УДК 004.056.5 **DOI:10.30837/rt.2024.1.216.02**

Д.Ю. ГОЛУБНИЧИЙ, канд. техн. наук, С.О. КАНДІЙ, М.В. ЄСІНА, канд. техн. наук, Д.Ю. ГОРБЕНКО

МЕТОДИ ТА ЗАСОБИ АНАЛІЗУ, ОЦІНКИ ТА ПОРІВНЯННЯ ВЛАСТИВОСТЕЙ ВИПАДКОВИХ ПОСЛІДОВНОСТЕЙ ТА ВИПАДКОВИХ ЧИСЕЛ

Вступ

Наразі випадкові послідовності (ВП) та випадкові числа (ВЧ), що виробляються фізично справжніми (PT RNG) та нефізично справжніми (NPT RNG) генераторам, широко застосовуються на практиці – вони по суті законодавчо визначають механізми генерування ключів у криптографічних системах [1 – 11]. У залежності від криптографічних перетворень, вони застосовуються для генерації довгострокових ключів та ключів сеансу симетричних криптоперетворень, довгострокових асиметричних пар ключів та пар сеансових ключів, загальних параметрів криптоперетворень та криптографічних протоколів, специфічних одноразових значень (nonces), викликів (challenges), засліплення та маскування значень тощо [2 – 6, 9, 12 – 18].

Серед множини вимог до таких генераторів є забезпечення у ряді, а можливо і більшості, криптографічних застосунків максимально можливого значення початкової ентропії. По суті, критерій максимуму початкової ентропії криптографічних застосунків є безумовним, і, при обґрунтовано вибраних розмірах ключових даних та параметрів, дозволяє забезпечити необхідні та достатні умови їх криптографічної живучості та криптографічної стійкості. Вказане однозначно визначає необхідність генерування ключів та параметрів криптографічних перетворень на основі тільки PT RNG та NPT RNG джерел шуму [2 – 5, 10, 11]. Іншими обов'язковими вимогами є обов'язковість стандартизації та сертифікації методів та засобів генерування ВП та ВЧ [2 – 6, 10 – 18] на основі відповідних джерел шуму.

Аналіз міжнародних та національних нормативно-правових документів щодо вимог до PT RNG та NPT RNG джерел та відповідно до генераторів показав, що вони, з урахуванням суттєвих викликів, що пов'язані з розширенням можливостей криптоаналізу на основі застосування, крім класичних, квантових та атак бічними каналами, в суттєвій мірі повинні бути удосконаленими [14 – 18] та оцінені з використанням комплексних методик з використанням системи безумовних критеріїв.

В подальшому у цій статті в якості основних будемо дотримуватись таких основних визначень [19]:

- випадкова послідовність (ВП) – послідовність незалежних та однаково розподілених змінних (з рівномірного розподілу);
- випадкові числа (ВЧ) – це дискретні значення (зазвичай біти, рядки бітів або цілі числа), які отримують у окремі моменти часу з джерела шуму генератори випадкових послідовностей;
- псевдовипадкова послідовність (ПВП) – послідовність символів, що обчислювально не відмінна від випадкової послідовності і згенерована детермінованим алгоритмом.

Пошук новітніх теоретичних та практичних досліджень та результатів щодо методів та засобів генерування ВП та ВЧ дозволив зупинитися на перспективній програмі німецьких загальних критеріїв (CC) протягом приблизно двох десятиліть – AIS 20 і AIS 31, що визначають як удосконалювати та оцінювати різні RNG генератори [21, 17 – 19]. Вони визначають класи функціональності для різних типів RNG. Щоб бути сумісним з певним класом функціональності, RNG мають відповідати всім вимогам класу. Крім того, AIS 20 і AIS 31 окреслюють методологію оцінювання детермінованих RNG (DRNG) і справжніх RNG (TRNG) [2 – 5].

Традиційно в ряді випадків вважається, що наявність статистичного тестування RNG генераторів та відповідність вимогам згідно з [6 – 8] вже гарантує випадковість. Але при застосуванні порушником класичних, квантових атак і атак бічними каналами, цього недостатньо. Однозначно визнано, що крім статистичного тестування, необхідно застосовувати і стохастичне тестування на основі аналізу та оцінки початкової ентропії ВП та ВЧ, наприклад на основі PTRNG та NPTRNG, в умовах застосування класичних, квантових атак та атак бічними каналами [12 – 18, 20].

Метою даної статті є обґрунтування, розробка та експериментальне підтвердження коректного застосування алгоритмів генерування ВП та ВЧ на основі PTRNG та NPTRNG, в тому числі при застосуванні класичної та квантової мікроелектроніки, а також розробка рекомендацій щодо їх застосування для генерування ключів та параметрів для квантово стійких методів та стандартів криптографічних перетворень.

1. Джерела шуму, що можуть застосовуватися для генерації ВП та ВЧ на основі PTRNG та NPTRNG

Наразі існує значне число ДШ, які можливо застосовувати для генерування ВП та ВЧ. Аналіз показав, що ДШ, які можуть задовольняти вимогам, можливо поділити на два класи [2, 4, 5, 10, 11]:

- фізичні ДШ, в основі таких лежить деяке непередбачуване ймовірнісне фізичне явище, яке містить певну кількість ентропії [2, 3];
- нефізичні ДШ, в основі яких також лежить в певній мірі передбачуване явище, кожне з яких містить певну кількість ентропії.

Причому на основі фізичних ДШ можливо генерувати фізично справжні PT RNG ВП та ВЧ, а на основі нефізичних ДШ можливо генерувати нефізичні справжні NPT RNG ВП ТА ВЧ [2, 10, 11].

Фізичні ДШ можливо класифікувати згідно з фізичним явищем, яке має ймовірнісний характер, що має значну кількість ентропії [2, 3]:

- ДШ на основі шуму, в цьому класі певний фізичний процес має ймовірнісний характер, який практично неможливо передбачити, та можливість генерувати фізично справжні PT RNG ВП та ВЧ;
- ДШ на основі хаосу, цей клас ДШ ґрунтується на наявності певної системи (можливо навіть детермінованої) з багатьох складових, що має хаотичну поведінку у цілому;
- ДШ на основі вільних осциляторів, що можуть мати ентропію, ця можливість ґрунтується на непередбачуваності явищ в цифровій електротехніці. Такі вільні осцилятори є популярним вибором фізично справжнього PT RNG генератора ВП чи ВЧ на основі різноманітних персональних пристроїв;
- ДШ на основі ефектів квантової мікроелектроніки, що складають відносно новий, перспективний клас ДШ, в ньому ентропія створюється з використанням квантових ефектів мікроелектроніки.

Джерелами нефізичного шуму, що можуть генерувати NPT RNG нефізично справжні ВП чи ВЧ можуть слугувати [2, 4, 5]:

- події, що ґрунтуються на процесах взаємодії з суб'єктом (наприклад, користувачем) чи об'єктом та які можливо завдати на основі ймовірнісного подання;
- події, що ґрунтуються на використанні переривання апаратних пристроїв (наприклад, мережевої карти), для яких можливо задати ентропію на основі ймовірнісного подання переривання тощо.

