УДК 004.056:519.254

ДВУХУРОВНЕВОЕ ТЕСТИРОВАНИЕ СЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ДЛИНОЙ 128 И 256 БИТ

Н.Г. КИЕВЕЦ[1], А.И. КОРЗУН[2]

[1]*Белорусская государственная академия связи, Республика Беларусь*

[2]*ЗАО «Центр новых интеллектуальных интегрированных систем», Республика Беларусь*

Поступила в редакцию 3 апреля 2017

Аннотация. Найдены теоретические распределения вероятностей превышения тестовыми статистиками значений, получаемых при тестировании случайных последовательностей длиной 128 и 256 бит по тесту на самые длинные подпоследовательности единиц в блоках. Приведены результаты двухуровневого тестирования случайных последовательностей, выработанных генераторами электронных пластиковых карт.

Ключевые слова: двухуровневое тестирование, случайная последовательность, теоретическое распределение.

Abstract. The theoretical distributions of probability of exceeding the test statistics values for test for the longest run of ones in a block corresponded to random 128-bit and 256-bit sequences were found. The results of two-level testing of random sequences produced by electronic plastic card random number generators were presented.

Keywords: two-level testing, random sequence, theoretical distribution.

Doklady BGUIR. 2017, Vol. 105, No. 3, pp. 78–84
Two-level testing of 128-bit and 256-bit random sequences
N.G. Kiyevets, A.I. Korzun

Введение

В криптографических системах для генерации ключевой информации требуется применение генераторов случайных чисел (ГСЧ), которые вырабатывают равномерно распределенные случайные последовательности (РРСП). По определению РРСП состоят из независимых элементов с равномерным распределением вероятностей [1]. Наиболее предпочтительными являются ГСЧ, в которых используются физические источники случайности [2]. Для выявления отказов, сбоев и изменений физических параметров при функционировании физических источников должна быть разработана процедура тестирования выходных последовательностей ГСЧ, которая должна включать оценку энтропии источника случайности генератора [2].

Оценка качества работы ГСЧ выполняется с использованием систем тестов. Наиболее известные системы тестирования требуют относительно длинных битовых последовательностей. Для тестов стандарта FIPS 140 необходимо 20 тыс. бит данных [3], для системы CRYPT-X – от 100 тыс. бит [4], для системы NIST – от 1 млн бит [5]. Однако на практике для шифрования требуются случайные последовательности (СП) определенных длин, которые применяются в качестве ключей. Например, в симметричных алгоритмах шифрования используются ключи следующих длин: 128, 192 и 256 бит в AES (Advanced Encryption Standard); 128 бит в IDEA; 256 бит в алгоритме ГОСТ 28147-89 [6]. СП определенной длины

можно получить, взяв отрезок длинной протестированной последовательности, если она является РРСП. Поскольку сложно доказать, что последовательность является РРСП, то результат ее тестирования следует с осторожностью распространять на ее отрезки [7]. Кроме того, тестирование длинных последовательностей не позволяет оценить энтропию ГСЧ как источника СП определенных длин. Так как энтропия источника максимальна, когда СП вырабатываются с одинаковыми вероятностями, то требуется проверка равновероятности генерируемых последовательностей.

Решить задачу проверки статистических свойств каждой из сгенерированных СП и проверить равновероятность СП позволяет применение методики двухуровневого тестирования [7]. Разработчики системы NIST рекомендуют использование этой методики применительно к СП длиной от 1 млн бит [5]. При двухуровневом тестировании СП длиной от 1 млн бит можно выделить два основных этапа.

1. Для каждой из СП рассчитывается тестовая статистика и определяется величина P_T – вероятность превышения полученного значения статистики при предположении о том, что СП состоит из равномерно распределенных и независимых элементов. На данном этапе выполняется проверка статистических свойств каждой СП.

