DOI: 10.34185/1991-7848.itmm.2024.01.062

МЕТОДИ ОЦІНКИ ВІДПОВІДНОСТІ ПСЕВДОВИПАДКОВОЇ ПОСЛІДОВНОСТІ ДО ВИПАДКОВОГО РОЗПОДІЛУ В ЗАДАЧАХ АВТОМАТИЗАЦІЇ, КІБЕРБЕЗПЕЦІ, КОМП'ЮТЕРНО-ІНТЕГРОВАНИХ ТЕХНОЛОГІЯХ ТА НЕРУЙНІВНОМУ КОНТРОЛІ

Клименко О.Д., Малайчук В.П.

Дніпровський національний університет імені Олеся Гончара, м. Дніпро, Україна

Анотація. *Створення методів визначення міри близкості заданої псевдовипадкової послідовності до випадкового розподілу є дуже актуальним завданням у багатьох галузях, де використовуються псевдовипадкові числа. Наприклад, у криптографії важливо, щоб псевдовипадкові послідовності були непередбачувані та мали властивості випадкових чисел. Методи визначення близкості до випадкового розподілу допомагають в оцінці якості таких послідовностей та виявленні в них недоліків. У наукових дослідженнях та інженерних проектах використовують псевдовипадкові числа для моделювання складних систем і симуляції випадкових явищ. Важливо мати засоби для визначення, наскільки добре такі числа відтворюють випадковий процес. У статистиці важливо, щоб псевдовипадкові послідовності відповідали вимогам різних статистичних властивостей. Методи визначення близкості до випадкового розподілу допомагають в оцінці властивостей цих послідовностей. У програмній інженерії та інших сферах важливо перевіряти якість генераторів псевдовипадкових чисел, щоб уникнути вразливостей та недоліків у програмах.*

Таким чином, розробка та вдосконалення методів визначення міри близкості псевдовипадкових послідовностей до випадкового розподілу є актуальним завданням з важливими практичними застосуваннями.

Ключові слова: *автоматизація, кібербезпека, комп'ютерно-інтегровані технології, неруйнівний контроль, розподіл, тест, псевдовипадкові числа, випадковий процес.*

Існує кілька методів для визначення міри близькості заданої псевдовипадкової послідовності до випадкового розподілу. Деякі з найпоширеніших методів включають: статистичні тести, тести на серії, спектральний тест, тести на складність, тести на кореляцію. Розглянемо кожен з цих методів більш детально.

Статистичні тести використовуються для перевірки властивостей псевдовипадкових чисел, таких як рівномірність та незалежність. Популярні

тести включають тест Колмогорова-Смірнова, тест хі-квадрат та інші. Статистичні тести - це методи аналізу, які використовуються для перевірки статистичних гіпотез про властивості вибірок даних. Вони застосовуються для визначення, наскільки добре деякі дані або результати досліджень відповідають очікуваному розподілу чи моделі. В контексті випадкових чисел, статистичні тести використовуються для оцінки, наскільки псевдовипадкова послідовність відповідає властивостям справжньої випадкової послідовності. Найпоширеніші статистичні тести для перевірки якості генераторів псевдовипадкових чисел включають: тест Колмогорова-Смірнова – тест використовується для перевірки, наскільки псевдовипадкові числа рівномірно розподілені на вказаному інтервалі; тест серій (тест Мільора) – допомагає виявити будь-які аномалії або закономірності в послідовності; тест хі-квадрат – тест використовується для перевірки, наскільки послідовність відповідає певному розподілу, наприклад, рівномірному або нормальному; тест довжин серій (тест Вальда-Вольфовіца) – оцінює, наскільки добре послідовність випадкових чисел відповідає рівномірному розподілу; тест на дисперсію – тест вимірює різницю між фактичною і очікуваною дисперсією послідовності чисел. Ці тести допомагають виявити відхилення від властивостей випадкових чисел і дозволяють вдосконалити генератори псевдовипадкових чисел.

Тести на серії. Ці тести перевіряють наявність патернів або серій у послідовності, що може свідчити про нелінійність або нерівномірність розподілу чисел. Тести на серії – це статистичні методи аналізу для виявлення аномалій або закономірностей у послідовності даних, де кожне значення залежить від попередніх значень. Вони часто використовуються для перевірки випадковості чи структури в даних. Основні тести на серії включають такі:

- тести на серії випадкових величин – тести визначають, наскільки послідовності даних містять серії, тобто послідовності однакових значень;
- тести на алеаторність – тести визначають, наскільки послідовність даних відповідає випадковому процесу, вони оцінюють рівномірність розподілу даних і виявляють будь-яку структуру або закономірність;

- тести на автокореляцію – тести визначають наявність кореляції між значеннями в послідовності даних і виявляють повторюваність або закономірності в даних; т
- ести на періодичність – тести виявляють наявність періодичності або циклічності в послідовності даних;
- тести на випадковість – тести визначають, наскільки послідовність даних відповідає випадковому процесу та чи містить вона будь-яку структуру або закономірність.