Більш детальні дані щодо цих ДШ та відповідно фізично справжніх PT RNG генераторів і нефізично справжніх генераторів ВП (ВЧ) можна знайти в [2, 3].

У табл. 1 наведено основні переваги, недоліки та особливості і результати порівняння фізично справжніх ДШ, що можуть застосовуватись для генерування ключових даних та

загальних параметрів для існуючих класичних та перспективних квантово стійких криптографічних перетворень та криптографічних протоколів.

Таблиця 1

Переваги, недоліки та особливості ДШ

Тип	Переваги	Недоліки	Особливості
На основі шуму	Багато конструкцій генераторів, практичних досліджень та наявних впроваджень	Рух частинок, які генерують шум, в певній мірі взаємокорельований. Шум не можна «перезапустити», щоб перервати кореляції між послідовними вимірюваннями генерування бітів. У них більшість процесів, наприклад у резисторах, стабілітронах та транзисторах, мають певний ефект пам'яті. Випадковість джерел шуму неможливо завдати, виміряти або навіть контролювати під час виготовлення пристрою. Щодо них необхідний захист від впливу зовнішніх електромагнітних полів та випромінювання генератора	Для таких конструкцій використовують переважно шуми електричної природи. Основні фізичні ефекти: ефект Джонсона та ефект Зенера
На основі хаосу	Різноманітність фізичних процесів, на яких ґрунтуються генератори	Важко довести, що система дійсно хаотична. ДШ на основі хаосу в довгостроковій перспективі не можуть виробляти нову ентропію, що неминуче закінчується виробленням не менше 1 біта ентропії на кожен новий згенерований випадковий біт	Основні типи конструкцій: оптичні, електричні, оптико-електричні та механічні
На вільних осциляторах	Перевірені та стандартизовані конструкції. Найрозповсюдженіший тип генераторів. RNG на основі вільних осциляторів є недорогими рішеннями, які можна легко реалізувати в звичайних програмованих або реконфігурованих логічних мікросхемах	Коли декілька осциляторів розташовані близько один до одного (наприклад, на одній мікросхемі), вони, як правило, синхронізуються через електромагнітну взаємодію, що сприяє високому посиленню підсилювачів вільних осциляторів, що робить ДШ вразливими до атак із зовнішнім електромагнітним випромінюванням. Якщо такі ДШ синхронізуються або принаймні частково синхронізуються, з'явиться шаблон зі стохастичним відхиленням (шумом). Окрім цього, ще одна дуже важлива проблема щодо ДШ на вільних осциляторах полягає в тому, що амплітуда вихідного сигналу ДШ залежить від деталей блукаючих реактивних опорів і затримок у ланцюзі. Складність процедур постобробки, необхідних для проходження статистичних тестів, у RNG на основі вільних осциляторів часто така, що будь-який доказ випадковості ускладнений	Часткове вирішення проблем було знайдено в новій синергетичній комбінації регістра зсуву лінійного зворотного зв'язку (LFSR), а також вільних осциляторів, яка називається кільцевим осцилятором Фібоначчі (FIRO) і кільцевим осцилятором Галуа (GARO)
Радіоактивний розпад	Дійсно випадковий квантовий процес. Добре досліджені конструкції. Доказова випадковість	Оскільки випадкове джерело є радіоактивним, воно вимагає особливої обережності (покращених заходів безпеки) та знань. Обмеженням є також «мертвий час» детектора через накопичення іонів усередині детектора. Низька швидкість генерації	Одне із перших квантових явищ, що були використані для генерації випадкових бітів та чисел. Можливо використовувати два основні методи – метод швидкої синхронізації та метод повільного годинника. Деякі сучасні генератори випадкових чисел, засновані на радіоактивному розпаді. У них замість GM-трубок використовують напівпровідникові пристрої

Продовження табл. 1

Тип	Переваги	Недоліки	Особливості
Атомарні системи	Використовується квантовий ефект. Забезпечується доказова випадковість	Експериментальні установки, необхідні для генерації випадкових чисел з використанням захоплених іонів, набагато складніші, генерація з дуже низькою швидкістю	Використовують спіновий ефект (шум)
Фотонні детектори	Використовуються квантові ефекти. Велика швидкість генерації. Доказова випадковість.	Після кожної події виявлення, детектори неактивні протягом певного періоду, протягом якого вони не можуть виявити фотони. Це призводить до кореляції між згенерованими бітами та збільшує час, необхідний для їх отримання. Цього можна уникнути, використовуючи лише один детектор	Швидкість генерації випадкових бітів також може бути покращена, якщо генератор вимірює кілька шляхів проходження фотонів
Вакуумний шум	Використовуються квантові ефекти. Доказова випадковість. Практичні реалізації досягли швидкості генерації до 3 Гбіт/с	Швидкість генерації ВП та ВЧ у цих пристроях обмежена швидкістю детектора в зоні дробового шуму, коли в загальному спостережуваному шумі домінує вакуумний шум	Використовують випадковість вакуумних флуктуацій електромагнітного поля

2. Аналіз вимог щодо ДШ та їх ентропії

Основоположні вимоги до джерел шуму та його ентропії було запропоновано в стандарті NIST SP800-90B. Метою вимог NIST США до джерела ентропії [2, 4] є надання розробнику допомоги в розробці/впровадженні джерела ентропії, яке може надати вихідні дані з постійною кількістю ентропії, а також надати необхідну документацію для перевірки джерела ентропії.

Аналіз показує, що стандарт NIST SP800-90B [4] висуває наступні вимоги до ДШ у вигляді PTRNG та NPTRNG:

1. Наявність обґрунтованої стохастичної моделі вихідних сигналів ДШ. Модель повинна включати опис того, як працює ДШ та яким чином створюється непередбачуваність, а також і обґрунтування того, чому джерело шуму забезпечує прийнятну вихідну ентропію.

2. Поведінка джерела шуму має бути стаціонарною, коли розподіли ймовірностей вихідних сигналів джерела шуму не змінюються з часом при роботі джерела в нормальних умовах. Для цього повинне бути обґрунтовано, звідки походить непередбачуваність, і приблизно описано поведінку ДШ щодо стаціонарності його поведінки.

3. Модель ДШ повинна надавати чітке визначення очікуваної ентропії, що забезпечується вихідними сигналами джерела шуму, і надавати технічну аргументацію, чому джерело шуму може підтримувати таку швидкість ентропії.

4. Стан джерела шуму має бути максимально захищений від впливу. Методи, що використовуються для цього, повинні бути задокументовані, щодо межі безпеки захисту ДШ від впливу.

5. Незважаючи на те, що джерело шуму не зобов’язане створювати неупереджені та незалежні вихідні сигнали, воно повинно демонструвати випадкову поведінку, коли вихід не може бути визначений жодним відомим алгоритмічним правилом.

6. Джерело шуму має генерувати випадкові значення фіксованої довжини та має опис вихідного простору джерела шуму.

7. Якщо для підвищення безпеки використовуються додаткові ДШ, необхідно мати документ, який описує додаткові джерела шуму.