2. Проверяется равномерность распределения значений P_T с использованием критерия согласия «хи-квадрат». Равномерность распределения значений P_T свидетельствует о равновероятности генерируемых СП.

В связи с тем, что в тестах вместо действительных теоретических распределений статистик используются аппроксимированные распределения, при относительно коротких длинах СП может иметь место отклонение значений P_T от равномерного распределения [7, 8]. Следовательно, при двухуровневом тестировании СП требуется находить теоретические распределения значений P_T. В работе [8] найдены теоретические распределения значений величины P_T для частотного теста и теста на подпоследовательности одинаковых бит системы NIST, приведены результаты тестирования СП длиной 128 и 256 бит. Данная статья посвящена нахождению теоретических распределений P_T для теста на самые длинные подпоследовательности единиц в блоках.

Нахождение теоретических распределений вероятностей P_T для теста на самые длинные подпоследовательности единиц в блоках

Для нахождения распределения случайной величины P_T необходимо знать все ее возможные значения p_{Ti} (i – номер значения) и соответствующие вероятности их появления $P(p_{Ti})$. Так как существует конечное число битовых последовательностей заданной длины n, то величина P_T является дискретной. Величина P_T однозначно связана с тестовой статистикой S, поэтому вероятности появления значений P_T совпадают с вероятностями появления соответствующих значений S. В связи с этим можно выделить следующие этапы нахождения распределения значений P_T.

1. Определяются все возможные значения статистики S: $s_1, s_2, ..., s_k$.

2. Определяются все возможные значения p_{Ti} величины P_T, которые являются вероятностями превышения соответствующих значений статистики $p_{Ti} = P(S > s_i)$.

3. Определяются значения вероятности $P(P_T)$ появления всех возможных значений P_T, которые равны вероятностям появления соответствующих значений статистики S: $P(p_{Ti}) = P(s_i)$.

Для проведения тестирования по тесту на самые длинные подпоследовательности единиц в блоках при длинах СП до 6272 бит, рекомендуется разбивать СП на подпоследовательности (блоки) длиной $M = 8$ бит [5]. При этом каждая из подпоследовательностей относится к одной из четырех категорий в зависимости от длины l самой длинной непрерывной подпоследовательности единиц в блоке.

К первой категории относят подпоследовательности, для которых $l \le 1$; ко второй категории относят подпоследовательности, для которых $l = 2$; к третьей категории – подпоследовательности с параметром $l = 3$; к четвертой категории относят подпоследовательности, для которых $l \ge 4$.

В соответствии с определенным порядком находим распределения вероятностей P_T для теста на самые длинные подпоследовательности единиц в блоках.

1. Определяем все возможные значения s_i тестовой статистики.

1.1. Формируем массив V, содержащий все возможные комбинации чисел подпоследовательностей, относящихся к каждой из четырех категорий.

$$V = \left\{v_k^{(j)}\right\} = \begin{pmatrix} v_0^{(1)} = N & v_1^{(1)} = 0 & v_2^{(1)} = 0 & v_3^{(1)} = 0 \\ v_0^{(2)} = N-1 & v_1^{(2)} = 1 & v_2^{(2)} = 0 & v_3^{(2)} = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{pmatrix}, \quad (1)$$

где $v_k^{(j)}$ – j-я комбинация чисел, принадлежащая к k-й категории.

Для чисел, содержащихся в массиве V, при всех j должно выполняться условие

$$\sum_{k=0}^{3} v_k^{(j)} = N. \quad (2)$$

1.2. Для каждой комбинации $v_k^{(j)}$ рассчитываем значение тестовой статистики [5]:

$$\chi_j^2 = \sum_{k=0}^{3} \frac{(v_k^{(j)} - N\pi_k)^2}{N\pi_k}, \quad (3)$$

где N – количество подпоследовательностей; π_k – вероятность отнесения подпоследовательности к k-й категории.

1.3. Формируем массив $S = \{s_i\}$, содержащий все возможные значения тестовой статистики, рассчитанные в соответствии с выражением (3).

