

On the practical secrecy of keys formed from instantaneous values of a stationary Gaussian process

Bogdanov Dmitriy Sergeevich

Lomonosov Moscow State University

June 2, 2025

Practical secrecy

In theory

Cryptographic keys

- Key bits are considered independent and equally probable
- Ensures maximum security.

Practical secrecy

In theory

Cryptographic keys

- Key bits are considered independent and equally probable
- Ensures maximum security.

In practice

- Implementations may not meet these ideals — bits can be dependent or have non-uniform distributions.
- This leads to the concept of "practical secrecy of key"

Practical secrecy

What is practical key secrecy?

It is the average number of attempts needed to find the correct key.

Founder

- Arbekov, I.M., «Key Secrecy Criteria», Mat. Vopr. Kriptogr., 7:1 (2016), 39–56.
- Arbekov, I.M., «Lower bounds for the practical secrecy of a key», Mat. Vopr. Kriptogr., 8:2

Mathematical Model of PRNG

Article

A. S. Logachev, V. O. Mironkin, “On the Influence of Probabilistic Characteristics of Discrete Sources Generating Cryptographic Keys on Practical Key Secrecy,” PDM, 2024, No. 65, pp. 66–83.

Mathematical Model of Physical Random Number Generators

- The proposed model accounts for **non-stationarity** of such devices.
- It also considers the **presence of statistical dependencies** between generated values.

Mathematical Model of PRNG

Probabilistic characteristic determining the practical secrecy of the key

A value ε such that, for any $k \in \mathbb{N}$, for any distinct $t_1, t_2, \dots, t_k \in \mathbb{Z}$, and for any $x_1, x_2, \dots, x_k \in \{0, 1, \dots, m-1\}$,

$$\begin{aligned} \left(\frac{1}{m} - \varepsilon\right)^k &\leq \\ &\leq \mathbf{P}_{t_1, t_2, \dots, t_k}(X_{t_1} = x_1, X_{t_2} = x_2, \dots, X_{t_k} = x_k) \leq \\ &\leq \left(\frac{1}{m} + \varepsilon\right)^k. \quad (1) \end{aligned}$$

Scheme of instantaneous values

Our goal

Find an upper bound for ε in physical random number generators constructed based on the scheme of instantaneous values.

What is the scheme of instantaneous values?

D. S. Bogdanov, A. S. Logachev, V. O. Mironkin, "Theoretical Probabilistic Models of Physical Random Number Generators," Problems of Information Security. Computer Systems, 61(3) (2024), pp. 9–19.

Scheme of instantaneous values

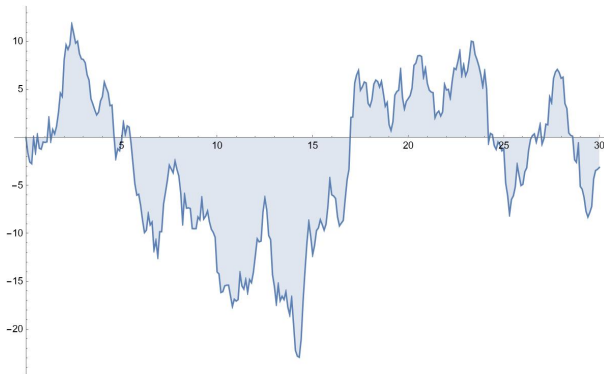
Definition

Let $\{X_t, t \in \mathbb{R}\}$ be a stochastic process with values in $\mathbb{R}$ over the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathcal{F}$ is the Borel σ -algebra on $\mathbb{R}^{\mathbb{R}}$. Let $T \in \mathbb{R}$. Then, denote by $\{\gamma_n, n \in \mathbb{Z}\}$ the random variables defined as

$$\gamma_n = g(X_{nT}), \quad (2)$$

where $g : \mathbb{R} \rightarrow \{0, 1, \dots, m-1\}$ is a measurable function.

Scheme of instantaneous values



Three stages

Three stages

We will achieve the goal in three stages.

- 1 Derive conditions under which the model assumptions hold — for all physical random number generators (PRNGs).
- 2 We will obtain bounds on ε , assuming the implementation of the PRNG based on a scheme of instantaneous values of a stationary Gaussian process
- 3 Show how to apply the results to specific functions g .

