A Chaos-Based Random Number Generator and Its Verification Through the Bootstrap

Salih Ergün†

†TÜBİTAK-National Research Institute of Electronics and Cryptology
PO Box 74, 41470, Gebze, Kocaeli, Turkey
Email: salih.ergun@tubitak.gov.tr

Abstract— A random number generation method based on a bipolar-transistor cross-coupled chaotic oscillator is introduced. Numerical model for the proposed design has been developed where bootstrap method is used which allows us to estimate the statistical characteristics of underlying chaotic signal. Numerical results, verifying the feasibility and correct operation of the random number generator are given such that numerically generated bit sequences fulfill FIPS-140-2 statistical test suite without any further post-processing. Proposed random number generator features much higher and constant throughput rates and allows for offset compensation.

1. Introduction

In the 20th century, because of the increasing demand of electronic official & financial electronic transactions and digital signature applications, the need for information secrecy has raised. In this manner, as an unseparable part of the secure systems, random number generators (RNGs) which have been used for only military cryptographic applications in the past got expanding usage for a typical digital communication equipment.

Random Number Generators are used for a variety of cryptographic applications and certain requirements on such generators are outlined in [1]. Four different types of random number generators are used in the literature and these are categorized as: amplification of a noise source [1] jittered oscillator sampling [2], discrete-time chaotic maps [3] and continuous-time chaotic oscillators [4, 5, 6].

The use of discrete-time chaotic maps for random number generation is well-known [3]. It was shown recently that continuous-time chaotic oscillators can also be used to realize RNGs [4, 5, 6]. In particular, we have reported preliminary results of a RNG using a novel continuous-time chaotic oscillator in [6]. In this work we recall this chaotic oscillator and further introduce the design of a RNG, which relies on generating non-invertible random bit sequences according to regional distributions from one of the waveform of the chaotic oscillator.

Furthermore, we develop numerical model for the proposed RNG design where bootstrap method is used which allows the estimation of statistical characteristics of underlying chaotic signal, thus provides determination of design parameters for the chaotic source appropriately. Proposed RNG offers some considerable advantages over the existing ones [4, 5, 6]. In comparison with the previous design [6], RNG introduced in this paper offers approximately sixfold rate expansion and constant output rate.

Moreover, proposed RNG has some other technical advantages. For example, although the design is capable of passing randomness tests without compensation, it allows for offset compensation for bias removal thus provides more robustness against external interference. Numerical results verifying the feasibility and the correct operation of the proposed RNG are presented such that numerically generated bit sequences fulfill FIPS-140-2 test suite [8] without any further post-processing.

2. Cross-Coupled Chaotic System

In this paper, we use a simple cross-coupled chaotic system as the core of the RNG, which was proposed in [6]. Routine analysis of the bipolar-transistor cross-coupled chaotic circuit yields the state equations given in [6] which transforms into the following equation by using the normalized quantities:

$$\begin{aligned} \dot{x}_1 &= a[e^{(x_1+x_2)} - e^{(x_2-x_1)}] - y \\ \dot{y} &= x_1 - z \\ 2\dot{z} &= y - 2z + b\tanh(x_1) \\ \dot{x}_2 &= c - a[e^{(x_1+x_2)} + e^{(x_2-x_1)}] \end{aligned} \tag{1}$$

where $a = I_s R/2V_T$, $b = I_0 R/2V_T$, $c = (2I_B - I_0)R/2V_T$, V_T and I_s are the bipolar transistor thermal voltage and saturation current, respectively.

The equations in 1 generate chaos for different sets of parameters. The chaotic attractor shown in Fig.1 is obtained from the numerical analysis of the system with $a = 0.5 \times 10^{-6}$, $b = 2$ and $c = 1$ using a 4^{th}-order Runge-Kutta algorithm with adaptive step size.