2. Для каждого из значений s_i определяем значение вероятности P_T [5]:

$$p_{Ti} = \left(\int_{s_i/2}^{\infty} t^{\frac{1}{2}} e^{-t} dt \right) \Big/ \left(\int_{0}^{\infty} t^{\frac{1}{2}} e^{-t} dt \right). \quad (4)$$

3. Находим значения $P(p_{Ti})$ вероятности $P(P_T)$, для чего найдем значения $P(s_i)$ вероятности $P(S)$. Значения $P(s_i)$ определяются из выражения

$$P(s_i) = \sum_j P\left(\chi_j^2\right). \quad (5)$$

В формуле (5) суммируются только значения вероятностей $P\left(\chi_j^2\right)$ тестовых статистик χ_j^2, которые равны значению s_i.

Значения $P\left(\chi_j^2\right)$ равны вероятностям $P\left(v^{(j)}\right)$ появления соответствующих комбинаций $v^{(j)} = \left\{v_k^{(j)}\right\}$ $\left(k = \overline{0,3}\right)$, поэтому необходимо найти значения $P\left(v^{(j)}\right)$.

Пусть событие A состоит в том, что $v_0^{(j)} = a$, событие B состоит в том, что $v_1^{(j)} = b$, событие C состоит в том, что $v_2^{(j)} = c$ и событие D состоит в том, что $v_3^{(j)} = d$. События A, B, C и D являются совместными событиями, поэтому вероятность их одновременного появления P_{ABCD} будем рассчитывать по формуле [9]:

$$P_{ABCD} = P(A)P_A(B)P_{AB}(C)P_{ABC}(D), \quad (6)$$

где $P(A)$ – вероятность события A при условии того, что события B, C и D еще не наступили; $P_A(B)$ – вероятность события B при условии того, что событие A уже наступило,

а события C и D еще не наступили; $P_{AB}(C)$ – вероятность события C при условии того, что события A и B уже наступили, а событие D еще не наступило; $P_{ABC}(D)$ – вероятность события D при условии, что события A, B и C уже наступили.

Вероятность $P_{ABC}(D)=1$, так как события A, B, C и D представляют полную группу событий, в которой D произойдет обязательно, если произошли A, B и C. Получаем:

$$P_{ABCD}=P(A)P_A(B)P_{AB}(C). \tag{7}$$

Вероятность появления последовательности с параметром $v_0^{(j)}=a$, т. е. наступления события A:

$$P(A)=\frac{\pi_0^a\left(1-\pi_0\right)^{a-1}N!}{a!\left(N-a\right)!}. \tag{8}$$

После наступления события A точно известно, что a подпоследовательностей принадлежит к первой категории и ко второй категории могут относиться максимум $(N-a)$ подпоследовательностей. Вероятность попадания подпоследовательности во вторую категорию становится равной $\pi_1^{/}=\pi_1/(1-\pi_0)$ и вероятность $P_A(B)$ определяется:

$$P_A(B)=\frac{\left(\pi_1^{/}\right)^b\left(1-\pi_1^{/}\right)^{N-a-b}(N-a)!}{b!(N-a-b)!}. \tag{9}$$

Аналогичным образом получаем выражение для вероятности $P_{AB}(C)$ с учетом того, что a и b подпоследовательностей из N уже принадлежат к первой и второй категориям соответственно:

$$P_{AB}(C)=\frac{\left(\pi_2^{/}\right)^c\left(1-\pi_2^{/}\right)^{N-a-b-c}(N-a-b)!}{c!(N-a-b-c)!}, \tag{10}$$

где $\pi_2^{/}=\pi_2/(1-\pi_0-\pi_1)$.

