество слов, сформированных ГСЧ в одном стохастическом цикле.

Мешающее слово – одно слово из области определения функции распределения случайной величины, сформированное генератором помехи.

Мешающее слово может быть как добавлено, так и изъято из потока случайных чисел.

Последовательность мешающих слов, сформированных генератором помехи, переводит гипотетическое распределение случайной величины в эмпирическое.

Для обоснованного выбора конструкции ГСЧ сформулируем предъявляемые к его выходной последовательности требования в следующем виде:

– равномерное распределение слов последовательности;

– отсутствие взаимосвязи (корреляции) между словами.

Будем считать, что корреляция между словами в стохастическом ГСЧ отсутствует при условии равной вероятности появления на его выходе любой группы из k ($k \geq 2$) слов, т.е. при

$$p(x_0 x_1 \ldots x_{k-1}) = 1/M^k .$$

Указанные требования определяют основное свойство «белого» шума. В настоящей работе ставится и решается задача определения количественной оценки точности воспроизведения «белого» шума дискретным стохастическим ГСЧ.

Рассмотрим методы построения стохастических ГСЧ, у которых одно мешающее слово приходится на 100, 1000 или более слов гипотетического потока. Такой способ оценки качества ГСЧ, названный h^2-критерием [6], связан с широко используемым в математической статистике критерием Пирсона (дальше χ^2-критерий) следующим образом:

$$\chi^2 h^2 = V , \qquad (2)$$

где V – объем выборки (число слов в выборке).

Рассмотрим принцип построения стохастического ГСЧ.

Как показано в работе [5], при простом M, определяющем область определения функции распределения случайной величины, граф состояний генератора конгруэнтной последовательности содержит d циклов по t слов каждый и один нуль-цикл. При этом $d \cdot t = M - 1$, что позволяет выбирать разные конструкции графа, зная разложения числа $M - 1$ на простые сомножители. Под заданную конструкцию графа необходимо подобрать численное значение параметра K, а затем, используя принцип решета Эратосфена, построить граф генератора конгруэнтных чисел как совокупность его циклов. Перечислим поочередно все слова, входящие в каждый из циклов, запишем их в столбцы матрицы, показанной на рис. 1. Эту матрицу будем называть dt-матрицей стохастического генератора или его основной матрицей.

Отметим, что матрица составлена из неповторяющихся $M - 1$ слов отрезка $\overline{0,(M-1)}$, отсутствует только одно слово, которое порождает нуль-цикл.

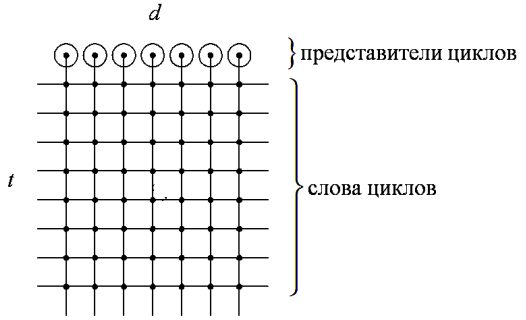

Рис. 1. Основная матрица стохастического ГСЧ

Представляется возможным расширить область определения функции распределения случайной величины до отрезка $\overline{0,(2^m-1)}$, где m – количество разрядов в двоичном представлении числа M. Для этого сформируем вектор-строку дополнительных слов множества, которое включает слово, определяющее нуль-цикл, и слова, доопределяющие функцию распределения. Поэтому, полный граф стохастического ГСЧ содержит основную dt-матрицу и строку дополнительных слов.

Таким образом, формирование стохастической последовательности включает два независимых процесса:

– вычисление слов основной dt-матрицы (в случайном порядке);

– расстановка дополнительных слов (в случайном порядке).

Порядок расположения слов в основной dt-матрице и порядок расстановки дополнительных слов изменяются для каждого стохастического цикла.

Совокупность указанных операций приводит к формированию стохастического цикла из 2^m слов, при этом каждое из слов отрезка $\overline{0,(2^m-1)}$ используется только один раз, что гарантирует равномерный закон распределения случайной величины на интервале, равном или кратном стохастическому циклу.

