УДК 519.2

М.С. АБРАМОВИЧ, А.В. ВЛАСОВ

ОБНАРУЖЕНИЕ ПЕРИОДИЧНОСТИ БИНАРНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ С ИСПОЛЬЗОВАНИЕМ ПРЕОБРАЗОВАНИЯ УОЛША - АДАМАРА

The test for detecting periodicity in a binary random sequence using the Walsh - Hadamard transform is proposed and its performance is investigated using the statistical modelling.

Обнаружение периодичности бинарных последовательностей является актуальной задачей при исследовании стойкости криптосистем [1]. Для изучения периодичности непрерывных последовательностей применяются спектральные тесты, основанные на преобразовании Фурье [1]. Для выявления особенностей дискретных последовательностей в [2] рекомендуется применять другое ортогональное преобразование - Уолша - Адамара. В настоящей работе предложен критерий обнаружения периодичности бинарной последовательности, основанный на преобразовании Уолша - Адамара.

Математическая модель

Рассмотрим бинарную случайную последовательность:

$$\Xi = \left\{\xi_j \mid j = \overline{1, N}\right\}, \quad \xi_j = \pm 1, \ N = 2^n. \tag{1}$$

Нулевая гипотеза о распределении вероятностей данной последовательности формулируется в виде: H_0:$\{\xi_j\}$ - независимы и одинаково распределены $P\{\xi_j=1\}=\text{P}\{\xi_j=-1\}=0{,}5$.

Альтернативная гипотеза H_1 для выявления периодичности в последовательности (1) записывается следующим образом:

H_1 : $\xi_j=z_j\varepsilon_j$; где $\{z_j\}$, $z_j=\pm1$, - неизвестная периодическая последовательность с неизвестным периодом $T=2^k$, $z_j=z_{j+T}$, $k<n$; последовательность $\{\varepsilon_j\}$: $\text{P}\{\varepsilon_j=1\}=\varepsilon$, $\text{P}\{\varepsilon_j=-1)=1-\varepsilon$ - «зашумляющая» исходную периодическую последовательность, ε - уровень «зашумления».

Статистические свойства последовательностй Уолша - Адамара

Преобразование Уолша - Адамара случайной бинарной последовательности (1) в матричном виде определяется как [3]:

$$X = \frac{1}{N} H(n)\Xi,$$

где $X = \left\{x_{Ni} \mid i = \overline{1, N}\right\}$, $H(n)$ - матрица Адамара порядка n.

Матрица Адамара вычисляется но рекуррентному правилу:

$$H(k) = \begin{bmatrix} H(k-1) & H(k-1) \\ H(k-1) & -H(k-1) \end{bmatrix}, \ k = 1, 2, \ldots, n,$$

где $H(0)=1$.

Элемент преобразования Уолша - Адамара записывается в виде:

$$x_{Ni} = \frac{1}{N}\sum_{j=1}^{N} H_{ij}(n)\xi_j, \ i = \overline{1, N},$$

где $H_{ij}(n)$ - элемент матрицы Уолша - Адамара.

Теорема 1. Пусть $\{\xi_j, \ j = \overline{1, N}\}$- последовательность независимых одинаково распределенных случайных величин $P\{\xi_j=1\}=p$, $P\{\xi_j=-1\}=1-p$, $E\{\xi_j\}=m=2p-1$, $D\{\xi_j\}=\sigma^2=4p(1-p)$, а $X_N = \left\{x_{Ni}, \ i = \overline{1, N}\right\}$ - ее преобразование Уолша - Адамара. Тогда для любого $i=i(N)\in\{1, 2, ..., N\}$ при $N\to\infty$ справедливо асимптотическое соотношение $\frac{\sqrt{N}}{\sigma}\left(x_{Ni} - m_i\right) \xrightarrow{d} N(0, 1)$, где $m_i=E\{x_{Ni}\}=\begin{cases} m, \text{ если } i = 1, \\ 0, \text{ если } i = 2, 3, \ldots, N. \end{cases}$

Доказательство. Положим $\eta_{Nj} = \frac{H_{ij}(n)(\xi_j - m)}{\sqrt{N}\sigma}$, $S_N = \sum_{j=1}^{N}\eta_{Nj}$.

