A. F. N i k o l a e v (Nizhny Novgorod, Intel Corp.). **On one algorithm for min-entropy estimation.**

Introduction. The industry has the clear demand for reliable entropy source, true random number generator (TRNG). Its numerous applications are in cryptography, simulations, engineering computations, and system software. With cloud computing moving into mainstream [1] the role of supportive security platforms and, thus, the need in the source of «true» random numbers will further increase.

TRNG which is based on some physical phenomena relies on hardware components (the example of entropy source which uses the Johnson noise in resistors underneath and its analysis are available in [2]). The key characteristics of the entropy source should have include good statistical properties, immunity to various types of attacks, resistance to changing environmental conditions. The quantitative measure of TRNG quality is entropy. While various definitions of entropy are available (e.g., Shannon, min, conditional, Renyi) below we will use the one suggested by ANS X9.82 [3]. Assume that TRNG produces one of n possible outputs or sequences of outputs at the given time interval. Each possible output is generated with probability p_i. The min-entropy is defined as $H = -\log_2(\max_i p_i)$.

Thus, the accurate methodology for the min-entropy estimation becomes the important problem on all stages of TRNG project that includes design, validation, and manufacturing. The solution to this problem can be connected to the adequate mathematical description of the specific TRNG design followed by the deduction of the entropy estimate. In this communication we will take a complementary approach which is related to the analysis of bit stream produced by the generator (which, in its turn, is considered as «black box»), and computing min-entropy estimate given some assumptions. Unlike some other schemes which give only the entropy approximation valid for the limited parameter domain under simplified conditions we present the methodology that produces the exact result in more general context (the additional schemes for min-entropy estimation are also available in [3]). Another advantage of the described scheme is that it admits extension to the more general case.

Problem statement. Let $(\Omega, \mathcal{F}, \mathbf{F}, \mathbf{P})$ be stochastic basis which follows the usual conditions [4]. At each time the observed Markov process $b = (b_n)_n = 0, 1, 2, \dots$ takes one of two possible values: $b_n = 1$ with probability $p \in (0, 1)$ and $b_n = 0$ with probability $1 - p$. Coefficient ρ_n defines the correlation between b_n and b_{n+1}. Thus, the transitional probabilities for the process b have the following format for all $n = 0, 1, \dots$

$$\mathbf{P}\{b_{n+1} = 1|b_n\} = \rho_n b_n + p(1 - \rho_n), \quad \mathbf{P}\{b_{n+1} = 0|b_n\} = 1 - \mathbf{P}\{b_{n+1} = 1|b_n\}.$$

Let $H = -\log_2(\max_i p_i)$ define min-entropy of a bit sequence of size n with probability p_i to have i-th bit pattern $(x_1, x_2, \dots, x_n)$, $p_i = \mathbf{P}\{b_1 = x_1, b_2 = x_2, \dots, b_n = x_n\}$. Our goal is to design the algorithm for computation of the value of H.

Result. Before we formulate the main result of the report we introduce the following notations. Let for all $k = 1, 2, \dots$ coefficient a_k be $a_k = \rho_{k-1}x_{k-1} + p(1 - \rho_{k-1})$. We also define the system of the Bellman equations $\Psi_i(x)$, $i = 1, 2, \dots, n+1$, in the following way (the coefficient a_k is defined earlier):

$$\Psi_{n+1}(x) = 1, \quad \Psi_k(x_{k-1}) = \max\{(1 - a_k)\Psi_{k+1}(0), a_k\Psi_{k+1}(1)\}, \quad k = n, n-1, \dots, 2,$$

$$\Psi_1 = \max\{(1 - p)\Psi_2(0), p\Psi_2(1)\}.$$

We also introduce the sequence of the thresholds B_k:

$$B_k = \begin{cases} 1/2, & \text{if} \quad k = n + 1, \\ \dfrac{\Psi_k(0)}{\Psi_k(0) + \Psi_k(1)}, & \text{if} \quad k = n, n-1, \dots, 2. \end{cases}$$

Theorem. *The bit pattern* $x^* = (x_1^*, x_2^*, \ldots, x_n^*)$ *which delivers the function* p_i *to the maximum is computed in the following way*:

$$x_k^* = \begin{cases} 1, & \textit{if} \quad a_k \geqslant B_{k+1}, \\ 0, & \textit{if} \quad a_k < B_{k+1}, \end{cases} \quad k = n, n-1, \ldots, 2, \quad x_1^* = \begin{cases} 1, & \textit{if} \quad p \geqslant B_2, \\ 0, & \textit{if} \quad p < B_2. \end{cases}$$

The value of min-entropy is $H = -\log_2(\Psi_1)$ *with*

$$\Psi_k(x_{k-1}) = \begin{cases} a_k(x_{k-1})\Psi_{k+1}(1), & \textit{if} \quad a_k \geqslant B_{k+1}, \\ (1 - a_k(x_{k-1}))\Psi_{k+1}(0), & \textit{if} \quad a_k < B_{k+1}, \end{cases} \quad k = n, n-1, \ldots, 2,$$

$$\Psi_1 = \begin{cases} p\Psi_2(1), & \textit{if} \quad p \geqslant B_2, \\ (1-p)\Psi_2(0), & \textit{if} \quad p < B_2. \end{cases}$$

Proof of the theorem is direct application of the dynamic programming approach, [5].

Notes. 1. If $\rho_k = 0$ then $B_k = 1/2$ for $k = n, n-1, \ldots, 2$ and Ψ_1 is reduced to p^n or $(1-p)^n$. 2. Computation of min-entropy can be generalized to case of non constant probability of one. 3. The theorem sets the base for the algorithm of the min-entropy estimation.

REFERENCES

1. IDC Worldwide Software 2010–2014 Forecast Summary.
2. *Nikolaev A.*, *Pradhan P.* On one modeling approach for true random number generator. — In: Proceedings of IADIS International Conference Informatics, Netherlands, 2008, p. 199–204.
3. *ANS X9.82.* Random Number Generation. P. 2: Entropy Sources (Draft), 2007.
4. *Jacod J.*, *Shiryaev A.* Limit Theorems for Stochastic Processes. Berlin–Heidelberg, Germany: Springer-Verlag, 2002.
5. *Bellman R.* Dynamic Programming. Princeton: Princeton University Press, 1957.