Randomness Properties of Stream Ciphers for Wireless Communications

Benny Y. Zhang and Guang Gong
Department of Electrical and Computer Engineering, University of Waterloo, Canada
{y277zhan, ggong}@uwaterloo.ca

***Abstract*—In this paper we evaluate the randomness properties of four synchronous stream ciphers for use in wireless communications. These ciphers are: SNOW 3G, ZUC, and AES, all of which are a part of the current LTE standard, and the WG-16 stream cipher. The randomness testing is performed with the NIST SP 800-22 test suite, and its results are accordingly tabulated. In addition, we also examine the autocorrelation properties of the keystreams.**

Keywords—randomness testing; NIST; synchronous stream ciphers

I. Introduction

Synchronous stream ciphers are a form of symmetric encryption and are ubiquitous in modern communication systems. In a synchronous stream cipher, each bit of the plaintext is encrypted with the keystream via the XOR operation individually and independently of the previous bits. An important criterion in the evaluation of stream ciphers is its randomness. Random sequences can more effectively remove statistical weaknesses in the plaintext, resulting in a much more secure transmission. The extent to which the keystream resembles a truly random source may be quantified with a statistical measure of randomness.

Several tests of randomness exist in the literature. Golomb originally proposed three randomness properties known as Golomb's postulates [1]. They are: the balanced property, the runs property, and the autocorrelation property. More extensive properties of randomness have since been developed, such as DIEHARD [2], and the NIST test suite [3]. We chose to evaluate the stream ciphers using the NIST SP 800-22 statistical test suite. The suite consists of 15 tests, namely: frequency, block frequency, runs, longest run, matrix rank, spectral (DFT), non-overlapping template, overlapping template, a "Universal Statistical" test based on compressibility, linear complexity, serial, approximate entropy, cumulative sums, and two random excursions tests.

The synchronous stream ciphers to be tested currently form the core of the 3GPP algorithms. Implementations of SNOW 3G, ZUC, and AES are readily available online. SNOW 3G forms a part of the UEA2 and UIA2 specifications in 3GPP [4]; ZUC belongs to the EEA3 and EIA3 specifications [5]. The implementation of WG-16 is detailed in this source [6]. AES has also been previously specified [7].

II. Statistical Randomness Testing

A. Golomb's Postulates of Randomness

Given a sequence $a = \{a_o, a_1, \dots, a_{N-1}\}$, $a_i \in \{0, 1\}$, the Golomb randomness properties proposed are specified below.

Balanced Property: this property states that the number of zeros in a period is nearly equal to the number of ones. Given a binary keystream of period N, the sum should be bounded as

$$\left| \sum_{i=0}^{N-1} (-1)^{a_i} \right| \leq 1. \tag{1}$$

The equality in (1) holds when N is odd.

Runs Property: this property states that in a period, half of all the runs (a run is a substring of consecutive ones or zeros) are of length 1, one fourth are of length 2, and so on. Moreover, for each length, the runs of zeros and ones are equal. Accordingly, the expected number of runs over all lengths will be approximately $N/2$.

Autocorrelation Property: this property is a re-statement of the ideal 2-level autocorrelation. The aperiodic autocorrelation of a binary sequence is given as

$$C_a(\tau) = \sum_{i=0}^{N-\tau-1} (-1)^{a_{i+\tau}+a_i}. \tag{2}$$

The property holds when $C_a(\tau) = N$ for $\tau = 0 \ (mod\ N)$, and $C_a(\tau) = K$ for $\tau \neq 0 \ (mod\ N)$, where K is a constant. The ideal 2-level correlation requires that $K = -1$ for odd N and $K = 0$ for even N.

These properties are often insufficient to fully quantify the randomness of sequences. A more extensive test is given by the NIST suite.

B. NIST Statistical Testing

The NIST test suite is based on the method of statistical hypothesis testing. The approach is to test a null hypothesis H_0. The H_0, as defined by NIST, is that the given sequence is random. A standard parameter, α, is then defined as the level of significance of the test. Each test will generate a real number called the *p*-value which ranges from 0 to 1, such that,

- If $p \leq \alpha$, then H_0 is rejected; the sequence is not random.
- If $p > \alpha$, then H_0 is accepted; the sequence is random.

In addition, α is also the probability of a false rejection, that is, the probability that we claim the sequence as non-random when it actually is. In NIST terminology, this scenario is a Type I Error. In all of our NIST tests, α was set to 0.01.