Подставив в формулу (7) выражения (8)–(10), получаем:

$$P_{ABCD}=\frac{\pi_0^a\cdot\left(1-\pi_0\right)^{N-a}N!}{a!\cdot(N-a)!}\cdot\frac{\left(\pi_1^{/'}\right)^b\cdot\left(1-\pi_1^{/}\right)^{N-a-b}\cdot(N-a)!}{(N-a-b)!\cdot b!}\times$$
$$\times\frac{\left(\pi_2^{/}\right)^c\cdot\left(1-\pi_2^{/}\right)^{N-a-b-c}\cdot(N-a)\cdot(N-a-b)!}{(N-a-b-c)!\cdot c!}. \tag{11}$$

Получили выражение (11), которое позволяет рассчитать значения $P\left(v^{(j)}\right)=P(v_0^{(j)}=a,v_1^{(j)}=b,v_2^{(j)}=c,v_3^{(j)}=d)=P_{ABCD}$. Значения $P\left(v^{(j)}\right)$ равны соответствующим значениям $P\left(\chi_j^2\right)$. Подставив в формулу (5) вместо $P\left(\chi_j^2\right)$ выражение для $P\left(v^{(j)}\right)$ в соответствии с формулой (11) и вместо значений $P(s_i)$ значения $P(p_{Ti})$, получаем:

$$P(p_{Ti})=\sum_j\left(\frac{\pi_0^{v_0^{(j)}}\cdot\left(1-\pi_0\right)^{N-v_0^{(j)}}N!}{(v_0^{(j)})!\cdot(N-v_0^{(j)})!}\cdot\frac{\left(\frac{\pi_1}{(1-\pi_0)}\right)^{v_1^{(j)}}\cdot\left(1-\frac{\pi_1}{(1-\pi_0)}\right)^{N-v_0^{(j)}-v_1^{(j)}}\cdot\left(N-v_0^{(j)}\right)!}{\left(N-v_0^{(j)}-v_1^{(j)}\right)!\cdot\left(v_1^{(j)}\right)!}\times\right.$$
$$\left.\times\frac{\left(\frac{\pi_2}{(1-\pi_0-\pi_1)}\right)^{v_2^{(j)}}\cdot\left(1-\frac{\pi_2}{(1-\pi_0-\pi_1)}\right)^{N-v_0^{(j)}-v_1^{(j)}-v_2^{(j)}}\cdot\left(N-v_0^{(j)}-v_1^{(j)}\right)!}{\left(N-v_0^{(j)}-v_1^{(j)}-v_2^{(j)}\right)!\cdot\left(v_2^{(j)}\right)!}\right). \tag{12}$$

На основе выражений (3)–(5) и (12) найдены распределения вероятностей P_T при длинах СП 128 и 256 бит. Полученные распределения представлены на рис. 1. Из рисунка

видно, при $n=128$ бит и $n=256$ бит для теста на самые длинные подпоследовательности единиц в блоках значения P_T распределены неравномерно.

