24. Feher C., Elovici Y., Moskovitch R., Rokach L., Schclar A. User Identity Verifi-cation via Mouse Dynamics // Information Sciences. 2012. Vol. 201. P. 19-36.
25. Stanić M. Continuous user verification based on behavioral biometrics using mouse dynamics // Proceedings of the ITI 2013 35th International Conference on Information Technology Interfaces. 2013. P. 251-256.
26. Идентификация пользователей Tor Browser через анализ особенностей работы с мышью. URL: https://www.opennet.ru/opennews/art.shtml?num=44027 (дата обращения 20.06.2019).

УДК 004.421.6: 519.23

ТЕСТИРОВАНИЕ КРИПТОГРАФИЧЕСКИХ ГЕНЕРАТОРОВ СЛУЧАЙНЫХ И ПСЕВДОСЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ НА ОСНОВЕ ЭНТРОПИЙНЫХ ПРОФИЛЕЙ

В.Ю. ПАЛУХА, Ю.С. ХАРИН
НИИ прикладных проблем математики и информатики,
Белорусский государственный университет
г. Минск, 220030, Республика Беларусь

Введение. Генераторы случайных и псевдослучайных последовательностей являются одним из элементов систем криптографической защиты информации (СКЗИ). Стойкость СКЗИ зависит от того, насколько близка генерируемая последовательность по своим свойствам к равномерно распределённой случайной последовательности (РРСП) [1], которая на практике называется «чисто случайной» последовательностью.

Для проверки качества криптографических генераторов используются статистические тесты, в которых проверяется гипотеза $H_* = \{\{x_t\}$ является РРСП$\}$ о том, что наблюдаемая последовательность $\{x_t\}$ является равномерно распределённой случайной последовательностью. В качестве тестовой статистики целесообразно использовать статистические оценки энтропии [2]. Одними из самых распространённых функционалов энтропии являются функционалы Шеннона, Реньи и Тсаллиса, которые и будут рассмотрены в дальнейшем.

Оценки энтропии**.** Пусть на вероятностном пространстве (Ω, F, P) с множеством состояний $\Omega = \{\omega_1, \ldots, \omega_N\}$ определена случайная величина $x = x(\omega) = \omega$ с дискретным распределением вероятностей $p = \{p_k\}$, $p_k = P\{x = \omega_k\}$, $p_k \geq 0$, $\sum_{k=1}^{N} p_k = 1$, $k = 1, \ldots, N$. В таблице 1 приведены формулы наиболее распространённых функционалов энтропии.

Таблица 1. Функционалы энтропии:

Энтропия Шеннона	$H(p) = -\sum_{i=1}^{N} p_i \ln p_i$
Энтропия Реньи	$H_r(p) = \frac{1}{1-r} \ln\left(\sum_{i=1}^{N} p_i^r\right), \quad r \in \mathbb{N}, r > 1.$
Энтропия Тсаллиса	$S_r(p) = \frac{1}{r-1}\left(1 - \sum_{i=1}^{N} p_i^r\right), \quad r \in \mathbb{N}, r > 1.$

Пусть имеется случайная последовательность $\{x_t : t = 1, \ldots, n\}$ объёма n из распределения вероятностей $\{p_k\}$, по которой будет оцениваться энтропия. Частотные оценки вероятностей имеют вид

$$\hat{p}_k = \frac{\nu_k}{n}, \quad \nu_k = \sum_{t=1}^{n} I\{x_t = \omega_k\} \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}, \quad I\{x_t = \omega_k\} = \begin{cases} 1, x_t = \omega_k; \\ 0, x_t \neq \omega_k. \end{cases} \tag{1}$$

Рассмотрим асимптотику совместного увеличения объёма выборки и сложности задачи:

$$n, N \to \infty, n/N \to \lambda, 0 < \lambda < \infty. \tag{2}$$

В асимптотике (2) для распределения вероятностей статистик $\{\nu_k\}$ справедлива аппроксимация законом Пуассона $\Pi(\lambda_k)$ с параметром $\lambda_k = np_k$. При истинной гипотезе H_* все элементарные вероятности равны: $p_k = 1/N$, $k = 1..., N$, поэтому все частоты $\{\nu_k\}$ имеют одинаковый параметр распределения Пуассона $\lambda = n/N$.

