DOI: 10.18372/2310-5461.46.14813

УДК 519.212.2:681.51: 621.317

С. В. Поперешняк, канд. фіз.-мат. наук, доц.
Київський національний університет імені Тараса Шевченка
http://orcid.org/0000-0002-0531-9809
spopereshnyak@gmail.com

ТЕСТУВАННЯ ГЕНЕРАТОРА ПСЕВДОВИПАДКОВИХ ЧИСЕЛ ЯК СКЛАДОВА БЕЗПЕКИ ІНТЕРНЕТУ РЕЧЕЙ

Вступ

Концепція Інтернету Речей (Internet of Things, IoT), яка стрімко розвивається останнім часом, приносить нові підходи до повсякденних проблем, а також промислових застосувань. Ці підходи покладаються на групи дешевих, ефективних та виділених мережевих пристроїв, які працюють та взаємодіють постійно [1]. Ці так звані «легкі» пристрої IoT мають обмежену потужність, простір та обчислювальні ресурси; отже, існує величезна потреба у розробці відповідних протоколів та методологій безпеки, пристосованих до них. Крім того, більшість сучасних протоколів безпеки не оптимізовані для легких середовищ і потребують більшої кількості джерел, ніж можливо на пристроях IoT. Наприклад, майже всі протоколи аутентифікації використовують генератори випадкових чисел для функціонування, а їх вдосконалення можуть сприяти використанню пристроїв IoT у тих областях, де є проблеми безпеки.

У цій статті пропонуємо рекомендацію до тестування генератора випадкових чисел для застосувань безпеки в обмежених пристроях. Працюючи з генераторами випадкових чисел (ГВЧ), найважливіше, що слід врахувати, — це те, що немає ідеального генератора, який би виконував усі умови [2]. Це в основному тому, що справжня випадковість — це недетермінований процес, який неможливо синтезувати за допомогою математичних методів у програмному середовищі, і це призводить до концепції псевдовипадковості. З іншого боку, навіть використовуючи апаратні джерела, практично важко створити ряд чисел, які мають очікувані характеристики справжньої випадковості.

У будь-якому випадку якість генератора корелює з його близькістю до справжньої випадковості та його обчислювальними вимогами [3]. Проте, висока ступінь випадковості може бути дуже коштовної або неефективною або просто непотрібною в деяких випадках. Отже, найкращий варіант слід обирати відповідно до потреб галузі, де буде застосовуватися дана випадкова послідовність.

Постановка проблеми

Сучасні інформаційні системи потребують особливого підходу до передачі даних по відкритих каналах зв'язку, що забезпечується з використанням систем безпеки. Ефективна система безпеки інформації повинна забезпечувати:

- таємність інформації або важливої її частини;
- автентичність суб'єктів і об'єктів інформаційної взаємодії;
- захист від несанкціонованого доступу;
- захист прав власників інформації;
- оперативний контроль процесів управління, обробки і передачі інформації.

Усі перераховані пункти, яким повинна задовольняти ефективна система безпеки інформації, вирішуються з використанням генератора псевдовипадкових послідовностей (ПВП). Такі генератори є детермінованими алгоритмами і застосовують як вхідні дані початкового значення, а на виході породжують послідовність значень, яка дуже схожі на випадкову [3].

У праці [4] запропонована така класифікація вибірок за чисельністю, виходячи з вимог представлених в програмі критеріїв:

- дуже малі вибірки — від 5 до 12;
- малі вибірки — від 13 до 40;
- вибірки середньої чисельності — від 41 до 100;
- великі вибірки — від 101 і вище.

Для виявлення закономірностей аналізованих ПВП (або до їх відрізків різної довжини) застосовують широкий спектр різних статистичних тестів, розроблених в останні десятиліття. Попередній аналіз розглянутих статистичних тестів дозволив визначити, які з них найбільш придатні для використання в різних завданнях розробки засобів криптографічного захисту інформації. Аналіз існуючих пакетів для статистичного тестування призводить до висновку, що багато тестів можуть успішно використовуватися для дослідження ПВП на випадковість. Разом з тим, особливості прикладних задач показують, що класична математична модель статистичного

тестування не цілком адекватно відображає потреби в дослідженні деяких об'єктів на випадковість. Така ситуація виникає, коли генеруються послідовності невеликої довжини (до 100 біт). У такому випадку досліджувану ПВП неможливо коректно протестувати за допомогою одновимірної статистики. У цій ситуації пропонуємо протестувати ланцюжок на випадковість використовуючи дво- і/або тривимірні статистики.

Аналіз останніх досліджень і публікацій

Апаратна безпека для Інтернету речей або кіберсистеми призводить до необхідності застосування криптографії для різних сенсорних інфраструктур в цих галузях. Зокрема, створення надійних криптографічних ключів на такому пристрої з обмеженими ресурсами залежить від криптографічно безпечного генератора випадкових чисел [5]

Праця [6] присвячена розробці та оцінці якості «легких» генераторів псевдовипадкових чисел. Описано принципи конструювання полегшеного генератора псевдовипадкових чисел (lightweight pseudorandom number generators, LW-PNRG). Сформульовано вимоги до якості LW-PNRG. Також у праці [6] наведено приклад перетворення класичного генератора в LW-PNRG за допомогою модифікації функції зворотного зв'язку (OFB). Пропонуються методи оцінки якості псевдовипадкових чисел, що формуються полегшеним генератором, також наведені результати статистичних випробувань розробленої LW-PNRG.

