A Study on Output Bit Tampering of True Random Number Generators Using Time-Varying EM Waves

Saki Osuka[1, 2], Daisuke Fujimoto[1], Arthur Beckers[2], Benedikt Gierlichs[2] Ingrid Verbauwhede[2] and Yuichi Hayashi[1]
[1] Nara Institute of Science and Technology, 8916-5 Takayama-cho, Ikoma, Nara 630-0192, Japan
[2] ESAT/COSIC, KU Leuven, Kasteelpark 10, B-3001 Heverlee, Belgium
Email: osuka.saki.ok1@is.naist.jp

***Abstract*— True random number generators (TRNGs) based on ring oscillators are employed in many devices because they can be constructed with logic gates only. If the randomness is degraded by the attack, it may affect the security of the cryptographic protocol. As a physical attack against TRNG, several attack methods that reduce its randomness have been proposed. On the other hand, TRNG is recommended to implement an online health test to guarantee the quality of random numbers, and if this module exists, conventional attacks would be difficult. In this paper, we have shown the possibility of tampering with a part of the random number sequence by controlling the electromagnetic waves applied to the TERO-based TRNG. Specifically, disturbance waves are applied to TRNG to change the power supply voltage of the device. By modulating the amplitude of the disturbance wave, some of the random number bits are tampered with.**

Keywords*— *true random number generator, intentional electromagnetic interference, health test

I. Introduction

True random number generators (TRNGs) generate random numbers based on a random physical phenomenon called an entropy source. Random numbers generated by TRNGs are used for various cryptographic systems as session keys, nonces, initialization vectors, and masks [1]. The security of these applications relies on the uniformity and unpredictability of the utilized random numbers. To ensure the randomness of the TRNG is very important for the security of the cryptographic protocol. Therefore, it has already been recommended for TRNG to implement health tests that evaluate randomness [1], [2].

TRNGs based on ring oscillators (ROs) are particularly used in many devices because their circuit can be implemented using only digital logic. Among them, the transition effect ring oscillator (TERO)-based TRNG is widely used because it can generate random numbers with a high entropy and bit rate [3]. Therefore, if the TERO-based TRNG is attacked and the randomness is reduced, many systems can be affected. As physical attack methods against TERO-based TRNG, a frequency injection attack [4], [5] and supply voltage manipulation [6] are known. These attacks cause a decrease in randomness due to intentional electromagnetic interference (IEMI) with the device. However, if a health test module is implemented, attacks may be detected or disabled. Therefore, it is difficult to affect TRNG with conventional IEMI-based attacks.

In this paper, we propose an attack method that reduces randomness even in TRNG equipped with a health test module. Specifically, disturbance waves that deterministically change the output bit are temporarily injected into the TERO-based TRNG. If we can know the output bits affected by the disturbance waves, it is possible to predict a part of the random number sequence. At this time, by setting the period of applying the disturbance waves to a constant multiple of the output rate of the TRNG, it becomes easy to predict the periodicity of the bit whose entropy is reduced.

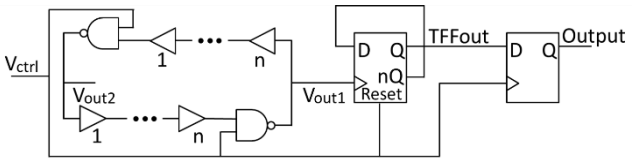

Fig. 1. Basic TERO-based TRNG.

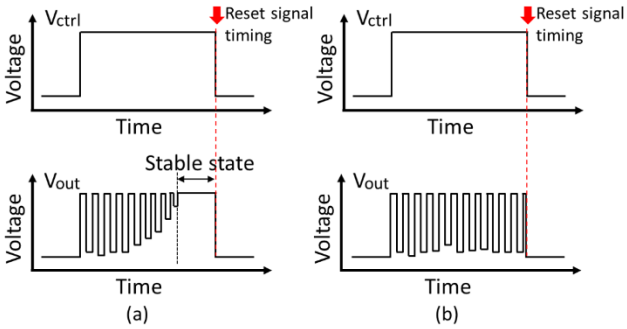

Fig. 2. TERO's input/output waveforms (a) when TERO has good entropy and (b) when the output bits are biased.

