УДК 004.056.55: 519.248

Савчук М.М.[1], д.ф.-м.н., доц.
Бурлака М.К.[2], студ.

Кодування і класифікація перестановок за спеціальним перетворенням з оцінками потужності класів

[1, 2] Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», 03056, м.Київ, Проспект Перемоги, будинок 37,
e-mail: mikhail.savchuk@gmail.com , maria.k.burlaka@gmail.com

M.M. Savchuk[1], Dr. Sci., docent
M.K. Burlaka[2], stud.

Encoding and classification of permutations by special conversion with estimates of class power

[1, 2] National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, 03056, Kyiv, Peremoga avenue, building 37,
e-mail: mikhail.savchuk@gmail.com , maria.k.burlaka@gmail.com

Вводиться спеціальне S-перетворення на множині перестановок та визначаються характеристики перестановки за цим перетворенням. Запропонована класифікація та кодування перестановок за класами еквівалентності та їх характеристиками відносно S-перетворення. Отримано точні значення числа перестановок в класах для деяких розмірів перестановки та оцінки потужності класів з різними характеристиками методом статистичного моделювання. Отримано інтервальні оцінки для потужності класів з різними характеристиками для перестановок порядку n = 11, 26, 30, 31, 32, 33, 45, 55. Запропоновано статистичний критерій згоди, який використовує характеристики перестановок за S-перетворенням, для перевірки генераторів випадкових перестановок та підстановок. Обговорюється застосування результатів в криптографії.

Ключові слова: *перестановки, підстановки, класифікація, статистичне моделювання.*

Scientific articles investigating properties and estimates of the number of so-called complete permutations are surveyed and analyzed. The paper introduces a special S-transform on the set of permutations and determines the permutation properties according to this transform. Classification and coding of permutations by equivalence classes according to their properties with respect to S-transformation is proposed. This classification and permutation properties, in particular, generalize known results for complete permutations regarding determining certain cryptographic properties of substitutions that affect the cryptographic transformations security. The exact values of the number of permutations in equivalence classes for certain permutation sizes are calculated and the estimates of the cardinality of classes with various properties are constructed by statistical modeling. The complete list of permutation classes with the exact values of their sizes for permutations of order n = 11 is presented. The interval estimates for the size of classes with various characteristics for permutations of order n = 11, 26, 30, 31, 32, 33, 45, 55 are obtained. Monte Carlo estimates and bounds of confidence intervals used the approximation of the binomial distribution by the normal and Poisson distributions, as well as the Python programming language package Scipy. Statistical tables have been calculated that can be used for further conclusions and estimates. The classification of permutations by their properties with respect to the introduced transform can be used in constructing high-quality cryptographic transformations and transformations with special features. The classes of complete permutations with their properties are selected as the best for rotary cryptosystems applications. The obtained results can be used, in particular, to search for permutations with certain characteristics and properties, to find the probability that the characteristic of the generated permutation belongs to a collection of given characteristics, to estimate the complexity of finding permutations with certain properties. A statistical criterion of consent, which uses the characteristics of permutations by S-transformation to test the generators of random permutations and substitutions is proposed.

Key Words: permutations, substitutions, classification, statistical modeling.

Статтю представив д.ф.-м.н., проф. Анісімов А.В.

Вступ

У статті досліджуються спеціальні властивості перестановок та підстановок, які важливі для побудови стійких криптографічних перетворень в системах захисту інформации. Вводиться спеціальне перетворення на множині перестановок та визначаються характеристики перестановки за цим перетворенням. Запропонована класифікація перестановок за їх характеристиками. Отримано точні значення перебором числа перестановок в таких класах для деяких розмірів перестановки та оцінки потужності класів з різними характеристиками методом статистичного моделювання. Введена характеристика перестановки дасть можливість, зокрема, оцінити якість підстановок в ключах електромеханічних роторних шифрувальних машин. Експериментально отримано статистичні оцінки потужностей класів для перестановок різного розміру, з різними характеристиками, проведено порівняння точності оцінок при використанні різних апроксимацій для ймовірнісних розподілів при застосуванні метода Монте-Карло. Наведено алгоритми для побудови класів перестановок за характеристиками методом повного перебору перестановок та методом статистичного моделювання. Виділено перестановки деяких довжин: 11, 26, 30, 31, 32, 33, 45 та 55, для яких побудовано довірчі інтервали для потужностей класів з різними характеристиками. Розраховано статистичні таблиці, які можна використовувати для подальших висновків і оцінок. Запропоновано статистичний критерій згоди для перевірки якості генераторів випадкових перестановок та підстановок, який використовує характеристики перестановок за введеним перетворенням та класифікацію перестановок за їх характеристиками.

1 Повні перестановки та оцінки їх кількості

Означення 1. Перестановка $\pi = (i_0, i_1, \dots, i_{n-1})$ елементів множини G={0, 1,…, n − 1} називається повною перестановкою порядку n, якщо числа $j_k = (i_k + k) \bmod n,\ k = \overline{0, n-1},$ попарно різні і утворюють перестановку $\pi_0 = (j_0, j_1, \dots, j_{n-1})$ елементів множини G [1].