Степень корреляции между словами определяется статистической зависимостью в порядке перемещения по графу состояний автомата. Для внесения неопределенности в порядок перемещения по графу состояний автомата и обеспечения статистической независимости между словами, генерируемыми ГСЧ, в каждом стохастическом цикле случайным образом устанавливается разный порядок обхода вершин полного графа состояний.

Для иллюстрации предлагаемого метода формирования стохастической последовательности рассмотрим пример построения стохастического ГСЧ, обеспечивающего генерацию равномерно распределенной на интервале (0,255) некоррелированной последовательности чисел с нулевой конструктивной ошибкой (не более одного мешающего слова на каждую тысячу слов полного потока, т.е. $h^2 > 10^6$).

Отметим, что при выбранном интервале (0,255) такой генератор реализуется на 8-разрядной однокристальной ЭВМ

Определим конструкцию ГСЧ. Выберем наибольшее, не превосходящее 256 простое число M: M=251. Отсюда, $(M-1) = d \cdot t = 250 = 5 \cdot 5 \cdot 5 \cdot 2$. Выберем конструкцию $d = 10,\ t = 25$. Такая конструкция графа состояний обеспечивается генератором конгруэнтных чисел с параметрами: M=251, K=5, C=5 и любымS(0)≠187. Назовем этот генератор основным конгруэнтным генератором (ОКГ) стохастического ГСЧ – генератором, который будет вычислять слова основной матрицы.

Используя принцип решета Эратосфена, рассчитаем несколько циклов по 25 слов в каждом, для чего зададим произвольный ВНЗ, например, s(0)=0 и рассчитаем 1 цикл: 5, 30, 155, 27, 140, 203, 16, 85, 179, 147, 238, 191, 207, 36, 185, 177, 137, 188, 192, 212, 61, 59, 49, 250, 0. Из натуральной последовательности чисел 0…250 вычеркнем все числа, входящие в первый цикл. Выберем одно из чисел из не вычеркнутой части последовательности, например, s(0)=1 и рассчитаем 2 цикл: 10, 55, 29, 150, 2, 15, 80, 154, 22, 115, 78, 144, 223, 116, 83, 169, 97, 239, 196, 232, 161, 57, 39, 200, 1. Из натуральной последовательности чисел 0…250 вычеркнем все числа, входящие во второй цикл (числа, входящие в первый цикл, вычеркнуты ранее). Так будем продолжать до тех пор, пока не останется не вычеркнутых чисел. Результаты сведем в табл. 1. Таким образом, основная dt-матрица содержит 250 слов из 251, отдельно хранится второй столбец таблицы, содержащий по одному представителю каждого из циклов и нуль-цикл (слово 187).

Таблица 1

Циклы генератора конгруэнтных чисел

Цикл	ВНЗ	Длина цикла
1	0	25
2	1	25
3	3	25
4	6	25
5	7	25
6	11	25
7	13	25
8	32	25
9	38	25
10	42	25
11	187	1

Расширим область определения равномерно распределенной случайной величины от значения $\overline{0,(M-1)} = \overline{0,249}$ до значения $\overline{0,(2^m-1)} = \overline{0,255}$.

С учетом нуль-цикла сформируем строку дополнительных слов: 187, 251, 252, 253, 254, 255, – которые подлежат вставке в каждом сверхцикле в поток случайных чисел, порожденных ГСЧ. В результате будет сформирован полный граф стохастического ГСЧ, содержащий 256 вершин (250 вершин основной dt-матрицы и 6 вершин дополнительных слов).

Запомним в постоянной памяти 10 слов представителей циклов. Для организации случайного обхода каждой из вершин графа ровно один раз в стохастическом цикле используем три вспомогательных генератора равномерно распределенных случайных величин (ВГ1, ВГ2, ВГ3). ВГ1 и ВГ2 обеспечивают задание координат основной матрицы $d_i,\ t_i$ ($d_i \in \overline{0,d-1},\ t_i \in \overline{0,t-1}$), ВГ3 необходим для указания расположения дополнительных слов в выходной последовательности.

Таким образом, формирование выходного слова стохастического ГСЧ происходит в следующей последовательности:

– ВГ1 и ВГ2 формируют по одному случайному числу d_i и t_i, по которым ОКГ формирует выходное слово;

– слова в стохастическом цикле подсчитываются счетчиком (от 0 до 255);

– производится сравнение состояния счетчика слов стохастического цикла с каждым из 6 слов, сформированных ВГ3. При совпадении текущего состояния счетчика слов стохастического цикла с одним из 6 слов, сформированных ВГ3, на выход стохастического ГСЧ выводится дополнительное слово, порядковый номер которого совпадает с его порядковым номером в последовательности значений ВГ3.

