Testing the reliability of a TRNG at run time

Florian Caullery

Institut de Mathématiques de Marseille
Aix-Marseille Université

BIRS workshop on Mathematics of Communications: Sequences, Codes and Designs, Banff, 25–30 January 2015

INSTITUT de MATHÉMATIQUES de MARSEILLE

Outline

- An embedded test for a TRNG
  - What is a TRNG ?
  - What is it used for and why should we test it ?
  - What are the constraints ?
- Classical measures of pseudo-randomness
  - Knuth's measure
  - Mauduit-Sarkozy measures
- A test based on Boolean functions
  - Nonlinearity based test
  - Autocorrelation based test

INSTITUT de MATHÉMATIQUES de MARSEILLE

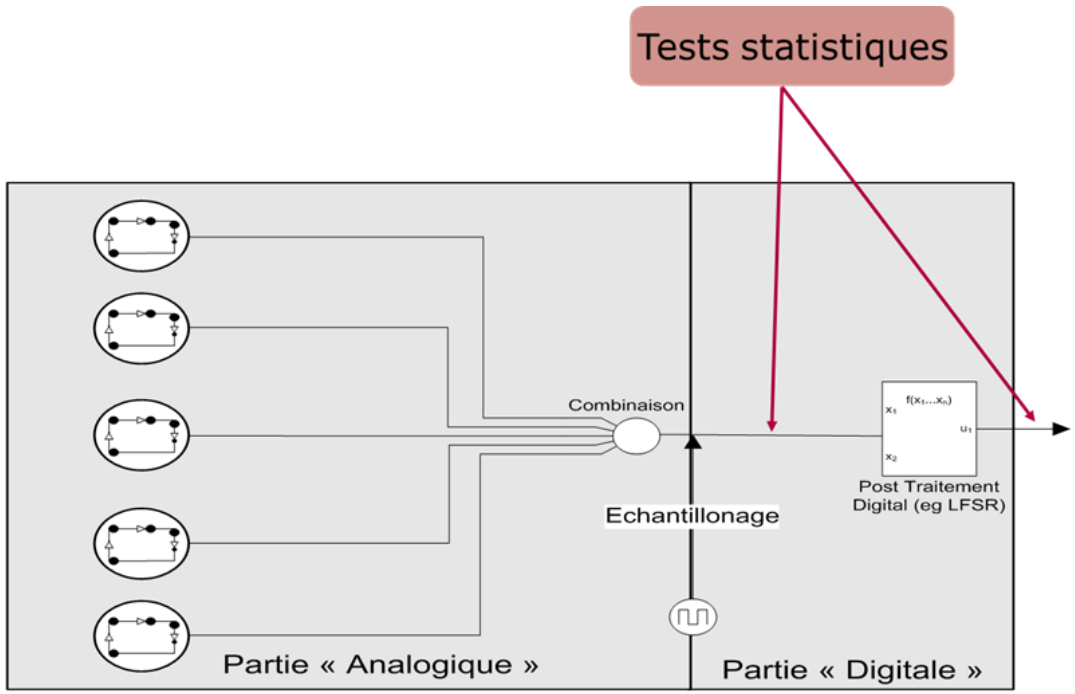
What is a TRNG ?

- TRNG stands for : **T**rue **R**andom **N**umbers **G**enerator
- An electronic device producing random string of bits
- Why is that **True** and not **Pseudo** ?
- It relies on a physical method (hardware) :
  - Quantum random properties (e.g. nuclear decay, photons traveling through a semi-transparent mirror, etc.)
  - Physical noises (e.g. heat, Power supply and clock glitches, etc.)

INSTITUT de MATHÉMATIQUES de MARSEILLE

What is a TRNG ?

- An example :



INSTITUT de MATHÉMATIQUES de MARSEILLE

What is it used for ?

- Mainly security applications :
  - Nonce in protocols
  - Keys in cryptography
  - Counter measures to physical attacks
- Simulations
- Video games

INSTITUT de MATHÉMATIQUES de MARSEILLE

How is it tested ?

- A good TRNG should produce bits sequences which should be :
  - Non predictable : the bits should be *iid* variables
  - Non manipulable : resistant to attacks and no context should make it predictable
- Test at the end of the production :
  - Classical statistical tests on string of length 10^4 or 10^5
  - Life cycle tests
- **No embedded test !**

INSTITUT de MATHÉMATIQUES de MARSEILLE

What are the constraints for the embedded test ?