Тогда $\{\eta_{Nj}\}_{j=1}^{N}$- последовательность серий независимых в каждой серии случайных величин. Случайные величины η_{N1}, ..., η_{NN} независимы как борелевские функции от независимых случайных величин ξ_1, ..., ξ_N. Эта последовательность обладает свойствами:

$$E\left\{\eta_{Nj}\right\} = \frac{H_{ij}(n)}{\sqrt{N}\sigma} E\{\xi_i - m\} = 0, \tag{2}$$

$$D\left\{\eta_{Nj}\right\}=\frac{1}{N\sigma^2}D\{\xi_i-m\}=\frac{1}{N}, \quad (3)$$
$$D\{S_N\}=1.$$

Покажем, что для последовательности серий $\{\eta_{Nj}\}_{j=1}^{N}$ выполняется условие Линдеберга [4], т. е. для любого $\tau>0$ $\sum_{j=1}^{N}E\left\{\eta_{Nj}^2 I\{|\eta_{Nj}|>\tau\}\right\}\to 0$ при $N\to\infty$.

Имеем $\sum_{j=1}^{N}E\left\{\eta_{Nj}^2 I\{|\eta_{Nj}|>\tau\}\right\}=\frac{1}{\sigma^2 N}\sum_{j=1}^{N}E\left\{(\xi_j-m)^2 I\left\{|\xi_j-m|>\tau\sqrt{N}\sigma\right\}\right\}$.

Так как $|\xi_j-m|\le 1+m=\text{const}$, то $I\{|\xi_j-m|>\tau\sqrt{N}\sigma\}=0$ при $N\ge N(\tau,p)$.

Таким образом, $\sum_{j=1}^{N}E\left\{\eta_{Nj}^2 I\{|\eta_{Nj}|>\tau\}\right\}\to 0$ при $N\to\infty$. Выполнение условия Линдеберга, а также условий (2), (3) позволяет применить центральную предельную теорему для схемы серий [4] и получить следующее соотношение:

$$S_N\xrightarrow{d}N(0,\ 1),\ \text{т. е.}$$
$$S_N=\sum_{j=1}^{N}\frac{H_{ij}(n)\xi_j}{\sigma\sqrt{N}}-\frac{m}{\sigma\sqrt{N}}\sum_{j=1}^{N}H_{ij}(n)=\frac{\sqrt{N}}{\sigma}(x_{Ni}-m_i)\xrightarrow{d}\mathcal{N}(0,\ 1).$$

Следствие. При выполнении нулевой гипотезы H_0 имеем:

$$\sqrt{Nx_{Ni}}\xrightarrow{d}\mathcal{N}(0,\ 1).$$

Для доказательства асимптотической независимости элементов преобразования Уолша - Адамара введем понятие диадической стационарности случайной последовательности (см. [2]).

Пусть $u,\ v$ - неотрицательные действительные числа, представимые в виде рядов:

$$u=\sum_{l=-\infty}^{\infty}u_l 2^l,\ \ u_l=0\ \text{или}\ 1,\ \ v=\sum_{l=-\infty}^{\infty}v_l 2^l,\ \ v_l=0\ \text{или}\ 1.$$

Определим $u\oplus v=\sum_{l=-\infty}^{\infty}(u_l\oplus v_l)2^l$, где $\oplus$ - операция побитового сложения (сложения по модулю 2).

Определение. Последовательность $Y=\left\{y_i\mid i=\overline{1,\ N}\right\}$ называется диадически стационарной, если ее математическое ожидание постоянно $E\{y_i\}=const$, а ковариационная функция $cov\ \{y_n y_m\}$ зависит только от суммы $n\oplus m$.

Утверждение 1. Случайная бинарная последовательность (1) при истинной нулевой гипотезе H_0 является диадически стационарной.

Доказательство. Поскольку в условиях нулевой гипотезы H_0: $p=0{,}5$ и члены последовательности (1) независимы, то $E\{\xi_i\}=2p-1=0$. При $i\neq j$: $\text{cov}\{\xi_i\xi_j\}=E\{\xi_i\xi_j\}=E\{\xi_i\}E\{\xi_j\}=0$. При $i=j$ будем иметь: $\text{cov}\{\xi_i\xi_j\}=E\{\xi_i^2)\}=1$.

Таким образом, ковариационная функция зависит только от величины $i\oplus j$, т. е. последовательность (1) является диадически стационарной.

Утверждение 2. Пусть элементы последовательности (1) независимы и одинаково распределены. Тогда элементы преобразования Уолша - Адамара $X=\left\{x_{Ni}\mid i=\overline{1,\ N}\right\}$ являются асимптотически независимыми.

Доказательство. Из утверждения 1 следует, что бинарная последовательность (1) является диадически стационарной. В [2] доказано, что элементы преобразования Уолша - Адамара диадически стационарной последовательности являются асимптотически независимыми.

Критерий обнаружения периодичности и исследование его эффективности

Построим последовательность $\left\{w_{Ni}=x_i^2,\ i=\overline{1,\ N}\right\}$. Заметим, что из теоремы 1 вытекает:

$$Nw_{Ni} \xrightarrow{d} \chi_1^2. \tag{4}$$

Определим максимум последовательности $\{w_{Ni}\}$:

$$w_{N\max} = \max_i w_{Ni}. \tag{5}$$

Найдем распределение статистик и (5).

