A Universal Statistical Test for Random Bit Generators

Ueli M. Maurer

Institute for Signal and Information Processing
Swiss Federal Institute of Technology
CH-8092 Zurich, Switzerland

Abstract. A new statistical test for random bit generators is presented that is universal in the sense that any significant deviation of the output statistics from the statistics of a perfect random bit generator is detected with high probability when the defective generator can be modeled as an ergodic stationary source with finite memory. This is in contrast to most presently used statistical tests which can detect only one type of non-randomness, for example, a bias in the distribution of 0's and 1's or a correlation between consecutive bits. Moreover, the new test, whose formulation was motivated by considering the universal data compression algorithms of Elias and of Willems, measures the entropy per output bit of a generator. This is shown to be the correct quality measure for a random bit generator in cryptographic applications. A generator is thus rejected with high probability if and only if the cryptographic significance of a statistical defect is above a specified threshold. The test is easy to implement and very fast and thus well-suited for practical applications.

1. Introduction

A random bit generator is a device whose output sequence can be modeled as a sequence of statistically independent and symmetrically distributed binary random variables (both values 0 and 1 are equally probable), i.e., as a so-called binary symmetric source. Random bit generators have many applications in cryptography, VLSI testing, probabilistic algorithms and other fields. Their major application in cryptography is as the secret-key source of a symmetric cipher system, but random bit generators are also required for generating public-key parameters (e.g., RSA-moduli) and for generating the keystream in the well-known one-time pad system. In these applications, the security crucially depends on the randomness of the source. In particular, a symmetric (secret-key) cipher whose security rests on the fact that an exhaustive key search is infeasible may be completely insecure when not all keys

A.J. Menezes and S.A. Vanstone (Eds.): Advances in Cryptology - CRYPTO '90, LNCS 537, pp. 409-420, 1991.

are equiprobable. Similarly, the security of the RSA public-key cryptosystem may be strongly reduced when, because of a statistical defect in the random source used in the procedure generating the primes, the two primes are chosen from a small set of primes only.

Randomness is a property of an abstract model. Whether such a model can give an exact description of reality is a philosophical question related to the question of whether the universe is deterministic or not, and seems to be impossible to answer to everyone's satisfaction. However, there do exist chaotic processes in nature, such as radioactive decay and thermal noise in transistors, that allow the construction of a random bit generator that is completely unpredictable for all practical applications. It is a non-trivial engineering task, however, to design an electronic circuit that explores the randomness of such a process in a way that guarantees the statistical independence and symmetrical distribution of the generated bits. It is therefore essential in a cryptographic application that such a device be tested intensively for malfunction after production, and also periodically during operation.

This paper is concerned with the application of random bit generators as the secret-key source of a symmetric cipher system. A new statistical test for random bit generators is presented that offers two major advantages over the presently used statistical tests (including the common frequency test, serial test, poker test, autocorrelation tests and run test which are described in [1] and [4]). First, unlike these tests, the new test is able to detect any one of a very general class of possible defects a generator may have, including all the defects the above mentioned tests are designed to detect. This class of defects consists of those that can be modeled by an ergodic stationary source and contains those that could reasonably be assumed to occur in a practical implementation of a random bit generator. Second, rather than measuring some parameter (like the relative frequency of 1's) of the output of a generator, the new test measures the actual cryptographic significance of a defect. More precisely, the test parameter is very closely related to the running time of the enemy's optimal key-search strategy when he exploits knowledge of the secret-key source's statistical defect, and thus to the effective key size of the cipher system (if there exists no essentially faster way than an exhaustive key search for breaking the system).

The paper is not concerned with tests for pseudo-random bit generators that stretch a short (randomly selected) seed deterministically into a long sequence of pseudo-random bits, i.e., it is not concerned with the security evaluation of practical keystream generators for stream ciphers. However, it is certainly a necessary (but far from sufficient) condition for security that such a generator pass the test presented here. Design criteria for practical keystream generators are discussed in [5].

