Evaluating Min-Entropy of Random Bits by Markov Chains

E. N. Allini[1], M. Skórski[2]

[1]Hubert Curien Laboratory
University of Lyon, France

[2]Hubert Curien Laboratory
University of Lyon, France

16th CryptArchi Workshop (Lorient), 2018

Outline

Markov chains

recup

A Markov chain of order d is a sequence of random variables $X = (X_n)_{n=1}^{\infty}$ over a finite space $\mathcal{S}$ such that

- transition probabilities depend only on last d states

$$\forall n : \quad \Pr[X_{n+d} | X_{n+d-1}, \ldots, X_1] = \Pr[X_{n+d} | X_{n+d-1}, \ldots, X_n]$$

- transition probabilities are time-invariant

$$\forall n \, \forall s_0, \ldots, s_d : \quad \Pr[X_{n+d} = s_d | X_{n+d-1} = s_{d-1}, \ldots, X_n = s_0] \\ = \Pr[X_d = s_d | X_{d-1} = s_{d-1}, \ldots, X_0 = s_0]$$

Modeling TRNGs by Markov Chains

Markov chains are convenient models for temporal (short-memory) dependencies for True Random Number Generators [TBKMB+18].

- we can model raw (not processed) bits with higher-order (longer memory)
- we can model output (processed) bits with low-order (short memory)
- the appropriate order can be assessed based on stochastic properties of the entropy source (e.g. observed autocorrelation)
- higher order gives more accuracy but is less efficient to evaluate

Estimating transition matrix

theory

Let bits $b_1, \ldots, b_N$ be samples from a Markov chain X of order d. Define for convenience $b_{i:i+d} = b_i b_{i+1} \ldots b_{i+d-1} \in \{0,1\}^d$. Then

$$\hat{P}_{s,t} = \frac{\#\{i : b_{i:i+d} = s, b_{i+1:i+d+1} = t\}}{\#\{i : b_{i:i+d} = s\}} \tag{1}$$

is the estimate of the transition matrix $P = P_X$. Note the matrix states are d-bit strings $\{0,1\}^d$. We have

$$\forall s,t \in \{0,1\}^d \quad \hat{P}_{s,t} \longrightarrow P_{s,t} \quad \text{when } N \to \infty \tag{2}$$

if the chain is irreducible and aperiodic.

Estimating transition matrix

complexity

Since patterns s and t share all but 1 digit, it suffices to:

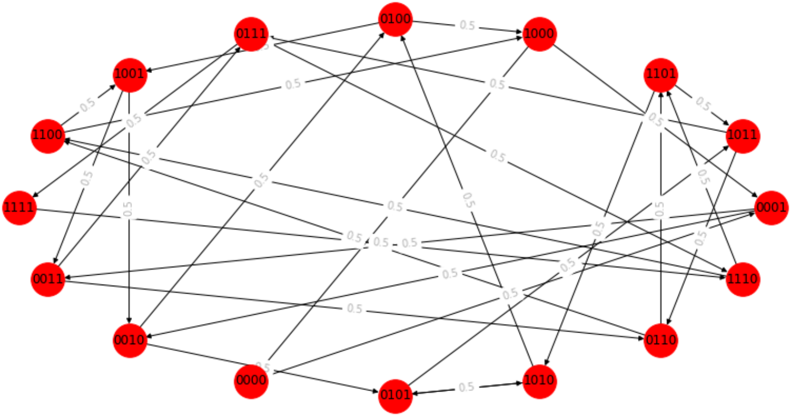
- use a sliding window of length $d+1$ (*one pass* over the data)
- populate only $2 \cdot 2^d = 2^{d+1}$ entries in the matrix

Therefore the time T and space S are given

$$T = O(N), \quad S = O(2^d) \tag{3}$$

Note the transition matrix is sparse.