Exploited chaotic system offers some considerable advantages over the existing ones. Considering that the necessary conditions for exhibiting chaos in an autonomous system are at least three variables which correspond to three energy storage components in implementation and one nonlinearity, chaotic attractor consists of as few components as possible. In conclusion, due to the absence of large blocks such as analog multiplication stage, the core

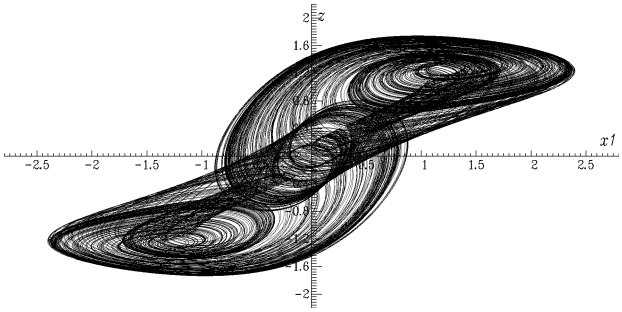

Figure 1: Numerical analysis results of the chaotic system for $a = 0.5 \times 10^{-6}$, $b = 2$ and $c = 1$.

chaotic system is simple and is easy to construct, which make the usage of it suitable for practical RNG applications.

3. Chaos-Based Random Number Generator

The method, introduced in this paper for random number generation, relies on generating non-invertible bit sequences according to regional distributions from one of the state of the chaotic system. It should be noted that non-invertibility is a key feature for generating random numbers.

In this method, in order to obtain random bit sequences from the chaotic attractor numerically, we used the periodic samples of the state x_1 in Eqn. 1, obtained at the rising edges of an external periodical pulse signal, that is at times t satisfying $wt \; mod \; 2\pi = 0$ where w is the frequency of the pulse signal. Note that, although 4-dimensional trajectories in the $x_1 - y - x_2 - z$ plane are invertible, one may obtain a non-invertible section by considering only the values corresponding to one of the states, say x_1.

We don't know much about an irregular signal used to generate random number but its distribution. At first, x_1 values have been numerically generated and the distribution of periodic samples have been examined to determine appropriate sections where the distributions look like random signal. Although, we could not find sections of which x_1 values have a single distribution, we determined various sections where the distribution of x_1 has at least two regions.

Distribution of x_1 having two regions, suggests us to generate random binary data from regional x_1 values for regional thresholds. Following this direction, we have generated the bit sequences $S_{(top)i} = sgn(x_{1i} - q_{top})$ *when* $x_{1i} \geq q_{middle}$ and $S_{(bottom)i} = sgn(x_{1i} - q_{bottom})$ *when* $x_{1i} < q_{middle}$, where $sgn(.)$ is the signum function, x_{1i}'s are the values of x_1 obtained from one of the above defined section, q_{top} and q_{bottom} are appropriately chosen thresholds for top and bottom distributions, respectively and q_{middle} is the boundary between the distributions.

In the previous designs [4, 5, 6], the well-known Von Neumanns deskewing technique is employed to eliminate the bias. In this paper, on the contrary to [4, 5, 6], another method $\bigotimes$ exclusive-or operation (*XOR*) is exploited in order not to decrease the throughput. The potential problem with the exclusive-or method is that a small amount of correlation between the input bits will add significant bias to the output. The correlation coefficient of generated bit sequences S_{top} and S_{bottom} of length 196 KBits is calculated as 0.00354 and it is determined that the generated bit sequences are independent.

According to this result, we have generated the new binary data $S_{(xor)i} = S_{(top)i} \bigotimes S_{(bottom)i}$. Using the above procedure, unbiased and uncorrelated bit sequences $S'_{xor}s$ have been obtained from the regional x_1 values. Then these bit sequences are subjected to the four tests (monobit, poker, runs and long-run) of FIPS-140-2 test suite. Furthermore, mathematical model of the proposed design has been developed, where bootstrap method is utilized to estimate the statistical characteristics of underlying chaotic signal

It is noteworthy that, the *XOR* corrector is not a sophisticated post-processor but a minor operation which sensibly combines the top and the bottom sequences generated from two regions which are separated according to distribution of the underlying chaotic signal.