Ці тести допомагають виявити аномалії або нерегулярності у послідовності даних, що може бути корисним при аналізі та обробці даних.

Спектральний тест. Використовується для визначення періодичності у послідовності, що може свідчити про недостатню складність генератора псевдовипадкових чисел.

Спектральний тест – це один з методів статистичного тестування послідовностей псевдовипадкових чисел на рівномірність та випадковість. Цей тест базується на властивостях перетворення Фур'є, яке можна використовувати для аналізу частотного складу послідовності. Основна ідея полягає в тому, щоб перетворити послідовність чисел у частотний домен і перевірити розподіл частот для виявлення будь-яких нерегулярностей або закономірностей.

Процес спектрального тестування можна умовно поділити на наступні кроки:

- перетворення Фур'є – це коли послідовність чисел перетворюється у частотний домен за допомогою перетворення Фур'є;
- аналіз спектра, тобто аналізується отриманий спектр для виявлення нерегулярностей або виокремлення характеристик, що свідчать про випадковість чи її відсутність;
- тестування гіпотез, який на основі аналізу спектра вирішує, чи відповідає послідовність чисел вимогам випадковості;
- робиться висновок про те, чи можна вважати дану послідовність випадковою або чи потрібно подальше тестування.

Спектральний тест є одним з багатьох методів для перевірки якості генерації псевдовипадкових чисел і використовується у криптографії,

моделюванні, статистичному аналізі даних та інших галузях, де важлива випадковість даних.

Тести на складність. Ці тести оцінюють складність псевдовипадкової послідовності, зокрема, ентропію та взаємну інформацію між символами. Тести на складність використовуються для оцінки ефективності алгоритмів управління, планування та рішення складних завдань. Ці тести включають в себе різні вимоги та завдання, які дозволяють визначити рівень складності алгоритмів та їх здатність до розв'язання складних проблем. Деякі з найпоширеніших тестів на складність включають такі:

- тест Тюрінга – дозволяє оцінювати здатність комп'ютерної системи до виконання завдань, які вимагають людського інтелекту, який базується на концепції Тюрінг-машини і оцінює, наскільки ефективно комп'ютер може імітувати людський мислення;
- тест Дейкстри – тест вимагає від алгоритму знаходження найкоротшого шляху між двома точками у графі та використовується для оцінки швидкодії та ефективності алгоритмів управління графами;
- тест Коллатца – тест вимагає від алгоритму генерації послідовності чисел за певними правилами та використовується для визначення рівня складності алгоритмів та їх здатності до генерації складних послідовностей;
- тест Дейвіса-Патнема – тест вимагає від алгоритму розв'язання булевої задачі заданої складності та використовується для оцінки швидкодії та ефективності алгоритмів управління булевими задачами.

Ці тести допомагають визначити рівень складності алгоритмів та їх придатність для рішення складних завдань у різних галузях науки та техніки.

Тести на кореляцію. Перевіряють наявність кореляційних залежностей між елементами послідовності. Тести на кореляцію використовуються для визначення наявності і сили зв'язку між двома або більше змінними.

Ці тести допомагають встановити, чи є статистично значуща залежність між змінними, тобто чи можна зробити висновок про наявність зв'язку в популяції на основі даних з вибірки.

Один з найпоширеніших тестів на кореляцію – це коефіцієнт кореляції Пірсона, який вимірює лінійний зв’язок між двома змінними. Цей тест дає

значення від (-1) до (+1), де (+1) – вказує на досконалий позитивний лінійний зв’язок, (-1) – на досконалий негативний лінійний зв'язок, а (0) – на відсутність лінійного зв’язку.

Інші тести на кореляцію включають рангові кореляції, такі як Спірмена і Кендалла, які вимірюють не лінійний, але монотонний зв’язок між змінними. Такі тести використовуються, коли дані не відповідають вимогам для застосування коефіцієнта кореляції Пірсона, наприклад, коли дані мають великі відхилення від нормального розподілу або мають викиди.