In section 2, an analysis of the enemy's optimal key-search strategy based on knowledge about the statistical defect of the secret-key source is presented. It is argued that the per-bit entropy of a bit generator is the correct measure of its cryptographic quality. Section 3 introduces the fundamentals of statistical testing and some of the presently used statistical tests are reviewed. The new universal statistical test is introduced in section 4 and the close relation between the test

parameter and the per-bit entropy of a generator is established.

2. Reduction of cipher security due to a statistical defect in the secret-key source

Throughout the paper, let $B = \{0,1\}$ and let $R^N = R_1, \ldots, R_N$ denote a sequence of N statistically independent and symmetrically distributed binary random variables. When a random bit generator based on a chaotic physical phenomenon like thermal transistor noise is either defective or not properly designed, then the generated bits may be biased and/or statistically dependent. The simplest example of such a statistical defect is modeled by a binary memoryless source whose output bits are statistically independent and identically (but not necessarily symmetrically) distributed. Let BMS_p denote the binary memoryless source that emits 1's with probability p and 0's with probability $1-p$. Another type of statistical defect is modeled by a binary source, denoted by ST_p, whose output bits are symmetrically distributed (0's and 1's occur with probability 1/2) but whose transition probabilities are biased: a binary digit is followed by its complement with probability p and by the same digit with probability $1-p$. This is an example of a binary stationary source with one bit of memory. In general, the probability distribution of the i-th bit of a generator's output may depend on the previous M output bits where M is the memory of the source. We argue that the statistical behavior of virtually every (even defective or badly designed) random bit generator can well be modeled by such a source with relatively small memory.

Consider a source S that emits a sequence $U_1, U_2, U_3, \ldots$ of binary random variables. If there exists a positive integer M such that for all $n > M$, the conditional probability distribution of U_n, given $U_1, \ldots, U_{n-1}$, depends only on the past M output bits, i.e., such that

$$P_{U_n|U_{n-1}\ldots U_1}(u_n|u_{n-1}\ldots u_1) \;=\; P_{U_n|U_{n-1}\ldots U_{n-M}}(u_n|u_{n-1}\ldots u_{n-M}) \qquad (1)$$

for $n > M$ and for every binary sequence $(u_1, \ldots, u_n) \in B^n$, then the smallest such M is called the *memory* of the source S and $\Sigma_n = [U_{n-1}, \ldots, U_{n-M}]$ denotes its *state* at time n. Let $\Sigma_1 = [U_0, \ldots, U_{-M+1}]$ be the initial state where $U_{-M+1}, \ldots, U_0$ are dummy random variables. If in addition to (1) the source satisfies

$$P_{U_n|\Sigma_n}(u|\sigma) \;=\; P_{U_1|\Sigma_1}(u|\sigma)$$

for all $n > M$ and for all $u \in B$ and $\sigma \in B^M$, then it is called *stationary*. A stationary source with memory M is thus completely specified by the probability distribution of the initial state, P_{Σ_1}, and the state transition probability distribution $P_{\Sigma_2|\Sigma_1}$. The state sequence forms a Markov chain with the special property that each of the 2^M states has at most 2 successor states with non-zero probability. See [3], chapters XV and XVI, for a treatment of Markov chains. We will denote the 2^M possible states of the source (or the Markov chain) by the integers in the interval $[0, 2^M - 1]$. ($\Sigma_n = j$ means that $U_{n-1}\ldots U_{n-M}$ is the binary representation of j.) For

the class of ergodic Markov chains (see [3] for a definition), which includes virtually all cases that are of practical interest, there exists an invariant state probability distribution, i.e.,

$$\lim_{n\to\infty} P_{\Sigma_n}(j) \;=\; p_j$$

for $0 \le j \le 2^M - 1$, where the p_j's are the solution, satisfying $\sum_{j=0}^{2^M-1} p_j = 1$, of the following system of linear equations

$$p_j \;=\; \sum_{k=0}^{2^M-1} P_{\Sigma_2|\Sigma_1}(j|k)\; p_k, \qquad 0 \le j \le 2^M - 1. \tag{2}$$

