Physical Unclonable Functions based on Temperature Compensated Ring Oscillators

Sha Tao and Elena Dubrova
KTH Royal Institute of Technology
Kistagången 16, 164 40 Stockholm, Sweden

Abstract. Physical unclonable functions (PUFs) are promising hardware security primitives suitable for low-cost cryptographic applications. Ring oscillator (RO) PUF is a well-received silicon PUF solution due to its ease of implementation and entropy evaluation. However, the responses of RO-PUFs are susceptible to environmental changes, in particular, to temperature variations. Additionally, a conventional RO-PUF implementation is usually more power-hungry than other PUF alternatives. This paper explores circuit-level techniques to design low-power RO-PUFs with enhanced thermal stability. We introduce a power-efficient approach based on a phase/frequency detector (PFD) to perform pairwise comparisons of ROs. We also propose a temperature compensated bulk-controlled oscillator and investigate its feasibility and usage in PFD-based RO-PUFs. Evaluation results demonstrate that the proposed techniques can effectively reduce the thermally induced errors in PUF responses while imposing a very low power overhead.

Keywords: Physical unclonable function (PUF), delay-based PUF, RO-PUF, temperature variation, hardware security, device authentication.

1 Introduction

Physical unclonable functions (PUFs) which extract inherent randomness in physical devices have recently emerged as a light-weight cryptographic primitive. Over the past decade, PUFs have found their usage in many practical applications such as device identification, authentication, secret key generation and storage [1, 2]. Among different solutions, the most popular and low-cost PUFs are based on uncontrollable and unpredictable process variations during silicon manufacturing. From an implementation perspective, most existing silicon PUFs can be widely divided into two categories: (1) delay-based, including arbiter PUFs [1] and ring oscillator (RO) PUFs [2], (2) bistable element-based, such as SRAM-PUFs [3] and bistable ring PUFs [4]. An extensive characterization and comparative study of different silicon PUFs fabricated on the same CMOS technology can be found in [5].
1.1 Prior Related Works

Among many PUF variants, the PUF structure based on comparing the frequency difference of ring oscillator pairs, called RO-PUF, has become a popular solution due to its ease of implementation and entropy evaluation [2]. For reliable identification and authentication, PUF designs are expected to provide stable responses under varying operating conditions. However, the frequency of ROs changes with varying environment. In addition, the rate of change in frequency can also vary for different ROs. These, in turn, lead to unstable responses.

In this subsection, we review existing reliability enhancement methods applied specifically to RO-PUFs. Error correction codes (ECCs) have been traditionally applied to RO-PUFs to produce error-free responses [6]. However, ECCs usually impose high overhead, which also scales up quickly with the increased number of error correction bits [7]. Alternatively, various architectural and circuit level techniques have been proposed in the literature to improve reliability of RO-PUFs.

From an architectural perspective, the first RO-PUF design [2] employs a 1-out-of-N masking scheme that selects only those pairs with frequency differences sufficiently large to overwhelm environmental noise. In [8], a temperature-aware cooperative RO-PUF is proposed, which defines a bit-generation rule so as to convert unreliable bits into reliable ones. A more hardware-efficient approach is suggested in [9], where a feedback-based supply control scheme is used to vary supply voltage according to operating temperature. Another approach, introduced in [10], improves the reliability of RO-PUFs against temperature variation through applying multiple supply voltages on different inverters in each RO.

From a circuit perspective, one way of increasing the stability of RO-PUFs against supply variation is to operate the transistors with a forward body bias [11]. In [12], two methods for reducing temperature sensitivity of delay based PUFs are proposed. The first finds an optimum power supply under which temperature effects are minimized. The second uses negative temperature coefficient resistance to compensate for temperature effects of inverters. Both methods are applied to a conventional RO-PUF and a phase differential RO-PUF proposed in [13]. Most recently, a hybrid RO-PUF is proposed [14], which uses the positive temperature coefficient of current starved inverters to offset the response instability due to the negative temperature coefficient of regular inverters. Furthermore, in [15], the impact of aging effect is addressed, and the first aging-resistant PUF design is presented.

1.2 Contributions and Organization

In this work, we aim to counteract the effect of temperature-induced frequency errors in ring oscillators, which is considered to be a major problem that can notably degrade the reliability of a RO-PUF. Our contributions can be summarized as follows: (1) We investigate and compare two circuit-level techniques for realizing temperature compensated ROs, namely, bulk-controlled oscillator (BCO) and current-starved voltage-controlled oscillator (CSVCO). (2) Inspired by the
Fig. 1. Block diagram of the proposed temperature aware PUF.