- Low computational capabilities
- Low memory resources
- A TRNG is a very slow working component
- This implies :
  - Only short bits strings can be tested ($\leq 10^3$)
  - No statistical method can be employed
  - The test should be "on the fly"
  - The test should detect the attacks
- The test cannot be perfect

INSTITUT de MATHÉMATIQUES de MARSEILLE

Multiple points of view

- No finite sequences can be "random" : we'll talk of pseudo-random (PR) sequences
- Multiple ways to see binary sequences :
  - Bits sequences
  - Binary sequences (i.e. $\{-1,1\}^N$ sequence)

Classical measures of pseudo-randomness

- Knuth [81]
- Mauduit - Sarkozy [97] :
  - Normality measure
  - Well-distribution measure
  - Correlation measure
- All these measures derive from the infinite binary sequences case

INSTITUT de MATHÉMATIQUES de MARSEILLE

Measures in infinite case

- Let $E_N = (e_1, e_2, \ldots) \in \{-1, 1\}^\infty$, $k, M \in \mathbb{N}$, $X \in \{-1, 1\}^k$, $a \in \mathbb{Z}$, $b \in \mathbb{N}$, $D = (d_1, \ldots, d_k) \in \mathbb{N}^k$, $d_1 < d_2 < \ldots < d_k$.
- Define the following quantities :

$$T(E, M, X) = |\{0 \leq n \leq M, (e_{n+1}, e_{n+2}, \ldots, e_{n+k}) = X\}|,$$

$$U(E, M, a, b) = \sum_{j=1}^{M} e_{a+jb},$$

and

$$V(E, M, D) = \sum_{n=0}^{M-1} e_{n+d_1} e_{n+d_2} \ldots e_{n+d_k}.$$

INSTITUT de MATHÉMATIQUES de MARSEILLE

Knuth's measure

- For Knuth :
  A finite binary sequence $E_N = (e_1, \ldots, e_N) \in \{-1, 1\}^N$ is said PR if for all $k \in \mathbb{N}$ with

$$k \leq \frac{\log N}{\log 2},$$

  and for all $X \in \{-1, 1\}^k$ we have

$$|T(E_N, N + 1 - k, X) - \frac{N + 1 - k}{2^k}| \leq \frac{1}{\sqrt{N}}.$$

INSTITUT de MATHÉMATIQUES de MARSEILLE

Knuth's measure

- A sequence is PR or not
- We cannot adapt this measure to different contexts
- It has a high computational complexity

INSTITUT de MATHÉMATIQUES de MARSEILLE

Mauduit Sarkozy measure 1 : Normality measure

- The normality measure of order k of a sequence $E_N \in \{-1,1\}^N$ is defined as :

$$N_k(E_N) = \max_{X\in\{-1,1\}^k} \max_{0<M\leq N+1-k} |T(E_N, M, X) - M/2^k|.$$

- Its normality measure is :

$$N(E_N) = \max_{k\leq \log N/\log 2} N_k(E_N).$$

INSTITUT de MATHÉMATIQUES de MARSEILLE

Mauduit Sarkozy measure 2 : Well-distribution measure

- The well-distribution measure of a sequence $E_N \in \{-1,1\}^N$ is defined as :

$$W(E_N) = \max_{a,b,t} |U(E_N,t,a,b)| = \max_{a,b,t} |\sum_{j=1}^{t} e_{a+jb}|$$

where the maximum is taken over all a, b, t such that $1 \leq a+b \leq a+tb \leq N$.

INSTITUT de MATHÉMATIQUES de MARSEILLE

Mauduit Sarkozy measure 3 : Correlation measure

- The correlation measure of order k of a sequence $E_N \in \{-1,1\}^N$ is defined as :

$$C_k(E_N) = \max_{M,D} |V(E_N, M, D)| = \max_{M,D} |\sum_{n=0}^{M-1} e_{n+d_1} e_{n+d_2} \dots e_{n+d_k}|.$$

- Its correlation measure is :

$$C(E_N) = \max_{k \leq log(N)/log2} C_k(E_N).$$

INSTITUT de MATHÉMATIQUES de MARSEILLE

Mauduit Sarkozy measures

- A sequence is considered PR if both well-distribution and correlation measures are small (i.e. less than $N \log(N)$)
- More modulation in the measures
- Computational complexity is again too high !

INSTITUT de MATHÉMATIQUES de MARSEILLE

Boolean function point of view

- Take a bit sequence of length $N = 2^n$
- See it as the truth table of a Boolean function $f : \mathbb{F}_{2^n} \to \mathbb{F}_2$:

e_0	e_1	e_2	$\dots$	e_{n-1}
0	1	1	$\dots$	1
$f(0)$	$f(1)$	$f(2)$	$\dots$	$f(n-1)$
$f(0,\dots,0)$	$f(0,\dots,1)$	$f(0,\dots,1,0)$	$\dots$	$f(1,\dots,1)$

- Now the question is : **is this boolean function random ?**
- We'll use the **nonlinearity** and **Absolute indicator** of f

INSTITUT de MATHÉMATIQUES de MARSEILLE

Nonlinearity of random Boolean functions

- The nonlinearity of a Boolean function is defined as :

$$NL(f) = 2^{n-1} - \frac{1}{2} \max_{v \in \mathbb{F}_2^n} \left| \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x)+v.x} \right|$$

- Main fact [Dib, Halász, Rodier, Schmidt] : Choose a random Boolean function $f : \mathbb{F}_2^n \to \mathbb{F}_2$, then

$$NL(f) \to 2^{n-1} - \sqrt{2^{n-1} \log 2}$$

almost surely when $n \to \infty$.