У праці [7] представлено рішення безпеки для бездротових мереж, який використовує новий недорогий генератор псевдовипадкових чисел (ГПВЧ). Генератор псевдовипадкових чисел створюється шляхом представлення певного числа або періодичних хвиль таким чином, що одне і те саме псевдовипадкове число генерується на двох різних пристроях, які спільно використовують конкретний ключ. Згенеровані псевдовипадкові числа використовуються для захисту повідомлень між двома пристроями. Показано, що цей ГПВЧ здатний до самосинхронізації, захищений від більшості атак і може бути реалізований як в апаратному, так і в програмному забезпеченні. Крім того, рішення є легковажним і ідеально підходить для бездротового зв'язку, де пристрої можуть бути обмежені в ресурсах. Проведено аналіз і всебічне тестування з використанням тестів DIEHARD.

Згенеровані псевдовипадкові числа також є криптографічно безпечні і тому підходять для додатків безпеки.

У дослідженні [8] пропонується метод генетичного програмування для створення генераторів псевдовипадкових чисел для легких пристроїв, особливо для пристроїв сімейства платформ бездротової ідентифікації та виявлення. Якість цих ГПВЧ перевіряються за допомогою набору статистичних тестів NIST, який являє собою комплексний інструмент, який оцінює статистичне якість вихідних даних даного ГПВЧ. Крім того, за допомогою цього методу генетичного програмування генерується набір ПВЧ, і результати перевіряються статистично.

Аналіз останніх досліджень і публікацій показує широкий інтерес до безпеки в IoT і, як наслідок, в останній час, запропоновано різні підходи до побудови полегшеного генератору псевдовипадкових чисел, який може застосовуватися в пристроях з певними обмеженими характеристиками. Як показує аналіз, якість даних ГПВЧ перевіряються за допомогою набору статистичних тестів, які були розроблені в останні десятиліття і набули досить широкої популярності.

В наступному розділі розглянемо найбільш відомі статистичні тести та обмеження щодо їх використання.

Огляд інструментів для статистичного тестування ПВП

Для виявлення закономірностей аналізованих ПВП (або до їх відрізків різної довжини) застосовують широкий спектр різних статистичних тестів, розроблених в останні десятиліття. Наведемо відомі набори статистичних тестів:

1. 11 тестів: Donald Knuth (Stanford University);
2. 15 тестів: NIST Statistical Test Suite;
3. 12 тестів: DIEHARD;
4. 11 тестів: TestU01;
5. 5 тестів: Crypt-XS.

Незважаючи на чималу кількість існуючих реалізацій статистичних тестів ПВП, даний напрямок постійно розвивається, і в даний час активно з'являються нові проекти, які пропонують нові реалізації розглянутих тестів. Розглянуті набори статистичних тестів ПВП становлять зручний і гнучкий інструмент дослідження генераторів ПВП, що застосовуються в криптографічних додатках.

Пакет NIST STS має більшу гнучкість, розширюваність і ефективність (з точки зору витрат часу на здійснення тестування) і є найбільш повним з наявних пакетів для статистичного тестування двійкових послідовностей. На практиці прийняття або відхилення нульової гіпотези ґрунтується на результатах застосування декількох незалежних тестів.

Коли незалежні тести призводять до різних висновків, використовується комбінування результатів тестів за допомогою статистик, які враховують сукупність результатів всіх використаних тестів. У табл. 1 наведені деякі тести для тестування ПВП різних довжин [9].

Таблиця 1

Статистичні тести для ПВП різних довжин

№ з/п	Пакети статистичних тестів	Тестування ПВП довжини порядку 300 біт	Тестування ПВП довжини порядку 10^6 біт
1	Тесты Д. Кнута	–	• Frequency criterion • Series criterion • max-t criterion • Criterion of monotony
2	NIST STS	• Frequency Test • Frequency Test within a Block • Runs Test • Non-overlapping Template Matching Test • Serial Test • Binary Matrix Rank Test • Approximate Entropy Test • Cumulative Sums (Cusum) Test • Test for the Longest Run of Ones in a Block	• Frequency Test within a Block • Runs Test • Overlapping Template Matching Test • Serial Test • Linear Complexity Test • Discrete Fourier Transform Test • Maurer Universal Test • Binary Matrix Rank Test • Approximate Entropy Test • Random Excursions Test • Random Excursions Variant Test • Test for the Longest Run of Ones in a Block
3	DIEHARD	–	• The Birthday Spacing Test
4	TestU01	• Hamming Weight Test • Autocorrelation • Random Walk Test • Longest Run of 1's Test	• Hamming Weight Test • Run and Gap Test • Lempel-Ziv Complexity Test • Autocorrelations Test • Longest Run of 1's Test • CAT Test
5	Crypt-XS	• Frequency Test • Binary Derivative Test	• Complexity Test

На жаль, ідеальних генераторів не існує, а список їх відомих властивостей поповнюється переліком недоліків.

Це призводить до ризику використання в комп'ютерному експерименті поганого генератора. Тому перед проведенням комп'ютерного експерименту необхідно або оцінити якість вбудованої в ЕОМ функції генерації випадкових чисел, або вибрати відповідний алгоритм генерації випадкових чисел.

Тому, крім тестування генератора, надзвичайно важливою є перевірка його за допомогою типових задач, що допускають незалежну оцінку результатів аналітичними або чисельними методами.

Попередній аналіз розглянутих статистичних тестів дозволив визначити, хто з них найбільш підходить для використання в різних завданнях розробки інструментів криптографічного захисту інформації. Аналіз існуючих пакетів для статистичного тестування призводить до висновку, що багато тестів можуть бути успішно використані для дослідження тесту послідовностей на випадковість.

