Entropy Estimation on the Basis of a Stochastic Model

Werner Schindler
Bundesamt für Sicherheit in der Informationstechnik (BSI)
Bonn, Germany

Presented by Peter Birkner

Gaithersburg, May 2, 2016

Introduction

- Motivation and Background
- Stochastic model
  - Definition and objective
  - Illustrating examples
  - Health tests (online tests)
- Experiences with the AIS 31
- Conclusion

NIST SP 800-90B [4]

- Entropy estimation is the most critical part of a security evaluation of a physical RNG.
- Among others [4], Subsection 3.2.2, demands that the documentation *... shall include a description of how the noise source works and rationale about why the noise source provides acceptable entropy output,...*

Entropy estimation

- Unfortunately, entropy cannot be measured like voltage and temperature.
- Instead, entropy is a property of random variables.
- In the following we interpret random numbers as realizations of (i.e. as values taken on by) random variables.
- We present a field-tested method for the estimation of the entropy of physical RNGs.

Notation

- In the following we use the terminology of SP 800-90B [4]. In particular,
  - digitized data = data after the digitization of the analog signals
  - raw data = data after (non-cryptographic) postprocessing
- NOTE: In the literature also other definitions are widespread. In particular,
  - raw random numbers (or digitized analog signals) = data after digitization
  - internal random numbers = data after (non-cryptographic / cryptographic) postprocessing

What is a stochastic model?

- Ideally, a stochastic model specifies a family of probability distributions, which contains the true (but unknown) distribution of the raw data (interpreted as realizations of random variables).
- In a second step therefrom the (average gain of) entropy per raw data bit is estimated.
- In most cases it is yet easier to develop and to verify a stochastic model for the digitized data (or, alternatively, for suitable 'auxiliary random variables').
  $\rightarrow$ entropy(digitized data) $\rightarrow$ entropy(raw data)

Example 1: Coin tossing

- A coin is tossed N times ('head'$\simeq 1$, 'tail'$\simeq 0$').
- We interpret the observed outcome $x_1, \ldots, x_N$ (= digitized data) of N coin tosses as realizations of random variables $X_1, \ldots, X_N$.
- The random variables $X_1, \ldots, X_N$ are assumed to be iid (independent and identically distributed).
  Justification: A coin has no memory.
- $p := \mathrm{Prob}(X_j = 1) \in [0, 1]$ with unknown parameter p.

Example 1: Entropy estimation

- $X_1, \ldots, X_N$ are iid $\Longrightarrow H(X_1, \ldots, X_N)/N = H(X_1)$ (= (average) entropy per coin toss) where

  $H(X_1) = -(p \log_2(p) + (1-p) \log_2(1-p))$ (Shannon entropy)

- Equivalently, $H_{min}(X_1, \ldots, X_N)/N = H_{min}(X_1)$ with

  $H_{min}(X_1) = min\{-log_2(p), -\log_2(1-p)\}$ (min entropy)

- $$\widetilde{p} := \frac{x_1 + \cdots + x_N}{N} \quad \text{(estimator for } p\text{)}$$

- Substituting $\widetilde{p}$ into the above formulae provides estimators for the Shannon entropy and for the min entropy per coin toss.

Example 1: Stochastic model

- A stochastic model is not a physical model. In Example 1 a physical model would consider the impact of the start conditions and the mass distribution within the coin etc. on the trajectory.
- It is much easier to develop and to verify a stochastic model than a physical model.
- In our coin tossing example the stochastic model defines a 1-parameter family of probability distributions.

Real world RNGs

- For real world physical RNGs the derivation of the stochastic model is more complicated. The stochastic model should be confirmed by engineering arguments and experiments.
- Typically, a stochastic model specifies a 1-, 2- or a 3-parameter family of distributions.
- If the digitized data are not iid *the increase of entropy* per random bit has to be considered.
- During the life cycle of the RNG the true distribution shall remain in the specified family of probability distributions, also if the quality of the random numbers goes down ($\rightarrow$ health tests).