The remainder of this paper is organized as follows: In Sec. 2, the architecture of TERO-based TRNG and the effect of frequency injection attack is explained. Next a proposed method that can deterministically tamper with the output bits is described. Sec. 3 demonstrates that the output bits can be tampered with by using the proposed method. We conclude this paper in Sec. 4.

II. Proposal of Output-bit Tampering Attack against TERO-based TRNG

A. TERO-based TRNG Architecture targeted for physical attack

The TERO-based TRNG was proposed in [3]. The TERO is a multi-event ring oscillator that oscillates temporarily. It is composed of an even number of inverters and a couple of NAND gates that restart temporary oscillations (Fig. 1). Following the rising edge of the V_{ctrl} input, the outputs V_{out1} and V_{out2} start to oscillate (Fig. 2 (a)). This state is called the metastable state. These signals circulate inside the loop until a collision occurs, during which the edge that moves faster reaches the slower one. The difference in the speed of these signals is caused by differences in delays between inverters in loop branches and noise. After the collision, these signals are stabilized, and this state is called the stable state. The number of oscillations before the outputs stabilize is not constant but varies due to the electronic noises.

TERO-based TRNG uses the least significant bit (LSB) of the number of oscillations as a random number, and the LSB is counted by a T flip-flop (TFF). The output of TFF is sampled at the falling edge of V_{ctrl}. If the output signal of the

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

TFF is low when V_{ctrl} falls, the output bit is 0. On the other hand, if the output signal of the TFF is high when V_{ctrl} falls, the output bit is 1. In order to harvest sufficient entropy, the TERO has to stabilize before V_{ctrl} falls. If TERO continues to oscillate until the fall of V_{ctrl}, the number of oscillations will be greatly biased (Fig. 2 (b)). As a general configuration, every time an output bit is generated, the TFF that counts the number of oscillations is reset [7], [8]. The reset signal is input after generating an output bit. This signal is generated by the clock, and the TFF is reset after a certain period from the rising edge of the V_{ctrl}.

B. Frequency injection attack against TERO-based TRNG

In a frequency injection attack, a resonator with a certain natural frequency will lock to a different frequency when driven by it, similar to the resonance phenomenon of a pendulum. Jitter is suppressed by applying a signal of the fundamental wave of the oscillation frequency of TERO or its harmonics to TERO-based TRNG. As a result, the collision between the edges of V_{out1} and V_{out2} is suppressed, and the TERO oscillation continues. Reference [5] shows that by applying a second harmonic of the oscillation frequency of TERO, oscillation continues until the fall of V_{ctrl} as shown in the Fig. 2 (b). At this time, the number of oscillations of TERO and the output bit which is its LSB depend on the period of V_{ctrl} and the oscillation frequency of TERO. Based on these conditions, it is determined whether the output of TRNG is biased to 0 or 1. On the other hand, RO generates self-heating due to a switching event during oscillation. Therefore, in the TERO, the oscillation frequency changes continuously during the oscillation process [9]. This change in oscillation frequency may affect the lock efficiency due to injection. Reference [10] shows that the decrease in randomness due to TERO locking due to frequency injection attacks does not occur in all bits. From this reason, it is considered difficult to deterministically tamper with the output bits of TERO-based TRNG by frequency injection attack.

C. Proposed method

In order to tamper with output bits of TERO-based TRNG, an attack method that always causes a change in TERO oscillation as shown in Fig.2 (b) is required. As mentioned in the previous section, deterministic output bit tampering with frequency injection attacks is difficult. On the other hand, [10] shows that the oscillation may be stabilized by self-locking in which the edges of V_{out1} and V_{out2} lock each other. If the oscillation frequencies and initial noise conditions of V_{out1} and V_{out2}, which are likely to cause self-locking, can be given by external interference at the TERO oscillation start time, it may be possible to efficiently tamper with TRNG random numbers.