Mathematical Model of Physical Random Number Generators

What are the disadvantages of this model?

$$\begin{aligned} \left(\frac{1}{2} - \varepsilon\right)^k &\leq \\ &\leq \mathbf{P}_{t_1, t_2, \dots, t_k}(X_{t_1} = x_1, X_{t_2} = x_2, \dots, X_{t_k} = x_k) \leq \\ &\leq \left(\frac{1}{2} + \varepsilon\right)^k. \end{aligned}$$

Mathematical Model of Physical Random Number Generators

What are the disadvantages of this model?

- The first is the multivariate distribution function, which is difficult to work with.
- The second is the need to study all possible finite subsets of integers. There are many of them.

Mathematical Model of Physical Random Number Generators

What are the disadvantages of this model?

- The first is the multivariate distribution function, which is difficult to work with.
- The second is the need to study all possible finite subsets of integers. There are many of them.

What are we going to do?

Let's tackle the first disadvantage!

Sufficient conditions

Sufficient conditions

Expression (1) is evidently satisfied if, for any $k \in \mathbb{N}$, for all pairwise distinct $t_1, t_2, \dots, t_k \in \mathbb{Z}$, and for all $x_1 \in \{0, 1, \dots, m-1\}$,

$$\left| \mathbf{P}(\gamma_{t_1} = x_1 | \gamma_{t_2}, \dots, \gamma_{t_k}) - \frac{1}{m} \right| \leq \varepsilon. \quad (3)$$

Here $\mathbf{P}(\gamma_{t_1} = x_1 | \gamma_{t_2}, \dots, \gamma_{t_k})$ is a random variable equal to $\mathbf{P}(\gamma_{t_1} = x_1 | \gamma_{t_2} = x_2, \dots, \gamma_{t_k} = x_k)$ for those $\omega \in \Omega$ such that $\gamma_{t_2}(\omega) = x_2, \gamma_{t_3}(\omega) = x_3, \dots, \gamma_{t_k}(\omega) = x_k$, where $x_2, \dots, x_k \in \{0, 1, \dots, m-1\}$.

New Problems

From Distributions to Random Variables

The condition

$$\left| \mathbb{P}(\gamma_{t_1} = x_1 \mid \gamma_{t_2}, \dots, \gamma_{t_k}) - \frac{1}{m} \right| \leq \varepsilon.$$

can be violated on sets of measure zero.

Solution

We move to the "essential supremum".

Essential supremum

Definition

The essential supremum of a random variable ξ , denoted by $\text{vraisup } \xi$, is the smallest number such that

$$\mathbb{P}(\omega : \xi(\omega) > \text{vraisup } \xi) = 0,$$

and for any $\varepsilon > 0$,

$$\mathbb{P}(\omega : \xi(\omega) > \text{vraisup } \xi - \varepsilon) > 0.$$

Lemma

What are we going to do?

Let's tackle the second disadvantage!

Lemma

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let ξ be a random variable on it. Suppose $\mathcal{B}_1, \mathcal{B}_2$ are two σ -algebras such that $\mathcal{B}_2 \subseteq \mathcal{B}_1 \subseteq \mathcal{F}$. Then

$$\sup_{\omega \in \Omega} \mathbb{E}[\xi \mid \mathcal{B}_2] \leq \sup_{\omega \in \Omega} \mathbb{E}[\xi \mid \mathcal{B}_1]. \quad (4)$$

Final conditions

Proposition

If

$$\sup_n \sup_{\omega \in \Omega} \max_{0 \leq r \leq m-1} \left| \mathbb{P}(\gamma_n = r \mid \gamma_{n+1}, \gamma_{n-1}, \gamma_{n+2}, \gamma_{n-2}, \dots) - \frac{1}{m} \right| \leq \varepsilon,$$

then the inequalities (1) hold for any $k \in \mathbb{N}$, for all pairwise distinct $t_1, t_2, \dots, t_k \in \mathbb{Z}$, and for all $x_1, x_2, \dots, x_k \in \{0, 1, \dots, m-1\}$.