Кожна повна перестановка π визначає бієкцію $\varphi: G \to G,\ \varphi(k) = j_k,\ k = \overline{0, n-1}$, як і кожна бієкція повну перестановку. Дослідники повних перестановок [1-4] називають повні перестановки також повними відображеннями, повними підстановками або «хорошими» перестановками.

Відомо, що якщо n парне, то кількість повних перестановок дорівнює нулю (наприклад, див. [1]). Для n непарних кількість повних перестановок швидко зростає з ростом *n.* Задача знайти число повних перестановок відома ще з часів панування серед засобів криптографічного захисту електромеханічних роторних шифрувальних машин в XX столітті. Кожен ротор (диск) такої машини реалізовував частину ключа - бієктивну підстановку на алфавіті, яка виконувалася з'єднанням електричних контактів (які відповідають літерам алфавіту) з одного боку диску з такими ж контактами з другого боку диску дротами. Якщо підстановка не була повною у сенсі означення 1, то серед цих з'єднань існували паралельні дроти. Повні перестановки тому називали також перестановками без паралельних перепайок. Якщо існували паралельні перепайки, то це давало можливість підвищити ефективність криптоаналізу. Тобто повні перестановку були кращими с точки зору стійкості і задача визначення числа повних перестановок і алгоритмів їх побудови була практично важливою. Криптографічні властивості перестановок і підстановок важливі і сьогодні, оскільки ці перетворення є основними в сучасних алгоритмах шифрування.

Далі будемо розглядати повні перестановки для непарного порядку n. У роботі [1] отримано верхню границю числа перестановок без паралельних перепайок: ймовірність того, що випадкова перестановка виявиться повною перестановкою $P(n)$ не перевищує e^{-cn} при достатньо великих непарних *n*, де $c \geq 0{,}08854$. Таке значення було отримано ймовірнісно-комбінаторними методами. З теорії стохастичних процесів отримано більш слабку оцінку $c \geq 0{,}06766$. Автори відзначили, що можливо буде вірна асимптотична границя $P(n) \leq e^{-(1-\varepsilon)n}$, для будь-якого $\varepsilon > 0$.

У статті [2] оцінку величини *c* було покращено: доведено, що $c \geq \frac{1}{2}\ln 2 \approx 0{,}35$. Це означає, що загальна кількість перестановок без паралельних перепайок не перевищує $n!\, 2^{-m+O(1/\sqrt{\ln m})}$, де $m = (n+1)/2$ - ціле число.

У [3] вводиться таке поняття, як «k-good permutation» (k-хороша перестановка): перестановка є k-хорошою, якщо на перших k

позиціях числа $j_l = (i_l + l) \bmod n$, $l = \overline{0,k}$, усі різні.

Очевидно, що перестановка не може бути повною, якщо усі її часткові представлення не k-хороші для усіх k. В [3] приводяться такі результати. Для n = 25 кількість 10-хороших перестановок починаючи з нуля була знайдена перебором. Виходячи з цього знайдено, що ймовірність того, що загальна перестановка довжини 25 є повною перестановкою оцінюється щонайменше числом $1.86785 \cdot 10^{-9}$ [3]. Позначивши $\pi(n) = e^{-c_n n}$, автори отримують деяку оцінку для c_n при n= 25. Відповідне значення для n= 25 оцінюється як $c_{25} \leq 0.8039$. Приводиться і більш точне значення, базоване на середній кількості повних перестановок на деяких проміжках, $c_{25} = 0{,}789$. У статті приводиться також таблиця точних значень $P(n)$ та c_n від 1 до 19. Також були приведені оцінки $P(n)$ та c_n для n = 25, 35, 45, 55 ([3, табл. 3]). У [3] приведений аналіз повних перестановок (без паралельних перепайок) для застосування у криптографії, зокрема у роторних шифрувальних машинах.

У [4] для оцінки кількості перестановок без паралельних перепайок пропонується використовувати метод пришвидшеного моделювання. Не зважаючи на те, що метод є досить простим для реалізації, він дозволяє при відносно невеликих витратах часу побудувати незміщені оцінки і відповідні довірчі інтервали для $P(n)$ при досить великих n (у статті [4] приводяться оцінки для $P(n)$ до n = 155 включно. Більш того, даний алгоритм дозволяє у практичних підрахунках підтвердити співвідношення $P(n) \sim ae^{-cn}$ і вказати більш точні границі для c , а саме у [4] приводяться така оцінка: $0.9825 \leq c \leq 0.9883$. Загалом, для $n \geq 75$ для $P(n)$ були знайдені та указані наступні межі: $413.099\exp\{-0.9883\, n\} \leq P(n) \leq 267.384\exp\{-0.9825\, n\}$, а наведені чисельні дані для деяких n свідчать про високу степінь точності нижніх та верхніх оцінок і можливість з їх допомогою прогнозувати значення $P(n)$.