A good practical cipher is designed such that no essentially faster attack is known than an exhaustive key search. The size of the key space is chosen large enough to ensure that to succeed in such an exhaustive search, even with only very small probability of success, requires an infeasible searching effort. If not all possible values of the secret key have equal *a priori* probability, then the enemy's optimal strategy in an exhaustive key search is to start with the most likely key and to continue testing keys in order of decreasing probabilities. Let Z denote the secret key, let n be its length in bits and let $z_1, z_2, \ldots, z_{2^n}$ be a list of the key values satisfying

$$P_Z(z_1) \;\ge\; P_Z(z_2) \;\ge\; \cdots \;\ge\; P_Z(z_{2^n}).$$

For a given source S and for δ satisfying $0 \le \delta \le 1$ let $\mu_S(n,\delta)$ denote the minimum number of key values an enemy must test (using the optimal key-searching strategy) in order to find the correct key with probability at least δ when S is used to generate the n-bit key Z, i.e.,

$$\mu_S(n,\delta) \;=\; \min\left\{k:\; \sum_{i=1}^{k} P_Z(z_i) \ge \delta\right\}. \tag{3}$$

We define the *effective key size* of a cipher system with key source S to be $\log_2 \mu_S(n, \frac{1}{2})$, i.e., the logarithm of the minimum number of keys an enemy must try in order to find the correct key with probability at least 50%. The choice $\delta = 1/2$ in this definition is somewhat arbitrary, but in general, for large enough n, $\log_2 \mu_S(n,\delta)/n$ is almost independent of δ when δ is not extremely close to 0 or 1. Note that when the key is truly random, i.e., when S is a binary symmetric source, then $\log_2 \mu_S(n, \frac{1}{2}) = n - 1$.

We now determine the effective key size of a cipher system whose key source is BMS_p. Without loss of generality assume that $0 < p \le 1/2$. Note that the source ST_p described above can be modeled by the source BMS_p with a summator at the output (summing modulo 2 the output bits of BMS_p). Therefore the set of probabilities of keys and hence also the effective key size is identical for both sources. The probability distribution of Z is given by

$$P_Z(z) \;=\; p^{w(z)}(1-p)^{n-w(z)},$$

where $w(z)$ denotes the Hamming weight of z. In order to succeed with probability approximately $1/2$ the enemy must examine all keys z with Hamming weight $w(z) \le$

pn. The effective key size is thus well approximated by

$$\log_2 \mu_{\mathrm{BMS}_p}(n, \tfrac{1}{2}) \approx \log_2 \sum_{i=0}^{pn} \binom{n}{i}. \tag{4}$$

It is well-known (e.g., see [8]) that the term on the right side of (4) is well approximated by $nH(p)$, where $H(x)$ is the binary entropy function defined by

$$H(x) = -x \log_2 x - (1-x)\log_2(1-x) \tag{5}$$

for $0 < x < 1$ and by $H(0) = H(1) = 0$. Note that $H(x) = H(1-x)$ for $0 \leq x \leq 1$. This approximation is asymptotically precise, i.e.,

$$\lim_{n\to\infty} \frac{\log_2 \mu_{\mathrm{BMS}_p}(n, \delta)}{n} = H(p) \quad \text{for } 0 < \delta < 1.$$

Note that the entropy per output bit of the source BMS_p, $H(p)$, is equal to the factor by which the effective key size is reduced. Shannon proved (see [6], theorem 4) that for a general ergodic stationary source S,

$$\lim_{n\to\infty} \frac{\log_2 \mu_S(n, \delta)}{n} = H_S \quad \text{for } 0 < \delta < 1,$$

where H_S is the per-bit entropy of S defined as

$$H_S = -\sum_{j=0}^{2^M-1} p_j \sum_{k=0}^{2^M-1} P_{\Sigma_2|\Sigma_1}(k|j) \log_2 P_{\Sigma_2|\Sigma_1}(k|j), \tag{6}$$

and where p_j is for $0 \leq j \leq 2^M - 1$ defined by (2). In other words, for the general class of ergodic stationary sources, the per-bit entropy H_S is the correct measure of their cryptographic quality when they are used as the secret-key source of a cipher system. Conversely, the per-bit redundancy, $1 - H_S$, is the correct measure of the cryptographic weakness of a key source.