Scheme of instantaneous values

Remark

Take an arbitrary $n \in \mathbb{Z}$ and $r \in \{0, \dots, m-1\}$. By the definition (2), $\gamma_i = g(X_{iT})$. Then, the σ -algebra $\mathcal{B}_1^{(n)}$, generated by the random variables $\gamma_{n+1}, \gamma_{n-1}, \gamma_{n+2}, \gamma_{n-2}, \dots$, is contained in the σ -algebra $\mathcal{B}^{(n)}$, generated by the random variables $X_{(n+1)T}, X_{(n-1)T}, X_{(n+2)T}, X_{(n-2)T}, \dots$. Then we obtain that

$$\begin{aligned} & \sup_n \sup_{\omega \in \Omega} \max_{r=0}^{m-1} \left| \mathbf{P}(\gamma_n = r | \gamma_{n+1}, \gamma_{n-1}, \gamma_{n+2}, \gamma_{n-2}, \dots) - \frac{1}{m} \right| \leq \\ & \leq \sup_n \sup_{\omega \in \Omega} \max_{r=0}^{m-1} \left| \mathbf{P}(\gamma_n = r | X_{(n+1)T}, X_{(n-1)T}, \dots) - \frac{1}{m} \right|. \end{aligned}$$



Scheme of instantaneous values

Remark

Moreover

$$\begin{aligned} \mathbf{P}(\gamma_n = r | X_{(n+1)T}, X_{(n-1)T}, X_{(n+2)T}, X_{(n-2)T}, \dots) = \\ = \mathbf{P}(f(X_{nT}) = r | \mathcal{B}^{(n)}) = \mathbf{P}(X_{nT} \in f^{-1}(r) | \mathcal{B}^{(n)}) . \end{aligned} \quad (5)$$

Stationary Gaussian Random Processes

Notations

Denote by $(X_t, t \in \mathbb{R})$ a stationary in the wide sense Gaussian random process with real values over the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with $\mathbb{E}[X_t] = 0$ for all $t \in \mathbb{R}$. Let $R(t)$ denote the covariance function of the process X_t .

Stationary Gaussian Random Processes

Notations

Denote by $(X_t, t \in \mathbb{R})$ a stationary in the wide sense Gaussian random process with real values over the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with $\mathbb{E}[X_t] = 0$ for all $t \in \mathbb{R}$. Let $R(t)$ denote the covariance function of the process X_t .

Remark

The conditional distribution of the random variable X_{nT} given $\mathcal{B}^{(n)}$

$$\mathbb{P}\left(\{\omega : X_{nT} \in B\} \mid \mathcal{B}^{(n)}\right),$$

can be explicitly determined.

Main Theorem

Theorem

For any $\omega \in \Omega$, the conditional distribution of X_{nT} given $\mathcal{B}^{(n)}$ is a Gaussian distribution with mean

$$\mathbb{E} \left[X_{nT} \mid \mathcal{B}^{(n)} \right] (\omega)$$

and variance

$$\mathbb{E} \left[\left(X_{nT} - \mathbb{E} \left[X_{nT} \mid \mathcal{B}^{(n)} \right] \right)^2 \right].$$

Interpolation Problem for a Random Sequence

Remark

The random variable

$$\mathbb{E} \left[X_{nT} \mid \mathcal{B}^{(n)} \right]$$

is the **projection** (interpolation) of X_{nT} onto the space generated by the random variables $X_{(n+1)T}, X_{(n-1)T}, X_{(n+2)T}, X_{(n-2)T}, \dots$

Moreover,

$$\mathbb{E} \left[\left(X_{nT} - \mathbb{E} \left[X_{nT} \mid \mathcal{B}^{(n)} \right] \right)^2 \right]$$

is called the "interpolation error."

Solution to the Interpolation Problem

Under certain conditions on the spectral density function $F_T(\lambda)$ of the sequence X_{nT} , the following holds:

$$\begin{aligned}\mathbb{E} \left[X_0 \mid \mathcal{B}^{(0)} \right] &= \int_{-\pi}^{\pi} \check{\varphi}(\lambda) Z(d\lambda), \\ \mathbb{E} \left[\left(X_0 - \mathbb{E} \left[X_0 \mid \mathcal{B}^{(0)} \right] \right)^2 \right] &= 2\pi\alpha,\end{aligned}$$

where $\check{\varphi}(\lambda) = 1 - \frac{\alpha}{F_T(\lambda)}$, $\alpha = \frac{2\pi}{\int_{-\pi}^{\pi} \frac{d\lambda}{F_T(\lambda)}}$, and $Z(d\lambda)$ is an orthogonal stochastic measure from the spectral representation of the sequence (X_{nT}) , $n \in \mathbb{Z}$.