2 Характеристики перестановок за спеціальним перетворенням, розбиття перестановок на класи еквівалентності

Визначимо поняття *S*-перетворення і характеристики перестановки відносно *S*-перетворення, яка узагальнює поняття повної перестановки. Можливості та рівень застосовності перестановки, зокрема для роторних шифрів, залежить від характеристики даної перестановки.

Розглянемо довільну перестановку $\pi = (i_0, i_1, \dots, i_{n-1})$ порядку n елементів множини G={0, 1,…, n − 1} та вектор (кортеж) чисел

$$\theta(\pi) = (j_0, j_1, \dots, j_{n-1}), \qquad (1)$$

де числа $j_k = (i_k + k) \bmod n$, $k = \overline{0, n-1}$, належать множині G.

Числа в векторі (1) не обов'язково усі різні. Кожен кортеж чисел $\theta(\pi)$ однозначно визначає деяке відображення $\varphi: G \to G$, $\varphi(k) = j_k$, $k = \overline{0, n-1}$, не обов'язково бієктивне.

Кортеж $\theta(\pi)$ визначає мультимножину $G(\theta) = \{j_0, j_1, \dots, j_{n-1}\}$. Очевидно, одна мультимножина може відповідати різним кортежам. Запишемо цю мультимножину у вигляді

$$G(\theta) = \{0^{k_0}, 1^{k_1}, 2^{k_2}, \dots (n-1)^{k_{n-1}}\}, \qquad (2)$$

де k_0 - кількість нулів серед $\{j_0, j_1, \dots, j_{n-1}\}$, k_1 - кількість одиниць,…, k_{n-1} - кількість чисел з $G(\theta)$, які рівні n–1; при цьому $\sum_{i=0}^{n-1} k_i = n$. Якщо $k_i = 0$, то число i в $G(\pi)$ не зустрічається. За термінологією В.Н. Сачкова [5] кортеж $[k_0, k_1, k_2, \dots k_{n-1}]$ – показники первинної специфікації мультимножини (2) $G(\theta)$. Розглянемо тепер $\{k_i,\ i = \overline{0, n-1}\}$ як мультимножину

$$\{0^{\alpha_0}, 1^{\alpha_1}, 2^{\alpha_2}, \dots (n-1)^{\alpha_{n-1}}, n^{\alpha_n}\}$$

та кортеж показників її первинної специфікації

$$[\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_n], \qquad (3)$$

де $\sum_{i=0}^{n} i\alpha_i = n$; за термінологією В.Н. Сачкова $[\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_n]$ – показники вторинної специфікації мультимножини $G(\theta)$, яка відповідає довільній перестановці $\pi = (i_0, i_1, \dots, i_{n-1})$. Позначимо S_n множину усіх перестановок порядку n.

Означення 2. Перетворення довільної перестановки $\pi = (i_0, i_1, \dots, i_{n-1})$ відповідно правилам (1-3) будемо називати *S*-перетворенням, яке задає відображення $\chi(\pi): S_n \to \{0,1,2, \dots, n\}^{n+1}$.

Означення 3. Характеристикою або кодом перестановки $\pi = (i_0, i_1, \dots, i_{n-1})$ відносно *S*-перетворення будемо називати значення $\chi(\pi)$, тобто вектор (кортеж) довжини $n + 1$

$$\chi(\pi) = (\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_n) \qquad (4)$$

Таким чином, для побудови характеристики (коду) $\chi(\pi)$ довільної перестановки $\pi = (i\ , i_1, \dots, i_{n-1})$ порядку n елементів множини G={0, 1,…, $n-1$} відносно

S-перетворення знаходимо кортеж чисел $\theta(\pi) = (j\ , j_1, \dots, j_{n-1})$, де числа $j_k = (i_k + k) \bmod n$, $j_k \in G$, $k = \overline{0, n-1}$, мультимножину

$$G(\theta) = \{0^{k_0}, 1^{k_1}, 2^{k_2}, \dots (n-1)^{k_{n-1}}\},$$

її первинну специфікацію (2) і показники вторинної специфікації (3).

Наприклад, для перестановки порядку $n = 7$ $\pi = (3,1,4,2,5,0,6)$ знаходимо:

$$\theta(\pi) = (3,2,6,5,2,5,5),$$
$$G(\theta) = \{0\ , 1\ , 2^2, 3^1, 4\ , 5^3, 6^1\},$$
$$\chi(\pi) = (3,2,1,1,0,0,0,0).$$

Характеристика повної перестановки порядку n має вигляд $\chi(\pi) = (0, n, 0,0, \dots, 0)$, тобто кожне число з множини G={0, 1,…, $n-1$} в векторі $\theta(\pi)$ зустрічається по одному разу. Характеристика перестановки π порядку n, у якої вектор $\theta(\pi) = (j, j, \dots, j)$ з однаковими координатами має характеристику $\chi(\pi) = (n-1, 0, \dots, 0, 1)$. Наприклад, для перестановки $\pi = (6,5,4,3,2,1,0)$ та усіх 7-ми її циклічних зсувів $\chi(\pi) = (6,0,0,0,0,0,0,1)$.