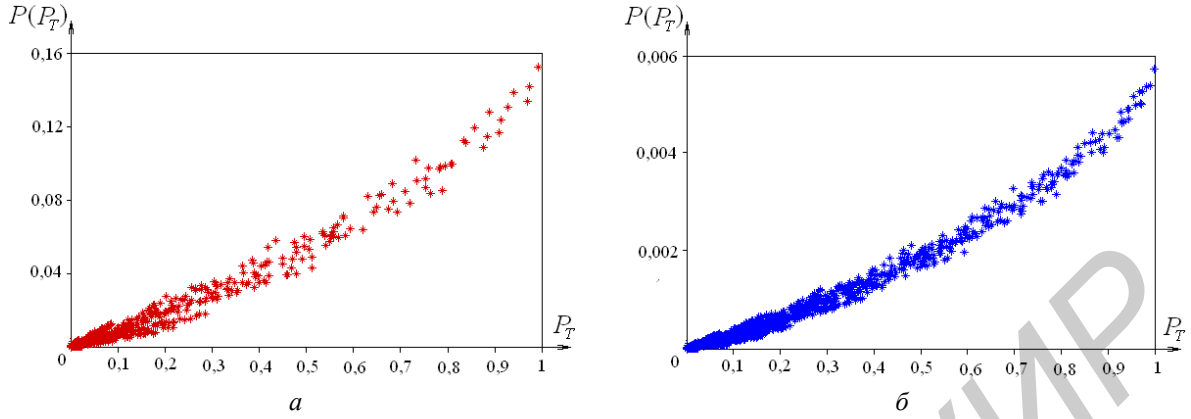

Рис. 1. Теоретические распределения вероятностей P_T : *а* – при *n* = 128 бит; *б* – при *n* = 256 бит

Диапазон возможных значений величины P_T был разбит на интервалы L: $l_1=[0;0,1]$, $l_2=(0,1;0,2]$, $l_3=(0,2;0,3]$, $l_2=(0,1;0,2]$, $l_2=(0,3;0,4]$, $l_5=(0,4;0,5]$, $l_6=(0,5;0,6]$, $l_7=(0,6;0,7]$, $l_8=(0,7;0,8]$, $l_9=(0,8;0,9]$, $l_{10}=(0,9;1]$. Далее рассчитаны значения p_i вероятности P попадания значений P_T в интервалы l_i $\left(i=\overline{1,10}\right)$. Значения величины P представлены в табл. 1, а соответствующие гистограммы на рис. 2.

Таблица 1. **Значения вероятностей *P* для СП различных длин *n***

L	*P* (*n* = 128 бит)	*P* (*n* = 256 бит)	*L*	*P* (*n* = 128 бит)	*P* (*n* = 256 бит)
[0-0,1]	0,0940	0,0943	(0,5–0,6]	0,1200	0,1127
(0,1-0,2]	0,1053	0,1045	(0,6–0,7]	0,0775	0,0964
(0,2-0,3]	0,1016	0,1058	(0,7–0,8]	0,1092	0,0973
(0,3-0,4]	0,1065	0,1030	(0,8–0,9]	0,0892	0,0983
(0,4-0,5]	0,1030	0,0894	(0,9–1]	0,0936	0,0981

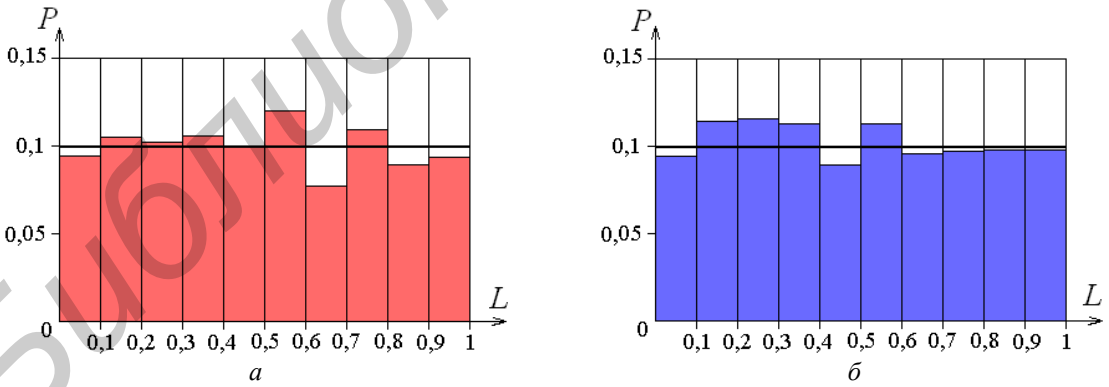

Рис. 2. Гистограммы вероятностей *P*: *а* – при *n* = 128 бит; *б* – при *n* = 256 бит

Из гистограмм на рис. 2 видна неравномерность распределения значений P_T как при длине СП *n* = 128 бит, так и при *n* = 256 бит. Видно, что при *n* = 256 бит гистограмма более равномерна, чем при *n* = 128 бит.