У той же час, особливості прикладних задач показують, що класична математична модель статистичного тестування не зовсім адекватно відображає необхідність вивчення певних об'єктів на випадковість [10; 11].

Така ситуація виникає, коли генерується послідовність невеликої довжини (до 100 біт).

Мета — Розглянути сумісні розподіли 2-ланцюжків і 3-ланцюжків фіксованого виду випадкової (0, 1)-послідовності, які дозволяють проводити статистичний аналіз локальних ділянок цієї послідовності.

Наукова новизна

У статті запропоновано критерій для перевірки на випадковість (0, 1)-послідовностей невеликої довжини (до 100 біт). Даний підхід доцільно використовувати для тестування полегшеного генератора псевдовипадкових чисел в мобільних пристроях або пристроях Інтернету Речей з певними обмеженнями на ресурси.

Виклад основного матеріалу

Сумісні розподіли числа 2-ланцюжків і числа 3-ланцюжків фіксованого виду

Розглянемо послідовність випадкових подій

$$\gamma_1 \gamma_2 \ldots \gamma_n, \qquad (1)$$

де $\gamma_i = \{0,1\}$, $i = 1, 2, \ldots, n$, $n > 0$.

Підпослідовність $\gamma_j \gamma_{j+1} \ldots \gamma_{j+s-1}$ послідовності (1) будемо називати s-ланцюжком, $s = 1, 2, \ldots, n$, $j = 1, 2, \ldots, n-s+1$. Позначимо $\eta(t_1 t_2 \ldots t_s)$ число s-ланцюжків в послідовності (1), які співпадають з $t_1 t_2 \ldots t_s$, де $t_i = \{0,1\}$, $i = 1, 2, \ldots, s$. Сформулюємо умову (У).

Умова (У): послідовність (1) складається з n, $n > 0$, незалежних однаково розподілених випадкових величин; ймовірність подій $\{\gamma_i = 1\}$, $\{\gamma_i = 0\}$ відомі та дорівнюють $P\{\gamma_i = 1\} = p$, $P\{\gamma_i = 0\} = q$, $p + q = 1$, $i = 1, 2, \ldots, n$.

У цій статті будуть розглянуті сумісні розподіли випадкових величин $\eta(t_1 t_2 \ldots t_s)$, $\eta(t'_1 t'_2 \ldots t'_{s'})$, де $t_j, t'_i \in \{0,1\}$, $j = 1, 2, \ldots, s$, $i = 1, 2, \ldots, s'$, для деяких значень s і s'. Перейдемо до точних формулювань.

Теорема. Нехай виконуються умови (У) n, k_1, k_2, k_3, t, t_1 — цілі числа такі, що $k_1 \geq 0$, $k_2 \geq 0$, $k_3 \geq 0$, $n \geq \max(2k_1, 3)$, t, $t_1 \in \{0,1\}$. Тоді

$$P\{\eta(t t^*) = k_1, \eta(t 1 t^*) + \eta(t 0 t^*) = k_2\} = \sum_{m_1 = k_1}^{n-k_1} p^{m_1} q^{m_0} \sum \prod_{i=0}^{1} C_{k_1}^{\delta_1} C_{m_i - k_1}^{k_1 - \delta_i}, \qquad (2)$$

де $m_0 = n - m_1$, символ $\sum$ позначає сумування по всім цілим невід'ємним числам δ_0 і δ_1 таким, що $\delta_0 + \delta_1 = 2k_1 - k_2$, $t^* = 1 - t$;

$$P\{\eta(t t^*) = k_1, \eta(t \alpha t^*) = k_2\} = \sum_{m_1 = k_1}^{n-k_1} p^{m_1} q^{m_0} C_{k_1}^{k_2} C_{m_t - k_1}^{k_2} C_{m_{t^*}}^{k_1}, \qquad (3)$$

де $\alpha \in \{0,1\}$;

$$P\{\eta(t t^*) = k_1, \eta(t 1 t^*) = k_2, \eta(t 0 t^*) = k_3\} = \sum_{m_1 = k_1}^{n-k_1} p^{m_1} q^{m_0} C_{k_1}^{k_2} C_{k_1}^{k_3} C_{m_1 - k_1}^{k_2} C_{m_0 - k_1}^{k_3}. \qquad (4)$$

Приклад. Для (0,1)-послідовності виду

0 1 1 0 0 1 1 0 1 1 1 0 0 0 0 1 0 1 0 1

значеннями випадкових величин $\eta(t t^*)$, $\eta(t 1 t^*)$, $\eta(t 0 t^*)$ і $\eta(t \alpha t^*)$, де $\alpha \in \{0,1\}$, для $t = 1$ є відповідно числа 5, 3, 2 і 5.

ПРИКЛАДИ ДО ТЕОРЕМИ

Ілюстрація використання рівності (2)

У табл. 2 і на рис. 1 приведено використання співвідношення (2) для $p = q = 1/2$, малої вибірки довжини n, $n = 32$, і деяких значень k_1, k_2.