Якщо характеристики кількох перестановок співпадають, то можна очікувати, що ці перестановки з точки зору криптоаналізу роторних шифраторів мають близькі властивості, а отже криптоаналіз систем, побудованих на них, має складність одного порядку. Очевидно, що чим більше значення другої позиції характеристики перестановки, тобто чим більше різних чисел у векторі $\theta(\pi)$, тим краще ця перестановка для використання у якості підключа шифрування у роторних системах.

Будь-яка характеристика $\chi(\pi)$ виділяє підмножину $S(n, \chi(\pi)) \subset S_n$ перестановок з однаковою характеристикою – клас еквівалентності, а π представник цього класу. Очевидно, для різних характеристик класи еквівалентності не перетинаються і тоді вся множина S_n розбивається на класи еквівалентності відповідно до їх характеристик. Повні перестановки будуть одним з класів еквівалентності.

Позначимо:

$S_n(\alpha\ , \alpha_1, \alpha_2, \dots, \alpha_n)$ - число перестановок π порядку n з характеристикою $\chi(\pi) = (\alpha\ , \alpha_1, \alpha_2, \dots, \alpha_n)$, скорочено $S_n(\chi(\pi))$,

$S_n^p = S_n(0, n, 0,0, \dots, 0)$ - число повних перестановок порядку n,

$P(n)$ - ймовірність випадкового вибору повної перестановки порядку n,

$P_n(\chi(\pi)) = P_n(\alpha\ , \alpha_1, \alpha_2, \dots, \alpha_n)$ ймовірність випадкового вибору перестановки порядку n з характеристикою $\chi(\pi) = (\alpha\ , \alpha_1, \alpha_2, \dots, \alpha_n)$.

Під випадковим вибором розуміється рівноймовірний вибір перестановки (з ймовірністю $1 / n!$) з множини S_n.

Потужності класів еквівалентності з характеристикою $\chi(\pi) = (\alpha\ , \alpha_1, \alpha_2, \dots, \alpha_n)$ задовольняють співвідношенням

$$S_n(\alpha\ , \alpha_1, \alpha_2, \dots, \alpha_n) = n!\, P_n(\alpha\ , \alpha_1, \alpha_2, \dots, \alpha_n),$$
$$S_n^p = n!\, P(n) \quad (5)$$

3 Класи повних та вироджених перестановок

Характеристика повної перестановки порядку n має вигляд $\chi(\pi) = (0, n, 0,0, \dots, 0)$. Відомі оцінки класу повних перестановок наведено у розділі 1. Для непарного n від 1 до 15 повним перебором точне число повних перестановок розраховано ще в 70-х роках минулого століття в Інституті кібернетики (Київ) на БЭСМ-6 (див. таблицю 1).

Таблиця 1

Число повних перестановок та ймовірність випадкового вибору повної перестановки

n	$n!$	S_n^p	$P(n)$
3	6	1×3=3	0,5
5	120	3×5=15	0,125
7	540	19×7=133	0,0263889
9	362 880	224×9=2 016	0,00555555
11	39 916 800	3441×11= 37 851	0,000948247
13	6 227 020 800	79259×13= 1 030 367	0,0001654670
15	1 307 674 368 000	2424195×15 =36 362 925	0,0000278073

З кожної перестановки порядку n з класу повних перестановок можна отримати n повних перестановок (включно з першою) циклічними зсувами. Повні перестановки, що відрізняються циклічними зсувами, утворюють підкласи еквівалентності відносно зсуву по n елементів в

кожному підкласі, що відмічено в третьому стовбці таблиці 1. Число підкласів рівно S_n^p / n.

Для $n = 17$ і 19 точні значення S_n^p обчислені в 90-х роках перебором англійськими дослідниками, швидкий розвиток обчислювальної техніки дав можливість знайти точні значення S_n^p для n =21, 23 і 25. Але безпосереднє обчислення для більших n вже недоступно сучасній обчислювальній техніці, бо на декілька порядків складніше. Для багатьох значень n>19 теоретично отримано границі зверху, наближені оцінки для числа повних перестановок та оцінки методом статистичного моделювання і прискореного статистичного моделювання [1-4]. Деякі результати робіт [1-4] наведено в розділі 1.

Означення 4. Перестановку π порядку n будемо називати виродженою, якщо її характеристика $\chi(\pi) = (n-1, 0, \dots, 0, 1)$.

У виродженої перестановки $\theta(\pi) = (j, j, \dots, j)$, тобто усі лишки суми $j_k = (i_k + k) \bmod n = j$ однакові. Використання такої перестановки в системах шифрування значно спрощує криптоаналіз роторного шифру.

