Використання Пакетів Тестування для Статистичного Дослідження Генераторів Випадкових Чисел

Марія Семаньків, Іляш Юрій
кафедра інформатики
Прикарпатський національний університет
Івано-Франківськ, Україна
dlyamarii@gmail.com, yurchukil@cym.org

Даріуш Саля
кафедра управління
Університет AGH
Краків, Польща
dsala@zarz.agh.edu.pl

Use of Test Packages for Statistical Research of Random Numbers Generators

Maria Semankiv, Yurii Iliash
dept. of Computer Science
Precarpathian National University
Ivano-Frankovsk, Ukraine
dlyamarii@gmail.com, yurchukil@cym.org

Dariusz Sala
dept. of Enterprise Management
AGH University
Krakow, Poland
dsala@zarz.agh.edu.pl

***Анотація*—подано результати дослідження послідовності псевдовипадкових чисел, що сформовані методом перестановки ваги розрядів двійкового коду.**

***Abstract*—the results of the investigation of the sequence of pseudorandom numbers, generated by the method of permutation of the bits digit bits code, are presented.**

Ключові слова—генератор випадкових чисел, пакет статстичних досліджень.

Keywords—random number generator, static studies pack.

I. Вступ

В залежності від сфери використання (наприклад, імітаційне моделювання, системи захисту інформації, чисельні методи, комп'ютерна графіка, криптографія) до генераторів випадкових та псевдовипадкових чисел ставлять ряд вимог, які вони повинні задовольняти:

- простота апаратної або програмної реалізації;
- низька собівартість;
- максимальна швидкодія;
- максимальна наближеність послідовності, яка отримана на виході генератора, до вибраного закону розподілу (наприклад рівномірного);
- можливість роботи генератора у широкому діапазоні частот.

Якщо такий генератор буде використовуватися в системах захисту інформації (а такі генератори часто є важливою ланкою в забезпеченні інформаційної безпеки, зокрема в криптографії), то до нього ставиться ще ряд додаткових вимог, що будуть жорсткішими порівняно з вимогами до генераторів, що використовуються в інших галузях [1,2]. Виконання таких вимог дає змогу передусім виявити генератори, що наперед задовольняють вимоги конкретної прикладної задачі.

Тому виникає питання дослідження статистичних характеристик та визначення типу розподілу послідовностей випадкових і псевдовипадкових чисел для оптимального їх використання у зазначених вище сферах застосування.