Таблиця 2

Значення ймовірності здійснення події для чисел *k* з використанням співвідношення (2)

k_1	k_2	P	P_c
8	8	0,07285	1
8	7	0,06475	0,92715
7	8	0,06459	0,8624
7	7	0,0604	0,79781
8	9	0,0518	0,73742
9	8	0,05099	0,68562
9	7	0,04768	0,63462
7	9	0,04306	0,58694
6	7	0,03752	0,54388
8	6	0,03626	0,50636
7	6	0,03523	0,4701
9	9	0,03399	0,43487
6	8	0,03388	0,40087
9	6	0,02781	0,367
6	6	0,02502	0,33918
10	7	0,02323	0,31417
8	10	0,02308	0,29094
10	8	0,02097	0,26786
6	9	0,01807	0,24689
7	10	0,01762	0,22883
10	6	0,01549	0,21121
9	10	0,01391	0,19572
5	7	0,0137	0,18182
7	5	0,01265	0,16812
8	5	0,01259	0,15547
5	6	0,01199	0,14288
10	9	0,01118	0,1309
6	5	0,01001	0,11971

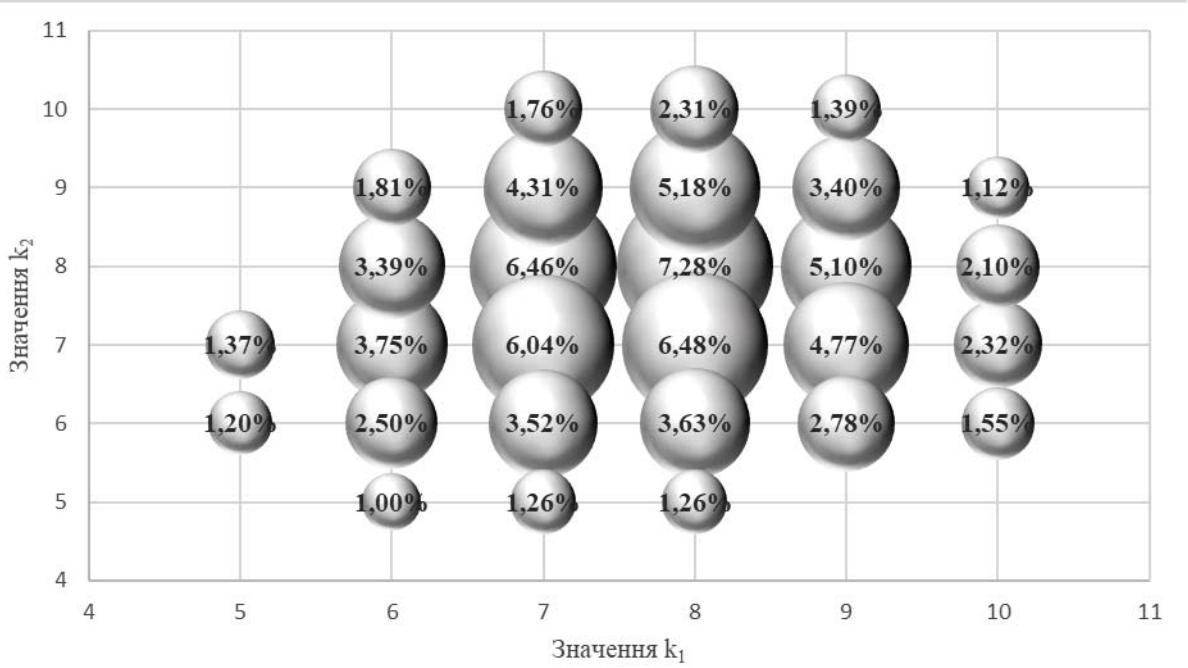

Рис. 1. Діаграма реалізації співвідношення (2)

У табл. 1 у першому та другому стовпчиках розташовані всі можливі варіанти значень k_1 і k_2, для яких імовірність

$$P\left\{\eta(tt^*)=k_1,\eta(t1t^*)+\eta(t0t^*)=k_2\right\}\geq 0{,}01.$$

У третьому стовпчику табл. 1 наведені ймовірності (в неспадному порядку) $P\left\{\eta(tt^*)=k_1,\eta(t1t^*)+\eta(t0t^*)=k_2\right\}$ для пар чисел $(k_1,\ k_2)$, які вказані в перших двох стовпцях.

У кожному рядку четвертого стовпчика розташована сума накопичених імовірностей до реалізації події $\left\{\eta(tt^*)=k_1,\eta(t1t^*)+\eta(t0t^*)=k_2\right\}$ включно, де k_1 і k_2 вказані в цьому самому рядку в першому та другому стовпцях.

Наприклад, для $k_1=9$ і $k_2=7$ маємо: $P\left\{\eta(tt^*)=k_1,\eta(t1t^*)+\eta(t0t^*)=k_2\right\}=0{,}04768$,

$$P_c=\sum \mathrm{P}\left\{\eta(tt^*)=k_1,\eta(t1t^*)+\eta(t0t^*)=k_2\right\}=0{,}63462,$$

де знак сумування $\sum$ розповсюджується на всі пари $(k_1,\ k_2)$, для яких

$$P\left\{\eta(tt^*)=k_1,\eta(t1t^*)+\eta(t0t^*)=k_2\right\}\leq 0{,}04768.$$

Рис. 1 являє собою бульбашкову діаграму, у якій перший параметр (горизонтальна вісь) — значення k_1, другий (вертикальна вісь) — значення k_2, третій (розмір бульбашки) — імовірність здійснення події $\left\{\eta(tt^*)=k_1,\eta(t1t^*)+\eta(t0t^*)=k_2\right\}$, яка представлена у відсотках.

Наприклад, на рис. 1 при $k_1=9$ і $k_2=7$ імовірність здійснення події $\left\{\eta(tt^*)=k_1,\eta(t1t^*)+\eta(t0t^*)=k_2\right\}$ у відсотках дорівнює 7,28 %.

Ілюстрація використання рівності (3)

У табл. 3 і на рис. 2 наведено використання співвідношення (3) для $p=q=1/2$, малої вибірки довжини n, $n=32$, і деяких значень k_1 і k_2.