4. Numerical Verification Through the Bootstrap

In comparison with RNGs based on continuous-time chaotic systems, the other common techniques [1, 2, 3], seem to be advantageous in the sense that true random behaviors can be mathematically proven thanks to analytical models that have been developed.

The numerical models used in [4, 5, 6], can provide intuition about real RNGs, and can also lead to improved models for RNG randomness analysis. However, current numerical models do not suffice to prove correct behavior of the chaos based RNGs, and additional theory and analytical methods are needed. If distribution of underlying chaotic signal is known with its mean and variance then its probability density function serves a basis to its analytical model of the generator. This then mathematically helps to prove true random behavior of the generator. Development of a general theory and associated analytical models for the randomness analysis of chaos based RNGs are now the other wide open problems.

In the last decade, the use of applicable, simple and more accurate models has become a fundamental requirement in signal processing applications, where the Gaussian distribution assumption is not valid. A real solution for these applications is the bootstrap. On the contrary to classical statistical analysis, which assumes that the data with a large sample size available is Gaussian thus is inapplicable to many signal processing problems, the bootstrap method introduced by Efron, provides a rigorous metric for finding confidence intervals for parameters such as variances or probability distributions of parameter estimators, while few observed data is available [7].

This useful capability makes the utilization of bootstrap method ideal for the numerical model of the proposed design where numerical data are finite, particularly short and distribution of underlying chaotic signal is unknown. The bootstrap method allows the estimation of statistical characteristics such as bias, variance, distribution functions and thus the estimation of confidence limits for parameters of interest. Its paradigm suggests substitution of the unknown probability model of the observed data in real world by the estimated probability model of the bootstrap samples in bootstrap world. Practically, bootstrap method approximates the distribution by reusing the original data resampled randomly with the corresponding statistics of interest, instead of applying the central limit theorem by assuming that the underlying distribution is Gaussian [7].

For the given a, b, and c values, the two regional distribution of the state x_1 obtained for $wt \ mod \ 2\pi = 0$ is shown in Fig. 2.

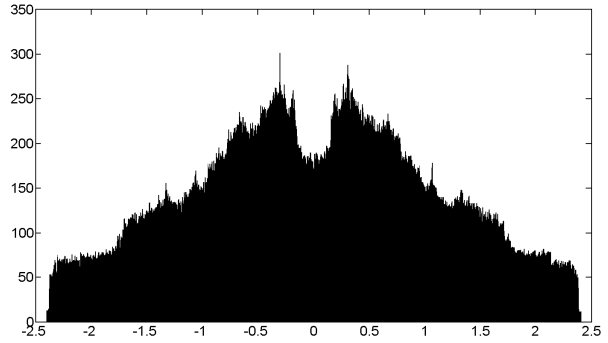

Figure 2: Histogram of x_1 obtained from the autonomous chaotic system for $wt \ mod \ 2\pi = 0$.

To be able to choose the thresholds appropriately, we examined top and bottom distributions as shown in Fig. 2 and then, thresholds q_{top} and q_{bottom} were determined as the means of the top and bottom distributions which were 0.8808 and −0.8808, respectively while the boundary between the distributions q_{middle} was determined as 0.

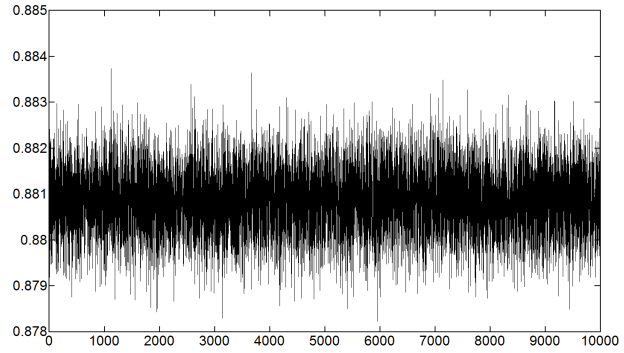

Figure 3: Confidence interval for the mean of top distribution.