3. Методи оцінки ентропій ВП та ВЧ, що згенеровані PTRNG та NPTRNG

Ентропійні методи оцінки ґрунтуються на використанні узагальненого поняття ентропії Реньї [2, 21]. Нехай X – випадкова змінна, що у найбільш узагальненому виді визначає ентропію Реньї. Для випадкової змінної X ентропію Реньї можливо розрахувати за формулою

$$H_\alpha(X)=\frac{1}{1-\alpha}\log_2\sum_{i=1}^{k}(Pr[X=\omega_i])^\alpha, 0\le\alpha<\infty. \quad (1)$$

На практиці значеннями параметра α є 1, 2 та ∞. При α є 1 із (1) отримуємо співвідношення для оцінки ентропії Шеннона, при α є 2 – для оцінки колізійної ентропії, а при α є ∞ – для оцінки мінімальній ентропії.

За визначенням ентропія Реньї $H_\alpha(X)$ залежить тільки від розподілу μ випадкової змінної X, тому будемо використовувати також нотацію $H_\alpha(\mu)$, показуючи залежність $H_\alpha(X)$ як від α, так і від μ.

Розглянемо більш детально кожен випадок і покажемо, що у приватних випадках відповідних α маємо збіг.

Врахуємо, що ентропія Шеннона є границею ентропії Реньї $H_\alpha(X)$ в точці $\alpha=1$, в результаті отримуємо

$$H_1(x)=\lim_{\alpha\to1}H_\alpha(x)=\lim_{\alpha\to1}\frac{1}{1-\alpha}\log_2\sum_{i=1}^{k}(Pr[X=\omega_i])^\alpha.$$

Для того щоб отримати класичну формулу ентропії Шеннона з цієї границі скористаємося правилом Лопіталя, згідно з яким для двох дійсних функцій $f(x)=\log_2\sum_{i=1}^{k}(Pr[X=\omega_i])^\alpha$ та $g(x)=1-\alpha$, що мають похідні $f'(x), g'(x)$ в околиці точки δ, має місце вираз

$$\lim_{x\to\delta}\frac{f(x)}{g(x)}=\lim_{x\to\delta}\frac{f'(x)}{g'(x)}.$$

У цьому випадку $f(x)=\log_2\sum_{i=1}^{k}(Pr[X=\omega_i])^\alpha$, $g(x)=1-\alpha$ і $\delta=\alpha=1$.

Для отримання похідних скористаємося такими формальними правилами взяття похідних складної функції

$$\frac{d}{dx}\log_2(x)=\frac{1}{x\ln 2}, \frac{d}{dx}a^x=a^x\cdot\ln a.$$

Підставляючи відповідні значення у формулу (1), отримуємо:

$$\lim_{\alpha\to1}H_\alpha(x)=\lim_{\alpha\to1}\frac{\log_2\sum_{i=1}^{k}(Pr[X=\omega_i])^\alpha}{1-\alpha}=\lim_{\alpha\to1}\frac{\frac{d}{dx}\left(\log_2\sum_{i=1}^{k}(Pr[X=\omega_i])^\alpha\right)}{\frac{d}{dx}(1-\alpha)}=$$

$$=\lim_{\alpha\to1}\frac{\left(\sum_{i=1}^{k}(Pr[X=\omega_i])^\alpha\right)^{-1}\sum_{i=1}^{k}Pr[X=\omega_i]^\alpha\cdot\ln(Pr[X=\omega_i])}{-\ln 2}$$

Використовуючи властивість натурального логарифму $\frac{\ln b}{\ln c}=\log_c b$, перетворюємо $\frac{\sum_{i=1}^{k}\ln(Pr[X=\omega_i])}{-\ln 2}$ і отримаємо наступний вираз:

$$\lim_{\alpha\to1}H_\alpha(x)=-\sum_{i=1}^{k}Pr[X=\omega_i]\cdot\log_2(Pr[X=\omega_i]).$$

Звідки, маємо класичну формулу Шеннона:

$$H_1(X) = H(X) = -\sum_{i=1}^{k} Pr[X = \omega_i] \log_2 (Pr[X = \omega_i]). \quad (2)$$

Якщо $Pr[X = \omega_i] = 0$, то, за домовленістю, $Pr[X = \omega_i] \log_2 (\Pr[X = \omega_i]) = 0$. Позначення H зазвичай використовується замість H_1 для ентропії Шеннона. Ентропію Шеннона $H = H_1$ іноді називають загальною ентропією, або просто ентропією через її важливість в теорії інформації [1].

У (1) вказано, що мінімальна ентропія є спеціальним випадком для якого $\alpha = \infty$. На основі виразу (1) отримаємо аналітичне співвідношення для мінімальної ентропії. Так як $\alpha = \infty$, то на основі обчислення границі скористаємося тим фактом, що для усіх i=1,2, …k, $0 \le Pr[X = \omega_i] \le 1$. При збільшенні α сума $\sum_{i=1}^{k} Pr[X = \omega_i]^{\alpha}$ буде наближатися до $\max_i Pr[X = \omega_i]$.

Позначимо $p_i = Pr[X = \omega_i]$.

Розглянемо на основі ентропії Реньї альтернативний варіант отримання співвідношення для мінімальної ентропії. Враховуючи прийняте позначення, маємо

$$\lim_{\alpha \to \infty} H_\alpha(x) = \lim_{\alpha \to \infty} \frac{\log_2 \sum_{i=1}^{k} p_i^{\alpha}}{1 - \alpha}.$$

Для виділення величини $\max_i p_i^{\alpha}$ виконаємо перетворення

$$\lim_{\alpha \to \infty} \frac{\log_2 \left(\sum_{i=1}^{k} \left(\left(\frac{p_i}{\max_i p_i} \right)^{\alpha} \cdot \max_i p_i^{\alpha} \right) \right)}{1 - \alpha}.$$

Оскільки добуток під логарифмом дає суму двох логарифмів, то справедливо, що

$$\lim_{\alpha \to \infty} \frac{\log_2 \sum_{i=1}^{k} \left(\frac{p_i}{\max_i p_i} \right)^{\alpha} + \log_2 \left(\max_i p_i^{\alpha} \right)}{1 - \alpha} =$$

$$= \lim_{\alpha \to \infty} \frac{\log_2 \sum_{i=1}^{k} \left(\frac{p_i}{\max_i p_i} \right)^{\alpha}}{1 - \alpha} + \lim_{\alpha \to \infty} \frac{\log_2 \left(\max_i p_i^{\alpha} \right)}{1 - \alpha}.$$

Позначимо, $\beta = \sum_{i=1}^{k} \left(\frac{p_i}{\max_i p_i} \right)^{\alpha}$, маємо:

$$\frac{\log_2 \beta}{1 - \alpha} + \lim_{\alpha \to \infty} \frac{\log_2 \left(\max_i p_i^{\alpha} \right)}{1 - \alpha}.$$

Оскільки, значення p_i пронормоване $\max_i p_i$, то під знаком суми одне $\max_i p_i = 1$, а сума всіх інших пронормованих ймовірностей – менше або дорівнює 1.