3. Fundamentals of statistical tests

Statistical tests are used to detect a possible statistical defect of a random bit generator, i.e., to detect when the statistical model describing the generator's behavior deviates significantly from a binary symmetric source. Such a test examines a sample sequence of a certain length N and rejects the generator when certain properties of the sample sequence indicate a possible non-randomness (e.g. when the number of 0's and 1's differ considerably). A statistical test T is a function $T : B^N \to \{\text{accept, reject}\}$ which divides the set B^N of binary length N sequences into a (small) set

$$S_T = \{s^N : T(s^N) = \text{reject}\} \subseteq B^N$$

of "bad" sequences and the remaining set of "good" sequences. The two main parameters of a statistical test are the length N of the sample sequence and the

rejection rate $\rho = |S_T|/2^N$, which is the probability that a binary symmetric source is rejected.

Note that a statistical defect of a random bit generator can only be detected with a certain detection probability, which depends on the seriousness of the defect and on the length N of the sample sequence. As in other detection problems, there exists a trade-off between the detection probability and the false alarm probability ρ. In a practical test, ρ should be small, for example $\rho \approx 0.001 \ldots 0.01$.

For reasons of feasibility, a statistical test for a reasonable sample length N cannot be implemented by listing the set of "bad" sequences. Instead, a statistical test T is typically implemented by specifying an efficiently computable function f_T that maps the binary length N sequences to the real numbers $\mathcal{R}$:

$$f_T : B^N \to \mathcal{R} : \; s^N \mapsto f_T(s^N) \; .$$

f_T must be such that the probability distribution of the real-valued random variable $f_T(R^N)$ can be determined, where R^N denotes a sequence of N statistically independent and symmetrically distributed binary random variables. A lower and an upper threshold t_1 and t_2, respectively, can then be specified such that

$$\Pr[f_T(R^N) \leq t_1] \; + \; \Pr[f_T(R^N) \geq t_2] \; = \; \rho \; .$$

Usually $\Pr[f_T(R^N) \leq t_1] \approx \Pr[f_T(R^N) \geq t_2] \approx \rho/2$. The set S_T of bad sequences with cardinality $|S_T| = \rho 2^N$ is thus defined by

$$S_T \; = \; \left\{ s^N \in B^N : \; f_T(s^N) \leq t_1 \text{ or } f_T(s^N) \geq t_2 \right\} .$$

Usually, f_T is chosen such that $f_T(R^N)$ is distributed (approximately) according to a well-known probability distribution, most often the normal distribution or the χ^2 distribution with d degrees of freedom for some positive integer d. Since extensive numerical tables of these distributions are available, such a choice strongly simplifies the specification of t_1 and t_2 for given ρ and N. The normal distribution results when a large number of independent and identically distributed random variables are summed. The χ^2 distribution with d degrees of freedom results when the squares of d independent and normally distributed random variables with zero mean and variance 1 are summed.

As an example, consider the most popular statistical tests for random bit generators, the *frequency test* T_F. It is used to determine whether a generator is biased. For a sample sequence $s^N = s_1, \ldots, s_N$, $f_{T_F}(s^N)$ is defined as

$$f_{T_F}(s^N) = \frac{2}{\sqrt{N}} \left(\sum_{i=1}^{N} s_i - N/2 \right) .$$

The number of 1's in a random sequence $R^N = R_1, \ldots, R_N$ is approximately distributed according to the normal distribution with mean $N/2$ and variance $N/4$ since $E[R_i] = 1/2$ and $\mathrm{Var}[R_i] = 1/4$ for $1 \leq i \leq N$. Thus the probability distribution of $f_{T_F}(R^N)$ is for large enough N well approximated by the normal distribution

with zero mean and variance 1, and reasonable values for the rejection thresholds are $t_2 = -t_1 \approx 2.5 \ldots 3$. The *serial test*, *run test* and *autocorrelation tests* can be defined by similar expressions for the corresponding test functions.