Очевидно, число перестановок із виродженою характеристикою дорівнює n і їх частка серед усіх можливих перестановок множини дуже мала, наприклад,

$$S_{11}\big((10,0,\dots,0,1)\big) = 11,$$
$$P_{11}((10,0,\dots,0,1)) = 0{,}000000275573192$$

Згідно з таблицями 1 та 2 повних перестановок (без паралельних перепайок) $S_{11}(0,11,0,0,\dots,0) = 37851$, що становить 0,000948247354497 від усіх можливих перестановок. На відміну від повних перестановок, які є найбільш вдалими для застосування у роторних шифраторах, перестановки з виродженим кодом χ =(0, 11, 0, 0, 0, 0, 0, 0, 0, 0, 0) є криптографічно неякісними (останній клас в таблиці 2).

4 Оцінки потужності класів перестановок з різними характеистиками

Для $n = 11$ множина S_{11} усіх перестановок порядку 11 розбивається на 50 класів еквівалентності за *S*-перетворенням згідно з своїми характеристиками. Точне значення потужності кожного класу знайдено повним перебором. Список потужностей класів та їх характеристик (кодів) в порядку спадання потужності наведено в таблиці 2.

Таблиця 2

Класи еквівалентності для n = 11

Клас №	Характеристика класу $\chi(\pi) = (\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_n)$	Потужність $S_n(\chi(\pi))$- точне значення
1	(4, 4, 2, 1, 0, 0, 0, 0, 0, 0, 0)	8252200
2	(3, 5, 3, 0, 0, 0, 0, 0, 0, 0, 0)	4783130
3	(4, 3, 4, 0, 0, 0, 0, 0, 0, 0, 0)	3746765
4	(3, 6, 1, 1, 0, 0, 0, 0, 0, 0, 0)	3522310
5	(5, 2, 3, 1, 0, 0, 0, 0, 0, 0, 0)	3274260
6	(2, 7, 2, 0, 0, 0, 0, 0, 0, 0, 0)	2518615
7	(5, 3, 1, 2, 0, 0, 0, 0, 0, 0, 0)	2390960
8	(4, 5, 1, 0, 1, 0, 0, 0, 0, 0, 0)	2024330
9	(5, 3, 2, 0, 1, 0, 0, 0, 0, 0, 0)	2017070
10	(4, 5, 0, 2, 0, 0, 0, 0, 0, 0, 0)	1202135
11	(6, 2, 1, 1, 1, 0, 0, 0, 0, 0, 0)	869990
12	(5, 4, 0, 1, 1, 0, 0, 0, 0, 0, 0)	744150
13	(6, 1, 2, 2, 0, 0, 0, 0, 0, 0, 0)	722370
14	(5, 4, 1, 0, 0, 1, 0, 0, 0, 0, 0)	559020
15	(2, 8, 0, 1, 0, 0, 0, 0, 0, 0, 0)	464035
16	(5, 1, 5, 0, 0, 0, 0, 0, 0, 0, 0)	445280
17	(6, 1, 3, 0, 1, 0, 0, 0, 0, 0, 0)	395670
18	(6, 2, 2, 0, 0, 1, 0, 0, 0, 0, 0)	335775
19	(3, 7, 0, 0, 1, 0, 0, 0, 0, 0, 0)	307340
20	(4, 6, 0, 0, 0, 1, 0, 0, 0, 0, 0)	192390
21	(6, 2, 0, 3, 0, 0, 0, 0, 0, 0, 0)	182710
22	(6, 0, 4, 1, 0, 0, 0, 0, 0, 0, 0)	166375
23	(6, 3, 0, 1, 0, 1, 0, 0, 0, 0, 0)	152460
24	(6, 3, 1, 0, 0, 0, 1, 0, 0, 0, 0)	119790
25	(6, 3, 0, 0, 2, 0, 0, 0, 0, 0, 0)	82280
26	(7, 1, 1, 1, 0, 1, 0, 0, 0, 0, 0)	65340
27	(5, 5, 0, 0, 0, 0, 1, 0, 0, 0, 0)	59048
28	(7, 0, 2, 1, 1, 0, 0, 0, 0, 0, 0)	58080
29	(7, 1, 0, 2, 1, 0, 0, 0, 0, 0, 0)	41140
30	(7, 1, 1, 0, 2, 0, 0, 0, 0, 0, 0)	38720
31	**(0, 11, 0, 0, 0, 0, 0, 0, 0, 0, 0)**	**37851**
32	(7, 1, 2, 0, 0, 0, 1, 0, 0, 0, 0)	22990
33	(7, 0, 1, 3, 0, 0, 0, 0, 0, 0, 0)	19360
34	(7, 2, 0, 0, 1, 1, 0, 0, 0, 0, 0)	19360
35	(7, 2, 0, 1, 0, 0, 1, 0, 0, 0, 0)	18150
36	(6, 4, 0, 0, 0, 0, 0, 1, 0, 0, 0)	18150
37	(7, 0, 3, 0, 0, 1, 0, 0, 0, 0, 0)	13310
38	(7, 2, 1, 0, 0, 0, 0, 1, 0, 0, 0)	10890
39	(8, 0, 0, 2, 0, 1, 0, 0, 0, 0, 0)	4235
40	(8, 0, 1, 0, 1, 1, 0, 0, 0, 0, 0)	3630
41	(8, 0, 0, 1, 2, 0, 0, 0, 0, 0, 0)	3025
42	(8, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0)	2420
43	(8, 0, 2, 0, 0, 0, 0, 1, 0, 0, 0)	2420
44	(7, 3, 0, 0, 0, 0, 0, 0, 1, 0, 0)	2420
45	(8, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0)	1210
46	(8, 1, 0, 0, 1, 0, 1, 0, 0, 0, 0)	1210
47	(8, 1, 1, 0, 0, 0, 0, 0, 1, 0, 0)	1210
48	(8, 2, 0, 0, 0, 0, 0, 0, 0, 1, 0)	605
49	(8, 1, 0, 0, 0, 2, 0, 0, 0, 0, 0)	605
50	**(10, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,1)**	**11**