In the proposed method, a disturbance wave that fluctuates in the same phase between V_{dd}/GND is applied, and mode conversion occurs in the imbalance and mismatch parts inside the device. In this way, the voltage difference between V_{dd}/GND can be varied with time. Manipulating the power supply voltage for TRNG changes the oscillation frequency of TERO, which is the source of entropy. Therefore, if disturbance waves can be applied to the device and the power supply voltage can be manipulated, it may be possible to manipulate the oscillation frequency of TERO and cause a decrease in random number due to self-locking.

In this paper, we propose a method to deterministically tamper with the output bit of TRNG by manipulating the potential between V_{dd}/GND using the applied electromagnetic

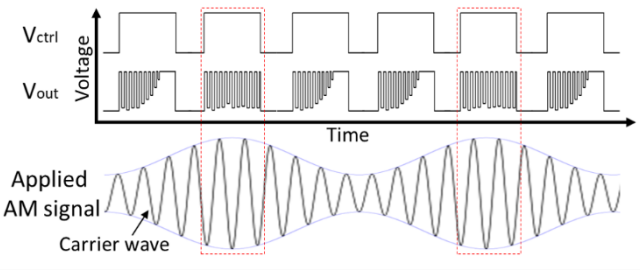

Fig.3. TRNG operation when amplitude-modulated signal is applied

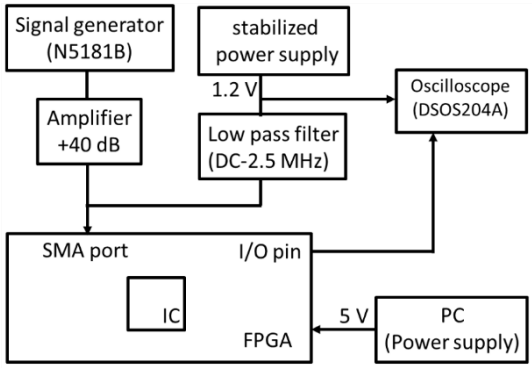

Fig.4. Experimental setup

(EM) wave. Specifically, the potential between V_{dd}/GND is changed by applying a time-varying EM wave to the device. And then, depending on the supply voltage of the device, TRNG generates a bit whose output is biased and a bit whose randomness is maintained. Amplitude modulated signals are used as time-varying EM waves. When an amplitude-modulated signal is applied, TRNG operates as shown in the Fig. 3. Using the injection frequency f_1 that reduces the randomness of TRNG as a carrier wave, a signal subjected to amplitude modulation (AM) at a different frequency f_2 is applied to the target TRNG. As shown in the figure, when the injection amplitude is high, TERO continues to oscillate until the fall of V_{ctrl}. At this time, since the number of oscillations is biased to a specific value according to the period of V_{ctrl} and the oscillation frequency of TERO, the output bit is also biased to either 0 or 1. When the injection amplitude is low, the output keeps its randomness. At this time, by setting the modulation frequency f_2 to a frequency of 1/n of the bit rate of TRNG, it is possible to periodically generate bits whose output is fixed by an attack (n is a natural number).

III. Experiment

This section shows the change in power supply voltage and the decrease in randomness due to the application of EM waves, and also shows the possibility that the randomness can be manipulated by applying an time-varying EM wave.

A. Experimental Setup

A target TERO-based TRNG was implemented in a field programmable gate array (FPGA). The TERO-based TRNG under attack was designed with reference to [8]. The clock signal that controls the TERO was generated by an 11-stage RO and an 8-bit asynchronous counter. When the rising edge of the clock was input, the TERO started oscillating and stabilized before the clock fell. A reset signal was input to the TFF at the falling edge of the clock pulse. The TERO consisted of two NAND gates and 20 inverters. The number of oscillations of the TERO was counted with the TFF, and the TFF output at the fall of V_{ctrl} was a 1-bit random number.

Regarding the implementation of TERO on FPGA, we implemented two TERO branches adjacent to each other in the same topology.