We computed a sample of 10000 bootstrapped means of random samples taken from the original data, and plotted these bootstrapped means. As shown in Fig. 3 and 4, the bootstrap method finds confidence intervals for the means of top and bottom distributions as $0.8782 \leq mean_{top} \leq 0.8837$ and $-0.8835 \leq mean_{bottom} \leq -0.8781$, respectively.

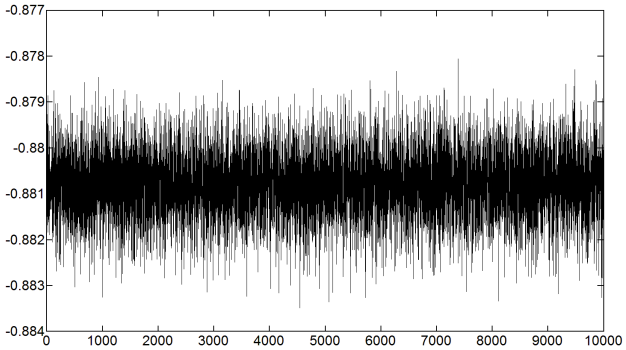

Figure 4: Confidence interval for the mean of bottom distribution.

In conclusion, we have numerically verified that the bit sequence S_{xor}, obtained for appropriate threshold values which are inside the given confidence intervals, passed the tests of FIPS-140-2 test suite without post-processing. Results for the uniformity of P-values [8] and the proportion of passing sequences are given in Table 1 for $q_{top} = 0.8837$ and $q_{bottom} = -0.8808$, where P-value ($0 \leq P - value \leq 1$) is a real number estimating the probability that a perfect RNG would have produced a sequence less random than the given sequence. It is reported that, for a sample size of $34 \times 20000 Bits$, the minimum pass rate for each statistical test with the exception of the random excursion (variant) test is approximately 0.938808.

Table 1: FIPS-140-2 Statistical Test Results.

STATISTICAL TESTS	S_{xor} Bit Sequence	
	$P - Value$	$Proportion$
Frequency	0.534146	0.9706
Block Frequency	0.100508	1.0000
Runs	0.213309	1.0000
Longest Run	0.000163	1.0000

In order to analyze output randomness with different set of threshold values and discuss their impact, the concept of approximate entropy ($ApEn$) [8] was employed as a measure of randomness. On the contrary to classical statistical tests, $ApEn$ provides a rigorous metric for proximity to randomness of a single finite sequence, particularly a very short sequence, without considering its underlying source [8]. This useful capability makes the utilization of $ApEn$ ideal for the numerical model of the proposed design. Shannon Entropy could be also used in the given model, however it should be noted that accurate calculation of Shannon Entropy requires the sequence to be infinite. The use of $ApEn$ is more appropriate for the devel-

oped model where numerical binary sequences are finite.

Table 2: $ApEn$ test results obtained for different threshold values.

Threshold Values		ApEn Test Results		
q_{top}	q_{bottom}	$P-Value$	$Proportion$	$ApEn$
-0.8781	0.8837	0.000569	1.0000	0.687446
-0.8781	0.8808	0.804337	0.9706	0.687714
-0.8781	0.8782	0.468595	1.0000	0.687718
-0.8808	0.8837	0.804337	0.9706	0.687602
-0.8808	0.8808	0.534146	1.0000	0.687823
-0.8808	0.8782	0.862344	1.0000	0.687616
-0.8835	0.8837	0.066882	0.9706	0.687225
-0.8835	0.8808	0.253551	0.9412	0.687596
-0.8835	0.8782	0.299251	0.9706	0.687629

In conclusion, we have numerically verified that the bit sequences $S_{xor}s$, obtained for the given confidence intervals, passed the $ApEn$ test [8] for a sample size of $34 \times 20000 Bits$ where $ApEn$ values and corresponding test results are given in Table 2.