В роботі реалізовано програмно алгоритми статистичного моделювання з випадковим генеруванням перестановки для отримання інтервальних оцінок потужностей класів з різними характеристиками з заданою довірчою ймовірністю. Під випадковим вибором перестановки розуміється рівноймовірний вибір перестановки (з ймовірністю $1 / n!$) з множини S_n. Проведено низку машинних експериментів для побудови верхніх і нижніх границь ймовірностей $P_n(\chi(\pi))$ з заданою довірчою

ймовірністю. Верхні та нижні границі потужності певних класів перестановок знаходяться за формулами (5). Для перевірки і налагодження програм використано точні дані з таблиць 1 і 2 та відомі оцінки статистичного моделюванням числа S_n^p та ймовірностей $P(n)$ для повних перестановок. Наведемо опис одного з таких алгоритмів.

Алгоритм 1. Алгоритм оцінки потужності класів еквівалентності методом статистичного моделювання.

1. Встановлюємо кількість експериментів N - велике число.

2. Будуємо множину G={0, 1, . . . , $n-1$}.

3. Будуємо випадкову перестановку $\pi = (i_0, i_1, \dots, i_{n-1})$ множини G={0, 1, . . . , $n-1$},

використовуючи якісний датчик випадкових чисел і $n-1$ випадкову інверсію [6].

4. Для кожної перестановки $\pi = (i_0, i_1, \dots, i_{n-1})$ отримуємо послідовність $\theta(\pi) = (j_0, j_1, \dots, j_{n-1})$ таку, що $j_k = (i_k + k) \bmod n, \quad k = \overline{0, n-1},$ та належать множині G.

5. За послідовністю $\theta(\pi) = (j_0, j_1, \dots, j_{n-1})$ находимо характеристику $\chi(\pi) = (\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_n)$.

6. Якщо характеристика $\chi(\pi) = (\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_n)$ не зустрічалася на попередніх ітераціях, встановлюємо лічильник $C(\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_n)$=1. Якщо така характеристика зустрічалася раніше, то збільшуємо лічильник $C(\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_n)$ (що вже існує) на 1.

7. Зменшуємо N на 1.

8. Повторюємо кроки 3) - 7) доти, поки $N \neq 0$. Коли $N = 0$, алгоритм закінчує роботу.

По закінченні роботи алгоритму 1 обчислюємо ймовірності $P_n(\chi(\pi))$ для кожного класу за формулою

$$\frac{1}{N} C(\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_n).$$

Точкову оцінку для потужності класу з кодом $\chi(\pi)$ отримаємо за формулою (5). Якщо після закінчення алгоритму не з'явилося ні однієї перестановки з деякою характеристикою $\chi(\pi)$, вважаємо, що число $S_n(\chi(\pi))$=0.

Для знаходження довірчих інтервалів для $P_n(\chi(\pi))$ - параметра в біноміальній моделі з числом випробувань N використовувалася також апроксимація біноміального розподілу нормальним розподілом, розподілом Пуассона та відповідні формули з робіт [7, 8], а також пакет Scipy мови програмування Python.

Наведемо вибірково для прикладу довірчі інтервали для потужності декількох великих класів для перестановок довжин n = 26 та n = 33 (таблиці 3 та 4 відповідно). Зауважимо, що потужності $|S_{26}|$ =26! =4.0329 · 10^{26} , а $|S_{33}|$ = 33!= 8.6833 · 10^{36}. Дані статистичного моделювання в таблицях 3 та 4 отримано за генеруванням N=10^7 випадкових перестановок та рівнем значущості $\alpha = 0{,}95$.

Таблиця 3

Характеристики перестановок порядку 26 та довірчі інтервали для потужності відповідних класів з рівнем значущості $\alpha = 0{,}95$.

Характеристика $(\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_{27})$	Довірчі інтервали
(9,10,5,2,0,...,0)	$[2.723086 \cdot 20^{25}, 73363 \cdot 20^{25}]$
(7,14,3,2,0,0,...,0)	$[6.06225 \cdot 20^{24}, 6.11344 \cdot 20^{24}]$
(3,20,3,0,0,0,...,0)	$[1.40603 \cdot 20^{21}, 1.01765 \cdot 20^{22}]$
(17,1,5,1,0,1,0,1,0,...0)	$[2.06861 \cdot 20^{18}, 1.91316 \cdot 20^{20}]$

Таблиця 4

Характеристики перестановок порядку 33 та довірчі інтервали для потужності відповідних класів з рівнем значущості $\alpha = 0{,}95$.

