О.Д. Азаров, Є.В. Яремчук

ДОСЛІДЖЕННЯ ГЕНЕРАТОРІВ ПОСЛІДОВНОСТЕЙ ПСЕВДОВИПАДКОВИХ ЧИСЕЛ НА ОСНОВІ p-ЧИСЕЛ ФІБОНАЧЧІ ТА ЦИКЛІЧНОЇ МАСКИ

Розглянуто генератор псевдовипадкових чисел на основі p-чисел, Фібоначчі та циклічної маски, досліджені дві різні схеми його побудови, а також описані два критерії, які дозволяють визначити криптографічну якість послідовності псевдовипадкових чисел.

Вступ

Генератори послідовностей псевдовипадкових чисел (ПВЧ) відіграють важливу роль в криптографічних алгоритмах захисту інформації. Вони застосовуються для генерування гами, секретних та відкритих ключів і для операції рандомізації [1].

Виділяють клас лінійних конгруентних генераторів, принцип роботи яких в загальному вигляді можна записати у вигляді:

$$X_{i+1} = f(\{X_{p+1}\}, [b_1, b_2, b_3, \ldots]), \quad (1)$$

де $\{X_{p+1}\} = X_{i-(p+1)}, X_{i-p}, \ldots, X_i$ – вектор ініціалізації генератора $(p \geq 0)$;

$[b_1, b_2, b_3, \ldots]$ – набір параметрів генератора.

Найбільш поширеними на сьогодні є генератори

$$X_{i+1} = (aX_i + c) \bmod m,$$

або згідно (1):

$$X_{i+1} = f(\{X_0\}, [b_1, b_2, b_3]);$$
$$\{X_0\} = X_0;$$
$$a = b_1,\ c = b_2,\ m = b_3.$$

Генератори такого типу докладно досліджені в роботі [2], але їх використання для генерації гами викликає певні труднощі: з одного боку, якщо m – максимальне число, яке можна записати у машинному слові, то необхідно кожний новий елемент послідовності розділяти на частини, по одному байту кожна, це зумовлено тим, що операція накладання гами виконується побайтно; а з іншого боку, якщо $m = 256$, то ми не зможемо отримати послідовність з періодом більшим за m. Ці труднощі зникають, якщо використати генератор, в якому використовується залежність не від одного попереднього елемента, а від двох і більше. Найпростішим з таких генераторів є:

$$X_{i+1} = (a_1 X_i + a_2 X_{i-p}) \bmod m. \quad (2)$$

Якщо $a_1 = a_2 = 1$, то цей генератор дозволяє генерувати ПВЧ, виконавши лише три операції: додавання, порівняння та, якщо потрібно, віднімання. Максимальний період такого генератора дорівнює $m^p - 1$ [3–4].

Оскільки генератор (2) побудований на основі узагальнених послідовностей p-чисел Фібоначчі, то назвемо його pF-генератором. Згідно (1) маємо:

$$X_{i+1} = f(\{X_{p+1}\}, b), \quad (3)$$

де $\{X_{p+1}\} = X_{i-(p+1)}, X_{i-p}, \ldots, X_i;\quad m = b.$

Результати дослідження послідовностей ПВЧ, отриманих за допомогою генератора (3), наведено в табл. 1.

Генератор (3) має значний недолік, який полягає у тому, що комбінація з трьох елементів $X_{i-1} < X_{i+1} < X_i$, яка повинна була б зустрітися у шостій частині усіх трійок, не зустрінеться жодного разу. Це підтверджується результатами тесту "Перевірка перестановок" [2]. Крім того, математично це можна довести так: X_{i+1} дорівнює або $X_{i-1} + X_i$, або $X_{i-1} + X_i - m$; якщо $X_{i+1} < X_i$, то ми повинні отримати $X_{i+1} = X_{i-1} + X_i - m$. Отже, $X_{i+1} < X_{i-1}$.