Теорема 2. Пусть элементы последовательности (1) независимы и одинаково распределены. Тогда статистика (5) имеет асимптотически распределение:

$$F_{w_{N\max}}(x) \xrightarrow{i} F_{\chi_1^2}^N(x),\ x \ge 0. \tag{6}$$

Доказательство. В силу утверждения 2 элементы преобразования Уол-ша - Адамара являются асимптотически независимыми. Следовательно, асимптотически независимы и элементы последовательности $\left\{w_{Ni},\ i=\overline{1,\ N}\right\}$ как борелевские функции от асимптотически независимых случайных величин $\left\{x_{Ni},\ i=\overline{1,\ N}\right\}$. С учетом (4) и по определению функции распределения имеем:

$$F_{w_{N\max}}(x) = P\left\{Nw_{N\max} \le x\right\} = P\left\{Nw_1 \le x,\ Nw_2 \le x,\ \ldots,\ Nw_N \le x\right\} =$$
$$= \prod_{i=1}^{N} P\left\{Nw_{N\max} \le x\right\} = F_{\chi_1^2}^N(x).$$

По аналогии со спектральным критерием обнаружения периодичности из [1] построим критерий обнаружения статистической периодичности последовательности (1):

$$\text{принимается гипотеза: } \begin{cases} H_0, & \text{если } P > \alpha, \\ H_1, & \text{если } P \le \alpha, \end{cases} \tag{7}$$

где $P = 1 - F_{\chi_1^2}^N\left(Nw_{N\max}\right)$ - P-значение, α - уровень значимости.

Замечание. В случае принятия альтернативной гипотезы H_1 за оценку длины периода последовательности (1) принимается величина $\hat{T} = 2^{k_{N\max}}$, где $k_{N\max} = \arg\max_i w_{Ni}$.

Эффективность критерия (7) исследовалась методом статистического моделирования. Для оценки вероятности ошибки первого рода моделировались симметричные последовательности Бернулли длиной N=64, 128, 256, 512, 1024, 2048, 4096. Количество экспериментов для каждого случая M=1000. Уровень значимости при проверке нулевой гипотезы H_0 полагался равным a=0,05.

В табл. 1 приведены оценки вероятности ошибки первого рода $\hat{\alpha}$ в зависимости от длины последовательности N, из которой видно, что оценка вероятности ошибки первого рода для всех исследуемых случаев не превышает 0,056, что сравнимо с уровнем значимости α=0,05.

Для оценки вероятности ошибки второго рода моделировались выборки длиной N=1024, имеющие периоды T=32, 64, 128, 256, 512 с различными уровнями «зашумления» e: 0,05; 0,1; 0,15; 0,2; 0,25; 0,3; 0,35; 0,4.

Таблица 1

Оценки вероятности ошибки первого рода в зависимости от длины последовательности

T	64	128	256	512	1024	2048	4096
$\hat{\alpha}$	0,042	0,041	0,056	0,036	0,056	0,049	0,053

В табл. 2 приведены оценки вероятности ошибки второго рода в зависимости от T.

Таблица 2

Оценки вероятности ошибки второго рода в зависимости от длины последовательности и уровня «зашумления»

T \ ε	0	0,05	0,1	0,15	0,2	0,25	0,3	0,35	0,4
32	0	0	0	0,001	0,004	0,215	0,761	0,979	0,992
64	0	0	0	0	0,001	0,238	0,892	0,991	0,994
128	0	0	0	0	0,019	0,642	0,969	0,996	0,996
256	0	0	0	0	0,305	0,92	0,995	0,998	0,998
512	0	0	0	0,078	0,776	0,966	0,994	1	0,999

Как видно из табл. 2, предложенный критерий позволяет обнаружить периодичность бинарной последовательности при уровнях «зашумления», не превышающих 0,15 для всех значений длины периода T. Кроме того, оценка вероятности ошибки второго рода возрастает с увеличением T.

1. Харин Ю.С., Агиевич С.В. Компьютерный практикум по математическим методам защиты информации. Мн., 2001.
2. Pedro A. Morettin // SIAM Review. 1981. Vol. 23. Iss. 3. P. 279.
3. Ахмед Н., Рао К.Р. Ортогональные преобразования при обработке цифровых сигналов. М, 1980.
4. Боровков А. А. Теория вероятностей. М., 1986.

Поступила в редакцию 23.09.04.

Михаил Семенович Абрамович - кандидат физико-математических наук, заведующий НИЛ статистического анализа и моделирования ННИЦ прикладных проблем математики и информатики БГУ.

Александр Владимирович Власов - магистрант кафедры математического моделирования и анализа данных.