Характеристика $(\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_{34})$	Довірчі інтервали
(11,13,7,2,0,...,0)	$[3.96965 \cdot 10^{35}, 3.98856 \cdot 10^{35}]$
(10,17,3,2,1,0,...,0)	$[4.56102 \cdot 10^{34}, 4.62664 \cdot 10^{35}]$
(9,19, 2, 2, 1, 0,…, 0)	$[7.470017 \cdot 10^{33}, 7.73807 \cdot 10^{33}]$
(9,21,0,2,0,0,1,0,...,0)	$[1.71074 \cdot 10^{30}, 9.12880 \cdot 10^{30}]$

При генеруванні в методі Монте-Карло 10^7 випадкових перестановок всього різних класів з ненульовою точковою оцінкою потужності для перестановок довжини n=26 отримано 688, для n=30 - 976, для n=31 - 1065, для n=32 - 1153, для n=33 - 1229, для n=45 - 2508, для n=55 - 3906. Роторні криптографічні системи та їх аналіз детально описано в монографії [9].

5 Статистичний критерій перевірки генератора випадкових перестановок та підстановок

Розбиття множини S_n перестановок на класи еквівалентності за *S*-перетворенням та дані про потужності класів з різними характеристиками дає можливість запропонувати статистичний критерій згоди для перевірки генераторів випадкових перестановок та підстановок.

Опишемо коротко критерій хі-квадрат. Нехай H_0- гіпотеза, яка полягає в тому, що перестановка π генерується з рівноймовірним розподілом, n – кількість експериментів (розмір виборки), $v=(v_1,...,v_N)$ - вектор частот влучення вибіркових точок у відповідні інтервали групування перестановок $E_1,...,E_N$, $v_1+...+v_N=n$, кожен з яких є класом еквівалентності або об'єднанням декількох класів,

$p^0=(p_1^0,...,p_N^0)$, $p_j^0=P(\pi\in E_j\,|\,H_0), j=1,....,N$.

У цьому випадку розподіл вектора частот v при умові гіпотези H_0 буде поліноміальним з параметрами n і p^0. Обраховується статистика хі-квадрат, що характеризує відхилення вибіркових даних (тобто частот v_j) від очікуваних середніх значень $np_j^0, j=1,....,N$:

$$X_n^2=X_n^2(v)=\sum_{j=1}^{N}(v_j-np_j^0)^2/(np_j^0), \qquad (6)$$

Для великих обсягів вибірок n статистика X_n^2 має при умові гіпотези H_0 граничний розподіл, що не залежить від чисел $p_j^0, j=1,2,...,N$. Має місце слабка збіжність розподілу статистики X_n^2

$$L(X_n^2\,|\,H_0)\to\chi^2(N-1),\ n\to\infty$$

де $\chi^2(N-1)$ — хі-квадрат розподіл зі $N-1$ степенем свободи, який затабульовано, а при великих N апроксимується за допомогою нормального розподілу.

Наприклад, для застосування цього критерію виокремимо $N-1$ найбільш ймовірних (з більшими $P_m(\chi(\pi))$ характеристик $\chi(\pi)$, класи еквівалентності за цими характеристиками позначимо $E_1,E_2,..E_{N-1}$ в порядку спадання ймовірностей, які позначимо $p_j^0,\ j=1,2,..,N-1$, і вважаємо їх інтервалами групування, а в інтервалом групування E_N вважаємо об'єднання усіх інших класів еквівалентності з меншими ймовірностями $P_m(\chi(\pi))$, сумарну ймовірність яких позначимо p_N^0.

Задаємся рівнем значущості, наприклад, $\alpha=0{,}95$. Знаходимо критичну область у вигляді $\Im_{1\alpha}=\{t\ge t_\alpha\}$, де t_α квантиль розподілу хі-квадрат або його нормальної апроксимації рівня α [7, 8].

Після проведення експерименту з n незалежно генерованими перестановками і підрахунку їх характеристик за *S*-перетворенням отримуємо вектор частот $(\nu_1,\nu_2,...\nu_N)$, ν_j - число перестановок, які потрапили після випадкового незалежного вибору n перестановок в інтервал групування $E_j,\ j=1,2,..,N$. Далі будуємо статистику хі-квадрат за формулою (6). Якщо значення статистики належить критичній області, то гіпотеза H_0 відхиляється. Якщо не належить, то гіпотеза не суперечить результатам експерименту.

Запропонований статистичний критерій згоди може краще відсіювати специфічні альтернативні гіпотези у деяких генераторах випадкових підстановок та перестановок.