INSTITUT de MATHÉMATIQUES de MARSEILLE

First test idea

- Fix your n (e.g. 9 or 10)
- Compute the range of the expected nonlinearity of random Boolean functions
- Compute the nonlinearity of the function defined by the sequence
- If this value is in the expected values range of the nonlinearity, declare your sequence PR

INSTITUT de MATHÉMATIQUES de MARSEILLE

Pro and cons

- The complexity of the test is $O(N)$!
- One can change the value range i.e. the test is adaptable
- **Question 1** : how do we determine the value range ?
- **Question 2** : Is there any links between this measure and the classical ones ?
- **Question 3** : Does the fact that we fixed the order of the truth table affect the test ?
- **Question 4** : Does this test work in real life ?

INSTITUT de MATHÉMATIQUES de MARSEILLE

Testing the test 1

- We first focused on sequences of length $2^9 = 512$.
- Computing the nonlinearity of $50,000$ random Boolean functions on Sage :

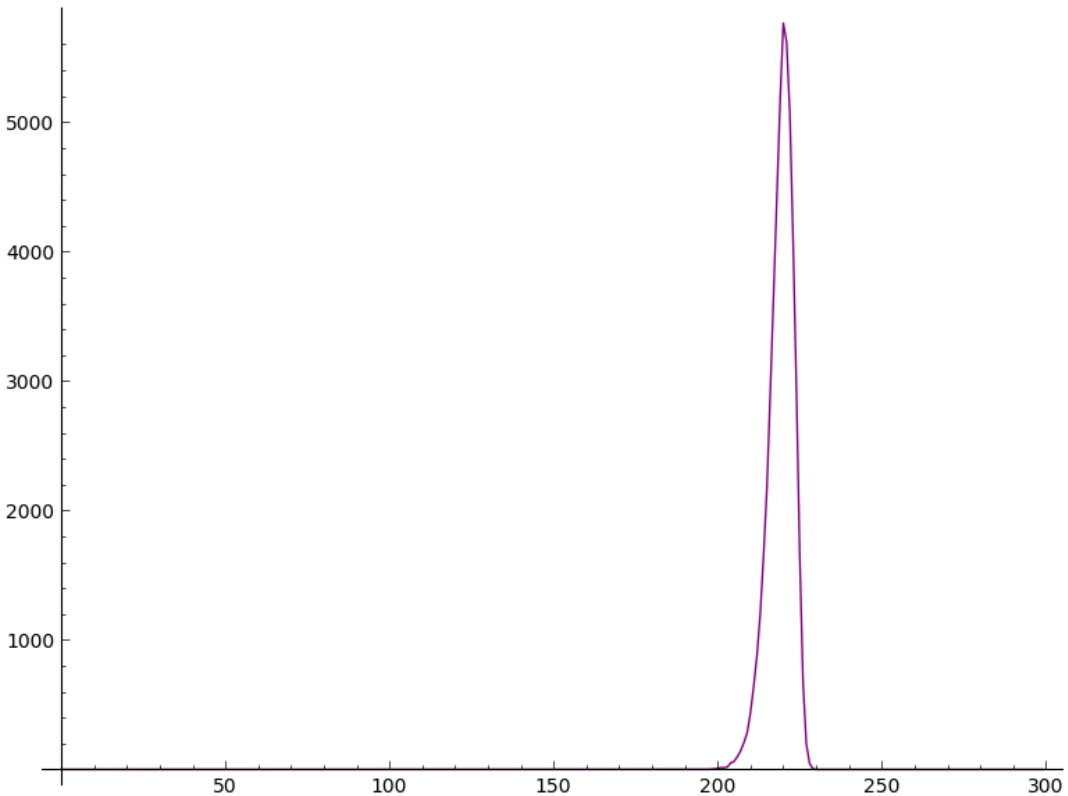
Mean	219.1854
Median	220
Variance	15.1837
Standard deviation	3.8966