Adjacency graph ($d = 4$,unbiased transitions)



Transition matrix ($d = 4$,unbiased transitions)

$$P = \begin{pmatrix}
0.50 & 0.50 & & & & & & & & & & & & & & \\
 & & 0.50 & 0.50 & & & & & & & & & & & & \\
 & & & & 0.50 & 0.50 & & & & & & & & & & \\
 & & & & & & 0.50 & 0.50 & & & & & & & & \\
 & & & & & & & & 0.50 & 0.50 & & & & & & \\
 & & & & & & & & & & 0.50 & 0.50 & & & & \\
 & & & & & & & & & & & & 0.50 & 0.50 & & \\
 & & & & & & & & & & & & & & 0.50 & 0.50 \\
0.50 & 0.50 & & & & & & & & & & & & & & \\
 & & 0.50 & 0.50 & & & & & & & & & & & & \\
 & & & & 0.50 & 0.50 & & & & & & & & & & \\
 & & & & & & 0.50 & 0.50 & & & & & & & & \\
 & & & & & & & & 0.50 & 0.50 & & & & & & \\
 & & & & & & & & & & 0.50 & 0.50 & & & & \\
 & & & & & & & & & & & & 0.50 & 0.50 & & \\
 & & & & & & & & & & & & & & 0.50 & 0.50
\end{pmatrix}$$

Estimating transition matrix

theory

- irreduciblity and aperiodicity are easily satisfied for reasonable TRNGs, for instance when the probability of next state being 0 is strictly between 0 and 1
- convergence is exponential and can be quantified by Chernoff-like bounds (see [Lez98; CLLM12])
- sparsity improves convergence intervals, as we have only $O(2^d)$ entries not $O(2^{2d})$

Estimating transition matrix

algorithm

Algorithm 1: Transition Matrix Estimation

Input: Samples $b_1, \ldots, b_N$ from a Markov chain of order d

Output: Sparse transition matrix of the equivalent first-order chain

counts $\leftarrow$ dict

$w \leftarrow 0 b_0 b_1 b_2 \ldots b_{d-1}$

for $i = d \ldots N$ **do**

  $w \leftarrow w_1 w_2 \ldots w_d b_i$

  if $w \in$ counts.keys **then**

    counts$[w] \leftarrow$ counts$[w] + 1$

  else

    counts$[w] = 1$

  end

end

for $s = s_0 \ldots s_{d-1}) \in \{0,1\}^d, t_{d-1} \in \{0,1\}$ **do**

  $t \leftarrow s_{1:d} t_{d-1}$

  $P_{s,t} \leftarrow \frac{\text{counts}[s_{0:d} t_{d-1}]}{\text{counts}[s_{0:d}0] + \text{counts}[s_{0:d}1]}$

end

return P

Entropy Rate for Markov Chains

Computing entropy rates is different depending on the entropy notion:

- Shannon entropy: computed from the transition matrix and stationary distribution (as conditional entropy)
- Renyi entropy: explicit formula, need to compute the spectral radius of Hadamard powers of the transition matrix
- Min-entropy: less explicit, involves optimization over graph cycles.

Min-Entropy Rate of Markov Chains

formula

Theorem (Min-entropy rate of Markov chains [KV16])

Let P be the transition matrix of an irreducible and aperiodic Markov chain with the state space S. Then

$$H_\infty(P) = \min_{\ell} \min_{(s_1,\ldots,s_{\ell+1})\in\mathcal{C}_\ell} \frac{1}{\ell}\sum_{k=1}^{\ell} \log \frac{1}{P_{s_k,s_{k+1}}} \tag{4}$$

where $\mathcal{C}_\ell$ denotes the set of all loops of length ℓ.