4. The universal entropy-related statistical test

The new statistical test T_U proposed in this section offers two main advantages over all the tests discussed in the previous section:

(1) Rather than being tailored to detecting a specific type of statistical defect, the new test is able to detect any one of the very general class of statistical defects that can be modeled by an ergodic stationary source with finite memory, which includes all those detected by the tests discussed in the previous section and all those that could realistically be assumed to occur in a practical implementation of a random bit generator.

(2) The test measures the actual amount by which the security of a cipher system would be reduced if the tested generator G were used as the key source, i.e., it measures the effective key size $\mu_G(n, \frac{1}{2})$ of a cipher system with key source G. Therefore, statistical defects are weighted according to their actual harm in the cryptographic application.

These two advantages are due to the fact that for the general class of binary ergodic stationary sources with finite memory $M \leq L$, where L is a parameter of the test, the resulting test quantity f_{T_U} is closely related to the per-bit entropy H_S of the source. This claim will be justified after the following description of the test.

The test T_U is specified by the three positive integer-valued parameters L, Q and K. To perform the test T_U, the output sequence of the generator is partitioned into adjacent non-overlapping blocks of length L. The total length of the sample sequence s^N is $N = (Q+K)L$, where K is the number of steps of the test and Q is the number of initialization steps. Let $b_n(s^N) = [s_{Ln}, \ldots, s_{Ln+L-1}]$ for $0 \leq n \leq Q+K-1$ denote the n-th block of length L. For $Q \leq n \leq Q+K-1$, the sequence is scanned for the most recent occurrence of block $b_n(s^N)$, i.e., the least positive integer $i \leq n$ is determined such that $b_n(s^N) = b_{n-i}(s^N)$. Let $A_n(s^N) = i$ if such an i exists and else let $A_n(s^N) = n$. $f_{T_U}(s^N)$ is defined as the average of the logarithm (to the base 2) of the K terms $A_Q(s^N), A_{Q+1}(s^N), \ldots, A_{Q+K-1}(s^N)$. More formally, the test function $f_{T_U} : B^N \to \mathcal{R} : \; s^N \mapsto f_{T_U}(s^N)$ is defined by

$$f_{T_U}(s^N) \;=\; \frac{1}{K} \sum_{n=Q}^{Q+K-1} \log_2 A_n(s^N) \tag{7}$$

where for $Q \leq n \leq Q+K-1$, $A_n(s^N)$ is defined by

$$A_n(s^N) \;=\; \begin{cases} n & \text{if there exists no positive } i \leq n \text{ such that } b_n(s^N) = b_{n-i}(s^N), \\ \min\{i : \; i \geq 1,\, b_n(s^N) = b_{n-i}(s^N)\} & \text{else.} \end{cases} \tag{8}$$

The test can be implemented by using a table (denoted below as Tab) of size 2^L that stores for each L-bit block the time index of its most recent occurrence. The main part of a program implementing the test is given below in a PASCAL-like notation:

```
FOR i := 0 TO 2^L - 1 DO Tab[i] := 0;
FOR n := 0 TO Q - 1 DO Tab[b_n(s^N)] := n;
sum := 0.0;
FOR n := Q TO Q + K - 1 DO BEGIN
    sum := sum + log_2(n - Tab[b_n(s^N)]);
    Tab[b_n(s^N)] := n;
END;
f_TU(s^N) := sum/K;
```

We recommend to choose L between 8 and 16, inclusive, $Q \geq 5 \cdot 2^L$ and K as large as possible (e.g., $K = 10^4$ or $K = 10^5$). This choice for Q guarantees that with high probability, every L-bit pattern occurs at least once in the first Q blocks of a random sequence, and thus that the table of $E[f_{T_U}(R^N)]$ and $\mathrm{Var}[\log_2 A_n(R^N)]$ given below for $Q \to \infty$ (Table I) are suitable for determining the threshold values t_1 and t_2. We also recommend to choose $\rho \approx 0.001 \ldots 0.01$, $t_1 = E[f_{T_U}(R^N)] - y\sigma$ and $t_2 = E[f_{T_U}(R^N)] + y\sigma$, where $\sigma = \sqrt{\mathrm{Var}[\log_2 A_n(R^N)]/K} \approx \sqrt{\mathrm{Var}[f_{T_U}(R^N)]}$ (see Table I) and where y is chosen such that $\mathcal{N}(-y) = \rho/2$. The function $\mathcal{N}(x)$ is the integral of the normal density function and is defined as

$$\mathcal{N}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\xi^2/2} d\xi.$$