External interference is a major concern in RNG design since interfered and random signals have comparable levels. To solve this problem and to be robust against parameter variations and attacks aimed to force throughput, we have proposed offset compensation loops that increase the statistical quality of the generated bit sequences. Offset compensations of q_{top} and q_{bottom} thresholds can be realized by implementing monobit test of FIPS-140-2 test suite for S_{top} and S_{bottom} binary sequences. For each sequence, bit streams of length 20000 bits are acquired, if the number of $0's$ is greater than 10275 then corresponding threshold is decreased and if the number of $0's$ is less than 9725 then corresponding threshold is increased until they reach and become stable at the means of top and bottom distributions.

In the previous RNG design [6], numerical simulations show that the system generates 209 bits per every 10,000 normalized time units after de-skewing while that is 1280 bits for the proposed design without any further post-processing. In conclusion, this indicates that the proposed design results in a sixfold rate expansion in comparison with the previous design [6].

On the contrary to the other chaos based RNGs reported in [4, 5, 6], RNG design proposed in this paper avoids the need of further post-processing which significantly decreases the throughput. Another disadvantage of the previous designs [4, 5, 6] is the disability to realize necessary offset compensation, which derives from the fact that instead of raw bit sequences, processed sequences can pass the statistical tests thanks to post-processing techniques.

As a result, in comparison with the previous RNGs [4, 5, 6], RNG proposed in this paper is an enhanced design which features much higher throughput rates, allows for compensation thus provides more robustness against external interference, parameter variations and tampering and fulfills the FIPS-140-2 statistical test suite without any further post-processing.

5. Conclusions

A random number generation method based on a bipolar-transistor cross-coupled chaotic oscillator is introduced which offers much higher and constant throughput rates, allows for offset compensation and fulfills the FIPS-140-2 statistical test suite without any further post-processing. Moreover, numerical models for the proposed design have been developed which use bootstrap method allowing the estimation of statistical characteristics of underlying chaotic signals. Numerical results presented in this paper not only verify the feasibilities and the correct operations of the proposed design, but also encourage its use as the core of a high-performance RNG as well.

References

[1] Göv, N.C., Mıhçak, M.K. and Ergün, S.: True Random Number Generation Via Sampling From Flat Band-Limited Gaussian Processes. IEEE Trans. Circuits and Systems I, Vol. 58. 5 (2011) 1044-1051

[2] Bucci, M., Germani, L., Luzzi, R., Trifiletti, A., Varanonuovo, M.: A High Speed Oscillator-based Truly Random Number Source for Cryptographic Applications on a SmartCard IC. IEEE Trans. Comput., Vol. 52. (2003) 403-409

[3] Callegari, S., Rovatti, R., Setti, G.: Embeddable ADC-Based True Random Number Generator for Cryptographic Applications Exploiting Nonlinear Signal Processing and Chaos. IEEE Transactions on Signal Processing, Vol. 53, 2 (2005) 793-805

[4] Yalcin, M.E., Suykens, J.A.K., Vandewalle, J.: True Random Bit Generation from a Double Scroll Attractor. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, Vol. 51(7). (2004) 1395-1404

[5] Ergün, S., Özoğuz, S., "Truly Random Number Generators Based on a Non-Autonomous Chaotic Oscillator," Int. J. Elec. Commun., vol. 61, (2007) 235-242

[6] Özoğuz, S., Elwakil A.S., and Ergün, S., "Cross-coupled Chaotic Oscillators and Application to Random Bit Generation," IEE Proc. Circ. Devices Syst., vol. 153, no. 5, (2006) 506-510

[7] A.M. Zoubir, and D.R. Iskander. "The bootstrap: A tutorial for the signal processing practitioner," IEEE Signal Processing Magazine, pp. 10-19, July 2007.

[8] National Institute of Standard and Technology, Information Technology Laboratory, http://csrc.nist.gov/groups/ST/toolkit/rng/index.html