Процесс формирования слов повторяется до завершения формирования стохастического цикла.

В каждом последующем стохастическом цикле должна производиться смена порядка обхода полного графа состояний стохастического ГСЧ. В силу конечности числа параметров, подлежащих модификации при переходе от одного стохастического цикла к другому, порядок обхода полного графа состояний периодичен. По этой причине задача создания стохастического ГСЧ сводится к задаче проектирования стохастического автомата, управляющего порядком обхода вершин полного графа состояний, который при максимальной простоте алгоритма дает максимальный период повторения генерируемой стохастической последовательности чисел. Для рассматриваемого здесь примера стохастического ГСЧ возможны два варианта построения ВГ1, ВГ2, ВГ3:

– использование генератора конгруэнтных чисел с $M = t = 10$ для ВГ1, $M = t = 25$ для ВГ2, $M = t = 256$ для ВГ3;

– декорреляция натуральной последовательности чисел (0,1,2,3…9) для ВГ1, декорреляция натуральной последовательности чисел (0,1,2,3…24) для ВГ2, декорреляция натуральной последовательности чисел (0,1,2,3…255) для ВГ3.

Максимальный период повторения у генератора, построенного на основе декорреляции натуральной последовательности чисел. В этом случае максимальный период повторения составит:

– для ВГ1 – $m_2 = d!$;

– для ВГ2 – $m_3 = t!$;

– для ВГ3 – $m_4 = C_b^z$,

где b – правая граница области определения функции распределения случайной величины (в данном примере b=256); z – количество дополнительных слов (в данном примере z=6).

Для декорреляции натуральной последовательности чисел используется сочетание операций перестановок слов и циклического сдвига слов. Циклом декорреляции будем считать сочетание перестановка-сдвиг, выполняемых над массивом или его частью. Принципы декорреляции достаточно хорошо известны (например, перестановкой слов в последовательности по заранее подготовленной таблице случайных чисел). Отметим лишь, что один из вариантов построения декоррелятора, созданного для решения поставленных здесь задач, приведен в [7].

Пусть, например, результатом декорреляции первых 10 чисел натуральной последовательности является последовательность 8, 4, 1, 2, 6, 3, 9, 0, 7, 5; а результатом декорреляции первых 25 чисел натуральной последовательности – последовательность 21, 15, 20, 17, 16, 14, 18, 24, 23, 19, 2, 6, 22, 7, 4, 0, 1, 3, 5, 8, 9, 10, 11, 12, 13.

Полученный результат допускает следующую трактовку, определяющую правило формирования слов основной матрицы (правило обхода вершин графа основной dt-матрицы):

– из 8-й ячейки хранения представителей циклов выбрать представителя 8-го цикла и загрузить его, как слово S(n – 1) в ОКГ;

– вычислить 21-е слово этого цикла, выдать результат – первое слово стохастического цикла.

Следующие два слова стохастического цикла формируются так:

– из 4-й ячейки хранения представителей циклов выбрать представителя 4-го цикла и загрузить его, как слово S(n – 1) в ОКГ;

– вычислить 15-е слово этого цикла, выдать результат;

– из 1-й ячейки хранения представителей циклов выбрать представителя 1-го цикла и загрузить как слово S(n – 1) в ОКГ;

– вычислить 20-е слово этого цикла, выдать результат.

Легко заметить, что если продолжать генерацию стохастической последовательности, то через 50 слов происходит зацикливание. Чтобы избежать этого явления, то через каждых 50 слов надо выполнить одну из следующих процедур:

– произвести перестановку массива указателей номеров циклов (6, 8, 7, 9, 4, 3, 5, 0, 1, 2) – это, по существу, перестановка столбцов основной матрицы;

– произвести циклический сдвиг массива указателей номеров циклов (4, 1, 2, 6, 3, 9, 0, 7, 5, 8) – это, по существу, циклический сдвиг столбцов основной матрицы;

– произвести перестановку массива указателей слов в цикле (14, 19, 10, 17, 16, 21, 18, 24, 23, 15, 2, 6, 22, 7, 4, 0, 1, 3, 5, 8, 9, 20, 11, 12, 13) – это, по существу, перестановка строк основной матрицы;

– произвести циклический сдвиг массива указателей слов в цикле (15, 20, 17, 16, 14, 18, 24, 23, 19, 2, 6, 22, 7, 4, 0, 1, 3, 5, 8, 9, 10, 11, 12, 13, 21) – это, по существу, циклический сдвиг строк основной матрицы.