Тобто, $1 < \beta \le k$. Так як $\alpha \to \infty$, то значення $\beta \to \infty$:

$$\frac{\log_2 \beta}{\infty} + \lim_{\alpha \to \infty} \frac{\log_2 \left(\max_i p_i^{\alpha} \right)}{1 - \alpha} = \frac{\log_2 \beta}{\infty} + \lim_{\alpha \to \infty} \frac{\alpha \log_2 \left(\max_i p_i \right)}{1 - \alpha}.$$

Оскільки β – скінченна величина, то $\frac{\log_2 \beta}{\infty} = 0$. Тому можна записати для похідних

$$0 + \lim_{\alpha\to\infty} \frac{\alpha \log_2 (\max_i p_i)}{1-\alpha} = \lim_{\alpha\to\infty} \frac{\frac{d}{d\alpha}\left(\alpha \log_2 (\max_i p_i)\right)}{\frac{d}{d\alpha}(1-\alpha)} =$$

$$= \lim_{\alpha\to\infty} \frac{\frac{d}{d\alpha}(\alpha)}{\frac{d}{d\alpha}(1-\alpha)} \log_2 (\max_i p_i) = \frac{1}{-1} \log_2 (\max_i p_i) = -\log_2 (\max_i p_i) \quad (3)$$

Отримаємо та перевіримо на основі використання виразу для ентропії Реньї вираз для ентропії колізій. Нехай H_2 позначає колізійну ентропію. Нехай також, X та X' – дві незалежні та однаково розподілені випадкові змінні з значеннями в деякій множині Ω. Тоді із виразу для ентропії Реньї (1), при $\alpha = 2$, маємо

$$H_2(X) = \frac{1}{1-2} \log_2 \left(\sum_{\omega\in\Omega} (Pr[X=\omega])^2 \right) = -\log_2 \left(\sum_{\omega\in\Omega} (Pr[X=\omega])^2 \right). \quad (4)$$

Практичні дослідження показують, що має місце таке співвідношення між енропією Шеннона, колізійною і мінімальною ентропією [2, 21]:

$$H_{min} \le H_2 \le H_1, \; H_{min} \le H_2 \le 2H_{min}. \quad (5)$$

Мінімальна ентропія є найбільш консервативною ентропією. На рис. 1 для прикладу зображені H_1, H_2 та H_∞ для бінарних випадкових змінних.

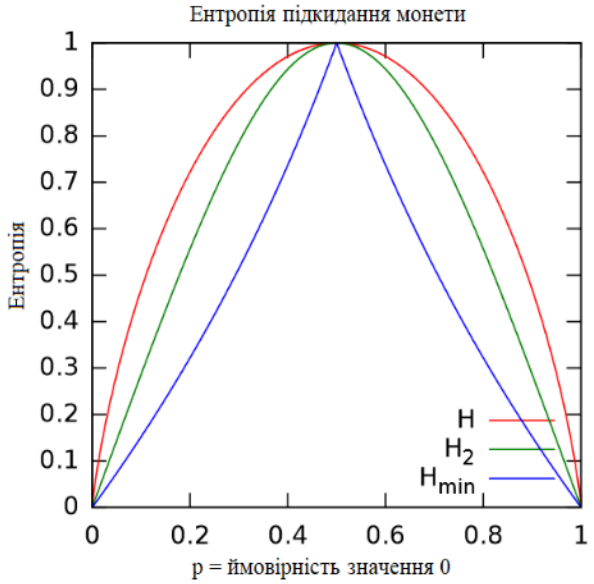

Рис. 1. Мінімальна ентропія, колізійна ентропія та ентропія Шеннона для випадкових бінарних змінних

Таким чином, задаючи значення ймовірностей $Pr[X=\omega_i]$, по кривим рис. 1 перевіряється, чи виконується практично співвідношення між $H_1, H_2, H_{\min}$ та $2H_{min}$. Якщо так, то приймається рішення, що послідовність відповідає ентропійним критеріям випадковості. Після цього етапу, для уточнення, рекомендується провести статистичне тестування, наприклад, з використанням NIST 800-22 [6] та DIEHARD [8].

4. Дослідження ентропій Шеннона, ентропії колізій та мінімальної ентропії ДШ згідно з NIST 800-90B

Методологічні основи оцінки ентропій ВП та ВЧ, що згенеровані PTRNG фізично справжніми генераторами, наведено в [2, 22 – 24]. Нижче наводяться обґрунтовані вище конкретизовані методики оцінки ентропій ВП та ВЧ – Шеннона, колізій та мінімальної ентропій. Вважається, що обґрунтовані та запропоновані нижче методики оцінки повинні (можуть) бути застосовані щодо ДШ (генераторів) як фізично справжніх (PTRNG), так і нефізично справжніх (NPT RNG) ДШ.

4.1. Оцінка ентропії Шеннона ДШ PT RNG та NPT RNG

Для дискретної випадкової змінної X з можливими значеннями $x_0, \mathrm{x}_1, ..., x_{k-1}$ та відповідними ймовірностями появи символів ($p_0, p_1, ..., p_{k-1}$), ентропія Шеннона визначена як

$$H(\mathrm{X}) = -\sum_{i=0}^{k-1} p_i \cdot \log p_i \,. \tag{6}$$

Надалі позначатимемо частоту появи символу x_i у вибірці розміру n як $n(\mathrm{x}_i)$ та відповідну емпіричну ймовірність $p_i = n(x_i)/n$.

Якщо застосовувати формулу (6) напряму до реальних статистичних даних, то оцінка буде зміщеною [25, 26]. Щоб отримати незміщену оцінку, можливо використовувати корекцію, зокрема наступні методи корекції, що задані формулами (7) – (9) [27]:

$$\hat{H}(X) = H(X) + \frac{m-1}{2n}, \tag{7}$$

де m – кількість різних символів, що зустрілися в статистичних даних, n – розмір вибірки;

$$\hat{H}(X) = n \cdot H(X) - \frac{n-1}{n} H_{-i}(X), \tag{8}$$

де $H(\mathrm{X})_{-i}$ – це $H(X)$ без доданка $p_i \cdot \log p_i$;

$$\hat{H}(X) = -\sum_{i=0}^{n-1} a_i \cdot h_i \,. \tag{9}$$

Також $h_i = \sum_{j=1}^{k-1} [[p_j \cdot n == i]]$ та $a_i = -\frac{i}{n}\log\frac{i}{n} + \left(1 - \frac{i}{n}\right)_{2n}$ (тут $[[\cdot]]$ позначає предикат).

Для зручності надалі будемо використовувати позначення MM для формули (8), JKF для формули (8) та BUB для формули (9).

Інший підхід до усунення зміщення полягає у використанні формули Баєса [28]:

$$\hat{H}(X) = -\sum_{i=0}^{k-1} \frac{p_i \cdot n + a_i}{n + A} \log\left(\frac{p_i \cdot n + a_i}{n + A}\right), \tag{10}$$

де $A = \sum_{i=0}^{k-1} a_i$, а значення a_i обираються в залежності від конкретного метода. Зокрема, часто $a_0 = a_1 = ... a_{k-1} = const$. Найбільш популярні вибори констант наступні [25, 28]:

- $a_i = 1/2$;
- $a_i = 1$;
- $a_i = 1/k$;
- $a_i = \sqrt{n}/k$.