Табл. 3 утворена зі стовпчиків, інтерпретація яких аналогічна інтерпретації вмісту стовпців табл. 2. Рис. 2 являє собою бульбашкову діаграму, у якій перший параметр (горизонтальна вісь) — значення k_1, другий (вертикальна вісь) — значення k_2, третій (розмір бульбашки) — імовірність здійснення події $\left\{\eta(tt^*)=k_1,\eta(t\alpha t^*)=k_2\right\}$, яка представлена у відсотках.

Ілюстрація використання рівності (4)

У табл. 4 наведено використання співвідношення (4) для $p=q=1/2$, малої вибірки довжини n, $n=20$, і деяких значень k_1, k_2 і k_3.

У табл. 3 в першому, другому та третьому стовпчиках розташовані всі можливі варіанти значень k_1, k_2 і k_3, для яких імовірність $P\left\{\eta(tt^*)=k_1,\eta(t1t^*)=k_2,\ \eta(t0t^*)=k_3\right\}\geq 0{,}01$. У четвертому стовпчику табл. 3 наведені ймовірності (в неспадному порядку) $P\left\{\eta(tt^*)=k_1,\eta(t1t^*)=k_2,\ \eta(t0t^*)=k_3\right\}$ для чисел k_1, k_2 і k_3, які вказані в перших трьох стовпцях.

Таблиця 3

Значення ймовірності здійснення події для чисел *k* з використанням співвідношення (3)

k_1	k_2	P	P_c
8	4	0,08476	1
7	4	0,0787	0,91524
8	3	0,0678	0,83654
7	3	0,06296	0,76874
8	5	0,05812	0,70578
9	4	0,05754	0,64766
7	5	0,05085	0,59012
9	3	0,04882	0,53927
6	4	0,04553	0,49045
6	3	0,03928	0,44492
9	5	0,03836	0,40563
8	2	0,02906	0,36727
7	2	0,02597	0,33822
6	5	0,02428	0,31224
10	4	0,02397	0,28796
10	3	0,02283	0,26399

k_1	k_2	P	P_c
9	2	0,02267	0,24115
8	6	0,02131	0,21849
6	2	0,01637	0,19718
5	3	0,01608	0,18081
7	6	0,01574	0,16473
5	4	0,01528	0,14899
9	6	0,01438	0,13371
10	5	0,01438	0,11933
10	2	0,01199	0,10494
5	2	0,00724	0,09296
11	3	0,00655	0,08572
8	1	0,00609	0,07917
11	4	0,00573	0,07308
9	1	0,00523	0,06735
7	1	0,00509	0,06211
5	5	0,005	0,05702

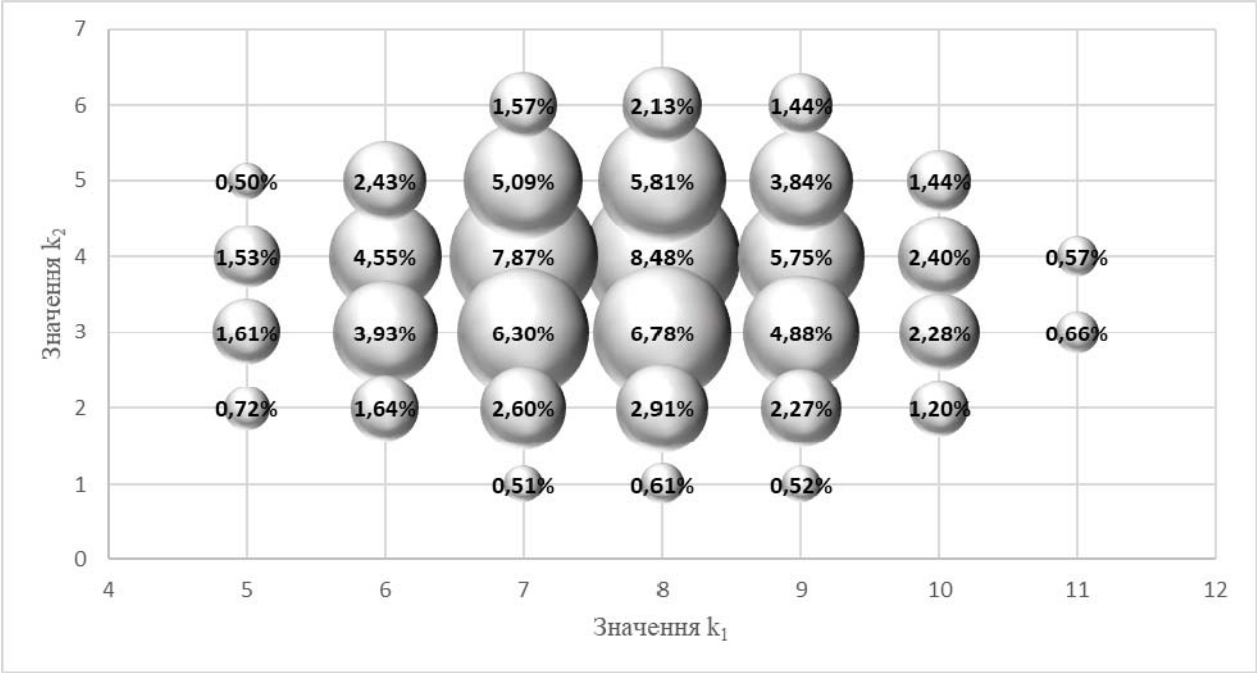

Рис. 2. Діаграма реалізації співвідношення (3)