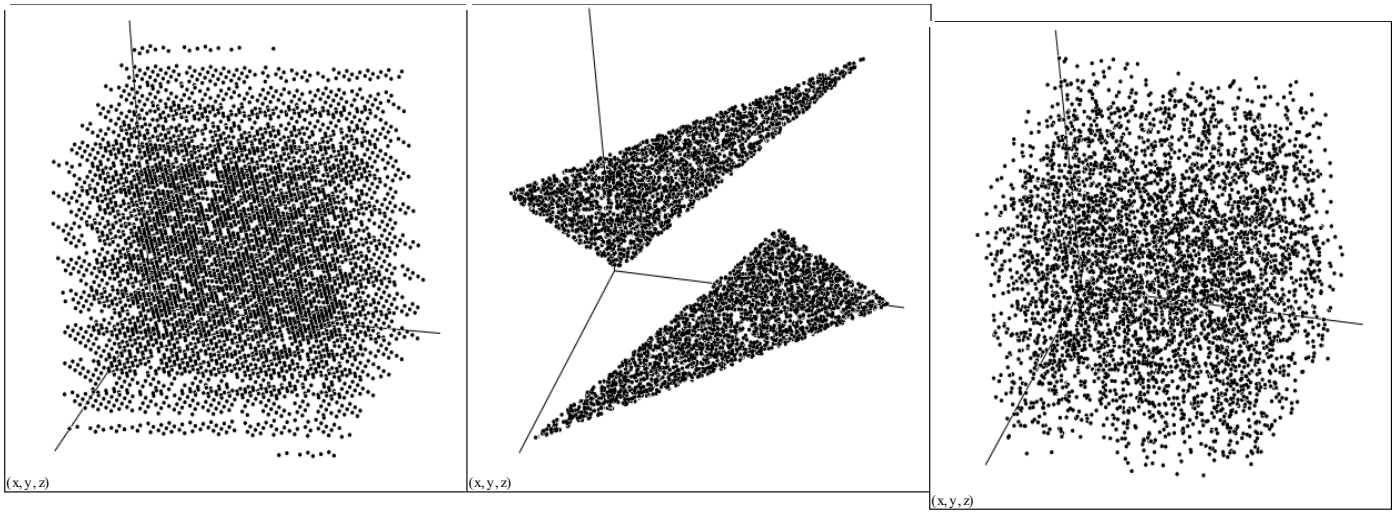
II. Графічні методи досліджень

Всі існуючі методи тестування послідовностей випадкових та псевдовипадкових чисел можна поділити на: евристичні, графічні та статистичні.

Рис. 1. Класифікація методів тестування послідовностей випадкових та псевдовипадкових чисел

Вибір необхідного тесту проводиться в залежності від потреб подальшого прикладного застосування послідовності випадкових чисел. Для застосування генератора випадкових чисел для методу Монте-Карло першочерговим стає питання рівномірності розподілу послідовності чисел, простота технічної та програмної реалізації та собівартість самого генератора.

Проведено дослідження рівномірності на площині запропоновано в статті [3] методу генерування псевдовипадкових чисел на основі перестановки ваги розрядів двійкового коду. Високі показники рівномірності дозволили стверджувати про доцільність використання даного методу формування псевдовипадкових чисел з метою забезпечення рівномірного «заповнення» значень на обраній площині. На рис. 2 подано результати дослідження рівномірності розподілу в просторі псевдовипадкових чисел даного методу генерування, та конгруентного генератора і генератора Фібоначчі для порівняння отриманих в графічному вигляді результатів дослідження.



а) конгруентний метод б) метод Фібоначчі в) метод перестановки ваги розрядів

Рис. 2. Графічне подання розподілу в просторі послідовностей псевдовипадкових чисел

Зауважимо, що система трьох випадкових чисел називається рівномірно розподіленою в просторі, якщо щільність ймовірності $f(x,y,z) = const$ в середині деякої області і дорівнює нулю за її межами:

$$f(x,y,z) = \begin{cases} \dfrac{1}{S_{пр}} = \dfrac{1}{(b-a)(d-c)(e-g)}, & всередині\ обл. \\ 0, & поза\ нею \end{cases}$$

простір виду $a \le x \le b, c \le y \le d, e \le z \le g$.

Подані на рис. 2 результати досліджень показують, що для конгруентного методу генерування послідовності псевдовипадкових чисел та методу Фібоначчі зберігається закономірність формування елементів, які заповнюють площини і фігури в просторі. Для методу перестановки ваги розрядів графічне представлення характеризується рівномірним заповненням простору.

Графічні методи статистичного аналізу послідовностей є досить ефективними для виявленні істотних недоліків послідовності псевдовипадкових чисел і для відображення результатів дослідження на рівномірність розподілу. З їх допомогою можна швидко відкинути генератори, чиї результати не задовольняють критеріям рівномірності розподілу [2]. Однак, графічні тести сприймаються людиною, що не гарантує їх однозначність. Для більш точних результатів використовуються статистичні тести,

які видають чисельну характеристику послідовності і дозволяють однозначно сказати, пройдений тест чи ні.

III. Статичтичні методи долідження

Статистичні тести дають можливість виконати чисельну оцінку якості генераторів випадкових чисел. Тести зазвичай об'єднуються в пакети тестування (серед них можна виділити тести DIEHARD, тести NIST і ін.). Однією з статистичних оцінок є оцінка помилки відтворення закону розподілу дискретної випадкової величини.