З експериментальних досліджень у роботах [25, 29, 30] також варто виділити наступні три перспективні підходи до оцінки ентропії Шеннона:

- Підхід, що був запропонований у роботі [31] (надалі – SHU-оцінка). Значення ентропії визначається за формулою

$$\hat{H}(X)=\psi(n)-\frac{1}{n}\sum_{i=0}^{k-1}\left(\psi(p_i\cdot n)+(-1)^{p_i\cdot n}\int_0^{1/\xi-1}\frac{t^{p_i\cdot n-1}}{1+t}dt\right). \quad (11)$$

- Підхід, що був запропонований у роботі [32] (надалі – CS-оцінка). Значення ентропії визначається за формулою

$$\hat{H}(X)=-\sum_{i=0}^{k-1}\frac{\tilde{p}_i\log\tilde{p}_i}{1-\left(1-\tilde{p}_i\right)^n}, \quad (12)$$

де $\tilde{p}_i=\left(1-\frac{m}{n}\right)p_i$.

- Підхід, що був запропонований у роботі [33] (надалі – SHR-оцінка). Значення ентропії визначається за формулою

$$\hat{H}(X)=-\sum_{i=0}^{k-1}\tilde{p}_i\log\tilde{p}_i, \quad (13)$$

де $\tilde{p}_i=\lambda/k+(1-\lambda)\mathrm{p}_i$, причому

$$\lambda=\frac{1-\sum_{i=0}^{k-1}\left(p_i\right)^2}{(n-1)\sum_{i=0}^{k-1}\left(1/k-p_i\right)^2}.$$

Також можливо виділити ряд інших методів оцінки ентропії Шеннона, що використовують більш складні статистичні методи, проте вони не набули широкого поширення через складність реалізації.

4.1.1. Експериментальна перевірка ентропії Шеннона

Для тестування методів оцінки ентропії Шеннона було написане програмне забезпечення, що генерує випадкові послідовності із заданою мінімальною ентропією (і відповідною ентропією Шеннона). Для цього генерується випадкова бінарна послідовність з рівномірного розподілу і до послідовності застосовується вибірка з відхиленням. Нехай p позначає ймовірність появи 1. Тоді, мінімальна ентропія, згідно з визначенням буде

$$H_{min}=\max(p,1-\mathrm{p}).$$

І відповідна ентропія Шеннона

$$H=-p\cdot\log_2 p-(1-p)\cdot\log_2(1-p).$$

На рис. 2 наведено експериментальні дані оцінки мінімальної ентропії бінарних послідовностей для $p=\{0.5,0.6,0.7,.0.8,0.9\}$ на послідовності довжини 10^6 біт.

З рис. 2 видно, що усі методи дають гарну оцінку ентропії як для високоентропійних послідовностей, так і для низькоентропійних послідовностей.

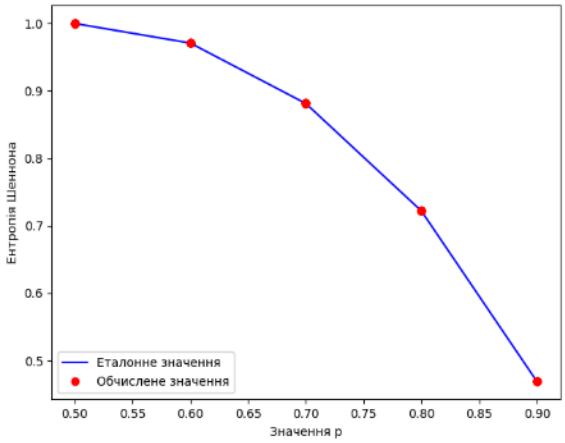

Рис. 2. Експериментальна оцінка ентропії Шеннона

4.2. Оцінка колізійної ентропії ДШ (ВП, ВЧ)

Колізійна ентропія визначена наступним чином:

$$H_2(X) = -\log\left(\sum_{i=0}^{k-1} p_i^2\right). \tag{14}$$

Для колізійної ентропії відомі наступні обмеження нерівностями:

$$H_{min} \le H_2 \le H\,, \tag{15}$$

$$H_{min} \le H_2 \le 2H_{min}\,. \tag{16}$$

Для практичних задач нерівностей, на нашу думку, достатньо, щоб оцінити значення колізійної ентропії, маючи оцінки мінімальної ентропії та ентропії Шеннона.

Для більш точних оцінок можливо скористатися методом, що описаний у роботі [34].

Метод залежить від двох параметрів – параметра точності оцінки δ та параметра статистичної помилки $\grave{o}$. Нехай деяка константа M, що гарантовано має значення більше за будь-яке можливе значення колізійної ентропії та довільна константа $c > 0$. Наприклад, $M = H(X)$ або $M = 2H_{min}(X)$. Тоді, можливо використовувати наступний алгоритм:

1. Обрати розмір блоку $N = \left\lceil c \cdot 2^{M/2} \delta^{-2} \right\rceil$
2. Обчислити кількість блоків $l = \lfloor n / N \rfloor$
3. Для $j = 1,\ldots,n/N$
   a. Для $i \in \{(j-1)N+1,\ldots,(j-1)N\}$
      i. $n(x_{i+1}) = n(x_i) + 1$
   b. $q_j = \frac{1}{m(m-1)}\left(\sum_{i=0}^{k-1} n(i)^2 - m\right)$
4. Знайти медіану q послідовності $q_1,\ldots,q_l$
5. Повернути $-\log q$

Метод доказово оцінює значення колізійної ентропії з похибками $(\delta, \grave{o})$, якщо $c \ge \log(1/\grave{o})$.

4.2.1. Експериментальна перевірка колізійної ентропії

Для експериментальної перевірки розглянутого методу для оцінки було використано таку ж методику, як і для ентропії Шеннона. Для заданого значення p колізійна ентропія визначається як

$$H_2 = -\log_2(\mathrm{p}^2 + (1-p)^2)\,.$$

На рис. 3 наведено експериментальні дані оцінки колізійної ентропії бінарних послідовностей для $p = \{0.5, 0.6, 0.7, .0.8, 0.9\}$ на послідовності довжини 10^6 біт.

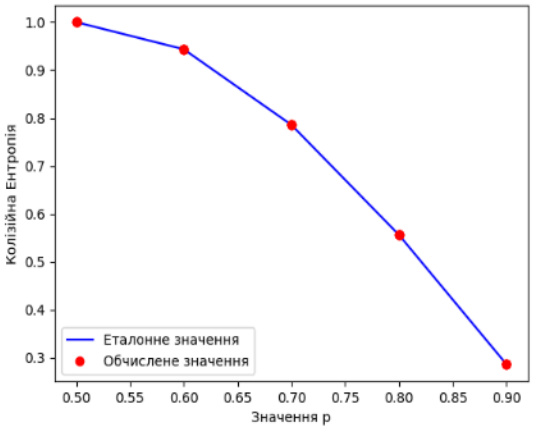

Рис. 3. Експериментальна перевірка оцінки колізійної ентропії

З рис. 3 видно, що усі методи дають гарну оцінку ентропії як для високоентропійних послідовностей, так і для низькоентропійних послідовностей.

4.3. Оцінка мінімальної ентропії ДШ (ВП, ВЧ)

Для дискретної випадкової змінної X з можливими значеннями $x_0, \mathrm{x}_1, ..., x_{k-1}$ (та відповідними ймовірностями $p_0, \mathrm{p}_1, ..., p_{k-1}$) мінімальна ентропія визначена як

$$H_{min}(X) = -\log_2\left(\max_{1\le i\le k}\{\mathrm{p}_i\}\right). \qquad (17)$$

Окрім позначення $H_{min}(\cdot)$ для мінімальної ентропії, також є популярним позначенням $H_{\infty}(\cdot)$.