Таблиця 1

Результат тестових перевірок послідовностей псевдовипадкових чисел

№, (m,p)	Тест на рівномірність	Тест "Перевірка серій"	Тест "Перевірка інтервалів"	Тест "Перевірка комбінацій"	Тест "Збирача купонів"	Тест "Перевірка перестановок"	Тест "Перевірка на монотонність"	Тест "Найбільше з t'	Довжина послідовності
1. (3,1)	0.25 так	2.78 так	4.984 так	4.00 так	99.00 так	4.00 так	1.00 так	0.99 так	8
2. (4,5)	1.62 так	10.40 так	13.37 так	**27*** ні	**284** ні	**99.9** ні	**13.08** ні	0.96 так	126
3. (6,4)	1.34 так	26.47 так	56.225 так	**94.991** ні	**322.540** ні	**311.451** ні	**206.528** ні	0.99 так	546
4. (4,1)	2.00 так	10.56 так	36.591 так	4.00 так	99.00 так	4.00 так	1.00 так	0.97 так	6
5. (8,6)	7.78 так	77.78 так	42.899 так	**59.347** ні	**171.00** ні	**241.379** ні	**125.767** ні	0.99 так	508
6. (10,1)	6.67 так	70.0 так	79.175 так	**21.33** ні	98.00 так	**26.2** ні	9.727 так	0.72 так	60
7. (19,3)	16.0 так	105.3 так	723.25 так	**3803.19** ні	**376.222** ні	**3236.98** ні	**12657.6** ні	0.79 так	49659541
8. (7,2)	3.79 так	34.114 так	60.730 так	13.091 так	**198.00** ні	**38.158** ні	5.857 так	0.99 так	57
9. (16,9)	**52.2** ні	**437.48** ні	294.97 так	**1483.24** ні	**189.375** ні	**1572.75** ні	**7752.58** ні	0.99 так	7112
10. (12,4)	**105** ні	174.32 так	76.052 так	**87.101** ні	121.333 так	**242.879** ні	**314.787** ні	0.96 так	546
11. (3,8)	**23.1** ні	**32.9** ні	9.410 так	14.625 так	117.00 так	**98.154** ні	16.25 так	0.99 так	80

* – Напівжирним виділено результати, які є незадовільними.

Описаний недолік та інші, які були виявлені при перевірці послідовностей, згенерованих (3) (див. табл. 1), не дозволяють використовувати його в криптографічних алгоритмах у такому вигляді. Тому запропонуємо генератор, побудований на основі pF-генератора і циклічної маски.

Опис генератора ПВЧ на основі циклічної маски

Циклічна маска – це 8-розрядне двійкове число $S_0 = s_0, s_1, \ldots, s_7$, яке можна циклічно зсувати в заданому напрямку на певну кількість розрядів. Очевидно, що період такої маски дорівнює 8.

Нехай маємо послідовність, отриману за допомогою pF-генератора. Побудуємо послідовність $\{X'_n\}$, елементи якої утворюються шляхом накладання циклічної маски на елемент послідовності $\{X_n\}$. Формально цей процес можна представити у вигляді:

$$X'_{i+1} = f(\{X_{p+1}\}, b, S_j), \tag{4}$$

де $\{X_{p+1}\} = X_{i-(p+1)}, X_{i-p}, \ldots, X_i$;

$m = b$;

$j = i \bmod 8$.

Операція накладання маски S_j на X_i виконується побітно. Після кожної операції накладання маска зсувається на L розрядів в бік молодших розрядів.

Операція накладання маски на елемент послідовності дозволяє послабити лінійну залежність між сусідніми елементами послідовності $\{X'_n\}$. Таким чином, якість останньої значною мірою залежить від вибору S_0.

Розглянемо дві схеми реалізації генератора (4):

1) з фіксованим значенням $L = 1, 2, \ldots, 7$;

2) значення L обирається в залежності від поточного елемента X_i:

$L = X_i mod 8$.

Результати досліджень послідовностей, отриманих за допомогою ***pF*** -генератора на основі циклічної маски (SM-генератор) для першої схеми наведені у табл. 2, а для другої – у табл. 3.

Таблиця 2

Результати тестових перевірок послідовностей ПВЧ для першої схеми генератора (4)