Example 2: Killmann, Schindler (CHES 2008)[6]

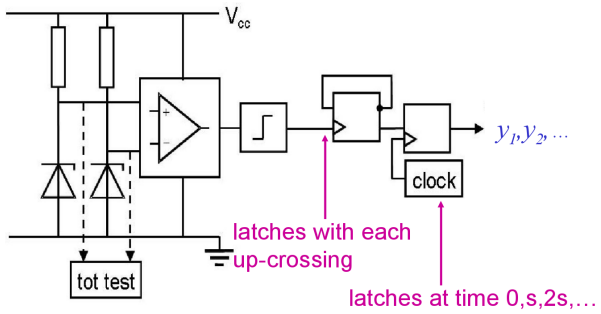


Abbildung: RNG with two noisy diodes, c.f. Fig. 1 in [6]

Stochastic model (for $y_1, y_2, \ldots$)

- t_n: time between the $(n-1)^{th}$ and the n^{th} upcrossing
- $T_1, T_2, \ldots$ is stationary (mild assumption) $\rightsquigarrow \ldots \quad \ldots \rightsquigarrow$ $Y_1, Y_2, \ldots$ is stationary
- 2-parameter family of distributions (depends on the expectation and the generalized variance of T_1)
- details: see [6]

Example 3: Haddad, Fischer, Bernard, Nicolai [5]

- Source of randomness: transient effect ring oscillator (TERO)
- Thorough analysis of the electric processes in the TERO structure
- $\rightarrow$ stochastic model of the TERO
- $\rightarrow$ stochastic model of the complete RNG
- Implementation of the RNG design on a 28 nm CMOS ASIC

Health tests (online tests)

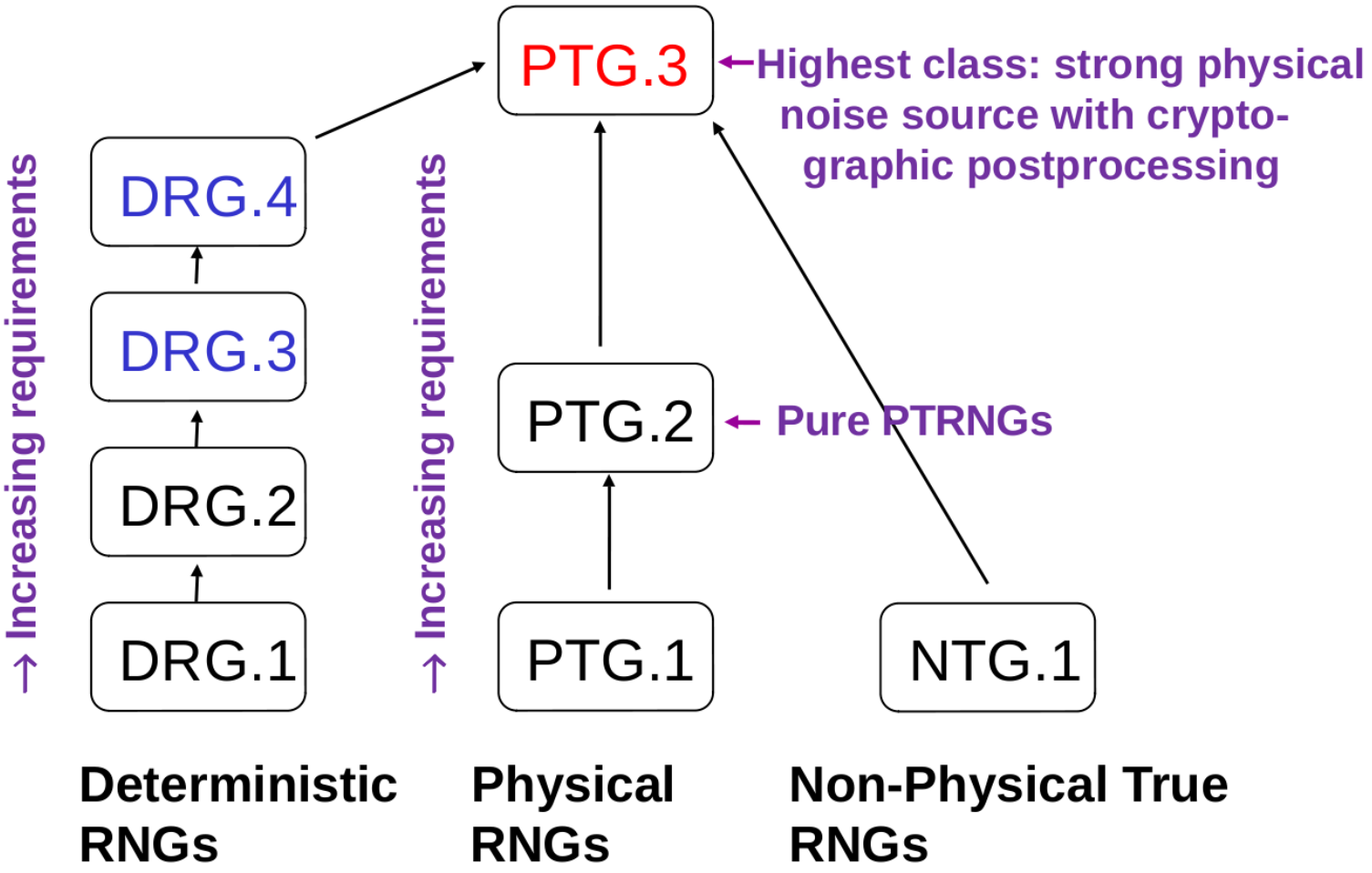
- Health tests, which are universally effective for any RNG design, do not exist.
- The health test (online test) should be tailored to the stochastic model. The health test should detect non-tolerable deficiencies of the random numbers sufficiently soon.
- Example 1: A monobit test would be suitable.
  If # '1's deviates significantly from sample size / 2
  $\rightsquigarrow$ indicator that p is (no longer) acceptable.