A short description of each NIST test is given below.

1) *Frequency*: balanced property.
2) *Block Frequency*: balanced property for *M*-sized blocks (subsequences) of the keystream.
3) *Runs Test*: tests whether number of runs is expected.
4) *Longest run in a block*: longest run of 1's in a block.
5) *Binary Matrix Rank*: checks for linear dependency among substrings of the sequence.
6) *DFT (Spectral)*: checks for percentage of peak values in the frequency components of the sequence.
7) *Non-Overlapping Template*: tests how many times a pre-defined template occurs in the sequence.
8) *Overlapping Template*: same as *7)*, except use a different windowing operation.
9) *Universal Statistical Test*: tests for the compressibility of the sequence.
10) *Linear Complexity*: checks for LFSR complexity by using the Berlekamp-Massey algorithm.
11) *Serial Test*: frequency of all *m*-bit patterns.
12) *Approximate Entropy*: compares frequency of *m* and *m*+1 blocks against a normally distributed sequence.
13) *Cumulative Sums*: maximum excursion from orgin of a random walk formed by the sequence.
14) *Random Excursions*: tests for number of cycles with *K* visits to origin in cumulative sum random walk.
15) *Random Excursions Variant*: number of times that a state is visited in a cumulative sum random walk.

III. TEST METHODOLOGY

A. Autocorrelation

The NIST statistical testing already covered the balanced and run properties. For each sequence, we found the worst-case normalized autocorrelation value, C_{max}, given as

$$C_{\max} = \max\left(\left|\frac{C_a(\tau)}{\sqrt{N-\tau}}\right|\right), \qquad \tau \neq 0. \tag{3}$$

To evaluate C_{max} for the keystreams, we used 100 random 128-bit key and IV pairs to generate 100 keystreams, and calculated the average C_{max}. The autocorrelation was run for lengths from 2^{18} to 2^{23}, or up to 1 megabyte (MB).

TABLE I. WORST-CASE APERIODIC AUTOCORRELATION

Stream Cipher	**Keystream Lengths**					
	2^{18}	2^{19}	2^{20}	2^{21}	2^{22}	2^{23}
WG-16	4.732	4.836	5.025	5.126	5.262	5.374
SNOW 3G	4.731	4.903	5.008	5.131	5.250	5.369
ZUC	4.753	4.858	4.972	5.158	5.224	5.413
AES	4.683	4.936	4.964	5.166	5.257	5.352

The AES, being a block cipher, was operated in counter mode to generate a stream cipher. A nonce (IV) was XOR-ed with a counter starting from zero. The counter was incremented by one for each 128-bit block of the AES cipher. The resulting 128-bit ciphertexts were concatenated to form the required 2^{18} to 2^{23} length keystream.

Due to the closeness of the worst-case autocorrelation values for any given length *N*, it appears that all the stream ciphers are sufficiently random for *N* of up to 1 MB.

B. NIST Randomness Testing

The NIST statistical suite used several user-defined parameters, which are detailed below. We used 100 pairs of keys and IVs to generate 100 keystreams of length 2^{20}.

- *Block Frequency*: given sequence length n, the block size M must be that, $M \geq 0.01n$, $n \geq MN$, and $N < 100$. For $n = 1048576$, we chose $M = 16384$
- *Templates Tests*: the default template size for non-overlapping and overlapping is $m = 9$. We kept this value in both tests.
- *Approximate Entropy Test*: given sequence length n, we need block length m such that $m < \lfloor \log_2 n \rfloor - 5$. We chose $m = 10$.
- *Serial Test*: we need to select block length m such that we have $m < \lfloor \log_2 n \rfloor - 2$. We chose $m = 16$.
- *Linear Complexity*: we need sequence length $n \geq 10^6$. Additionally, $500 \leq M \leq 5000$. In order to not discard any bits, we chose M to be a power of 2, as $M = 512$.

Provided that we have a length $n \geq 10^6$ and number of sequences $S \geq 55$, all the results of NIST is regarded as statistically valid. For us, having $n = 2^{20}$ and $S = 100$ satisfied these conditions.