Scheme of instantaneous values

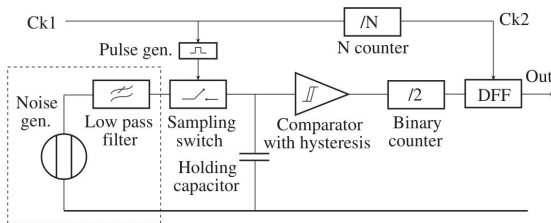


Рис.: PRNG from *Bagini, V, Bucci, M.*, «A Design of Reliable True Random Number Generator for Cryptographic Applications»

PRNG form *Bagini, V, Bucci, M.*

Notations

Thus, the function g from (2) is defined as follows:

$$g(x) = \begin{cases} 0, & \text{if } x < d; \\ 1, & \text{if } x \geq d. \end{cases}$$

PRNG form *Bagini, V, Bucci, M.*

Notations

Thus, the function g from (2) is defined as follows:

$$g(x) = \begin{cases} 0, & \text{if } x < d; \\ 1, & \text{if } x \geq d. \end{cases}$$

Lemma

Let $\xi \sim \mathcal{N}(\theta, \sigma_1^2)$, where $\theta \sim \mathcal{N}(a_2, \sigma_2^2)$, and $d \in \mathbb{R}$. Then,

$$\mathbb{P}(\xi \leq d) = \text{Erfc} \left(\frac{a_2 - d}{2\sqrt{\sigma_1^2 + \sigma_2^2}} \right),$$



Example (1/4)

Example

Let the covariance function $R(n)$ of the sequence $X_{nT}, n \in \mathbb{Z}$, be

$$R(n) = \frac{1}{T^n},$$

with $T > 1$. Then, the spectral density function $F_T(\lambda)$ of the sequence $X_{nT}, n \in \mathbb{Z}$, is given by

$$F_T(\lambda) = \frac{1}{2\pi} \frac{1 - T^{-2}}{\left|1 - \frac{e^{-i\lambda}}{T}\right|^2}.$$

Example (2/4)

Example

$$\begin{aligned}\mathbb{E} \left[X_0 \mid \mathcal{B}^{(0)} \right] &= \int_{-\pi}^{\pi} \frac{T^{-1}}{1 + T^{-2}} \left(e^{i\lambda} + e^{-i\lambda} \right) Z(d\lambda) = \\ &= \frac{T^{-1}}{1 + T^{-2}} (X_{-T} + X_T), \\ \mathbb{E} \left[\left(X_0 - \mathbb{E} \left[X_0 \mid \mathcal{B}^{(0)} \right] \right)^2 \right] &= \left(\frac{T^{-1}}{1 + T^{-2}} \right)^2 \mathbf{D} (X_{-T} + X_T) = \\ &= \left(\frac{T^{-1}}{1 + T^{-2}} \right)^2 (2 + T^{-2}).\end{aligned}$$

Example (3/4)

Example

Take $T = 5$, $d = -0.883$. Then, $\alpha \approx 0.146912$, and

$$\sup_n \sup_{\omega \in \Omega} \max_{r=0}^1 \left| \mathbb{P}(\gamma_n = r \mid \gamma_{n+1}, \gamma_{n-1}, \gamma_{n+2}, \gamma_{n-2}, \dots) - \frac{1}{2} \right| \leq 2.11 \cdot 10^{-4}.$$

Example (4/4)

Example

Note that in this example, the practical security of the key is extremely *sensitive* to the parameter d . For instance, with $T = 5$, $d = -0.88$, we obtain

$$\sup_n \sup_{\omega \in \Omega} \max_{r=0}^1 \left| \mathbb{P}(\gamma_n = r \mid \gamma_{n+1}, \gamma_{n-1}, \gamma_{n+2}, \gamma_{n-2}, \dots) - \frac{1}{2} \right| \leq 1.66 \cdot 10^{-3}.$$