№, (*m*,*p*)	Тест на рівномірність	Тест "Перевірка серій"	Тест "Перевірка інтервалів"	Тест "Перевірка комбінацій"	Тест "Збирача купонів"	Тест "Перевірка перестановок"	Тест "Перевірка на монотонність"	Тест "Найбільше з *t*"	Довжина послідовності
1. (3,1)	3.25 так	9.50 так	11,76 так	4.00 так	99.00 так	10.00 так	1.00 так	0.86 так	8
2. (4,5)	1.62 так	4.84 так	27.19 так	**20.8 ні**	**271.66 ні**	**107.99 ні**	**27.6 ні**	0.96 так	126
3. (6,4)	**14.7 так**	58,36 так	19,25 так	**92.05 ні**	**277.78 ні**	**286.32 ні**	**150.19 ні**	0.88 так	546
4. (4,1)	3.33 так	13.00 так	18,78 так	4.00 так	Так	10.00 так	4.00 так	0.94 так	6
5. (8,6)	5.95 так	61.46 так	42,12 так	**73.80 ні**	**167.66 ні**	**207.93 ні**	**120.05 ні**	0.99 так	508
6. (10,1)	4.00 так	96,66 так	4.42 так	13.83 так	97.99 так	**49.6 ні**	13.07 так	0.72 так	60
7. (19,3)	**ні**	**ні**	723.252 так	**3803.19 ні**	**197.49 ні**	**3076.42 ні**	**5226.80 ні**	1.00 так	14480
8. (7,2)	5.02 так	44.43 так	20.85 так	7.63 так	97.99 так	**33.11 ні**	11.25 так	0.78 так	57
9. (16,9)	16.1 так	**262,11 ні**	300.99 так	**1647.92 ні**	**184.79 ні**	**1861.75 ні**	**2590.89 ні**	1.00 так	7112
10. (12,4)	**32.90 ні**	**217.02 ні**	42.86 так	**120.95 ні**	114.57 так	**187.30 ні**	**133.25 ні**	0.99 так	546
11. (3,8)	2.57 так	**30.2 ні**	20.27 так	**26.5 ні**	179.90 так	**70.46 ні**	10.57 так	0.98 так	80

Порівняємо результати табл. 1 та табл. 2. Загальна кількість незадовільних результатів суттєво не змінилася, проте якісні показники для деяких тестів значно покращились. Наприклад, тест "Перевірка на рівномірність" у першому випадку не пройшли 9, 10 та 12 послідовності, а у випадку SM-генератор – лише 7 та 10 послідовності, причому значення критерію узгодженості зменшилось з 105 до 35, тобто в 3 рази. Аналогічна ситуація спостерігається і для інших тестів.

Можна зробити висновок, що послідовності, отримані за допомогою SM-генератора, мають кращі якісні показники у порівнянні з ***pF*** -генератором (3).

Результати, наведені в табл. 3, показують, що друга схема SM-генератора в деяких випадках дає кращі показники, наприклад, якщо для першої схеми тест "Перевірка серій" не пройшли 4 послідовності, то у випадку другої схеми лише 3, але в цілому результати гірші за наведені в табл. 2.

Розглянемо можливу причину цього явища. Друга схема реалізує перетворення, які можна представити у вигляді:

$$X'_{i+1} = (X_i + X_{i-(p+1)}) \bmod m;$$
$$S = f(X'_{i+1});$$
$$X_{i+1} = X'_{i+1} \oplus S_j.$$

В даному випадку маска S_j є деякою функцією від $p+1$ попередніх елементів послідовності, іншими словами SM-генератор, побудований за другою схемою, реалізує операцію підстановки відносно послідовності $\{X'_n\}$.

Для отримання послідовностей ПВЧ з добрими статистичними характеристиками необхідно виконати такі вимоги:

а) обрати за модуль m просте число, оскільки для таких послідовностей спостерігається більш-менш стабільна закономірність збільшення періоду із збільшенням m;

б) значення S відчутно впливає на якість послідовності, і експериментально встановлено, що задовільні характеристики мають місце при $S = \left| m \pm \frac{m}{2} \right|$, тобто коли S дорівнює найближчому цілому числу, яке віддалене від значення модуля на половину його величини.

Наприклад, при $m = 13$, $p = 3$ маємо:

$S = 11$, $\chi^2 = 97{,}15$;

$S = 7$, $\chi^2 = 51{,}82$;

$S = 5$, $\chi^2 = 71{,}11$;

$S = 17$, $\chi^2 = 200{,}95$;

$S = 19$, $\chi^2 = 125{,}58$.