Очевидно, что если указанные операции выполнять не одновременно, а в соответствии с некоторым правилом чередования, можно добиться максимального периода повторения расстановки слов основной матрицы, который равен $\tau_4 = 250 \cdot d! \cdot t!$. В рассматриваемом примере максимальный период повторения порядка размещения слов основной матрицы

$\tau_4 = 250 \cdot d! \cdot t! = 250 \cdot 3,6 \cdot 10^6 \cdot 1,54 \cdot 10^{25} = 1,386 \cdot 10^{34}$.

Допустим, что стохастический генератор имеет производительность $\upsilon = 10^{12}$ слов / сек.

Учтем, что год примерно равен $3,15 \cdot 10^7$ сек, тогда за это время стохастический генератор произведет $V = 10^{12} \cdot 3,15 \cdot 10^7 = 3,15 \cdot 10^{19}$ слов, а период повторения слов основной матрицы составит $\tau_4' = \frac{1,386 \cdot 10^{34}}{3,15 \cdot 10^{19}} \cong 4,4 \cdot 10^{14}$ лет. Обратим внимание на то обстоятельство, что в стохастическом генераторе объем выполняемых вычислений существенно больше, чем при вычислении последовательности линейным генератором конгруэнтных чисел.

В генераторе конгруэнтных чисел с максимальной длиной цикла M для вычисления всех слов цикла необходимо ровно M операций вычисления по уравнению (1).

В стохастическом ГСЧ вычисление смежного слова включает выполнение $t_i \left(i \in \overline{0,(t-1)}\right)$ операций вычисления по уравнению (1), поэтому для формирования всех слов цикла конгруэнц-генератора объем выполняемых вычислений равен $v_1 = \sum_{j=0}^{t-1} t_j = \frac{t(t-1)}{2}$, а для всех d циклов стохастического цикла объем вычислений $v = dt(t-1)/2$.

В целом, объем вычислений в стохастическом генераторе будет больше объема вычислений в генераторе конгруэнтных чисел в

$$\sigma = dt(t-1)/(2M) \qquad (3)$$

раз. Для рассматриваемого примера $\sigma = 12$.

Таким образом, стохастический генератор (M=251) имеет в 12 раз меньшую производительность, чем генератор конгруэнтных чисел, что приводит к необходимости поиска путей повышения производительности стохастического ГСЧ и (или) уменьшения числа выполняемых операций, необходимых для вычисления следующего слова.

Поскольку в стохастическом ГСЧ два случайным образом выбранных числа (d_i, t_i) определяют цикл d_i и смещение t_i следующего слова относительно выбранного представителя, то, строго говоря, в стохастическом генераторе необходимо вычислять не слово $S(n)$ по слову $S(n-1)$, а слово $S(0+t_i)$ по $S(0)$ в выбранном цикле d_i.

Тогда для слова $S(m) = S(0+t_i)$ по [5]:

$$S(m) = K^m S(0) + C\sum_{j=0}^{m-1} K^j - q_m M. \qquad (4)$$

Учтем, что $\sum_{j=0}^{m-1} K^j = \frac{K^m - 1}{K-1}$. Тогда (4) может быть представлено в виде:

$$S(m) = K^m S(0) + C\frac{K^m - 1}{K-1} - q_m M. \qquad (5)$$

Учтем, что параметры K, C, $S(0)$, $m = t_i$ известны, что позволяет представить выражение под знаком вычета и сам вычет равенств (1), (4), (5) в виде:

$$S(m) = |A_1|_M, \qquad (6.1)$$

где $A_1 = KS(m-1) + C$;

$$S(m) = |A_2|_M, \qquad (6.2)$$

где $A_2 = K^m S(0) + C\sum_{j=0}^{m-1} K^j$;

$$S(m) = |A_3|_M, \qquad (6.3)$$

где $A_3 = K^m S(0) + C\frac{K^m - 1}{K-1}$.