Figure 4 shows the experimental setup. In this paper, we constructed an experimental setup that allows two experiments to evaluate the decrease in randomness due to the manipulation of the power supply voltage, and the decrease in randomness due to the application of EM waves. The target TERO-based TRNG was implemented on the FPGA (Spartan6 LX9 Microboard). In order to confirm the feasibility of the proposed method, modifications were made to reduce the immunity of the FPGA. The modifications were the removal of the decoupling capacitor around the IC and the implementation of the SMA port for applying EM waves near the IC. The FPGA was powered from a PC at 5 V, and the regulator on the board supplies the IC with 1.2 V. In order to show the decrease in randomness due to the change in power supply voltage, the SMA port attached to the FPGA was branched by a T connector and connected to the stabilized power supply via the low-pass filter (DC-5 MHz). The wiring between the stabilized power supply and the low-pass filter was probed and measured with an oscilloscope (Keysight, DSOS204A) to evaluate the power supply voltage supplied to the IC. Further, in order to apply EM waves, a continuous wave generated from a signal generator (Keysight, N5181B) was applied to the FPGA from the SMA port via a 10W amplifier. The effect of TERO-based TRNG on attack was evaluated by measuring the TFF waveform and output bit waveform of the TERO-based TRNG output from the I/O pins.

B. *Decrease in randomness due to changes in power supply voltage*

In this section, we investigate the changes that occur in TERO-based TRNG by directly manipulating the power supply voltage supplied to the FPGA. And then, it is shown that the potential between V_{dd}/GND changes due to the introduction of EM waves, which causes a decrease in randomness.

First, the decrease in randomness due to the change of the power supply voltage is investigated. In the experiment, the power supplied from the stabilized power supply to the IC via the SMA port was manipulated. The attack target TRNG uses TERO with an oscillation frequency of 109.4 MHz to generate random numbers. The signal that controls TERO is a 286.8 kHz clock signal obtained by dividing RO (oscillation frequency: 145 MHz) with a 9-stage asynchronous counter. In addition, by using an oscilloscope the TFF waveform and the clock waveform was observed, and also the power supply voltage was evaluated by an oscilloscope. At this time, the inverted clock signal was output.

Fig. 5 shows the result of supplying a voltage of 1.2 V, which is the standard voltage, and the result of supplying a voltage (1.43 V) in which TERO continued to oscillate until V_{ctrl} fell, which was searched by sweeping the power supply voltage. When a standard voltage is supplied, it can be confirmed that the TERO's oscillation has stopped by the time the reset signal comes, according to the TERO design shown in Sec. 2. At this time, the output bitstreams showed randomness. The potential between V_{dd}/GND measured with an oscilloscope was 1.10 V. On the other hand, as shown in Fig. 5 (b), it can be confirmed that the TFF output waveform when a voltage of 1.43 V is supplied stops oscillating at the same time as the rising edge of the clock signal. Since the

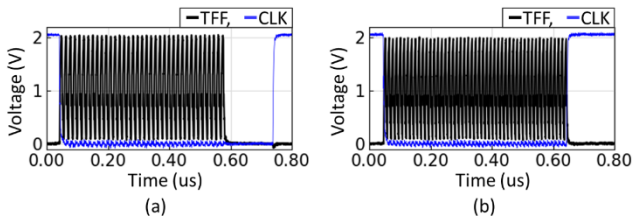

Fig. 5. Changes in TFF output waveform when the supply voltage is changed. (a) TFF waveform (black) and inverted clock waveform (blue) when the supply voltage is 1.2 V. (b) TFF waveform (black) and inverted clock waveform (blue) when the supply voltage is 1.43V.

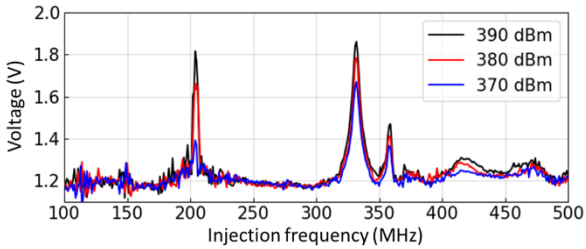

Fig. 6. Change in potential between V_{dd}/GND with respect to the frequency and amplitude of the applied EM wave.

number of oscillations of TERO was greatly biased, all output bits took a value of 1. At this time, the potential between V_{dd}/GND was 1.33 V.

