НАХОЖДЕНИЕ ТЕОРЕТИЧЕСКИХ РАСПРЕДЕЛЕНИЙ СТАТИСТИКИ ТЕСТА АППРОКСИМИРОВАННОЙ ЭНТРОПИИ

Киевец Н. Г., Ярук А. М.
Кафедра радио и информационных технологий, Белорусская государственная академия связи
Минск, Республика Беларусь
E-mail: kievetss@mail.ru

В работе найдены теоретические распределения статистики теста аппроксимированной энтропии для случайных последовательностей длиной 128 и 256 бит. С использованием полученных распределений выполнено двухуровневое тестирование последовательностей, выработанных генераторами электронных пластиковых карт.

Введение

В настоящее время для различных приложений широко используются электронные пластиковые карты (ЭПК), содержащие физические генераторы случайных чисел (ГСЧ). ГСЧ вырабатывают ключевую информацию для криптографических преобразований, в связи с чем задача оценки качества работы генераторов является актуальной.

Для оценки качества работы ГСЧ применяется двухуровневое тестирование вырабатываемых ими случайных последовательностей (СП) [1,2]. ГСЧ ЭПК вырабатывают СП только определенных длин, как правило, равных длинам практически используемых криптографических ключей. В связи с тем, что в тестах используются асимптотические распределения тестовых статистик, двухуровневое тестирование относительно коротких СП может привести к неверным выводам о случайности последовательностей и качестве работы ГСЧ [1,2].

Таким образом, чтобы выполнить корректное двухуровневое тестирование СП длин, равных длинам практически используемых ключей, требуется нахождение теоретических распределений тестовых статистик.

Данная работа посвящена нахождению теоретических распределений статистики теста аппроксимированной энтропии при длинах СП 128 и 256 бит.

I. Алгоритм нахождения теоретических распределений статистик

В соответствии с порядком тестирования по тесту аппроксимированной энтропии в СП заданной длины n подсчитываются количества v_i пересекающихся m-битных серий различных типов i и количества w_k пересекающихся $(m+1)$-битных серий различных типов k. Далее рассчитывается тестовая статистика [1,3]

$$\chi^2 = 2n\big(\ln 2 - \varphi^{(m)} + \varphi^{(m+1)}\big), \qquad (1)$$

где

$$\varphi^{(m)} = \sum_{i=0}^{2^m-1} \frac{v_i}{n} \ln \frac{v_i}{n}; \varphi^{(m+1)} = \sum_{k=0}^{2^{m+1}-1} \frac{w_k}{n} \ln \frac{w_k}{n}.$$

Предлагается следующий алгоритм нахождения теоретических распределений статистики теста аппроксимированной энтропии при выбранном параметре $m = 1$.

1. Определяются все возможные комбинации числа $n1$ единиц и числа r непрерывных подпоследовательностей бит в последовательности. Величина $n1 = \overline{0, n}$, а величина r зависит от $n1$ и может принимать следующие значения:

$$\begin{cases} r = 1, & \text{если } n1 = 1 \text{ или } n1 = n; \\ r = \overline{2, 2n1}, & \text{если } n1 = n/2; \\ r = \overline{2, 2n1 + 1}, & \text{если } 0 < n1 < n/2; \\ r = \overline{2, 2|n - n1| + 1}, & \text{если } n/2 < n1 < n. \end{cases}$$

2. Для каждой комбинации значений $n1$ и r рассчитывается их вероятность [4]:

если r – четное число:

$$P(n1, r) = 2 \binom{n1 - 1}{r/2 - 1} \binom{n - n1 - 1}{r/2 - 1} / 2^n;$$

если r – нечетное число:

$$P(n1, r) = \binom{n1 - 1}{(r-1)/2} \binom{n - n1 - 1}{(r-3)/2} / 2^n + \binom{n1 - 1}{(r-3)/2} \binom{n - n1 - 1}{(r-1)/2} / 2^n.$$

3. Определяются значения v_0, v_1, w_0, w_1, w_2 и w_3 для каждой комбинации значений $n1$ и r: $v_0 = n - n1$, $v_1 = n1$, w_1 и w_2 определяются из системы:

$$\begin{cases} w_1 = w_2 = r/2, & \text{если } r \text{ – четное число;} \\ w_1 = w_2 = (r-1)/2, & \text{если } r \text{ – нечетное число;} \end{cases}$$

$$w_3 = n1 - w_2;$$

$$w_0 = n - w_3 - w_1 - w_2.$$

4. Для каждой комбинации значений величин $n1$ и r рассчитываются значения тестовой статистики χ^2 из выражения (1). Вероятности $P(\chi^2)$ значений статистики равны соответствующим вероятностям $P(n1, r)$.

