

ДОСЛІДЖЕННЯ РІВНОМІРНОСТІ РОЗПОДІЛЕННЯ ПСЕВДОВИПАДКОВИХ ЧИСЕЛ

Мироненко А. А.

*студент 1-го курсу факультету кібербезпеки та інформаційних технологій
Національного університету «Одеська юридична академія»*

Сучасний розвиток обчислювальної техніки суттєво підвищує значимість використання методів комп'ютерного моделювання у дослідженні процесів управління інфраструктурою підприємств, а також складних технологічних процесів.

Комп'ютерне моделювання ґрунтується на імітації стохастичних процесів з використанням арифметичних датчиків псевдовипадкових чисел (ПВЧ) і відбувається у два етапи: спочатку створюється послідовність рівномірно розподілених на інтервалі $[0, 1]$ випадкових величин, а потім із них формується послідовність, яка відповідає заданому закону розподілення ймовірностей.

Послідовності, сформовані на основі обчислювального алгоритму, циклічно повторюються і, для того щоб вони мали властивості випадкових, період їх повторення має бути якомога більшим. На даний час ця проблема є вирішеною, але, що стосується рівномірності розподілення ймовірностей, то, не зважаючи на заяви розробників, переважна більшість вбудованих у популярні програмні середовища генераторів не відповідає сучасним вимогам [1].

Щоб можна було прийняти рішення по придатності того або іншого генератора псевдовипадкових чисел для моделювання числового потоку з відомим розподіленням ймовірностей, необхідно взяти за основу алгоритм формування рівномірно розподілених чисел, якість якого не нижче, ніж у основного потоку. З урахуванням цього ціллю роботи є розроблення алгоритму для оцінки якості програмних генераторів псевдовипадкових чисел.

Оцінка якості генераторів ПВЧ в програмних середовищах

Сучасні середовища програмування використовують для потреб моделювання лінійний конгруентний генератор (Linear congruent generator LCG) [2], і алгоритм, відомий під назвою Вихор Мерсенна (Mersenne twister, MT), який забезпечує надзвичайно великий період повторення ($2^{19937} - 1$ біт) [3]. Вони обидва входять до складу бібліотек майже всіх відомих спеціалізованих програмних середовищ, призначених для вирішення дослідницьких та інженерних задач.

Найчастіше для перевірки таких датчиків використовують критерій Карла Пірсона χ^2 [3]. Є й інші, більш складні, критерії, але вважається, що якщо перевірка за критерієм χ^2 не дає позитивний результат, то й інші критерії теж не пройдуть.

Суть цього критерію полягає в тому, що спочатку формується досить велика кількість псевдовипадкових чисел розміру N , рівномірно розподілених на інтервалі $[0, 1]$. Цей інтервал розділяється на однакових підінтервалів. У разі якісного

рівномірного розподілення, очікується, що у кожний інтервал попадає $N_i = N/n$ чисел, як це показано на рисунку 1, а частота попадання числа у той або інший інтервал дорівнює $p_i = 1/n$.

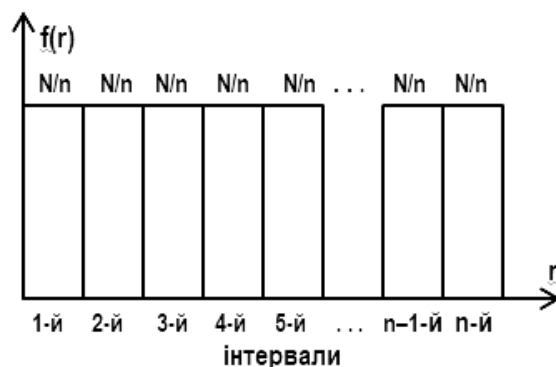


Рис. 1. Частотна діаграма ідеального генератора ПВЧ

Оскільки в реальній ситуації розподілення чисел на виході генератора відрізняється від рівномірного частоти будуть відрізнятися одна від одної і розподілення буде мати вигляд, наведений на рисунку 2.

Як показано у [3], показник рівномірності $\chi^2_{\text{екс}}$, обчислений на основі сформованої генератором ПВЧ чисел з використанням критерію Пірсона визначається як

$$\chi^2_{\text{екс}} = \frac{(N_i - p_i N)^2}{p_i N}. \quad (1)$$

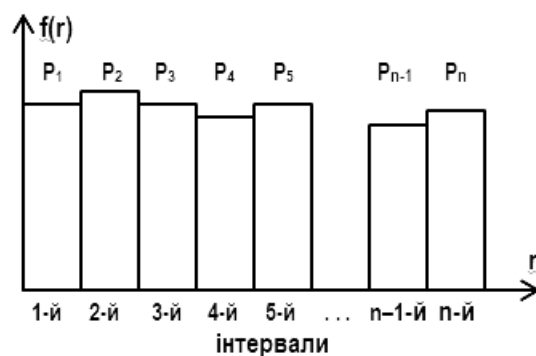


Рис. 2. Частотна діаграма реального генератора ПВЧ

Критичне значення показника $\chi^2_{\text{кр}}$ визначається із спеціальних таблиць або обчислюється з урахуванням кількості підінтервалів на яку розбитий інтервал $[0, 1]$ і величина допустимого відхилення від теоретичного значення N/n у процентах.

Для вирішення поставленої задачі було створення програмний продукт, який створює потік псевдовипадкових чисел заданого об'єму, записує його у текстовий файл, а потім переписує його вміст у тимчасовий динамічний масив для подальшого дослідження.

У якості вхідних даних експерименту обрані кількість підінтервалів $n = 16$, довжина послідовності ПВЧ вище тисячі символів і ступінь відхилення від рівномірності не більше 15%.