Next, the change in potential between V_{dd}/GND and the decrease in randomness caused by the application of EM waves are shown. In the experiment, the potential between V_{dd}/GND and the TFF output waveform were acquired while sweeping the frequency and amplitude of the applied EM wave. Fig. 6 shows the change in potential between Vdd/GND with respect to the injected frequency and amplitude. The horizontal axis shows the injection frequency, and the vertical axis shows the electric potential. The injection amplitude is as shown in the legend. There are large peaks at 205 MHz, 332 MHz and 358 MHz, and it can be confirmed that the potential changes significantly. It can also be confirmed that the change in potential increases as the injection amplitude increases. At this time, the behavior of TRNG under some attack conditions that matches the potential of 1.33 V between V_{dd}/GND was evaluated. It was confirmed that all output bits became 1 even when 332 MHz and 358 MHz, in which the potential between V_{dd}/GND fluctuated greatly, were applied, as in the case of operating power supply voltage. On the other hand, applying a 205 MHz sine wave reset the FPGA and TRNG did not work.

From the above, it was shown that the output bit can be fixed by changing the potential between V_{dd}/GND by applying EM waves to TRNG.

C. *Tampering with output bits using amplitude-modulated signal*

In this section, as an example of using the attack method presented in the previous section, it is shown that applying an amplitude-modulated signal can tamper with a part of the output bits.

In the experiment, as an experimental system that can apply signals efficiently, an FPGA was modified to bypass the regulator and power the IC directly. The attack target TRNG uses TERO with an oscillation frequency of 115.6 MHz to generate random numbers. The signal that controls TERO is a 602 kHz clock signal obtained by dividing RO with a 8-stage asynchronous counter. The injection frequency was swept and a sine wave (frequency: 237 MHz, amplitude: 38.5 dBm),

which efficiently causes a decrease in randomness, was selected as the carrier wave. As a modulated wave, a sine wave of 120.4 kHz, which is 1/5 times the frequency of the clock signal, was used. The depth of modulation was set to 100%.

Fig. 7 shows the output bit stream when an amplitude-modulated EM wave is applied. This bit stream is displayed as a two-dimensional black and white image corresponding to ones and zeros, respectively, raster scanning left-to-right, then top-to-bottom. In other words, the upper left cell indicates the 1st bit, and the lower right cell indicates the 100th bit. It can be confirmed that the cells in the 2nd and 3rd columns all show 0. Similarly, the 7th and 8th columns all show 0. Here, the TRNG used in this section has a different TERO and clock oscillation frequency from the TRNG used in 3.B, so the bit whose randomness is reduced by the attack outputs 0. In contrast, it can be confirmed that 0 and 1 appear randomly in the other columns. From this, the application of the amplitude modulated wave using the signal of the frequency 1/5 of the output rate of TRNG caused a decrease in randomness with a 5-bit period. This result shows the possibility of tampering with the output bit of 2 bits per 5 bits.

Therefore, we have demonstrated the possibility of biasing a part of the output bits of TRNG by applying the amplitude modulated wave to TRNG as a disturbance wave. Since the bits whose output is biased appear periodically, it is considered that a part of the random number sequence can be estimated. In this attack, by changing the modulated wave, the period of the bit whose output is biased could be changed.

We consider the reason why the power supply voltage supplied to the IC changes due to the application of EM waves. In the device used in this paper, we observed the phenomenon that the DC value of the power supply voltage rises due to the injection of sinusoidal waves. IC power supplies are often equipped with circuits for electro-static discharge (ESD) protection[11]. Generally, a diode structure is used for ESD protection, and when a potential difference exceeding the breakdown voltage occurs between the V_{dd}/GND, a current flows through the diode. It is possible that this phenomenon caused distortion in the potential fluctuation due to the sinusoidal waves, and the DC value of the power supply voltage changed.