Результаты двухуровневого тестирования случайных последовательностей длиной 128 и 256 бит

Для оценки качества работы ГСЧ четырех электронных пластиковых карт (ЭПК) было выполнено двухуровневое тестирование сгенерированных СП по тесту на самые длинные подпоследовательности единиц в блоках. Из каждой ЭПК получены наборы по $N_1=8000$ СП

длиной 128 бит и по $N_2 = 4000$ СП длиной 256 бит. Далее было выполнено тестирование каждого набора СП по следующему алгоритму.

1. Каждая СП протестирована по тесту на самые длинные подпоследовательности единиц в блоках, в результате чего получен массив значений P_T.

2. Рассчитаны количества m_i значений P_T, попадающих в соответствующие интервалы l_i $\left(i = \overline{1,10}\right)$.

3. Рассчитана случайная величина χ^2 [5]:

$$\chi^2 = \sum_{i=1}^{10} \frac{\left(m_i - Np_i\right)^2}{Np_i}, \qquad (13)$$

где N – количество значений P_T; p_i – вероятность попадания значения P_T в интервал l_i $\left(i = \overline{1,10}\right)$.

4. Рассчитана вероятность превышения полученного по формуле (13) значения χ^2 [5]:

$$P_c = 1 - \Gamma_{\chi^2/2}(9/2) / \Gamma(9/2), \qquad (14)$$

где $\Gamma_x(a) = \int_x^\infty e^{-t} t^{a-1} dt$ – неполная гамма-функция; $\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt$ – гамма-функция.

5. Выполнено сравнение значения P_c с уровнем значимости $\alpha = 0,0001$ [5]. При значениях $P_c \geq 0,0001$ делается вывод о том, что СП генерируются с одинаковыми вероятностями, т.е. о прохождении последовательностями двухуровневого тестирования.

В табл. 2 для каждого из наборов СП приведено значение вероятности P_c и количество *K* СП, прошедших тест на самые длинные подпоследовательности единиц в блоках.

Таблица 2. **Результаты двухуровневого тестирования СП**

№ ЭПК	1		2		3		4	
n, бит	128	256	128	256	128	256	128	256
P_c	0,3563	0,9206	0,3642	0,0141	0,0306	0,0829	0,5359	0,5755
K	7933	3961	7913	3967	7911	3966	7925	3970

Из табл. 2 видно, что наборы СП, полученных из всех четырех ЭПК прошли двухуровневое тестирование, так как все значения $P_c > 0,0001$.

Заключение

Найдены теоретические распределения значений вероятности P_T для теста на самые длинные подпоследовательности единиц в блоках при длинах СП 128 и 256 бит, что позволило корректно использовать методику двухуровневого тестирования. Выполнено двухуровневое тестирование СП длиной 128 и 256 бит, полученных из четырех ЭПК. Двухуровневое тестирование позволило проверить каждую сгенерированную СП и исключить использование в качестве ключевой информации тех СП, свойства которых не соответствуют РРСП, а также подтвердило, что ГСЧ ЭПК вырабатывают равновероятные СП указанных длин.

Список литературы

1. Харин Ю.С., Берник В.И. Математические основы криптологии. Минск: БГУ, 1999. 319 с.
2. Информационные технологии и безопасность. Требования безопасности к программным средствам криптографической защиты информации: СТБ 34.101.27-2011. Введ. 01.03.12. Минск: Госстандарт, 2012. 33 с.
3. Federal Information Processing Standards Publication 140-2. Security Requirements for Cryptographic Modules. [Электронный ресурс]. – Режим доступа: http://mayor.fri.utc.sk/v731/04/fips140-2.pdf – Дата доступа: 13.10.2014.