A table of $\mathcal{N}(x)$ can be found in almost every book on statistics or probability theory (e.g., see [3], p. 176). For example, to obtain a rejection rate of $\rho = 0.01$ or $\rho = 0.001$, one must choose $y = 2.58$ or $y = 3.30$, respectively. Note that σ decreases as $1/\sqrt{K}$ when K increases.

The definition of T_U is based on the idea, which has independently been suggested by Ziv [9], that a universal statistical test can be obtained by application of a universal source coding algorithm. A generator should pass the test if and only if its output sequence cannot be compressed significantly. However, instead of actually compressing the sample sequence we only need to compute a quantity that is related to the length of the compressed sequence. The formulation of our test was motivated by considering the universal source coding algorithms of Elias [2] and of Willems [7], which partition the data sequence into adjacent non-overlapping blocks of length L. For $L \to \infty$, these algorithms can be shown to compress the output of every discrete stationary source to its entropy. The universal source coding algorithm due to Ziv and Lempel [10] seems to be less suited for application as a statistical test because it seems to be difficult to define a test function f_T such that the distribution of $f_T(R^N)$ can easily be determined, i.e., to specify a concrete implementation of a statistical test based on [10]. No indication of the suitability of the Ziv-Lempel algorithm for a practical implementation of a statistical test is given in [9].

The expectation of $f_{T_U}(R^N)$ and a good approximation to the variance of $f_{T_U}(R^N)$, which are needed in order to describe a practical implementation of the proposed test, are determined in the following under the admissible assumption that $Q \to \infty$. For a source emitting the sequence of random variables $U^N = U_1, U_2, \ldots, U_N$ we have

$$\Pr[A_n(U^N)=i] \;=\; \sum_{b\in B^N} \Pr\left[b_n(U^N)=b,\, b_{n-1}(U^N)\neq b, \ldots, b_{n-i+1}(U^N)\neq b, b_{n-i}(U^N)=b\right].$$

for $i \geq 1$. When the blocks $b_n(U^N)$ for $Q \leq n \leq Q+K-1$ are statistically independent and identically distributed, then the above probability factors:

$$\Pr[A_n(U^N)=i] \;=\; \sum_{b\in B^N} (\Pr[b_n(U^N)=b])^2 \cdot (1-\Pr[b_n(U^N)=b])^{i-1}. \qquad (9)$$

for $i \geq 1$ and $Q \leq n \leq Q+K-1$. For a binary symmetric source we thus have

$$\Pr[A_n(R^N)=i] \;=\; 2^{-L}(1-2^{-L})^{i-1}$$

for $i \geq 1$. The expected value of the sum of random variables equals the sum of their expected values. Therefore

$$E[f_{T_U}(R^N)] \;=\; E[\log_2 A_n(R^N)] \;=\; 2^{-L}\sum_{i=1}^{\infty}(1-2^{-L})^{i-1}\log_2 i \qquad (10)$$

For sufficiently large L (i.e., $L \geq 8$), the terms $A_n(R^N)$ are virtually statistically independent and therefore

$$\begin{aligned} K\cdot \mathrm{Var}[f_{T_U}(R^N)] \;&\approx\; \mathrm{Var}[\log_2 A_n(R^N)] \\ &=\; E[(\log_2 A_n(R^N))^2] - (E[\log_2 A_n(R^N)])^2 \\ &=\; 2^{-L}\sum_{i=1}^{\infty}(1-2^{-L})^{i-1}(\log_2 i)^2 - (E[f_{T_U}(R^N)])^2. \end{aligned} \qquad (11)$$

Table I summarizes $E[f_{T_U}(R^N)]$ and $\mathrm{Var}[\log_2 A_n(R^N)]$ for $1 \leq L \leq 16$. Note that $E[f_{T_U}(R^N)]$ is closely related to the entropy of a block, which is L bits. In fact, it will be shown below that $E[f_{T_U}(R^N)] - L$ converges to the constant -0.8327 as $L \to \infty$.