Таблиця 3

Результати тестових перевірок послідовностей ПВЧ для другої схеми генератора (4)

№, (m,p)	Тест на рівномірність	Тест "Перевірка серій"	Тест "Перевірка інтервалів"	Тест "Перевірка комбінацій"	Тест "Збирача купонів"	Тест "Перевірка перестановок"	Тест "Перевірка на монотонність"	Тест "Найбільше з t'	Довжина послідовності
1. (3,1)	1.00 так	5.00 так	1.94 так	4.00 так	97.99 так	10.00 так	1.00 так	0.87 так	8
2. (4,5)	9.38 так	15.47 так	19.42 так	**28.4 ні**	**334.00 ні**	**87.71 ні**	**55.36 ні**	0.99 так	126
3. (6,4)	**26.39 ні**	51.52 так	86.40 так	**81.96 ні**	**194.27 ні**	**320.94 ні**	**292.45 ні**	0.99 так	546
4. (4,1)	11.33 так	23.67 так	1.86 так	4.00 так	так	10.00 так	4.00 так	0.99 так	6
5. (8,6)	13.83 так	75.07 так	41.90 так	**61.72 ні**	123.45 так	**212.41 ні**	**107.44 ні**	0.74 так	508
6. (10,1)	8.33 так	83.33 так	6.42 так	13.83 так	99.00 так	16.00 так	13.00 так	0.65 так	60
7. (19,3)	**1072.3 ні**	**1543.2 ні**	425.22 так	**3593.04 ні**	130.52 так	**3298.51 ні**	**5239.80 ні**	1.00 так	14480
8. (7,2)	5.26 так	54.75 так	4.93 так	**20.36 ні**	97.99 так	**17.31 ні**	**10.75 ні**	0.92 так	57
9. (16,9)	20.93 так	**374.56 ні**	252.81 так	**1625.24 ні**	**201.35 ні**	**1520.99 ні**	**3477.56 ні**	1.00 так	7112
10. (12,4)	**33.60 ні**	186.43 так	38.49 так	**91.69 ні**	98.33 так	**179.91 ні**	**270.29 ні**	0.99 так	546
11. (3,8)	**19.08 так**	**29.30 ні**	100.02 так	**16.05 ні**	**171.33 ні**	**97.03 ні**	**42.00 ні**	0.98 так	80

Дослідження послідовностей ПВЧ

Сформулюємо два критерії, згідно з якими будемо визначати криптографічну якість послідовності *ПВЧ*:

$$\{X_n\} = X_0,\ X_1,\ldots,\ X_{n-1}. \tag{5}$$

Критерій непредикативності вліво K_1 характеризує складність знаходження $i+1$ елемента послідовності ПВЧ з періодом T, якщо відомо k $(k \le i)$ елементів.

Припущення 1. Якщо $K_1 \to T$, то послідовність (5) буде достатньо сильною.

Обгрунтуємо припущення 1.

Маючи k елементів послідовності, криптоаналітик може встановити залежність між її елементами і побудувати ефективний алгоритм для відновлення усієї послідовності. Іншими словами, якщо послідовність непредикативна вліво, то скільки б елементів послідовності не знав криптоаналітик, цього було б замало для побудови ефективного алгоритму відтворення усієї послідовності. Якщо $k = T$, то криптоаналітик уже має усі елементи послідовності.

Критерій непредикативності вліво K_1 має такий самий зміст, що й лінійна складність для двійково-адитивних генераторів [1].

Нехай

$$B = b_0 b_1 \ldots b_{L-1} - \tag{6}$$

двійкова послідовність ПВЧ, якій відповідає характеристичний многочлен:

$$h(x) = x^{p+1} + x^p + 1. \tag{7}$$

Якщо цей многочлен є примітивним, то період послідовності (6) буде максимальним і дорівнюватиме $2^{p+1} - 1$. З визначення лінійної складності випливає, що для будь-якої послідовності виду (6) можна побудувати найкоротший двійковий регістр зсуву [1].

Оскільки послідовності, отриманій за допомогою pF-генератора, теж відповідає характеристичний многочлен (7), то за аналогією можна сформулювати таке припущення.

Припущення 2. Для будь-якої послідовності виду (5) можна побудувати m-ний найкоротший регістр зсуву.

Існує декілька алгоритмів визначення мінімальної довжини двійкового регістра зсуву [5]. Запропонуємо алгоритм для знаходження мінімальної довжини m-го регістра зсуву.

Алгоритм 1. Знаходження мінімальної довжини регістра зсуву.

Крок 1. Встановлюємо $I \leftarrow 0$, $J \leftarrow 1$, $M_1 \leftarrow 0$, $M_2 \leftarrow 0$.

Крок 2. $T \leftarrow Sequence[I] + Sequence[J]$.

Крок 3. Якщо $T > Sequence[J+2]$ і $M_1 = 0$, то $M_1 \leftarrow T - Sequence[J+2]$ і перейти на **Крок 6**, інакше $M_2 \leftarrow T - Sequence[J+2]$.