ТАБЛИЦЯ I. Пакети статистичних досліджень

Тести	Переваги	Недоліки	Кіль-кість тестів
DIEHARD	найбільш строгі з відомих тестів, найпростіші для автоматизації тестування за допомогою інтерфейсу командного рядка	нема детального опису тестів і методики трактування їх результатів, крім того більшість тестів є евристичними	13
NIST STS	велика збірка тестів	незручний для використання інтерфейс	16
RaBiGeTe	здійснюють підтримку багатопотокового тестування	Необхідність знань з статистики	24

Проведено дослідження послідовностей псевдовипадкових чисел, утворених методом перестановки ваги розрядів за допомогою пакету тестів DIEHARD.

Тести Diehard формують на виході числа p-значення, які рівномірно розподілені в інтервалі [0;1], якщо вхідний потік чисел дійсно випадковий. Після сортування отриманих p-значень, побудовано графік, що показує відхилення від X=Y діагональної лінії отриманих p-значень (рис. 3).

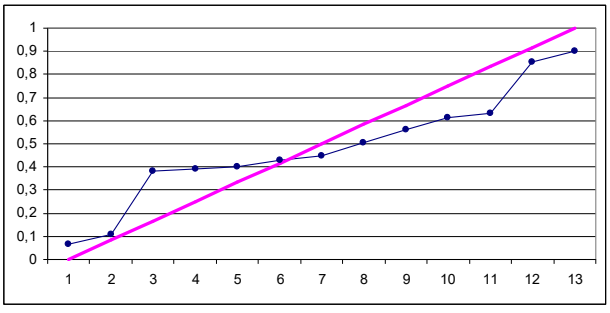


Рис. 3. Відхилення від X=Y діагональної лінії отриманих p-значень

Результати проведених досліджень на визначення типу розподілу послідовності псевдовипадкових чисел, генерованих зазначеним методом, показали високу якість рівномірності розподілу та ефективність використання даного генератора в складі аналого-цифрового перетворювача Монте-Карло [4,5].

Слід відзначити пакет статистичних тестів NIST STS, що є зручним та гнучким інструментом дослідження генератора випадкових (псевдовипадкових) чисел, що застосовуються в криптографічних додатках.

На відміну від пакета DIEHARD пакет NIST STS має більшу гнучкість, розширюваністю і ефективністю (з точки зору витрачається часу на здійснення тестування генератора). Крім того, пакет NIST STS має велику криптографічний спрямованість, яка досягається шляхом введення в пакет таких тестів як перевірка лінійної складності і універсального тесту Маурера. Результати застосування даного пакету тестів для дослідження характеристик методу перестановки ваги розрядів подано в статті [6].

RaBiGeTe враховуючи недоліки попередніх пакетів статистичних досліджень включила в себе основні тести вказаних пакетів. Програма включає 24 тести, в склад яких потрапили вибрані тести пакетів NIST DFT, Diehard, тести Д.Кнута, Маурера та додаткові статистичні тести. Користувач надає двійковий файл з генерованими числами та змінює параметри тестування в залежності від потреб. RaBiGeTe має зручний інтерфейс та надає можливості налаштування параметрів тестування, отримані результати подаються у чисельному та графічному вигляді. Використання даного пакету тестів підтвердило, що запропонований метод дозволяє отримати послідовності псевдовипадкових чисел з досить високою статистичною якістю розподілу p-значень за результатами тестування RaBiGeTe і може бути використаний для побудови генератора псевдовипадкових чисел з рівномірним розподілом[7].

IV. Кореляційний аналіз

Кореляційний аналіз використовується для кількісної оцінки взаємозв'язку двох наборів даних. Коефіцієнт кореляції використовується для визначення наявності взаємозв'язку між елементами. Коефіцієнт кореляції являється індексом в інтервалі від –1 до 1 включно та відображає ступінь лінійної залежності між двома множинами даних (табл.2).

ТАБЛИЦЯ II. Розподіл величини кореляції

Величина кореляції	0,1-0,3	0,3-0,5	0,5-0,7	0,7-0,9	0,9-1,0
Характеристика сили зв'язку	Слабка	Помірна	Помітна	Висока	Дуже висока

Проведено дослідження на визначення сили взаємозв'язку між псевдовипадковими числами, що сформовані методом перестановки ваги розрядів двійкового коду, результати подано в таблиці 3.