Оценка энтропии Шеннона на основе частотных статистик (1) имеет вид:

$$\hat{H} = \hat{H}(n, N) = -\sum_{k=1}^{N} \hat{p}_k \ln \hat{p}_k = -\sum_{k=1}^{N} \frac{\nu_k}{n} \ln \frac{\nu_k}{n} = \ln n - \frac{1}{n} \sum_{k=1}^{N} \nu_k \ln \nu_k. \tag{3}$$

Теорема 1 [3]. В асимптотике (2) статистика (3) при гипотезе H_* имеет асимптотически нормальное распределение с параметрами

$$\mu_H = \ln n - e^{-\lambda} \sum_{k=1}^{+\infty} \frac{\ln(k+1)\lambda^k}{k!}, \tag{4}$$

$$\sigma_H^2 = \frac{e^{-\lambda}}{n} \sum_{k=1}^{+\infty} \frac{(k+1)\lambda^k}{k!} \ln^2(k+1) - \frac{e^{-2\lambda}}{N} \left(\sum_{k=1}^{+\infty} \frac{\ln(k+1)\lambda^k}{k!} \right)^2 - \frac{e^{-2\lambda}}{n} \left(\sum_{k=1}^{+\infty} \ln(k+1) \frac{\lambda^k}{k!} (k+1-\lambda) \right)^2. \tag{5}$$

Как следует из теоремы 1, в асимптотике (2) оценка (3) является смещённой, и с уменьшением λ смещение растёт. Для функционалов энтропии Реньи и Тсаллиса можно построить несмещённую оценку в асимптотике (2), в т.ч. и при $\lambda < 1$. Как видно из таблицы 1, эти функционалы являются функциями от величины

$$P_r(p) = \sum_{k=1}^{N} p_k^r. \tag{6}$$

Определим r-ую нисходящую факториальную степень x:

$$x^{\underline{r}} = x(x-1)\ldots(x-r+1) = \frac{x!}{(x-r)!} = \sum_{i=0}^{r} s(r,i) x^i, \tag{7}$$

где $s(r, i)$ – число Стирлинга первого рода; при $x < r$ полагают $x^{\underline{r}} ::= 0$.

Несмещённая оценка для (6) основана на (7) [4]:

$$\tilde{P}_r(p) = \sum_{k=1}^{N} \frac{\nu_k^{\underline{r}}}{n^r}. \tag{8}$$

Полагая, что случайная величина ν имеет распределение Пуассона с параметром λ, т.е. $\mathcal{L}\{\nu\} = \Pi(\lambda)$, получим согласно [5] $E\{\nu^{\underline{r}}\} = \lambda^r$. Кроме того, согласно [5], $E\{\nu^r\} = \sum_{i=0}^{r} S(r,i)\lambda^i$, где $S(r, i)$ – число Стирлинга второго рода.

Статистические оценки энтропии Реньи и Тсаллиса, построенные с использованием оценки (8), имеют вид

$$\hat{H}_r(n, N) = \frac{1}{1-r}\ln\left(\sum_{k=1}^{N}\frac{\nu_k^r}{n^r}\right) = \ln n + \frac{1}{r-1}\left(\ln n - \ln\sum_{k=1}^{N}\nu_k^r\right), \tag{9}$$

$$\hat{S}_r(n, N) = \frac{1}{r-1}\left(1-\sum_{k=1}^{N}\frac{\nu_k^r}{n^r}\right) = \frac{1}{r-1}\left(1-\frac{1}{n^r}\sum_{k=1}^{N}\nu_k^r\right). \tag{10}$$