Учтем, что равенства (1) и (4) и, следовательно, равенства (6.1) и (6.2) справедливы для m>0, в то время как в стохастическом цикле обязательно присутствует одно слово с координатой $t_i = m = 0$, что соответствует вставке в поток слова $S(0)$. Отсюда следует, что для вычисления $S(m)$, $m \in \overline{0,(t-1)}$, необходимо пользоваться выражением (6.3). Отметим, что затраты времени на выполнение операций по (6.3) соизмеримы с затратами времени на выполнение операций по (6.1), что обеспечивает возможность создания стохастического ГСЧ с производительностью, близкой к производительности генератора конгруэнтных чисел той же размерности.

Рассмотрим вопрос расстановки дополнительных слов. Возьмем две последовательности чисел:

– последовательность (1, 2, 3, 4, 5, 0), которая соответствует позициям шести дополнительных слов (187, 251, 252, 253, 254, 255);

– последовательность (1, 2, 3, 4, 5…255, 0), которая определяет место вставки дополнительных слов в сверхцикл.

Выполним перестановку элементов первой из последовательностей, в результате чего, например, получим 4, 1, 3, 5, 0, 2. Для последовательности из 256 слов покажем фрагмент, например, из первых 24 слов, сгруппированных по 6 слов: 43, 138, 225, 240, 119, 54; 237, 92, 67, 98, 121, 72; 143, 14, 133, 180, 91, 58; 17, 160, 167, 230, 29, 112.

Полученный результат трактуется так:

– четвертое дополнительное слово (это число 253) размещается на 43 позиции в стохастическом цикле;

– первое дополнительное слово (это число 187) размещается на 138 позиции в стохастическом цикле;

– третье дополнительное слово (это число 252) размещается на 225 позиции в стохастическом цикле;

– пятое дополнительное слово (это число 254) размещается на 240 позиции в стохастическом цикле;

– нулевое дополнительное слово (это число 255) размещается на 119 позиции в стохастическом цикле;

– второе дополнительное слово (это число 251) размещается на 54 позиции в стохастическом цикле.

Получим следующий порядок размещения дополнительных слов в первом стохастическом цикле: слово 43 – число 253; слово 54 – число 251; слово 119 – число 255; слово 138 – число 187; слово 225 – число 252; слово 240 – число 254.

Во втором стохастическом цикле порядок размещения дополнительных слов будет следующий: слова 253, 251, 255, 187, 252, 254 на 237, 92, 67, 98, 121, 72 позициях стохастического цикла, соответственно.

Период последовательности порядка размещения дополнительных слов определяется наименьшим общим кратным чисел 6 и 256 и составит 256×3=768 стохастических циклов. Во избежание зацикливания вводятся перестановки дополнительных слов и перестановки порядка размещения. Тогда мощность множества комбинаций составит $\tau_5 = 768 \cdot 250 \cdot 6! \cdot 256! \cong 10^{515}$ слов. Из сравнения τ_4 и τ_5 видно, что период повторения слов для рассматриваемого стохастического ГСЧ определяется периодом повторения слов основной матрицы τ_4.

Существуют практически важные случаи, когда требуется стохастический генератор, порождающий слова большей размерности – 16, 24 или 32 бит. Увеличение размерности слова за счет соответствующего увеличения модуля M равенства (1) приводит к полной переработке стохастического ГСЧ и снижению его быстродействия за счет увеличения размерности слова. Этого можно избежать, если выходное слово ГСЧ размерности 16 образовывать конкатенацией двух слов размерности 8, а выходное слово размерности 32 – конкатенацией четырех слов размерности 8. Такой подход приводит к необходимости решения задачи создания группы из 2,3…q, синхронно работающих 8-разрядных стохастических ГСЧ. Здесь под синхронной работой нескольких ГСЧ будем понимать режим, когда вывод сгенерированного слова со всех ГСЧ происходит под одним тактом, а затраты времени на вычисление следующего слова не превосходят допустимого значения, которое меньше тактового интервала вывода слов из ГСЧ. Для организации группы стохастических ГСЧ можно:

– использовать разные таблицы перестановок для массивов ГСЧ;

– использовать различные значения параметров К и С для ОКГ (при выполнении основного требования – длина цикла L=251);

– использовать разные сочетания этих приемов.