Таблиця 4

Значення ймовірності здійснення події для чисел *k* з використанням співвідношення (4)

k_1	k_2	k_3	P	P_c
4	2	2	0,04419	1
5	2	2	0,04406	0,95581
5	2	3	0,04406	0,91175
5	3	2	0,04406	0,86769
4	2	3	0,03928	0,82364
4	3	2	0,03928	0,78436
5	3	3	0,03147	0,74508
6	2	2	0,02704	0,71361
4	3	3	0,02618	0,68657
3	2	2	0,02577	0,66039
6	2	3	0,02403	0,63462
6	3	2	0,02403	0,61058
5	1	3	0,02203	0,58655
5	3	1	0,02203	0,56452
4	1	3	0,01964	0,54249
4	3	1	0,01964	0,52285
4	1	2	0,01637	0,50321
4	2	1	0,01637	0,48685
5	1	2	0,01574	0,47048
5	2	1	0,01574	0,45475
5	2	4	0,01574	0,43901
5	4	2	0,01574	0,42328
6	1	3	0,01442	0,40754
6	3	1	0,01442	0,39312
3	2	3	0,01432	0,3787
3	3	2	0,01432	0,36438
6	3	3	0,01373	0,35006
3	1	2	0,01172	0,33633
3	2	1	0,01172	0,32462
5	1	4	0,01101	0,3129
5	4	1	0,01101	0,30188
6	1	2	0,01081	0,29087
6	2	1	0,01081	0,28006

У кожному рядку п'ятого стовпчика розташована сума накопичених імовірностей до реалізації події $\left\{ \eta(tt^*) = k_1,\ \eta(t1t^*) = k_2,\ \eta(t0t^*) = k_3 \right\}$ включно, де k_1, k_2 і k_3 указані в цьому самому рядку в першому, другому та третьому стовпцях.

Результат порівняння тестування ПВП невеликої довжини з використанням статистичного пакету NIST та багатовимірних статистик

Переважна більшість статистичних тестів справді просто порівнюють декілька вибірок на одній змінній або порівнюють кілька змінних в одній сукупності.

Тепер, якщо додати до цього порівняння кількість змінних чи сукупності та вимірювання декількох змінних, то вибір відповідного статистичного тесту раптом стає надзвичайно простим.

Розглянемо відомі приклади, які наведено у працях [12; 13], і обчислимо ймовірності відповідних подій використовуючи розподіли 2-ланцюжків і 3-ланцюжків фіксованого виду випадкової (0, 1)-послідовності.

Покладемо $t=0$. Таким чином розглянемо наступні випадкові величини: $\eta(01)$, $\eta(011)$, $\eta(001)$ і $\eta(0\alpha1)$, де $\alpha \in \{0,1\}$.

Проаналізуємо табл. 5, де використані такі позначення за умови, що $t=0$:

- P, отримана за допомогою відповідного тесту статистичного пакету NIST;
- P_1 — імовірність $P\left\{\eta(tt^*) = k_1, \eta(t1t^*) + \eta(t0t^*) = k_2\right\}$, яку можна обчислити використовуючи співвідношення (2);
- P_2 — імовірність $P\left\{\eta(tt^*) = k_1, \eta(t0t^*) = k_2\right\}$, яку можна обчислити використовуючи співвідношення (3);
- P_3 — імовірність $P\left\{\eta(tt^*) = k_1, \eta(t1t^*) = k_2\right\}$, яку можна обчислити використовуючи співвідношення (3);
- P_4 — імовірність $P\left\{\eta(tt^*) = k_1, \eta(t1t^*) = k_2,\ \eta(t0t^*) = k_3\right\}$, яку можна обчислити використовуючи співвідношення (4).

Таблиця 5

Порівняльна таблиця ймовірностей з використанням тестового пакету NIST

Test	Input Size Recommendation, *n* more than	length	Sequences	$\eta(t\,t^*)$	$\eta(t\,1t^*)$	$\eta(t\,0t^*)$	$\eta(t\,1t^*)+\eta(t\,0t^*)$	P	P_1	P_2	P_3	$\underline{P}_4$
Frequency (Monobit) Test	100	10	1011010101	4	1	0	1	0,53	0,0234	0,0205	0,0273	0,0117
Frequency Test within a Block	100	10	0110011010	3	2	1	3	0,8	0,0977	0,1641	0,0821	0,0439
Runs test	100	10	1001101011	3	2	1	3	0,15	0,0977	0,1641	0,0821	0,0439
Binary Matrix Rank Test	38000	N=20 M = Q = 3	0101100 1001010 101101	8	2	2	4	0,74	0,0017	0,0021	0,0021	0,0008
Discrete Fourier Transform (Spectral) Test	1000	N=10	0001010011	3	1	2	3	0,11	0,0977	0,0821	0,1641	0,0439
Non-overlapping Template Matching Test	200	N=20, 2 blocks of length 10	1010010 0101110 010110	6	2	3	5	0,34	0,0635	0,0573	0,0716	0,0241
Maurer's "Universal Statistical" Test	380000	20	0101101 0011101 010111	7	3	1	4	0,77	0,0201	0,0134	0,0121	0,0049
Serial test	100	10	0011011101	3	2	1	3	0,91	0,0977	0,1641	0,0821	0,0439
Approximate Entropy test	100	10	0100110101	4	1	1	2	0,26	0,0273	0,0273	0,0273	0,0156
Cumulative Sums (Cusum) Test	100	N=10	1011010111	3	2	0	2	0,41	0,1465	0,0684	0,0821	0,0293
Random Excursions Test	10^6	N=10	0110110101	4	2	0	2	0,5	0,0273	0,0205	0,0059	0,0059
Random Excursions Variant Test	10^6	N=10	0110110101	4	2	0	2	0,7	0,0273	0,0205	0,0059	0,0059

Як можна побачити з табл. 5 використання двовимірних або тривимірних статистик дає більший точний результат для подальшого аналізу двійкової послідовності невеликої довжини (до 100 біт).