Кожне випадкове число, записане у масив, перевіряється окремо на предмет номеру інтервалу у який воно попадає. Одночасно ведеться облік кількості чисел у кожному інтервалі. Отримання цих даних дає можливість обчислення величини $\chi^2_{\text{екс}}$, скориставшись виразом (1).

В окремий файл програмного продукту винесена процедура обчислення величини критичного значення $\chi^2_{\text{кр}}$. Методику її обчислення можна знайти на сайті NIST.

Програмний продукт має графічний інтерфейс, наведений на рисунку 3.

Він дозволяє вводити розмір послідовності N , ступінь відхилення від рівномірності α та ім'я файлу, у якому зберігається сформована послідовності.

На форму проекту у вигляді таблиць виведено дані про вміст файлу з псевдовипадковими числами, кількість випадкових чисел у кожному із 16-ти інтервалів, а також діаграма розподілення ПВЧ в інтервалах. Окремо виводяться дані про змінні $\chi^2_{\text{екс}}$ і $\chi^2_{\text{кр}}$, а також про очікувану кількість чисел N/n в кожному інтервалі.

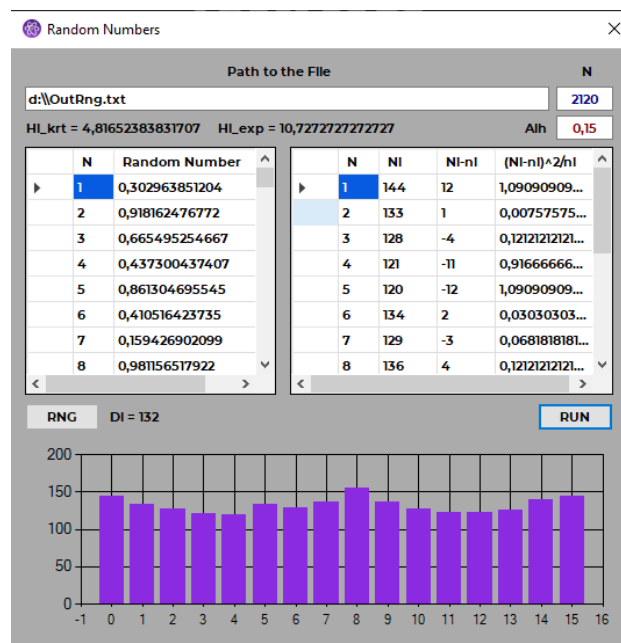


Рис. 3. Зовнішній вигляд графічного інтерфейсу програми

Розроблений програмний продукт дозволяє багаторазово формувати послідовність випадкових чисел, оцінювати кожну з них і порівнювати отриманий показник рівномірності з критичним значенням і, таким чином, накопичувати необхідний статистичний матеріал, на основі якого робиться висновок про придатність даного генератора ПВЧ.

Висновки

Проведені випробування лінійного конгруентного генератора, а також

генератора, побудованого на основі використання чисел Мерсенна, показали, що, не зважаючи на великий період повторення, обидва генератори не забезпечують рівень рівномірності псевдовипадкових чисел на виході алгоритму. При чому, на результати випробувань майже не впливають ні розмір величини сформованої послідовності N , ні обраний показник точності α .

Наведені недоліки не дозволяють використовувати ці генератори для точного моделювання випадкових процесів, які описуються іншими видами законів розподілення випадкових величин. Одною з причин такого явища є те, що арифметичні генератори формують натуральні числа, а вже потім їх перетворюють у дійсні числа шляхом приведення до максимального значення. Саме використання операції ділення дає нерівномірність розподілення вихідних чисел.

Список використаних джерел:

1. Averill M. Law. Simulation modeling and analysis, 5th. ed., McGraw-Hill Education, 2 Penn Plaza, New York, 2015. URL: <https://industri.fatek.unpatti.ac.id/wp-content/uploads/2019/03/108-Simulation-Modeling-and-Analysis-Averill-M.-Law-Edisi-5-2014.pdf>
2. Bruce Schneier, Applied Cryptography, Second Edition: Protocols, Algorithm, and Source Code in C, 20. ed., Boston, Mass, USA : Addison-Wesley, Longman Publishing, Addison-Wesley, Reading, Mass, 1998. doi:10.1002/9781119183471 URL: <https://lost-contact.mit.edu/afs/adrake.org/usr/rkh/Books/books/Schneier%20-%20Applied%20Cryptography%20ed%20-%20Wiley.pdf>.
3. D. E. Knuth, The Art of Computer Programming, Volume 2: Seminumerical Algorithms, 3rd. ed., Boston, Mass, USA : Addison-Wesley, Longman Publishing, Addison-Wesley, Reading, Mass, 1998. URL: https://doc.lagout.org/science/0_Computer%20Science/2_Algorithms/The%20Art%20of%20Computer%20Programming%20%28vol.%202_%20Seminumerical%20Algorithms%29%20%283rd%20ed.%29%20%5BKnuth%201997-11-14%5D.pdf

Науковий керівник: доцент Щербина Ю. В.

ВІЗУАЛІЗАЦІЯ АЛГОРИТМУ ДЕЙКСТРИ

Трофименко О. Г.

*кандидат технічних наук, доцент, доцент кафедри інформаційних технологій
Національного університету «Одеська юридична академія»*

Новіков Д. І.

*студент 2-го курсу факультету кібербезпеки та інформаційних технологій
Національного університету «Одеська юридична академія»*

Сьогодні для ІТ-освіти, і як наслідок для ІТ-галузі, характерні проблеми якості освіти і дефіцит спеціалістів водночас. Здобувачі ІТ-освіти мають багато вчитися і під час навчання, і під час професійної діяльності, щоб встигати за змінами