Крок 4. Якщо $M_1 = M_2$, то перейти на **Крок 8**.

Крок 5. $P \leftarrow P+1$, $I \leftarrow 0$, $J \leftarrow I+P$, $M_1 \leftarrow 0$ і перейти на **Крок 7**.

Крок 6. $I \leftarrow I+1$, $J \leftarrow J+1$.

Крок 7. Якщо $J < N$, то перейти на **Крок 2**.

Крок 8. J – довжина регістра зсуву, M_1 – модуль послідовності.

Кінець алгоритму.

Обчислювальна складність Алгоритму А1 становить $O(N^2)$ для послідовності з N елементів. Для послідовностей, згенерованих pF-генератором, $K_1 = p+1$.

Розглянемо інший критерій. Нехай маємо послідовність ПВЧ (5) і $n-1$ пару сусідніх елементів.

Підрахуємо статистики k^+ – кількість пар, для яких виконується рівність $X_i > X_{i+1}$, та k^- – кількість пар, для яких $X_i < X_{i+1}$. Випадки, коли $X_i = X_{i+1}$ до уваги не беруться, оскільки їх кількість буде незначна або рівна нулю і послідовність повинна відповідати умові рівномірності.

Критерій нестабільності K_2 характеризує послідовність з точки зору зміни знака між сусідніми елементами. Числове значення критерію нестабільності будемо знаходити за формулою:

$$K_2 = \frac{k^+ - k^-}{n-1}, \quad 0 \le K_2 \le 1. \tag{8}$$

Якщо $k^- > k^+$ або $k^- < k^+$, то $0 < K_2 < 1$. Якщо $k^- = k^+ = \frac{n-1}{2}$, то $K_2 = 0$.

З усіх $n!$ можливих послідовностей лише дві будуть мати $K_2 = 1$:

$$X_0 < \ldots < X_{\frac{n}{2}-1} < X_{\frac{n}{2}} < X_{\frac{n}{2}+1} < \ldots < X_n;$$

$$X_0 > \ldots > X_{\frac{n}{2}-1} > X_{\frac{n}{2}} > X_{\frac{n}{2}+1} > \ldots > X_n.$$

Також існує дві послідовності, для яких $K_2 = 0$, але які не є "випадковими":

$$X_0 < \ldots < X_{\frac{n}{2}-1} < X_{\frac{n}{2}} > X_{\frac{n}{2}+1} > \ldots > X_n;$$

$$X_0 > \ldots > X_{\frac{n}{2}-1} > X_{\frac{n}{2}} < X_{\frac{n}{2}+1} < \ldots < X_n.$$

Будемо вважати, що послідовність задовольняє критерій нестабільності, якщо значення K_2 належить проміжку $[-0{,}05;\ 0{,}05]$.

Висновки

1. Запропоновано генератор ПВЧ, який має кращі статистичні характеристики, ніж звичайний генератор на основі p-чисел Фібоначчі, має високу швидкодію і досить просто реалізується як програмно, так і апаратно.

2. Сформульовано два критерії (непредикативності вліво K_1 та нестабільності K_2), які в поєднанні з статистичними тестами дозволяють краще дослідити якість послідовності ПВЧ.

ЛІТЕРАТУРА:

1. *Месси Дж.М.* Введение в современную криптологию // ТИИЭР, 1988. – № 5. – С. 24–42.
2. *Кнут Д.* Искусство программирования для ЭВМ. Т. 2. – М.: Мир, 1977. – 724 с.
3. *Макуильямс Ф.Дж., Слоан Н.Дж.* Псевдослучайные последовательности и таблицы // ТИИЭР, 1976. – № 12. – С. 80–95.
4. *Берлекэмп Э.* Алгебраическая теория кодирования. – М.: Мир, 1971. – 477 с.
5. *Menezes A., van Oorschot P., Vanstone S.* Handbook of Applied Cryptography, CRC Press 1996.

АЗАРОВ Олексій Дмитрович – доктор технічних наук, завідувач кафедри обчислювальної техніки, декан факультету інформаційних технологій та комп'ютерної інженерії Вінницького державного технічного університету.

Наукові інтереси:

– цифро-аналогові та аналогово-цифрові перетворювачі інформації.

ЯРЕМЧУК Євген Вікторович – магістр комп'ютерних наук, аспірант кафедри обчислювальної техніки Вінницького технічного університету.

Наукові інтереси:

– криптографічний захист інформації.