Висновки

Вводиться спеціальне *S*-перетворення на множині перестановок та визначаються характеристики перестановки за цим перетворенням. Запропонована класифікація перестановок по класам еквівалентності за їх характеристиками відносно *S*-перетворення. Отримано точні значення перебором числа перестановок в таких класах для деяких розмірів перестановки та оцінки потужності класів з різними характеристиками методом статистичного моделювання. Наведено результати для знаходження оцінок потужності виділених класів перестановок різного порядку.

Наведено повний список класів перестановок з точним значенням їх потужності для перестановок порядку n = 11. Отримано інтервальні оцінки для потужності класів з різними характеристиками для перестановок порядку n = 11, 26, 30, 31, 32, 33, 45, 55, приклади яких наведені в статті. Окремо виділено повні перестановки з їх характеристиками, як найкращі для застосування у роторних криптосистемах. Характеристики та властивості кожного класу впливають на

стійкість роторних шифрів та шифрів, які використовують як базові елементи перестановки та підстановки. Властивості перестановок та їх кількість важливі для побудові таких криптосистем та проведення криптоаналізу.

Отримані статистичні значення (частина яких наведена в таблицях розділу) можна використовувати для знаходження ймовірності того, що згенерована перестановка має певну характеристику, або її характеристика належить групі заданих характеристик та оцінювати складність пошуку перестановок з певними властивостями. Запропоновано також статистичний критерій згоди, який використовує характеристики перестановок за *S*-перетворенням, для перевірки генераторів випадкових перестановок та підстановок, який може краще відсіювати специфічні альтернативні гіпотези у деяких генераторах випадкових підстановок та перестановок.

Список використаних джерел

1. *Коваленко І.М., Купер К.* Верхня границя для числа повних відображень / І.М.Коваленко, К. Купер // Теорія ймовірностей і математична статистика. – 1995. – Т. 53. – С. 69-75.

2. *Коваленко И.Н.* Об одной верхней оценке числа полных отображений / И.Н.Коваленко // Кибернетика и системный анализ. – 1996. – № 1. – С. 81-85.

3. *Cooper C., Gilchrist R., Kovalenko I.N., Novacovic D.* Deriving the number of good permutations with applications to cryptography / C.Cooper, R. Gilchrist, I.N. Kovalenko, D. Novacovic // Кибернетика и системный анализ. – 1999. – № 5. – Р. 10-16.

4. *Кузнецов Н.Ю.* Применение ускоренного моделирования к нахождению количества “хороших” перестановок / Н.Ю. Кузнецов // Кибернетика и системный анализ. – 2007. – № 6. – С. 80-89.

5. *Сачков В.Н.* Введение в комбинаторные методы дискретной математики. – 2-е изд., испр. и доп. / В.Н. Сачков. – М.: МЦНМО,2004. – 424 с.

6. *Рейнгольд Э., Нивергельт Ю., Део Н.* Комбинаторные алгоритмы / Э. Рейнгольд, Ю. Нивергельт, Н. Део. – М.: Мир, 1980. – 478 с.

7. *Шор Я. Б.* Таблицы для анализа и контроля надежности / Я. Б. Шор, Ф. И. Кузьмин. – М.: Советское радио, 1968. – 288 с.

8. *Кобзарь А. И.*Прикладная математическая статистика. Для инженеров и научных работников. / А. И. Кобзарь. – М.: ФИЗМАТЛИТ, 2006. – 816 с.

9. *Konheim A. G.* Computer security and cryptography / Alan G. Konheim. – New Jersey: John Wiley and Sons, Inc., 2007. – 521 с.

References

1. KOVALENKO, I.N. & COOPER, C. (1995) The upper bound for the number of complete mappings. *Probability theory and mathematical statistics.*Vol. 53. p. 69-75.

2. KOVALENKO, I.N. (1996) On one upper bound for the number of complete mappings. *Kibernetika i sistemnyj analiz.* Vol. 1. p. 81-85.

3. COOPER, C. & GILCHRIST, R. & KOVALENKO, I.N. & NOVACOVIC, D. (1996) Deriving the number of good permutations with applications to cryptography. *Kibernetika i sistemnyj analiz.* Vol. 5. p. 10-16.

4. KUZNETSOV, N.YU. (2007) Applying accelerated modeling to finding the number of “good” permutations. *Kibernetika i sistemnyj analiz.* Vol. 6. p. 80-89.

5. SACHKOV, V.N. (2004) Introduction to combination methods of a discrete motherboard. – 2-nd ed., Rev. and add. *M., MCCNMO*. 424 p.

6. REINGOLD, E. & NIEVERGELT, JU. & DEO, N. (1977) Combinatorial algorithms. Theory and practice. *Prent-Hall, Inc., Inglewood Cliffs, New Jersey 07632.*

7. SHOR, YA. B. & KUZMIN, F.I. (1968) Tables for analysis and control of reliability. *M.,Soviet Radio.* 288 p.

8. KOBZAR, A. I. (2006) Applied mathematical statistics. For engineers and scientists. *M., FIZMATLIT*. 816 p.

9. KONHEIM, A. G. (2007) Computer security and cryptography. *New Jersey: John Wiley and Sons, Inc.* 521 p.

Надійшла до редколегії 25.06.2019.