Numerical experiments on nonlinearity

- Detects sequences with period up to 260

INSTITUT de MATHÉMATIQUES de MARSEILLE

Testing the test 1 - 2



Nonlinearity distribution

INSTITUT de MATHÉMATIQUES de MARSEILLE

Absolute indicator of random Boolean functions

- Let $f : \mathbb{F}_2^n \mapsto \mathbb{F}_2$ be a Boolean function. The auto-correlation function AC_f of f is defined as

$$AC_f(a) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x)+f(x+a)}.$$

- Let $f : \mathbb{F}_2^n \mapsto \mathbb{F}_2$ be a Boolean function. The absolute indicator $AI(f)$ of f is defined as

$$AI(f) = \max_{a \in \mathbb{F}_2^n - \{0\}} AC_f(a).$$

INSTITUT de MATHÉMATIQUES de MARSEILLE

Absolute indicator of random Boolean functions

- Same test than before but compute the absolute indicator instead of nonlinearity
- The complexity is $O(N^2)$
- **Question 1** : how do we determine the value range ?
- **Question 2** : Is there any links between this measure and the classical ones ?
- **Question 3** : Does this test work in real life ?
- **Question 4** : Is it complementary with test 1 ?

INSTITUT de MATHÉMATIQUES de MARSEILLE

Testing the test 2

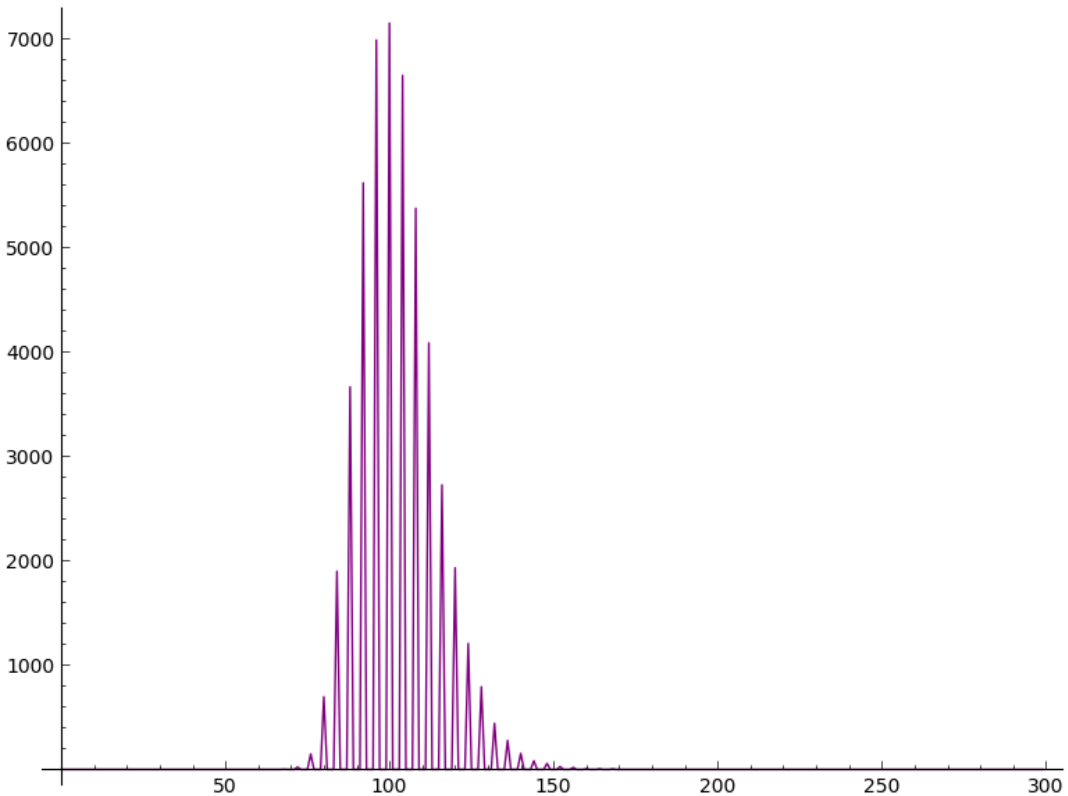
- We focused on sequences of length $2^9 = 512$. We computed the absolute indicator of $10,000$ random Boolean functionson Sage :

Mean	100.0625
Median	100
Standard deviation	24.1242

Numerical experiments on absolute indicator

- We could detect sequences with period up to 280

Testing the test 2 - 2



Absolute indicator distribution

INSTITUT de MATHÉMATIQUES de MARSEILLE

Testing the tests

- Next step : test the real TRNG
- Get the values and disturb them to check the test tolerance
- Solve the theoretical questions

INSTITUT de MATHÉMATIQUES de MARSEILLE

Informations

Florian Caullery
Institut de Mathematiques de Marseille
Aix Marseille Université
www.univ-amu.fr

Contact : fcaullery@gmail.com
@presquepartout
http ://presquepartout.hypotheses.org

INSTITUT de MATHÉMATIQUES de MARSEILLE