ТАБЛИЦЯ III. . КОЕФІЦІЇНТИ КОРЕЛЯЦІЇ ДЛЯ МЕТОДУ ПЕРСТАНОВКИ ВАГИ РОЗРЯДІВ

Розряд генератора	**Коефіцієнт кореляції між утворюючими та генерованими числами**	**Коефіцієнт кореляції між частинами генерованої послідовності**
8	0,010683	0,76479
10	0,007723	0,76531
12	0,002675	0,71485
14	0,007723	0,998712
16	0,004605	0,998712

Високий коефіцієнт кореляції (0,7-0,9) вказує на велику ймовірність взаємозв'язку між елементами послідовності псевдовипадкових чисел даного методу генерування, що обмежує їх практичне застосування в задачах, що вимагають «незалежності» сформованих псевдовипадкових чисел.

Це не знижує практичну цінність даного методу генерування псевдовипадкових чисел в області математичного моделювання, завдяки його високій якості рівномірності розподілу, що є вирішальним критерієм при відбору генератора для методу Монте-Карло.

Висновки

На сьогоднішній день є велике різноманіття пакетів для дослідження статистичних характеристик генераторів випадкових чисел. Незважаючи на їх велику кількість користувач стикається з проблемою не тільки незручного інтерфейсу, але і неоднозначності трактування результатів, внаслідок відсутності детальних описів тестів, що входять до обраного пакету. Незважаючи на вказаний недолік дані пакети статистичних тестів дозволяють здійснити швидке автоматизоване дослідження характеристик послідовностей випадкових та псевдовипадкових чисел, а деякі з пакетів дозволяють здійснити багатопотокове тестування.

Подані результати дослідження методу генерування псевдовипадкових чисел внаслідок перестановки ваги розрядів двійкового коду за допомогою графічного методу та пакетів статистичних тестів визначили високу ефективність використання даного методу генерування для формування послідовностей, що задовольняють вимоги рівномірності розподілу випадкових чисел.

Література References

[1] М.А. Иванов, И.В. Чугунков, Теория, применение и оценка качества генераторов, М.: КУДИЦ-ОБРАЗ, 2003.

[2] Donald Knuth The Art of Computer Programming, Seminumerical Algorithms / K. Donald // Volume 2, 3rd edition, Addison Wesley, Reading, Massachusetts, 1998.

[3] М. В. Лаврів, Аналіз ефективності застосування методів генерування сигналів з псевдовипадковим розподілом у системах статистичних досліджень / М.В. Лаврів, Л.Б. Петришин // Наукові вісті інституту менеджменту та економіки "Галицька академія". – Івано-Франківськ, 2007. – №2 (12). – С.61-66.

[4] М.В. Лаврів, Генератори рівномірно розподілених псевдовипадкових величин / М.В. Лаврів, Л.Б. Петришин // Вісник Прикарпатського національного університету. Фізика. – 2007. Вип. 3. – С. 112-118.

[5] М.В. Лаврів, Методи і засоби генерування псевдовипадкових сигналів із рівномірним розподілом та аналіз результатів дослідження їх статистичних характеристик / М.В. Лаврів, Л.Б. Петришин // Інформаційні технології та комп'ютерна інженерія. – 2009. – №2 (15). – С. 56-62.

[6] М.В. Семаньків, Генератори випадкових чисел в складі аналого-цифрового перетворювача Монте-Карло / М.В. Семаньків // Інформаційні технології: наука, техніка, технологія, освіта, здоров'я: Тези доповідей XXIV міжнародної науково-практичної конференції, Ч.IV (18-20 травня 2016 р., М.Харків, НТУ «ХПІ»). – С. 173

[7] М.В. Семаньків, Оцінка статистичних характеристик систем випадкових чисел / М.В. Семаньків // Вісник національного технічного університету «Харківський політехнічний інститут». Збірник наукових праць. Серія: Інформатика та моделювання. – Харків: НТУ «ХПІ». – 2016. – С.109-117.