AIS 20 [1] / AIS 31 [2]

- In the German evaluation and certification scheme the evaluation guidance documents
  - AIS 20: Functionality Classes and Evaluation Methodology for Deterministic Random Number Generators
  - AIS 31: Functionality Classes and Evaluation Methodology for Physical Random Number Generators

  have been effective since 1999, resp. since 2001.
- NOTE: The mathematical-technical reference [3] was updated in 2011.

Functionality classes



Miscellaneous

- The AIS 20 and the AIS 31 are technically neutral.
- For physical RNGs (PTG.2, PTG.3) a stochastic model is mandatory. The digitized data shall be stationary distributed.
- The applicant for a certificate and the security lab have to give evidence that the RNG meets the class-specific requirements.
- Further documents support the tasks of the developer and the evaluator.
- For sensitive applications the BSI prefers RNGs, which belong to the functionality classes PTG.3, DRG.4 or DRG.3.

The functionality classes PTG.3, DRG.4, DRG.3

- PTG.3 (highest class):
  - strong physical RNG (possibly with mathematical postprocessing), effective online test and total failure test
  - DRG.3-conformant postprocessing algorithm with memory; output rate(postprocessing) $\leq$ input rate(postprocessing)

  information theoretical security + computational security
- DRG.4:
  - DRG.3-conformant deterministic RNG
  - the internal state can be updated / reseeded (time-dependent, event-driven or on demand)

  substantially only computational security.
- DRG.3: deterministic RNG (backward secrecy, forward secrecy, enhanced backward secrecy)

Conclusion

- A sound stochastic model of a physical RNG allows to derive a reliable lower bound for the entropy per raw data bit.
- We explained the concept of a stochastic model by an elementary example.
- Elaborated stochastic models of real world RNGs can be found in the literature.
- In the German certification scheme (Common Criteria) the concept of stochastic models has proved successful for many years.

Contact

Bundesamt für Sicherheit in der Informationstechnik (BSI),
Bonn, Germany

Werner Schindler

P.O. Box 200363, 53133 Bonn, Germany
Tel.: +49 (0)228-9582-5652
Fax: +49 (0)228-10-9582-5652

Werner.Schindler@bsi.bund.de
https://www.bsi.bund.de
https://www.bsi-fuer-buerger.de

 [1] Bundesamt für Sicherheit in der Informationstecnik (BSI): Anwendungshinweise und Interpretationen zum Schema (AIS) AIS 20. Version 3, 15.05.2013; https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Zertifizierung/Interpretationen/AIS_20_pdf.html

 [2] Bundesamt für Sicherheit in der Informationstecnik (BSI): Anwendungshinweise und Interpretationen zum Schema (AIS) AIS 31. Version 3, 15.05.2013; https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Zertifizierung/Interpretationen/AIS_31_pdf.html

 [3] W. Killmann, W. Schindler: A Proposal for: Functionality Classes for Random Number Generators. Mathematical-Technical Reference to [1] and [2], Version 2, 18.09.2011;

https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Zertifizierung/Interpretationen/AIS_31_Functionality_classes_for_random_number_generators_e.pdf?__blob=publicationFile

[4] NIST Special Publication 800-90B (Second Draft): Recommendation for the Entropy Sources Used for Random Bit Generation.

[5] P. Haddad, V. Fischer, F. Bernard, J. Nicolai: A Physical Approach for Stochastic Modeling of TERO-Based TRNG. In: CHES 2015, Springer, LNCS 9293, 357–372

[6] W. Killmann, W. Schindler: A Design for a Physical RNG with Robust Entropy Estimators. In: CHES 2008, Springer, LNCS 5154, 146–163.