In order to show that for $L \to \infty$, $E[f_{T_U}(R^N)] - L$ and $\mathrm{Var}[\log_2 A_n(R^N)]$ converge (exponentially fast) to constants, let

$$v(r) \;\stackrel{\text{def}}{=}\; r\sum_{i=1}^{\infty}(1-r)^{i-1}\log_2 i \qquad (12)$$

$$\text{and} \quad w(r) \;\stackrel{\text{def}}{=}\; r\sum_{i=1}^{\infty}(1-r)^{i-1}(\log_2 i)^2. \qquad (13)$$

One can show that

L	$E[f_{T_U}(R^N)]$	$\mathrm{Var}[\log_2 A_n(R^N)]$	L	$E[f_{T_U}(R^N)]$	$\mathrm{Var}[\log_2 A_n(R^N)]$
1	0.73264948	0.690	9	8.17642476	3.311
2	1.53743829	1.338	10	9.17232431	3.356
3	2.40160681	1.901	11	10.1700323	3.384
4	3.31122472	2.358	12	11.1687649	3.401
5	4.25342659	2.705	13	12.1680703	3.410
6	5.21770525	2.954	14	13.1676926	3.416
7	6.19625065	3.125	15	14.1674884	3.419
8	7.18366555	3.238	16	15.1673788	3.421

Table I. Expectation of $f_{T_U}(R^N)$ and variance of $\log_2 A_n(R^N)$ for the test T_U with parameters L, Q and K. For $L \geq 8$, $\mathrm{Var}[f_{T_U}(R^N)]$ is very well approximated by $\mathrm{Var}[\log_2 A_n(R^N)]/K$.

$$\lim_{r\to 0}\left[v(r)+\log_2 r\right] \;=\; \lim_{r\to 0}\int_r^\infty e^{-\xi}\log_2\xi\,d\xi \;\stackrel{\text{def}}{=}\; C \;=\; -0.832746 \tag{14}$$

and

$$\begin{aligned}\lim_{r\to 0}\left[w(r)-(\log_2 r)^2+2C\log_2 r\right] \;&=\; \lim_{r\to 0}\int_r^\infty e^{-\xi}(\log_2\xi)^2 d\xi \\ &\stackrel{\text{def}}{=}\; D \;=\; 4.117181. \end{aligned}\tag{15}$$

Equations (10), (12), (13), (14), (11) and (15) imply that

$$\lim_{L\to\infty}\left(E[f_{T_U}(R^N)]-L\right) \;=\; C$$

$$\text{and}\qquad \lim_{L\to\infty}\mathrm{Var}[\log_2 A_n(R^N)] \;=\; D-C^2 \;=\; 3.423715,$$

which can both be verified numerically by considering Table I.

Let $U^N_{\mathrm{BMS}_p}$ be the output of the binary memoryless source BMS_p. The blocks are independent and thus using (9), (14) and the fact that for $L\to\infty$, $\Pr[b_n(U^N_{\mathrm{BMS}_p}) = b]\to 0$ for all $b\in B^N$ one can show that

$$\lim_{L\to\infty}\left(E[f_{T_U}(U^N_{\mathrm{BMS}_p})]-Lh(p)\right) \;=\; C \tag{16}$$

for $0<p<1$. Equation (16) demonstrates that the test T_U measures the entropy of any binary memoryless source. Table II summarizes $E[f_{T_U}(U^N_{\mathrm{BMS}_p})]$, $Lh(p)+C$ and $\mathrm{Var}[\log_2 A_n(U^N_{\mathrm{BMS}_p})]$ for $L=8$ and $L=16$ and for several values of p. Note that all entries of Tables I and II are computed precisely rather than obtained by simulations.