Теорема 2 [3]. В асимптотике (2) статистика (10) является состоятельной асимптотически несмещённой оценкой энтропии Тсаллиса и при истинной гипотезе H_* имеет асимптотически нормальное распределение с параметрами:

$$\mu_{S,r} = \frac{1}{r-1}\left(1-\frac{1}{N^{r-1}}\right), \tag{11}$$

$$\sigma^2_{S,r} = \frac{\lambda^{r-1}}{(r-1)^2 n^{2r-1}}\left(\sum_{i=1}^{r} s(r,i)\sum_{j=1}^{i-1} C_i^j r^{i-j}\sum_{k=1}^{j} S(j,k)\lambda^k - r^2\lambda^{r-1} + r!\right). \tag{12}$$

Следствие 1. *При $r = 2$ для математического ожидания и дисперсии асимптотического распределения оценки* (10) *справедливы выражения:*

$$\mu_{S,2} = 1-\frac{1}{N}, \quad \sigma^2_{S,2} = \frac{2}{Nn^2}.$$

Теорема 3 [3]. *В асимптотике* (2) *статистика* (9) *является состоятельной оценкой энтропии Реньи и при истинной гипотезе H_* имеет асимптотически нормальное распределение с параметрами:*

$$\mu_{H,r} = \ln N, \tag{13}$$

$$\sigma^2_{H,r} = \frac{\sum_{i=2}^{r} s(r,i)\sum_{j=1}^{i-1} C_i^j r^{i-j}\sum_{k=1}^{j} S(j,k)\lambda^k - r^2\lambda^{r-1} + r!}{(r-1)^2 n\lambda^{r-1}}. \tag{14}$$

Следствие 2. При $r = 2$ для дисперсии асимптотического распределения вероятностей оценки (9) справедливо выражение:

$$\sigma^2_{H,2} = \frac{2}{n\lambda}.$$

Тестирование генераторов. Пусть $\alpha \in (0,1)$ – заданный уровень значимости. Введём обозначения: $\hat{h}$ – статистическая оценка энтропии Шеннона (3), Реньи (9) или Тсаллиса (10), μ_h – асимптотическое математическое ожидание статистической оценки энтропии Шеннона (4), Реньи (13) или Тсаллиса (11), σ_h^2 – асимптотическая дисперсия статистической оценки энтропии Шеннона (5), Реньи (14) или Тсаллиса (12) при истинной гипотезе H_*. Вычислим для наблюдаемой последовательности статистику $\hat{h}$. Решающее правило, основанное на статистике $\hat{h}$, имеет вид [3]:

$$\text{принимается} \begin{cases} H_*, & \text{если } t_- < \hat{h} < t_+; \\ \overline{H_*}, & \text{в противном случае,} \end{cases} \qquad t_\pm = \mu_h \pm \sigma_h \Phi^{-1}\left(1-\frac{\alpha}{2}\right). \tag{15}$$

где $\Phi(\cdot)$ – функция распределения стандартного нормального закона.

Пусть генератор порождает двоичную выходную последовательность $\{y_\tau\}$, $\tau = 1, \ldots, T$. «Нарежем» её на непересекающиеся подряд идущие фрагменты длины s (s-граммы): $X^{(t)} = (X_j^{(t)}) = (y_{(t-1)s+1}, \ldots, y_{ts}) \in \{0,1\}^s$, $t = 1, \ldots, n = [T / s]$. Из полученных s-грамм

сформируем новую последовательность $\{x_t\}$ из алфавита мощности $N = 2^s$ по правилу

$$x_t = \sum_{j=1}^{s} 2^{j-1} X_j^{(t)} + 1.$$

На основе критерия (15) мы можем вычислить последовательность нормированных отклонений оценки энтропии от математического ожидания в зависимости от s, которые назовём **энтропийными профилями**:

$$\chi(s) = \frac{\hat{h}(s) - \mu_h(s)}{\sigma_h(s)\Phi^{-1}\left(1-\alpha/2\right)}, \; s = 1, \ldots, s_+. \tag{16}$$

Тестирование с помощью профиля позволяет выносить решение о принятии или отклонении гипотезы H_* на основе решающего правила (15) по совокупности значений $\chi(s)$ для различных s; такое решение видится более аргументированным, чем при принятии его по результатам применения теста (15) для отдельного значения s.