Min-Entropy Rate of Markov Chains

algorithm

Algorithm 2: Min-Entropy Rate of Markov Chains

Input: Transition matrix P of dimension $2^d \times 2^d$

Output: min-entropy rate of a Markov chain with transition matrix P

for $i, j \in \{0,1\}^d$ **do**
    $\mathrm{Q}[i,j] \leftarrow \log P_{i,j}$ // put $\log 0 = -\infty$
end
// intialize heaviest path weights for zero length
for $i, j \in \{0,1\}^d$ **do**
    $\mathrm{HeaviestPath}[i,j] \leftarrow \log[i = j]$ // put $\log 0 = -\infty$
end
entropy $\leftarrow 1$ // initial entropy per bit
// update heaviest paths for every next length ℓ
for $\ell \in \{0,1\}^d$ **do**
    for $i, j \in \{0,1\}^d$ **do**
        $W \leftarrow -\infty$
        for $k\{0,1\}^d$ **do**
            $\mathrm{W} \leftarrow \max(\mathrm{W}, \mathrm{HeaviestPath}[i,k] + \mathrm{Q}[k,j])$ // longer path
        end
        $\mathrm{NewHeaviestPath}[i,j] \leftarrow W$
    end
    for $i, j\{0,1\}^d$ **do**
        $\mathrm{HeaviestPath}[i,j] \leftarrow \mathrm{NewHeaviestPath}[i,j]$
    end
    // compute entropy for current length
    $w \leftarrow -\infty$
    for $i \in \{0,1\}^d$ **do**
        $w \leftarrow \max(w, \mathrm{HeaviestPath}[i,i])$
    end
    entropy $\leftarrow \min\left(\text{entropy}, -\frac{w}{\ell}\right)$
end

Complexity

space complexity

We need to store

- transition matrix P
- matrices used in dynamic programming NewHeaviestPath,HeaviestPath
- few auxiliary variables (scalar)

Therefore the memory costs is

$$S = O(2^{2d}) \tag{5}$$

(assuming finite precision)

Complexity

time complexity

As for the running time, consider that

- the execution time is dominated by the 4-fold loop over ℓ, i, j, k.
- it is enough to consider only k such that $P_{k,j} > 0$ - at most two explicit values (we know them from the shape of P)

Therefore the running time is

$$T = O(2^{3d}) \tag{6}$$

Evaluation on real device

- TRNG built out of two ring oscillators, raw bits are counters of jittery clock periods
- post-processing is done by taking first differences and extracting least significant bits

Comparison with AIST tests (standards require rate at least 0.997)

τ	Markov chain	AIS Test procedure B	AIS T8
periods of s_2	min-entropy per bit		Shannon entropy per bit
100000	0.9909	passed	0.9999
25000	0.9908	passed	0.9999
20000	0.9893	passed	0.9999
15000	0.9783	passed	0.9998
10000	0.8087	failed	0.9865
2000	0.2816	failed	0.0981

Table: Entropy estimation using two internal ROs and extracting the least significant bits of the first differences of counter values. Dependencies are modeled by Markov chains of eighth order.

Conclusion

- min-entropy is more conservative and suitable for cryptography than Shannon entropy
- min-entropy of bit sequences generated by TRNGs can be efficiently evaluated by fitting Markov chains
- we discussed theoretical and implementation details

References I

K. Chung, H. Lam, Z. Liu, and M. Mitzenmacher. "Chernoff-Hoeffding Bounds for Markov Chains: Generalized and Simplified". In: *29th International Symposium on Theoretical Aspects of Computer Science, STACS 2012, February 29th - March 3rd, 2012, Paris, France*. 2012. URL: https://doi.org/10.4230/LIPIcs.STACS.2012.124.

S. Kamath and S. Verdú. "Estimation of entropy rate and Rényi entropy rate for Markov chains". In: *IEEE International Symposium on Information Theory, ISIT 2016, Barcelona, Spain, July 10-15, 2016*. 2016. URL: https://doi.org/10.1109/ISIT.2016.7541386.

References II

P. Lezaud. "Chernoff-type bound for finite Markov chains". In: *Ann. Appl. Probab.* 8.3 (Aug. 1998). URL: https://doi.org/10.1214/aoap/1028903453.

M. S. Turan, E. Barker, J. Kelsey, K. A. McKay, M. L. Baish, and M. Boyle. 2018. URL: https://doi.org/10.6028/NIST.SP.800-90B.

Thank you for your attention!

Questions?