5. Определяются все возможные значения статистики χ^2 и их вероятности.

II. Двухуровневое тестирование по тесту аппроксимированной энтропии

При двухуровневом тестировании тест применяется к каждой из сгенерированных последовательностей, после чего проверяется соответствие эмпирического распределения значений вероятности P_T превышения полученных значений тестовых статистик теоретическому распределению. При использовании асимптотических распределений статистики χ^2 в качестве теоретического распределения P_T принимается равномерное распределение.

Значения вероятности P_T однозначно связаны со значениями статистики [1]:

$$P_T = \text{igamc}\left(2^{m-1}, \frac{\chi^2}{2}\right).$$

Вероятности значений величины P_T равны вероятностям соответствующих значений χ^2.

На рис. 1 приведены гистограммы вероятностей P попадания значений P_T в интервалы L при длинах СП 128 и 256 бит.

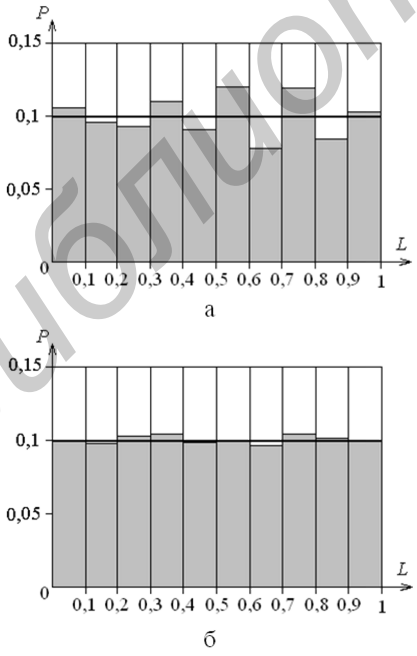

Рис. 1 – Гистограмма вероятностей P: а – при $n = 128$ бит; б – при $n = 256$ бит

Для оценки качества ГСЧ ЭПК было извлечено 8000 СП длиной 128 бит и 4000 СП длиной 256 бит. К каждой случайной последовательности был применен тест аппроксимированной энтропии с параметром $m = 1$, в результате чего для всех СП получены значения χ^2 и P_T. Произведено сравнение эмпирических распределений значений P_T с найденными теоретическими и равномерным распределениями по критерию согласия «хи-квадрат». В соответствии с данным критерием рассчитывается случайная величина, имеющая распределение «хи-квадрат» и рассчитывается вероятность P_0 превышения значения случайной величины. В табл. 1 приведены значения вероятности P_0 для наборов СП длиной $n = 128$ бит и $n = 256$ бит.

Таблица 1 – Значения P_0

n, бит	P_0 (найденные распределения P_T)	P_0 (равномерное распределение)
128	0,8899	0,0000
256	0,8232	0,5621

Из табл. 1 видно, что при уровне значимости $\alpha = 0,001$ оба набора СП прошли тестирование при использовании найденных распределений P_T, набор СП длиной 128 бит не прошел и набор СП длиной 256 бит прошел тестирование при использовании равномерного распределения в качестве теоретического распределения P_T.

III. Заключение

В работе приведен алгоритм нахождения теоретических распределений статистик теста аппроксимированной энтропии. Найдены теоретические распределения статистики и соответствующие распределения вероятности P_T превышения значений статистики при длинах СП 128 и 256 бит. Представлены результаты двухуровневого тестирования СП, выработанных ГСЧ ЭПК, с использованием найденных распределений P_T и равномерного распределения в качестве теоретических распределений. Показано, что использование равномерного распределения в качестве теоретического распределения величины P_T приводит к неверным выводам о качестве работы ГСЧ при двухуровневом тестировании СП длиной 128 бит по тесту аппроксимированной энтропии.

IV. Список литературы

1. A statistical test suite for random and pseudorandom number generators for cryptographic applications [Электронный ресурс]. – 2010. – Режим доступа: http://csrc.NIST.gov/publications/NISTpubs/800-22-rev1a/SP800-22rev1a.pdf. – Дата доступа: 30.06.2017.
2. L'Ecuyer, P. Testing random number generators / P. L'Ecuyer // Winter Simulation Conference. – 1992. – P. 305–313.
3. Rukhin, A. Approximate entropy for testing randomness / A. Rukhin // Journal of Applied Probability. – 2000. – Vol. 37. – P. 88–100.
4. Nonparametric statistical interference // Marcel Dekker, Inc. [Electronic resource]. – 2003. – Mode of access: http://f3.tiera.ru/2/M_Mathematics/MV_Probability/MVsa_Statistics and applications/Gibbons J. Nonparametric statistical inference (Dekker, 2003)(ISBN 0824740521)(O)(672s)_MVsa_.pdf. – Date of access: 25.02.17.