Також зауважимо, що відповідно до обмежень на довжину послідовності, які наведено в праці [12], рекомендована довжина послідовності *n* повинна бути, в більшості випадків, більша, ніж 10^6 біт.

Висновки

Аналіз ефективності генераторів псевдовипадкових послідовностей є нагальною проблемою кібербезпеки за умов використання більш досконалих методів шифрування та захисту інформації. Наявні способи показують низьку гнучкість та універсальність у засобах знаходження прихованих шаблонів у даних. Для вирішення цієї проблеми запропоновано використовувати алгоритми на основі багатовимірних статистик. Дані алгоритми поєднують усі переваги статистичних методів та є єдиною альтернативою для аналізу послідовностей короткої та середньої довжини.

В роботі встановлені сумісні розподіли числа 2-ланцюжків і числа 3-ланцюжків фіксованого виду випадкової бітової послідовності. Наведені приклади використання цих розподілів. Можливим застосуванням отриманих формул може бути перевірка гіпотези випадковості розташування нулів і одиниць в (0, 1)-послідовний скінченної довжини.

ЛІТЕРАТУРА

1. **Airehrour D**., Gutierrez J., Ray S. K., ‘‘Secure routing for Internet of Things: A survey,’’ J. Netw. Comput. Appl., vol. 66, 2016. Pp. 198–213 (eng).

2. **Mouha N.** “The Design Space of Lightweight Cryptography. NIST Lightweight Cryptography Workshop.” 2015 [Online]. Available: https://hal.inria.fr/hal-01241013 (eng)

3. **Popereshnyak S.** “The technique for testing short sequences as a component of cryptography on the Internet of Things”, CEUR-WS.org/vol/ 2516/paper 11.

4. **Гайдышев И. П.** Программное обеспечение анализа данных AtteStat. Руководство пользователя. Версия 13. 2012. 505 с.

5. **Bakiri M.,** Guyeux C., Couchot J., Marangio L., Galatolo S. “A Hardware and Secure Pseudorandom Generator for Constrained Devices” IEEE Transactions on Industrial Informatics Special Sections: Applied Cryptography, Security, and Trust Computing for Industrial Internet-of-Things(99) · March 2018 DOI: 10.1109/TII.2018.2815985 (eng)

6. **Chugunkov I. V.,** Novikova O. Yu., Perevozchikov V. A., Troitskiy S. S. “The development and researching of lightweight pseudorandom number generators” Conference Paper · February 2016 with 8 Reads DOI: 10.1109/EIConRusNW.2016.7448150 Conference: 2016 IEEE NW Russia Young Researchers in Electrical and Electronic Engineering Conference (EIConRusNW) (eng)

7. **Ramakrishnan K.,** Balasubramanian A., Mishra S., Sridhar R. “Wireless security protocol using a low cost pseudo random number generator” View All Authors Conference Paper · November 2005 with 45 Reads DOI: 10.1109/MILCOM.2005.1605863 · Source: IEEE Xplore Conference: Military Communications Conference, 2005. MILCOM 2005. (eng)

8. **Kösemen C.,** Aydın Öm., Dalkılıç G. “The Pseudorandom Number Generator Generation Method with Genetic Programming for Lightweight Devices” Conference Paper (PDF Available) · September 2018 with 41 Reads DOI: 10.1109/UBMK.2018.8566484 Conference: 2018 3rd International Conference on Computer Science and Engineering (UBMK) (eng)

9. **Busireddygari P.,** Kak S «Pseudorandom tableau sequences», IEEE 51st Asilomar Conference on Signals, Systems, and Computers, 2017, P. 1733–1736.

10. **Popereshnyak S,** Dimitrov GP “The Testing of Pseudorandom Sequences using Multidimensional Statistics” CEUR-WS.org/vol/ 2533/paper14.pdf (eng)

11. **Masol V.,** Popereshnyak S. “Statistical Analysis of Local Sections of Bits Sequences” Journal of Automation and Information Sciences, Volume 51, 2019 Issue 10, pp. 31-45 DOI: 10.1615/JAutomatInfScien.v51.i10.30 (eng)

12. **Special** Publication 800-22. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications. [Online]. Available: http://csrc. nist.gov (eng)

13. **Moody D.** “Post-quantum cryptography: NIST’s plan for the future,” Proceedings of the Seventh International Conference on Post Quantum Cryptography, Japan, 2016. [Online]. Available: https://pqcrypto2016.jp (eng)

Поперешняк С. В.

ТЕСТУВАННЯ ГЕНЕРАТОРА ПСЕВДОВИПАДКОВИХ ЧИСЕЛ ЯК СКЛАДОВА БЕЗПЕКИ ІНТЕРНЕТУ РЕЧЕЙ

Пристрої Інтернету речей (IoT) збирають деякі обсяги даних, які потребують захисту. Рішення захисту даних IoT повинні охоплювати хмарні технології, забезпечувати масштабоване шифрування та управління ключами, а не перешкоджати аналізу даних. Аналіз останніх досліджень і публікацій показує великий інтерес до пошуку різноманітних шляхів розробки полегшених генераторів псевдовипадкових чисел, які знайшли широке застосування в пристроях Інтернету речей або в мобільних пристроях. Ці так звані полегшені пристрої IoT мають обмежену потужність, простір та обчислювальні ресурси. Отже, існує величезна потреба у розробці та тестуванні якості полегшених генераторів псевдовипадкових чисел безпеки, що є важливою складовою кібербезпеки. Наявні підходи до тестування випадкових чи псевдовипадкових послідовностей демонструють низьку гнучкість та універсальність у способі пошуку прихованих шаблонів у даних.