При использовании стохастического ГСЧ для криптографических приложений к формируемой последовательности дополнительно предъявляются следующие требования [8]:

– статистическая устойчивость – последовательность чисел на выходе ГСЧ должна быть равномерно распределена, а корреляции должны отсутствовать;

– непредсказуемость – невозможность надежно предсказать значение последующего слова по одному или группе предшествующих слов при отсутствии знаний некоторых параметров формирования псевдослучайных чисел (играющих роль ключа);

– воспроизводимость – возможность повторить ту же самую последовательность псевдослучайных чисел в разных устройствах или в разные моменты времени при известном ключе.

Для обеспечения указанных свойств важно определить параметры стохастического ГСЧ, которые могут играть роль ключа, при этом важнейшую роль играет размерность ключевого пространства. В целом, данный вопрос применительно к предложенному стохастическому ГСЧ является предметом отдельного исследования. Отметим лишь, что существует возможность создания ключевого пространства достаточно большой размерности, используя принципы, изложенные выше для организации группы стохастических ГСЧ.

Полученные результаты

Используя предложенную методику построения ГСЧ, создан программный стохастический ГСЧ с равномерным распределением некоррелированных чисел в диапазоне $\overline{0,255}$. Внешний вид лицевой панели программного устройства приведен на рис. 2.

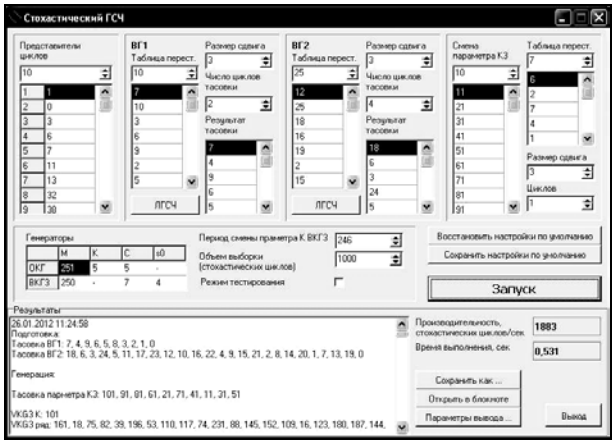

Рис. 2. Панель управления стохастическим ГСЧ

Предусмотрены три режима работы ГСЧ:

– настройки ГСЧ;

– генерации заранее заданного числа стохастических циклов;

– тестирования полученной стохастической последовательности.

Режим настройки предусматривает настройки «по умолчанию» и настройки с изменением параметров ГСЧ.

Настройки «по умолчанию» – установка параметров ГСЧ в соответствии с изложенными выше параметрами (интервал распределения случайной величины $\overline{0,255}$, M=251,K=5,C=5, представители циклов – в соответствии с табл. 1, нуль-цикл 187).

Настройки с изменением параметров позволяют изменять практически все параметры ГСЧ – параметры ОКГ, величину сдвига при декорреляции, число циклов «сдвиг-перестановка» и таблицы перестановок. Предусмотрена возможность использования линейного ГСЧ, решающего уравнение (1), с последующим использованием полученных последовательностей в качестве таблиц перестановок.

Программный ГСЧ в режиме генерации предусматривает возможность формирования стохастической последовательности заранее заданного объема и сохранения параметров ГСЧ и самих последовательностей в текстовом файле. Режим тестирования обеспечивает возможность проверки полученной стохастической последовательности по двум показателям:

– последовательности двух смежных стохастических циклов должны быть несовпадающими (последовательности не совпадают – тест пройден; тест не пройден, если имеет место совпадения хотя бы двух слов, т.е. одинаковое размещение одного и того же слова в двух смежных стохастических циклах);

– каждое из слов интервала $\overline{0,255}$ должно встречаться в стохастическом цикле только один раз.

Отметим, что проверка на отсутствие корреляции между словами стохастической последовательности, проверка периода повторения стохастической последовательности не проводилась, поскольку являются объемной самостоятельной задачей и предметом дальнейшего исследования.