Представляют теоретический и практический интерес задачи исследования стохастической зависимости статистик $\chi(s)$ в (16). Исследуем зависимость соседних статистик $\chi(s)$ и $\chi(s + 1)$. Для этого вначале исследуем зависимость частотных оценок (1). Пусть $\{\hat{p}_i(s)\}$ и $\{\hat{p}_k(s+1)\}$ – частотные оценки (1), вычисленные для s- и $(s + 1)$-грамм соответственно, $i = 0, \ldots, 2^s - 1, k = 0, \ldots, 2^{s+1} - 1$. Справедлива следующая теорема о коэффициенте корреляции частотных оценок вероятностей s- и $(s + 1)$-грамм.

Теорема 4. *При истинной гипотезе H_* для коэффициента корреляции* $Corr_*\{\hat{p}_i(s), \hat{p}_k(s+1)\} = \dfrac{\mathrm{cov}_*\{\hat{p}_i(s), \hat{p}_k(s+1)\}}{\sqrt{D_*\{\hat{p}_i(s)\} D_*\{\hat{p}_k(s+1)\}}}$ *частотных оценок вероятностей произвольной пары s- и (s + 1)-грамм* $(i, k), i = 0, \ldots, 2^s - 1, k = 0, \ldots, 2^{s+1} - 1$, *справедлива двусторонняя оценка*

$$-2\sqrt{\frac{s}{(s+1)\left(2^s-1\right)\left(2^{s+1}-1\right)}} \le Corr_*\{\hat{p}_i(s), \hat{p}_k(s+1)\} \le \frac{2^{s+2}-2s-4}{\sqrt{s(s+1)\left(2^s-1\right)\left(2^{s+1}-1\right)}}.$$

На рисунке 1 представлена зависимость верхней и нижней границы коэффициента корреляции частотных оценок вероятностей s- и $(s + 1)$-грамм от s. Из рисунка видно, что модуль коэффициента корреляции стремится к 0 с ростом s, причём нижняя граница приближается к нулю экспоненциально быстро. Следовательно, зависимость между оценками энтропии $\hat{h}_s$ и $\hat{h}_s$ по s- и $(s + 1)$-граммам также будет слабой, что позволяет выносить решение о качестве генератора по его энтропийному профилю, пренебрегая этой зависимостью.

На рисунке 2 представлен энтропийный профиль Реньи выходной двоичной последовательности физического генератора [6] длиной $T = 125 \cdot 2^{25}$ бит. На рисунке 3 представлен энтропийный профиль Тсаллиса выходной последовательности программного самосжимающего генератора на основе линейного регистра сдвига с многочленом $x^{24} + x^{11} + x^5 + x^2 + 1$ при фиксированном значении $\lambda = 2$ для $r = 2$ на уровне значимости $\alpha = 0.05$.

Как видно из рисунков 2, 3, для физического генератора гипотеза H_* принимается, для самосжимающего – отклоняется.

В настоящее время в НИИ ППМИ ведётся разработка программного комплекса энтропийного анализа выходных последовательностей криптографических генераторов, который позволит автоматизировать процесс принятия решений и визуализации энтропийных профилей последовательностей.