4. Evaluation of TOYOCRYPT-HR1. [Электронный ресурс]. – Режим доступа: http://www.cryptrec.go.jp/estimation/report_ID1090.pdf. – Дата доступа: 06.07.14.
5. A statistical test suite for random and pseudorandom number generators for cryptographic applications [Электронный ресурс]. – Режим доступа: http://csrc.NIST.gov/publications/NISTpubs/800-22-rev1a/SP800-22rev1a.pdf – Дата доступа: 30.06.2014.
6. Панасенко С.П. Алгоритмы шифрования. Специальный справочник. СПб.: БХВ-Петербург, 2009. 576 с.
7. L'Ecuyer P. Testing random number generators // Winter Simulation Conference. 1992. P. 305–313.
8. Киевец Н.Г., Корзун А.И. Методика нахождения эталонных законов распределения вероятностей, получаемых при статистическом тестировании последовательностей ключей // Докл. БГУИР. 2014. № 5 (83). С. 38–43.
9. Справочник по вероятностным расчетам / Г.Г. Абезгауз [и др.]. М.: Воениздат, 1970. 536 с.

References

1. Harin Ju.S., Bernik V.I. Matematicheskie osnovy kriptologii. Minsk: BGU, 1999. 319 s. (in Russ.)
2. Informacionnye tehnologii i bezopasnost'. Trebovanija bezopasnosti k programmnym sredstvam kriptograficheskoj zashhity informacii: STB 34.101.27-2011. Vved. 01.03.12. Minsk: Gosstandart, 2012. 33 s. (in Russ.)
3. Federal Information Processing Standards Publication 140-2. Security Requirements for Cryptographic Modules [Electronic resource]. – Access mode: http://mayor.fri.utc.sk/v731/04/fips140-2.pdf – Date of access: 13.10.2014.
4. Evaluation of TOYOCRYPT-HR1 [Electronic resource]. – Access mode: http://www.cryptrec.go.jp/estimation/report_ID1090.pdf. – Date of access: 06.07.14.
5. A statistical test suite for random and pseudorandom number generators for cryptographic applications [Electronic resource]. – Access mode: http://csrc.NIST.gov/publications/NISTpubs/800-22-rev1a/SP800-22rev1a.pdf – Date of access: 30.06.2014.
6. Panasenko S.P. Algoritmy shifrovanija. Special'nyj spravochnik. SPb.: BHV-Peterburg, 2009. 576 s. (in Russ.)
7. L'Ecuyer P. Testing random number generators // Winter Simulation Conference. 1992. P. 305–313.
8. Kievec N.G., Korzun A.I. Metodika nahozhdenija jetalonnyh zakonov raspredelenija verojatnostej, poluchaemyh pri statisticheskom testirovanii posledovatel'nostej kljuchej // Dokl. BGUIR. 2014. № 5 (83). S. 38–43. (in Russ.)
9. Spravochnik po verojatnostnym raschetam / G.G. Abezgauz [i dr.]. M.: Voenizdat, 1970. 536 s. (in Russ.)

Сведения об авторах

Киевец Н.Г., старший преподаватель кафедры радио и информационных технологий Белорусской государственной академии связи.

Корзун А.И., к.т.н., доцент, директор ЗАО «Центр новых интеллектуальных интегрированных систем».

Адрес для корреспонденции

220013, Республика Беларусь,
г. Минск, ул. П. Бровки, 14
Белорусская государственная академия связи
тел. +375-17-202-12-81
e-mail: kievetss@mail.ru
Киевец Наталья Григорьевна

Information about the authors

Kiyevets N.G., senior lecturer of radio and information technology department of Belarusian state academy of telecommunications.

Korzun A.I., PhD, assistant professor, head of closed company «Center of new intellectual integrated systems».

Address for correspondence

220013, Republic of Belarus,
Minsk, P. Brovka st., 14
Belarusian state academy of telecommunications
tel. +375-17-202-12-81
e-mail: kievetss@mail.ru
Kiyevets Natallia Grigorevna