Выводы

Выполненное исследование позволило сформулировать ряд важных выводов:

– существует возможность преобразования генератора конгруэнтной последовательности чисел в стохастический генератор;

– изложенный принцип построения стохастического ГСЧ гарантирует равномерный закон распределения и отсутствие корреляции между словами стохастической последовательности;

– при производительности стохастического ГСЧ, равной 10^{12} слов/сек, максимальный период повторения последовательности составляет сотни лет;

– показана возможность увеличения размерности слова ГСЧ и, соответственно, мощности множества слов путем конкатенации слов синхронно работающих стохастических ГСЧ.

Список литературы

1. Бокс Дж. Анализ временных рядов, прогноз и управление: Пер. с англ. под ред. В.Ф. Писаренко / Дж. Бокс, Г. Дженкинс. – Кн. 1. – М.: Мир, 1974. – 406 с.

2. Фауре Э.В. Выбор параметров генератора конгруэнтных чисел / Э.В. Фауре, Д.В. Фауре, И.Н. Коротєєв // Сучасна спеціальна техніка. – 2010. – №1 (20). – С. 30-35.

3. Иванов М.А. Теория, применение и оценка качества генераторов псевдослучайных последовательностей / М.А. Иванов, И.В. Чугунков. – М.: КУДИЦ-ОБРАЗ, 2003. – 240 с. – (СКБ – спец/ по компьютерной безопасности).

4. Кнут Д.Э. Искусство программирования. Том 2. Получисленные алгоритмы / Дональд Э. Кнут. – М.: Вильямс, 2007. – 832 с.

5. Митянкина Т.В. Рандомизация последовательности конгруэнтных чисел / Т.В. Митянкина, В.В. Швыдкий, А.И. Щерба // Вестник Инженерной академии Украины. – 2008. – №2. – С. 107-111.

6. Митянкина Т.В. Оценка качества генераторов случайных чисел / Т.В. Митянкина, В.В. Швыдкий, А.И. Щерба, М.А. Митянкин // Вісник ЧДТУ. – 2009. – №1. – С. 41-46.

7. Пат. 40649 Україна, МПК G 06 F 7/58. Пристрій декореляції випадкової послідовності чисел / Мітянкіна Т.В., Швидкий В.В., Мітянкін М.О.; заявник та патентовласник ЧДТУ. – №и200811384; заявл. 22.09.2008; опубл. 27.04.2009, Бюл. № 8.

8. Бараш Л.Ю. Генерация случайных чисел и параллельных потоков случайных чисел для расчетов Монте-Карло / Л.Ю. Бараш, Л.Н. Щур // Труды Семинара по вычислит/ технологиям в естественных науках. – М., 2011.

Поступила в редколлегию 23.02.2012

Рецензент: д-р техн. наук, проф. В.Н. Рудницкий, Черкасский государственный технологический университет, Черкассы.

ГЕНЕРАЦІЯ СТОХАСТИЧНОЇ ПОСЛІДОВНОСТІ ГЕНЕРАТОРОМ КОНГРУЕНТНИХ ЧИСЕЛ

В.В. Швидкий, Е.В. Фауре, В.В. Веретельник, В.О. Щерба

Розглядається спосіб генерації стохастичної послідовності чисел конгруентним генератором. Дається оцінка максимального значення періоду повторення стохастичної послідовності. Показана можливість отримання некорельованої рівномірно розподіленої випадкової послідовності конгруентних чисел з максимальним періодом повторення, що складає десятки років (за продуктивності генератора випадкових чисел до 10^{12} випадкових чисел/сек).

Ключові слова: *генератор випадкових чисел, генератор конгруентних чисел, стохастична послідовність, рівномірно розподілена послідовність чисел.*

GENERATION OF STOCHASTIC SEQUENCE BY CONGRUENTIAL GENERATOR

V.V. Shvidkiy, E.V. Faure, V.V. Veretelnik, V.A. Shcherba

In the article the way of stochastic sequence generation by congruential generator is examined. The estimation of the maximum period of repetition of a stochastic sequence is done. The possibility of getting the uncorrelated uniformly distributed random sequence of congruent numbers of maximum period, which amounts to tens of years (if the random number generator productivity is up to 10^{12} random numbers per second).

Keywords: *random number generator, congruential generator, stochastic sequence, uniformly distributed sequence.*