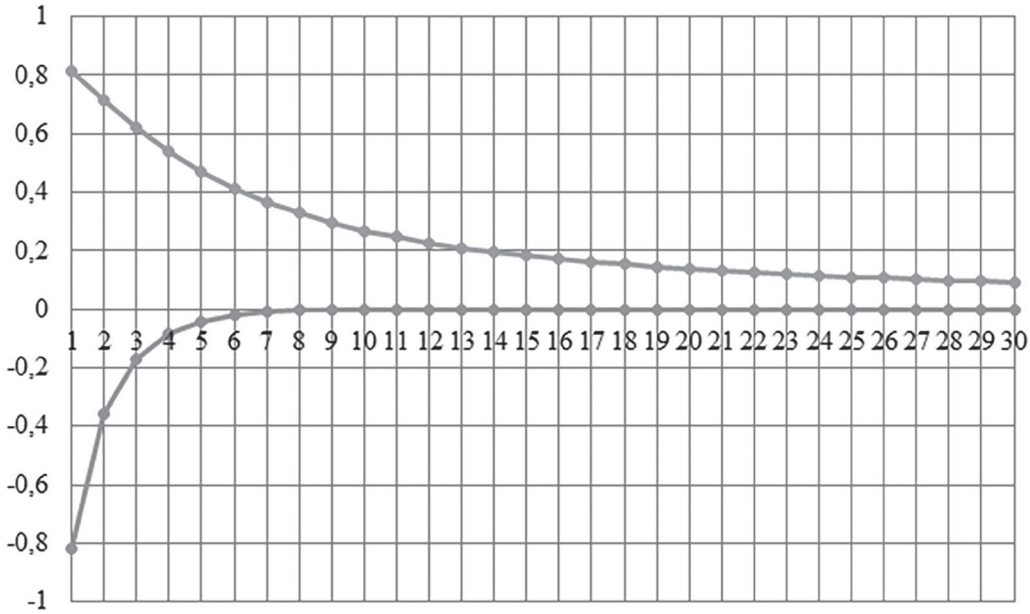

Рисунок 1. Границы коэффициента корреляции частотных оценок

Рисунок 2. Энтропийный профиль Реньи физического генератора

Рисунок 3. Энтропийный профиль Тсаллиса самосжимающего генератора

Список литературы

1. Криптология / Ю. С. Харин [и др.]. – Минск: БГУ, 2013. – 512 с.
2. Харин, Ю. С. Энтропийный анализ криптографических генераторов случайных и псевдослучайных последовательностей / Ю. С. Харин, В. Ю. Палуха // Веснік сувязі. – 2017. – № 6 (146). – С. 40–43.
3. Палуха, В. Ю. Статистические тесты на основе оценок энтропии для проверки гипотез о равномерном распределении случайной последовательности / В. Ю. Палуха // Весці НАН Беларусі. Серыя фізіка-матэматычных навук. – 2017. – № 1. – С. 79–88.
4. *Acharya, J.* Estimating Renyi Entropy of Discrete Distributions / J. Acharya, [et al.] – IEEE Transactions on Information Theory. – Vol. 63. – No. 1, 2017. – P. 38–56.
5. *Riordan, J.* Moment recurrence relations for binomial, Poisson and hypergeometric frequency distributions / J. Riordan // Annals of Mathematical Statistics. – 1937. – Vol. 8, № 2. – P. 103–111.
6. speedtest-500MB.bin [Electronic resource] // Humboldt Berlin University, Faculty of Mathematics and Natural Sciences, Department of Physics. – Mode of access: http://qrng.physik.hu-berlin.de/files/speedtest-500MB.bin.

УДК 004.056.5

АНАЛИЗ СТОЙКОСТИ КОМБИНИРОВАННОГО МЕТОДА ФОРМИРОВАНИЯ КРИПТОГРАФИЧЕСКОГО КЛЮЧА С СЕКРЕТНОЙ МОДИФИКАЦИЕЙ РЕЗУЛЬТАТОВ СИНХРОНИЗАЦИИ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

М.Л.РАДЮКЕВИЧ

Научно-производственное республиканское унитарное предприятие
«Научно-исследовательский институт технической защиты информации»,
г.Минск, Республика Беларусь

Введение. Использование синхронизируемых искусственных нейронных сетей (СИНС) для формирования общего криптографического ключа предложено В. Кантером, И. Кинцелем и описано в [1-5]. В работах [6-8] предложены методы повышения конфиденциальности формируемого общего секрета и уменьшения количества обменов информации по сравнению с технологией Neural key generation. Одним из них является комбинированный метод формирования криптографического ключа с секретной модификацией результатов синхронизации искусственных нейронных сетей (ИНС) (далее – комбинированный метод с секретной модификацией). Предлагается рассмотреть анализ стойкости данного метода.

Комбинированный метод с секретной модификацией**.** Комбинированный метод с секретной модификацией состоит из следующих шагов:

1. Задание входных параметров СИНС: n – количество входов каждого персептрона; K – количество персептронов; ±L – интервал возможных значений весовых коэффициентов персептронов; r – количество строк для функции свертки; $d_{ус}$ – количество тактов синхронизации; V – количество инвертируемых бит;