Для оцінки мінімальної ентропії можливо виділити два основних підходи. Перший підхід базується на ентропійній статистиці, вперше описаний в [35]. Другий підхід базується на предикторах (англ. predictor) (на основі прогнозування), вперше описаних у [36].

Ентропійна статистика призначена для обчислення окремої статистики на вибірках. До методів, що використовують ентропійну статистику, належать:

- колізійний тест;
- тест на стиснення;
- тест Маркова.

Хоча оцінювачі ентропії (за винятком тесту Маркова) спочатку були розроблені для застосування до незалежних виходів, тести показали хороші результати при застосуванні до даних із залежностями.

Оцінювачі ентропії припускають, що розподіл ймовірностей описує вихід випадкового джерела шуму, але розподіл ймовірностей невідомий. Метою кожного оцінювача є виявлення інформації про невідомий розподіл на основі статистичних вимірювань.

Тести колізій та стиснення розв'язують рівняння для невідомого параметра, де рівняння є різними для кожного оцінювача. Ці рівняння походять із очікуваного значення цільової статистики з використанням майже рівномірного розподілу, який забезпечує нижню межу мінімальної ентропії. Майже рівномірний розподіл є прикладом однопараметричного сімейства розподілів ймовірностей, параметризованих p, P_p:

$$P_p(i) = \begin{cases} p, \text{ якщо i=0} \\ \dfrac{1-p}{k-1}, \text{ інакше} \end{cases}, \quad (18)$$

де k – кількість станів у вихідному просторі, а $p \geq \dfrac{1-p}{k-1}$, що має місце, коли $p \geq 1/k$. Іншими словами, один вихідний стан має максимальну ймовірність, а решта вихідних станів рівноймовірні.

Підхід на основі предикторів використовує два показники для отримання оцінки. Перший показник базується на глобальній продуктивності предиктора P_{global}, яка в літературі з машинного навчання називається точністю. По суті, предиктор фіксує частку правильних припущень. Це приблизно вказує на те, наскільки добре можна очікувати, що предиктор вгадає наступний вихід із джерела шуму на основі результатів довгої послідовності припущень. Другий показник P_{local} базується на найбільшій кількості правильних передбачень у рядку, який називається локальним показником ефективності. Ця метрика корисна для виявлення випадків, коли джерело шуму переходить у дуже передбачуваний стан протягом деякого часу, але предиктор може не працювати добре на довгих послідовностях. Розрахунки для оцінки локальної ентропії походять з теорії ймовірностей пробігів і повторюваних подій [37]. Для отримання додаткової інформації про оцінку мінімальної ентропії за допомогою предикторів див. [11].

Для того щоб оцінки предиктора схилялися до консервативної недооцінки мінімальної ентропії, P_{global} замінюється на P'_{global}, що відповідає 99-му квантилю кількості правильних прогнозів на основі спостережуваної кількості правильних прогнозів. Зауважимо, що порядок, у якому відбуваються правильні прогнози, не впливає на оцінку мінімальної ентропії на основі P_{global}. Наприклад, прогноз завжди може бути правильним для першої половини вихідних даних у наборі даних і завжди неправильним для другої половини вихідних даних. Оцінка мінімальної ентропії цієї послідовності на основі P_{global} становить половину довжини даних у бітах. З іншого боку, для іншої послідовності предиктор може мати 50 % шанс бути правильним для кожного зразка в цій послідовності. Мінімальна оцінка ентропії цієї другої послідовності, заснована на P_{global}, така ж, як і для першої послідовності. Однак типова тривалість успішного прогнозування для цих двох послідовностей дуже різна. Таким чином, цей підхід враховує ефективність локального передбачення, щоб консервативно зменшити оцінку мінімальної ентропії, якщо спостережувана поведінка локального передбачення є статистично значущою, враховуючи глобальний рівень успіху передбачення. Оцінки предикторів досягають цього, базуючи оцінку мінімальної ентропії на $\max(P'_{global}, \mathrm{P}_{local})$, де P_{local} – це частка успішного прогнозу, для якої спостережуваний найдовший ряд правильних прогнозів становить 99 %. Фактично це одностороння перевірка гіпотези, яка відхиляє P'_{global} на користь P_{local}, якщо спостережуваний найдовший пробіг, враховуючи ймовірність успіху P'_{global}, перевищує 99 %.

4.3.1. Експериментальна перевірка методів оцінки мінімальної ентропії

Для експериментальної перевірки використовувалася така ж сама методологія як і для ентропії Шеннона та колізійної ентропії.

На рис. 4. наведено експериментальні дані оцінки колізійної ентропії бінарних послідовностей для $p = \{0.5, 0.6, 0.7, .0.8, 0.9\}$ на послідовності довжини 10^6 біт.

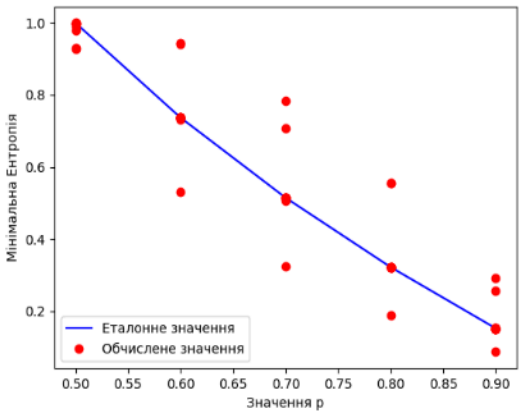

Рис. 4. Експериментальна перевірка оцінки мінімальної ентропії

З експериментальних даних видно, що для бінарних низькоентропійних даних оцінка стиснення може давати занижені значення мінімальної ентропії. У той же час тести оцінки прогнозування затримки та оцінки найдовшого повторюваного підрядка дають завищене значення ентропії. Причому ці зміщення в оцінках не зникають зі збільшенням вибірки.

4.4. Загальні рекомендації до оцінки ентропії Шеннона, колізійної мінімальної ентропії

На рис. 5 зведено дані експериментальних досліджень ентропії, що наведені вище на рис. 2 – 4.

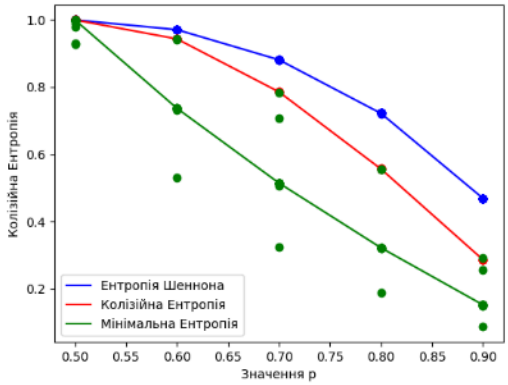

Рис. 5. Експериментальна оцінка ентропії

З отриманих даних випливає, що ентропію Шеннона та колізійну ентропію можливо оцінити доволі точно, у той час як мінімальну ентропію оцінити складніше. Проте, мінімальна ентропія є важливим показником для багатьох практичних застосувань, тому має сенс доповнювати оцінки мінімальної ентропії оцінками ентропії Шеннона та колізійної ентропії, якщо спостерігається розбіжність в отриманих оцінках мінімальної ентропії. Це є актуальним здебільшого для низькоентропійних джерел шуму.