Виявлено, що для послідовностей довжиною до 100 біт недостатньо існуючих статистичних пакетів. Наявні методи демонструють низьку гнучкість та універсальність в засобах пошуку прихованих шаблонів у даних. Розглянуто перспективний напрямок дослідження — статичне тестування послідовностей за допомогою багатовимірної статистики. Для вирішення цієї проблеми пропонується використовувати алгоритми, засновані на багатовимірній статистиці. У роботі наведені формули та сформульована теорема для тестування послідовностей на випадковість, використовуючи дво- чи тривимірну статистику, яка може бути використана для малих та середніх послідовностей. У статті запропонована нова методика тестування псевдовипадкових чисел, розглянуто декілька критеріїв тестування бітової послідовності невеликої довжини, що порівняно з одновимірною статистикою дає більш точний результат. В результаті впровадження цієї методики може бути створена інформаційна система, яка дозволить проаналізувати послідовність псевдовипадкових чисел невеликої довжини та вибрати якісний генератор псевдовипадкових чисел для використання в безпеці Інтернету речей.

Ключові слова: Інтернет Речей; алгоритми; багатовимірної статистики; випадкові послідовності; s-ланцюжки; криптографія; псевдовипадкова послідовність; статистичне тестування.

Popreshnyak S.

TESTING THE PSEUDORANDOM NUMBER GENERATOR AS A COMPONENT OF THE SECURITY OF THE INTERNET OF THINGS

Internet of Things (IoT) devices collect some volumes of data, some of which will require protection based on sensitivity or compliance requirements. IoT data protection solutions must span edge to cloud, provide scalable encryption and key management, and not impede data analysis. An analysis of recent research and publications shows a great deal of interest in finding various ways to develop lightweight pseudorandom number generators that have been widely used on the Internet of Things or mobile devices. These so-called lightweight IoT devices have limited power, space, and computing resources. Therefore, there is a huge need to develop and test the quality of lightweight pseudorandom number generators, which is an important component of cybersecurity. The available approaches to testing random or pseudorandom sequences show low flexibility and versatility in the means of finding hidden patterns in the data. It is revealed that for sequences of length up to 100 bits there are not enough existing statistical packets. The available techniques show low flexibility and versatility in the means of finding hidden patterns in the data. Perspective direction of research — static testing of sequences using multidimensional statistics is considered. To solve this problem, it is suggested to use algorithms based on multidimensional statistics. In the work, formulas are given and theorem for testing sequences for randomness, using two or three-dimensional statistics that can be used for small and medium-sized sequences is formulated. The new technique of PRS testing is proposed in the paper, and several criteria for testing bit sequence of small length are considered, which, in comparison with one-dimensional statistics, gives a more accurate result. As a result of the implementation of this technique, an information system can be created that will allow analyzing the PRS of a small length and choosing a quality PRS for use in the Internet of Things Security.

Keywords: Internet of Things; algorithms; multidimensional statistics; random sequences; s-chains; cryptography; pseudorandom sequence; statistical testing.

Поперешняк С. В.

ТЕСТИРОВАНИЕ ГЕНЕРАТОРА ПСЕВДОСЛУЧАЙНЫХ ЧИСЕЛ КАК СОСТАВЛЯЮЩАЯ БЕЗОПАСНОСТИ ИНТЕРНЕТ ВЕЩЕЙ

Устройства Интернета вещей (IoT) собирают некоторые объемы данных, которые нуждаются в защите. Решение защиты данных IoT должны охватывать облачные технологии, обеспечивать масштабируемое шифрования и управления ключами, а не препятствовать анализу данных. Анализ последних исследований и публикаций показывает большой интерес к поиску различных путей разработки облегченных генераторов псевдослучайных чисел, которые нашли широкое применения в устройствах Интернета вещей или в мобильных устройствах. Эти так называемые облегченные устройства IoT имеют ограниченную мощность, пространство и вычислительные ресурсы. Итак, существует огромная потребность в разработке и тестировании качества облегченных генераторов псевдослучайных чисел безопасности, что является важной составляющей кибербезопасности. Имеющиеся подходы к тестированию случайных или псевдослучайных последовательностей демонстрируют низкую гибкость и универсальность в способе поиска скрытых шаблонов в данных. Выявлено, что для последовательностей длиной до 100 бит недостаточно существующих статистических пакетов. Имеющиеся методы демонстрируют низкую гибкость и универсальность в средствах поиска скрытых шаблонов в данных. Рассмотрено перспективное направление исследования — статическое тестирование последовательностей с помощью многомерной статистики. Для решения этой проблемы предлагается использовать алгоритмы, основанные на многомерной статистике.

В работе приведены формулы и сформулирована теорема для тестирования последовательностей на случайность с использованием двух- или трехмерной статистики, которая может быть использована для малых и средних последовательностей. В статье предлагается новая методика тестирования псевдослучайных чисел, рассмотрены несколько критериев тестирования битовой последовательности небольшой длины, что по сравнению с одномерными статистиками дает более точный результат. В результате внедрения предлагаемой методики может быть создана информационная система, которая позволит проанализировать последовательность случайных чисел небольшой длины и выбрать качественный генератор псевдослучайных чисел для использования в безопасности Интернета Вещей.

Ключевые слова: Интернет Вещей; алгоритмы; многомерные статистики; случайные последовательности; s-цепочки; криптография; псевдослучайная последовательность; статистическое тестирование.

Стаття надійшла до редакції 04.042020 р.
Прийнято до друку 15.06.2020 р.