L	p	$E[f_{T_U}(U^N_{\mathrm{BMS}_p})]$	$Lh(p)+C$	$\mathrm{Var}[\log_2 A_n(U^N_{\mathrm{BMS}_p})]$
8	0.50	7.18367	7.16725	3.239
8	0.45	7.12687	7.10945	3.393
8	0.40	6.95557	6.93486	3.844
8	0.35	6.66617	6.63980	4.561
8	0.30	6.24950	6.21758	5.472
16	0.50	15.16738	15.16725	3.421
16	0.45	15.05179	15.05165	3.753
16	0.40	14.70169	14.70246	4.741
16	0.35	14.09853	14.11234	6.409
16	0.30	13.22556	13.26791	8.614

Table II. Performance of the universal statistical test for the binary memoryless source BMS_p for $L = 8$ and $L = 16$ and for different values of p.

We have devised an algorithm (which is not described here) for computing $E[f_{T_U}(U_S^N)]$ and $\mathrm{Var}[\log_2 A_n(U_S^N)]$ for an arbitrary stationary source S with memory $M \leq L$, where U_S^N is the output sequence of S. For all examples of stationary sources, the very close relation between $E[f_{T_U}(U_S^N)]$ and $LH_S + C$ could be verified, where H_S is the per-bit entropy of S defined by (6). It is possible to prove, by arguments similar to those used in [7], that for every binary ergodic stationary source S,

$$\lim_{L\to\infty} \frac{E[f_{T_U}(U_S^N)]}{L} = H_S.$$

5. Conclusions

A new universal statistical test for random bit generators has been proposed that measures the per-bit entropy of the generator, which has been argued to be the cryptographically significant quality measure for a secret-key source. The test parameter is virtually normally distributed since it is the average of K identically distributed and virtually independent random variables. Its expected value has been shown to be closely related to the per-bit entropy of the generator when it can well be modeled as an ergodic stationary source. For $1 \leq L \leq 16$, expectation and variance of the test parameter have been tabulated for a binary symmetric source. A practical implementation nas been proposed that makes use of these tables to specify the interval of acceptance for the test parameter as the interval between the expected value minus and plus a certain number of standard deviations. An implementation of our statistical test by Omnisec AG for testing random bit generators used in their equipment has confirmed the theoretical results and the practical usefulness.

Acknowledgement

The problem of designing efficient statistical tests for random bit generators was suggested to the author by Omnisec AG, Trockenloostrasse 91, 8105 Regensdorf, Switzerland. In particular, it is a pleasure to thank M. Benninger and P. Schmid for stimulating discussions and for their generous support. I am also grateful to J. Massey for many suggestions improving the presentation of the results, and to H.-A. Loeliger for suggesting an improvement on the original implementation of the presented statistical test.

References

[1] H. Beker and F. Piper, *Cipher Systems*, London: Northwood Books, 1982.

[2] P. Elias, *Interval and recency rank source coding: Two on-line adaptive variable-length schemes*, IEEE Trans. Inform. Theory, vol. IT-33, pp. 3-10, Jan. 1987.

[3] W. Feller, *An Introduction to Probability Theory and its Applications*, third ed., vol. 1, New York, NY: Wiley, 1968.

[4] D.E. Knuth, *The art of computer programming*, vol. 2, 2nd edition, Reading, MA: Addison-Wesley, 1981.

[5] R.A. Rueppel, *Analysis and Design of Stream Ciphers*, New York, NY: Springer, 1986.

[6] C.E. Shannon, *A mathematical theory of communication*, Bell Syst. Tech. J., vol. 27, pp. 379-423, 623-656, Oct. 1948.

[7] F.M.J. Willems, *Universal data compression and repetition times*, IEEE Trans. Inform. Theory, vol. IT-35, pp. 54-58, Jan. 1989.

[8] J.M. Wozencraft and B. Reiffen, *Sequential Decoding*, Cambridge, MA: Techn. Press of the M.I.T., 1960.

[9] J. Ziv, *Compression, tests for randomness and estimating the statistical model of an individual sequence*, in: Sequences (Ed. R.M. Capocelli), New York, NY: Springer Verlag, 1990.

[10] J. Ziv and A. Lempel, *A universal algorithm for sequential data compression*, IEEE Trans. Inform. Theory, vol. IT-23, pp. 337-343, May 1977.