Since we will have 100 *p*-values for each test, NIST further calculates a second statistic to measure the uniformity of those *p*-values. Non-uniformity implies non-randomness. To do this, the interval from 0 to 1 is partitioned into 10 bins, and a goodness-of-fit is performed on all 100 *p*-values to see if they are uniformly distributed into those bins. This second statistic is also a *p*-value, and must be greater than 0.0001 for the uniformity assumption to hold.

Finally, we gave AES highly correlated input as the plaintext to test the non-linear filtering properties of AES. To do this, we used a primitive polynomial of $f(x) = x^{20} + x^3 + 1$. This is periodic with period $2^{20} - 1$. The internal bits and output bits of the 20-degree LFSR were collected to form the 128-bit plaintext. After encryption, one keystream was formed by concatenating the first bit of each 128-bit ciphertext, then another keystream by concatenating the second bit, and so on. In all, we formed 100 keystreams, each of length $n = 10^6$.

Note that non-overlapping templates actually consists of 148 tests, for each of the 148 templates corresponding to $m = 9$. Also, there are 8 tests for random excursion (corresponding to *K* values of -4, -3, -2, -1, 1, 2, 3, 4), and 18 for random excursion variant (for states from -9 to 9). The *p*-values for those tests were averaged in our results.

TABLE II. NIST STATISTICAL TESTING RESULTS

Test Name	***Pass Rate***				
	WG-16	***SNOW***	***ZUC***	***AES***	***AES LFSR***[a]
Frequency	100	99	100	100	100
Block Frequency	99	97	99	100	99
Runs Test	99	98	99	99	100
Longest Run in Block	99	98	100	99	98
Matrix Rank	99	100	100	99	100
DFT (Spectral)	100	100	100	100	98
Non-Overlapping	98.8	98.8	99.0	98.9	99.1
Overlapping	99	99	99	99	98
Universal Statistical	100	99	100	99	100
Linear Complexity	99	99	99	100	100
Serial Test	98	100	99	99	98
Approximate Entropy	99	99	97	99	97
Cumulative Sums	100	99	100	100	99
Random Excursions	98.6	98.3	98.9	99.0	98.8
Random Excursions Variant	98.9	98.7	98.6	99.7	98.4

[a] AES LFSR is AES with correlated degree-20 LFSR input

The minimum pass-rate defined by NIST, for a sample size of $S = 100$, is approximately 96. This lower bound is defined from a 99% confidence interval, given that we expect, on average, 99 keystreams to pass. This expectation follows from the level of significance, α, being 0.01, meaning that 1 in 100 keystreams was known to fail. The formula for minimum pass rate percentage is given below [3].

$$\textit{Pass Rate} = (1-\alpha) - 3\sqrt{\frac{\alpha(1-\alpha)}{S}} \tag{4}$$

In our experiment, $\alpha = 0.01$, and $S = 100$, giving us *Pass Rate* as approximately 96. Finally, all 100 p-values of every test for all the ciphers were found to be uniformly distributed.

IV. CONCLUSION

In this paper we have presented several techniques to measure the randomness of a sequence. We found the aperiodic autocorrelation and the NIST statistical test results of WG-16, ZUC, SNOW 3G, and AES. It was found that for lengths of around 2^{20} all ciphers were able to produce keystreams that are indistinguishable from a random source.

ACKNOWLEDGMENT

The first author wishes to acknowledge the ComSec group for their support and resources in the performance of these experiments.

REFERENCES

[1] L. Chen, and G. Gong, *Communication System Security*. Boca Raton, FL: CRC Press, 2012.

[2] G. Marsaglia. *DIEHARD Battery of Tests of Randomness* [Online]. Available: *http://www.stat.fsu.edu/pub/diehard*

[3] National Institute of Standards and Technology, “A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications”, Special publication 800-22, April 2010.

[4] 3GPP. Specification of the 3GPP Confidentiality and Integrity Algorithms UEA2 & UIA2. Document 2: SNOW 3G Specification. September 2006.

[5] 3GPP. Specification of the 3GPP Confidentiality and Integrity Algorithms 128-EEA3 & 128-EIA3. Document 2: ZUC Specification. January 2011.

[6] X. X. Fan, T. Wu, and G. Gong, “Specification of the Stream Cipher WG-16 Based Confidentiality and Integrity Algorithms”, Technical Report, University of Waterloo, CACR 2013-06, 2013.

[7] National Institute of Standards and Technology, “Annoucing the Advanced Encryption Standard (AES)”, Federal Information Processing Standards Publication 197, November 2001.