Кожен з цих методів має свої переваги та обмеження і може бути використаний для оцінки відповідності псевдовипадкової послідовності до випадкового розподілу з різних точок зору. Розглянемо на прикладах деяких сучасних галузей знань, яким чином застосовуються псевдовипадкові числа.

У неруйнівному контролі (НК) псевдовипадкові числа можуть використовуватися для генерації випробувальних сигналів або шаблонів, які допомагають виявляти дефекти чи аномалії у тестових об'єктах без їх пошкодження. Прикладами можуть служити: генерація випробувальних сигналів (сигнали допомагають виявляти дефекти у внутрішніх структурах об'єктів), шумоподібні сигнали для оцінки шуму (допомагають оцінити рівень шуму в системі НК та визначити його вплив на якість контролю), симуляція дефектів (дозволяє тестувати ефективність методів НК у виявленні цих дефектів) та ін.

Загалом, використання псевдовипадкових чисел у неруйнівному контролі дозволяє покращити якість контролю, забезпечуючи більш точне виявлення дефектів та аномалій у тестових об'єктах.

У кібербезпеці псевдовипадкові числа використовуються для багатьох цілей, зокрема для генерації криптографічних ключів, створення випадкових ідентифікаторів, захисту від атак на основі відомостей про структуру випадковості та інших аспектів захисту інформації.

Наведемо деякі способи використання: генерація криптографічних ключів, генерація випадкових ідентифікаторів, захист від атак на основі відомостей про структуру випадковості; створення випадкових токенів та кодів

доступу, захист від атак методом брутфорсу, створення випадкових векторів ініціалізації.

Загалом, використання псевдовипадкових чисел у кібербезпеці є важливим елементом для забезпечення високого рівня захисту інформації та запобігання атакам.

Таким чином, можна стверджувати, що на сьогодні методи оцінки відповідності псевдовипадкової послідовності до випадкового розподілу в складних задачах багатьох різних галузей знань науки і техніки, особливо таких як автоматизація, кібербезпека, комп'ютерно-інтегровані технології, неруйнівний контроль – є найактуальнішими задачами та потребують постійного підвищення ефективності цих методів.

ЛІТЕРАТУРА

1. Pierre L'Ecuyer. Random Number Generation // Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice. – 2007. – P. 93 -137
2. ДСТУ ISO/IEC 10118-1:2018 Інформаційні технології. Методи захисту. Хеш-функції. Частина 1. Загальні положення (ISO/IEC 10118-1:2016, IDT). На заміну ДСТУ ISO/IEC 10118-1:2003. – Чинний від 2019-01-01. – Вид. офіц. Київ : УкрНДЦ. – 2018. – 18 с.

METHODS OF ASSESSING THE CONFORMITY OF A PSEUDO-RANDOM SEQUENCE TO A RANDOM DISTRIBUTION IN THE PROBLEMS OF AUTOMATION, CYBER SECURITY, COMPUTER-INTEGRATED TECHNOLOGIES AND NON-DESTRUCTIVE CONTROL

Klymenko O., Malaichuk V.

Abstract. *Creation of methods for determining the degree of closeness of a given pseudorandom sequence to a random distribution is a very relevant task in many fields where pseudorandom numbers are used. For example, in cryptography it is important that pseudorandom sequences are unpredictable and have the properties of random numbers. Methods for determining proximity to a random distribution help in assessing the quality of such sequences and identifying flaws in them. In scientific research and engineering projects, pseudorandom numbers are used to model complex systems and simulate random phenomena. It is important to have a means of determining how well such numbers reproduce a random process. In statistics, it is important that pseudorandom sequences meet the requirements of various statistical properties. Methods of determining proximity to a random distribution help in evaluating the properties of these sequences. In software engineering and other fields, it is important to*

test the quality of pseudorandom number generators to avoid vulnerabilities and flaws in programs.

Thus, the development and improvement of methods for determining the degree of closeness of pseudorandom sequences to a random distribution is an urgent task with important practical applications.

Keywords*: automation, cyber security, computer-integrated technology, nondestructive testing, distribution, test, pseudorandom numbers, random process.*

REFERENCE

1. Pierre L'Ecuyer. Random Number Generation // Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice. – 2007. – P. 93 -137

2. State standards of Ukraine ISO/IEC 10118-1: 2018 Information technology. Methods of protection. Hash functions. Part 1: General provisions (ISO/IEC 10118-1:2016, IDT). Replaces DSTU ISO/IEC 10118-1:2003. - Valid from 2019-01-01. Kyiv : UkrNDC. - 2018. - 18 p.