IV. Conclusion

In this paper, assuming an attack on TRNG with a health test module implemented, we proposed an attack method that can predict a part of the output by tampering with random number bits. It was shown that an attack that degrades the randomness of TERO-based TRNG is possible by manipulating the power supply voltage of the device by applying EM waves. Since this decrease in random number deterministically changes the output bit, it is possible to tamper with a part of the output bit by changing the applied EM wave. In the experiment, the output bits of TRNG were reduced in random number with a 5-bit cycle, but it is considered that the output bits can be tampered with in different patterns by changing the attack conditions. Therefore, it is considered that the attack by the proposed method can give a decrease in random number that is difficult to detect even if the health test is implemented. By tampering with a part of the output bits, the pattern of random numbers output by TRNG may be reduced, and the entropy of the output bit string may be greatly reduced.

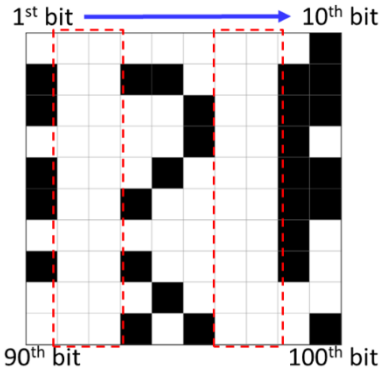

Fig. 7. TRNG output bit when an amplitude-modulated signal is applied. A continuous 100-bit TRNG output is expressed on a two-dimensional map, with black as 1 and white as 0. The upper left cell indicates the 1st bit, and the lower right cell indicates the 100th bit.

Acknowledgment

This work was supported in part by JPSP KAKENHI Grant Number 21J14528.

References

[1] M.S. Turan, E. Barker and J. Kelsey, K. Mckay, M.Baish and M. Boyle, “Recommendation for the entropy sources used for random bit generation,” Technical Report SP800-90B, National Institute of Standards and Technology, USA (2018).

[2] W. Killmann and W. Schindler. A proposal for: Functionality classes for random number generators, version 2.0. Technical report, Bundesamt fur Sicherheit in der Informationstechnik (BSI), Bonn (2011).

[3] M. Varchola and M. Drutarovsky, “New high entropy element for FPGA based true random number generators," in Cryptographic Hardware and Embedded Systems, CHES 2010, pp. 351–365, Springer (2010).

[4] A.T. Markettos, S.W. Moore, “The frequency injection attack on ring-oscillator-based true random number generators,” CHES, vol. 5747, pp. 317–331, 2009.

[5] P. Haddad, C. Kasmi, J.L. Esteves, V. Houchouas, “Electromagnetic harmonic attack on transient effect ring oscillator based true random number generator,” Hardwear.io Conference, Den Haag, Netherlands, 2016.

[6] Y. Cao, V. Rožić, B. Yang, J. Balasch and I. Verbauwhede,. Exploring active manipulation attacks on the TERO random number generator. In 2016 IEEE 59th International Midwest Symposium on Circuits and Systems (MWSCAS) (pp. 1-4). IEEE.

[7] P. Haddad, V. Fischer, F. Berdnard, and J. Nicolai, "A Physical Approach for Stochastic Modeling of TERO-based TRNG," In International Workshop on Cryptographic Hardware and Embedded Systems, pp. 357-372, Springer, Berlin, Heidelberg (2015).

[8] O. Petura, U. Mureddu, N. Bochard, V. Fischer and L. Bossuet, "A survey of AIS-20/31 compliant TRNG cores suitable for FPGA devices," In 2016 26th international conference on field programmable logic and applications (FPL), pp.1-10, IEEE, (2016).

[9] J. Delvaux, "Refutation and Redesign of a Physical Model of TERO-based TRNGs and PUFs." IACR Cryptol. ePrint Arch. 2019 (2019): 810.

[10] U. Mureddu, N. Bochard, L. Bossuet and V. Fischer, “Experimental study of locking phenomena on oscillating rings implemented in logic devices,” IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 66(7), pp. 2560-2571, 2019.

[11] Ming-Dou Ker and Kuo-Chun Hsu, "Latchup-free ESD protection design with complementary substrate-triggered SCR devices," in IEEE Journal of Solid-State Circuits, vol. 38, no. 8, pp. 1380-1392, Aug. 2003, doi: 10.1109/JSSC.2003.814434.