Висновки

1. Ентропія Шеннона та колізійна ентропія можуть бути легше оцінені на практиці, ніж мінімальна ентропія. Проте, мінімальна ентропія має багато важливих застосувань. Тому для більш точних оцінок має сенс комбінувати оцінки мінімальної ентропії та ентропії Шеннона і колізійної ентропії.

2. Окремі статистичні тести NIST SP 800-90B на низькоентропійних даних можуть давати значно занижені або завищені значення. Проте, експериментальні оцінки показують, що зміщене значення не перевищує значення колізійної ентропії.

3. Для оцінки ентропії важливо, щоб кількість статистичного матеріалу була достатньо великою. Це важливо враховувати при тестуванні джерел шуму. Ентропія Реньї є найбільш загальною формою ентропії, проте для практичних застосувань є цікавими її властивості при значеннях параметра $\alpha = 1, 2, \infty$, що відповідають випадкам ентропії Шеннона, колізійної ентропії та мінімальної ентропії відповідно.

4. Існує два основних підходи до оцінки мінімальної ентропії: на основі ентропійної статистики та на основі передбачення. Тести ентропійної статистики призначені для обчислення окремої статистики на вибірках. До методів, що використовують ентропійну статистику, належать: колізійний тест, тест на стиснення, тест Маркова. Такі оцінювачі ентропії припускають, що розподіл ймовірностей описує вихід випадкового джерела шуму, але розподіл ймовірностей невідомий. Метою кожного оцінювача є виявлення інформації про невідомий розподіл на основі статистичних вимірювань.

5. Підхід прогнозування використовує два показники для отримання оцінки. Перший показник базується на глобальній продуктивності предиктора, яка в літературі з машинного навчання називається точністю. По суті, предиктор фіксує частку правильних припущень. Це приблизно вказує на те, наскільки добре можна очікувати, що предиктор вгадає наступний вихід із джерела шуму на основі результатів довгої послідовності припущень. Другий показник базується на найбільшій кількості правильних передбачень у рядку, який називається локальним показником ефективності. Ця метрика корисна для виявлення випадків, коли джерело шуму переходить у дуже передбачуваний стан протягом деякого часу, але предиктор може не працювати добре на довгих послідовностях.

6. Ентропія Шеннона та колізійна ентропія можуть бути легше оцінені на практиці, ніж мінімальна ентропія. Проте, мінімальна ентропія має багато важливих застосувань. Тому для більш точних оцінок має сенс комбінувати оцінки мінімальної ентропії та ентропії Шеннона і колізійної ентропії. Окремі статистичні тести NIST SP 900-80B на низькоентропійних даних можуть давати значно занижені або завищені значення. Проте, експериментальні оцінки показують, що зміщене значення не перевищує значення колізійної ентропії.

7. Головним завданням оцінок як PTRNG, так і NPTRNG є перевірка того, що (середня) ентропія на один біт внутрішнього випадкового числа перевищує задану нижню межу. Важливою відмінністю між PTRNG і NPTRNG, яка впливає на глибину оцінювання, є те, що джерело фізичного шуму в PTRNG знаходиться «під контролем» розробника RNG, а джерело нефізичного шуму NPTRNG зазвичай не може контролюватися.

8. Наразі, у більшості випадків застосування, перевага надається джерелам шуму на основі PTRNG, аніж на основі NPTRNG. По-перше, джерела шуму, які використовують NPTRNG, часто працюють добре лише за певних обставин, а NPTRNG часто не в змозі перевірити, чи виконуються ці умови. По-друге, оцінка ентропії зазвичай базується на складних припущеннях щодо знань і можливостей зловмисника і оперативного середовища.

9. Значення мінімальної ентропії, ентропії Шеннона та колізійної ентропії є основними критеріями якості джерела шуму. Причому, для визначення точних оцінок має сенс використовувати декілька різних оцінок.

10. Генератор ВП ОС Linux для генерації свіжої ентропії використовує події переривань, блочних пристроїв та пристроїв вводу. Основну ентропію містять мітки часу відповідних подій.

11. Набір статистичних тестів NIST STS 800-22 та методика проведення статистичного тестування, що орієнтовані на використання у задачах криптографічного захисту інформації, на даний момент найкраще відповідає потребам усіх сторін.

12. Набір статистичних тестів DIEHARD розроблений для аналізу якості послідовності випадкових чисел. Тести DIEHARD вважаються одним з найжорсткіших існуючих наборів тестів. Набір тестів DIEHARDER дозволяє прийняти однозначне рішення про відмову від «слабкого генератора» (наприклад, на рівні 0,0001 %), а не з ймовірністю відмови 1 або 5 %.

Список літератури:

1. Закон України «Про електронні довірчі послуги» від 1 січня 2024 року N 2155-VIII. [Електронний ресурс]. Режим доступу: https://zakon.rada.gov.ua/laws/show/2155-19#Text.
2. BSI AIS 31. A Proposal for Functionality Classes for Random Number Generators, September 2022. Режим доступу: https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Certification/Interpretations/AIS_31_Functionality_classes_for_random_number_generators_e.pdf?__blob=publicationFile&v=7.
3. NIST SP 800-90 A, Revision 1: E. Barker, J. Kelsey: Recommendation for Random Number Generators Using Deterministic Random Bit Generators. June 2015. Режим доступу: http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-90Ar1.pdf.
4. NIST SP 800-90 B: M. Turan, E. Barker, J. Kelsey, K. McKay, M. Baish, M. Boyle: Recommendation for the Entropy Sources Used for Random Bit Generation. January 2018. Режим доступу: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-90B.pdf.
5. NIST SP 800-90 C, Third Draft: E. Barker, J. Kelsey: Recommendation for Random Bit Generator (RBG) Constructions. September 2022. Режим доступу: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-90C.3pd.pdf.
6. Federal Information Processing Standard (FIPS) 140-2. Security Requirements for Cryptographic Modules 2002. Режим доступу: https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.140-2.pdf.
7. NIST SP 800-22, Revision 1a: A. Rukhin, J. Soto, J. Nechvatal, M. Smid, E. Barker, S. Leigh, M. Levenson, M. Vangel, D. Banks, A. Heckert, J. Dray, S. Vo, (revision) L. Bassham: A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications, April 2010. Режим доступу: https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-22r1a.pdf.
8. DIEHARDER – A testing and benchmarking tool for random number generators. [Електронний ресурс]. Режим доступу: https://manpages.ubuntu.com/manpages/focal/man1/dieharder.1.html.
9. Горбенко Ю. І. Побудування та аналіз систем, протоколів і засобів криптографічного захисту інформації : монографія. Ч. 1: Методи побудування та аналізу, стандартизація та застосування криптографічних систем ; за заг. ред. І. Д. Горбенко. Харків : Форт, 2016. 960 с.
10. Urandom – Linux main page. [Електронний ресурс]. Режим доступу: https://linux.die.net/man/4/urandom.
11. Microsoft Documentation. CryptGenRandom function (wincrypt.h). 2021. [Електронний ресурс]. Режим доступу: https://learn.microsoft.com/en-us/windows/win32/api/wincrypt/nf-wincrypt-cryptgenrandom.
12. ДСТУ ISO/IEC 14888-3:2019 Інформаційні технології. Методи захисту. Цифрові підписи з доповненням. Ч. 3. Механізми на основі дискретного логарифмування (ISO/IEC 14888-3:2018, IDT).
13. ДСТУ 4145-2002 Інформаційні технології. Криптографічний захист інформації. Цифровий підпис, що ґрунтується на еліптичних кривих. Формування та перевіряння.
14. ДСТУ 7564:2014 Інформаційні технології. Криптографічний захист інформації. Функція гешування. З поправкою.
15. ДСТУ 7624:2014 Інформаційні технології. Криптографічний захист інформації. Алгоритм симетричного блокового перетворення. З поправкою.
16. ДСТУ 8845:2019 Інформаційні технології. Криптографічний захист інформації. Алгоритм симетричного потокового перетворення.
17. ДСТУ 8961:2019 Інформаційні технології. Криптографічний захист інформації. Алгоритми асиметричного шифрування та інкапсуляції ключів.
18. ДСТУ 9212:2024 Інформаційні технології. Криптографічний захист інформації. Алгоритм електронного підпису на алгебраїчних решітках із відхилами.
19. Goldreich O. Foundations of Cryptography: Vol. 1. Cambridge University Press, September 2006. [Електронний ресурс]. Режим доступу: https://dl.acm.org/doi/10.5555/1202577.
20. «Меморандум про національну безпеку з просування лідерства США в галузі квантових обчислень при одночасному зниженні ризиків для вразливих криптографічних систем». [Електронний ресурс]. Режим доступу: https://www.whitehouse.gov/briefing-room/statements-releases/2022/05/04/national-security-memorandum-on-promoting-united-states-leadership-in-quantum-computing-while-mitigating-risks-to-vulnerable-cryptographic-systems/.
21. A. Rényi On measures of information and entropy // Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability 1960, p. 547; Probability Theory, North-Holland, Amsterdam, 1970.
22. ISO/IEC 20543 Information technology – Security Techniques. Test and Analysis Methods for Random Bit Generators within ISO/IEC 19790 and ISO/IEC 15408. 2019.
23. Common Methodology for Information Technology Security Evaluation, Evaluation methodology, Version 3.1, Revision 4, September 2012. Режим доступу: https://www.commoncriteriaportal.org/files/ccfiles/CEMV3.1R4.pdf.
24. Common Methodology for Information Technology Security Evaluation, Evaluation Methodology, Version 2.3, August 2005. Режим доступу: https://www.commoncriteriaportal.org/files/ccfiles/cemv2.3.pdf.
25. Rodríguez L., Madarro-Capó E.,Legón-Pérez C., Rojas O., Sosa-Gómez G. Selecting an Effective Entropy Estimator for Short Sequences of Bits and Bytes with Maximum Entropy. [Електронний ресурс]. Режим доступу:

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8147137/.
26. Skorski M. Improved Estimation of Collision Entropy in High and Low-Entropy Regimes and Applications to Anomaly Detection. Режим доступу: https://eprint.iacr.org/2016/1035.pdf.
27. Paninski L. Estimation of entropy and mutual information. Neural Comput. 2003;15:1191–1253. Режим доступу: doi: 10.1162/089976603321780272.
28. Trybula S. Some problems of simultaneous minimax estimation. Ann. Math. Stat. 1958;29:245–253. Режим доступу: doi: 10.1214/aoms/1177706722.
29. Hausser J., Strimmer K. Entropy inference and the james-stein estimator, with application to nonlinear gene association networks // J. Mach. Learn. Res. 2009;10:1469–1484.
30. Valiant G., Valiant P. Estimating the unseen: Improved estimators for entropy and other properties // J. ACM. 2017; 64:1–41. Режим доступу: doi: 10.1145/3125643.
31. Schürmann T. Bias analysis in entropy estimation // J. Phys. A. Math. Gen. 2004;37:L295. Режим доступу: doi: 10.1088/0305-4470/37/27/L02.
32. Chao A., Shen T.J. Nonparametric estimation of Shannon's index of diversity when there are unseen species in sample // Environ. Ecol. Stat. 2003;10:429–443. Режим доступу: doi: 10.1023/A:1026096204727.
33. Hausser J., Strimmer K. Entropy inference and the james-stein estimator, with application to nonlinear gene association networks // J. Mach. Learn. Res. 2009; 10:1469–148.
34. Skorski M. Improved Estimation of Collision Entropy in High and Low-Entropy Regimes and Applications to Anomaly Detection. Режим доступу: https://eprint.iacr.org/2016/1035.pdf.
35. P. Hagerty and T. Draper, Entropy Bounds and Statistical Tests, NIST Random Bit Generation Workshop, December 2012. Режим доступу: https://csrc.nist.gov/csrc/media/events/random-bit-generation-workshop-2012/documents/hagerty_entropy_paper.pdf.
36. J. Kelsey, Kerry A. McKay, M. Sonmez Turan, Predictive Models for Min-Entropy Estimation // Proceedings of the Workshop on Cryptographic Hardware and Embedded Systems 2015 (CHES 2015), France. Режим доступу: https://doi.org/10.1007/978-3-662-48324-4_19.
37. U. Maurer, A Universal Statistical Test for Random Bit Generators // Journal of Cryptology. 1992. Vol. 5, No. 2. P. 89–105.
38. C.E Shannon, Prediction and Entropy of Printed English // Bell System Technical Journal. Vol. 30, January 1951. [Електронний ресурс]. Режим доступу: https://archive.org/details/bstj30-1-50.
39. Quantis QRNG USB. [Електронний ресурс]. Режим доступу: https://www.idquantique.com/random-number-generation/products/quantis-random-number-generator/.

Надійшла до редколегії 15.01.2024

Відомості про авторів:
Голубничий Дмитро Юрійович – канд. техн. наук, доцент, АТ "Інститут Інформаційних Технологій", начальник наукового відділу; Україна; e-mail: goldim1971@gmail.com; ORCID: https://orcid.org/0000-0002-6873-7004
Кандій Сергій Олегович – Харківський національний університет імені В. Н. Каразіна, аспірант кафедри безпеки інформаційних систем і технологій, факультет комп'ютерних наук; АТ "Інститут Інформаційних Технологій", науковий співробітник-консультант; Україна; e-mail: sergeykandy@gmail.com
Єсіна Марина Віталіївна – канд. техн. наук, доцент, Харківський національний університет імені В. Н. Каразіна, доцент кафедри безпеки інформаційних систем і технологій, факультет комп'ютерних наук, АТ «Інститут Інформаційних Технологій», науковий співробітник-консультант; Україна; e-mail: m.v.yesina@karazin.ua; ORCID: https://orcid.org/0000-0002-1252-7606
Горбенко Дмитро Юрійович – Харківського національного університету імені В. Н. Каразіна. студент факультету комп'ютерних наук, АТ "Інститут Інформаційних Технологій", молодший інженер-програміст; Україна; e-mail: jsciitua@gmail.com