2 Proposed PUF Design

The concept of the proposed temperature aware PUF is illustrated in Fig. 1. To generate a response, one pair of temperature compensated ROs is enabled and selected by two multiplexers that are controlled by a random challenge. Unlike conventional RO-PUFs that use power- and time-consuming counters and comparators to compare oscillating frequencies, the proposed PUF employs a phase/frequency detector (PFD) followed by a charge pump (CP) which are widely used in phase-locked loops (PLLs) for clock generation [16].

2.1 RO-PUFs based on Phase/Frequency Detector

Fig. 2 shows the schematic and timing diagram of the designed PFD and CP. The PFD, consisting of two D flip-flops (DFF), compares the leading edges of two signals, ‘IN1’ and ‘IN2’, coming from two selected ROs. The outputs of PFD, ‘UP’ and ‘DN’, depend on both phase and frequency differences between two input signals. For instance, Fig. 2 (b) illustrates the case of ‘IN1’ leading ‘IN2’: ‘UP’ goes high when there is a rising edge on ‘IN1’ and is reset by a rising edge on ‘IN2’. The transition is similar for the case shown in Fig. 2 (c) where ‘IN2’ is faster. The ‘UP’ and ‘DN’ signals control the CP consisting of two switched current sources: when ‘UP’ is high and ‘DN’ is low, the load capacitor is charged thus raising the voltage of ‘OUT’; when ‘DN’ is high and ‘UP’ is low, the load
Fig. 2. PFD and CP: (a) circuit schematics, (b) timing diagram when $f_{IN1} > f_{IN2}$, and (c) timing diagram when $f_{IN1} < f_{IN2}$.

capacitor is discharged thus reducing the voltage of ‘OUT’; when both ‘UP’ and ‘DN’ are low, both switches are off and thus the voltage of ‘OUT’ remains. By the end of each enable period, ‘OUT’ can be captured by a logic analyzer in practice with thresholds $V_{TH} = V_{TL} = V_{DD}/2$.

Compared to the phase differential RO-PUF [13] that uses a latched comparator to sense an instant output voltage difference of two ROs at end of an enable signal ‘EN’, our PFD-based approach not only consumes less power, but also provides better comparison accuracy, since the phase/frequency difference of two ROs are accumulated for a period and thus gets “amplified” rather than being decided at once. Moreover, this PFD-based approach has no concern of reliability issue induced by the jitter effect of the enable signal ‘EN’. On the other hand, in this approach, the result is not directly varied by the duration of ‘EN’. Therefore, we cannot apply the area-efficient method in [13], where ‘EN’ with various pulse widths are used as challenges feeding into the same RO pair.

Like in conventional RO-PUFs, to generate $n$ output responses with maximum entropy, the total number of ROs, $N$, needs to satisfy: $\log_2(N!) \geq n$. For instance, to generate $n = 128$ independent bits, $N = 36$ ring oscillators are required. We can use the redundant $\log_2(36!) - 128 \approx 10$ bits to perform an optional data screening (DS) at post-processing. The DS eliminates 10 out of 138 output responses while the candidates are determined according to the detected phase/frequency difference. We remove the responses generated by 10 RO
pairs with the minimum absolute value of ‘OUT’, which can be recorded by an oscilloscope in practice.

2.2 Temperature Compensated Bulk-Controlled Oscillators

Voltage controlled oscillator (VCO) is the core building block of PLL circuits. VCOs which employ current starvation to control their output frequencies are called current-starved VCO (CSVCO) [17]. However, the relationship between output frequency and input control voltage is very non-linear in CSVCOs. Alternatively, controlling the frequency through tuning the bulk voltage of transistors can alleviate the linearity issue [18]. In this subsection, we introduce a bulk-controlled oscillator (BCO) with dynamic temperature correction to be used in the proposed PUF. We investigate its hardware overhead and its feasibility in minimizing frequency errors due to environmental changes. We also demonstrate the advantages of this BCO based approach over its CSVCO based counterpart.

Fig. 3 (a) shows the schematic of a CSVCO consisting of a 10-stage CS inverter chain whose delay is controlled by $V_{\text{Ctrl}}$, buffer inverters whose delays are
made negligible compared to the CS inverter chain, and a NAND gate to enable or disable the RO. It has been shown in [19] that using CS inverters in delay based PUF circuits is beneficial, as they can increase the standard deviation of delay compared to regular inverters. For ROs, this advantage remains. The control voltage $V_{\text{Ctrl}}$ modulates the on-resistance of current source transistors $M_{BN1-10}$ and $M_{BP1-10}$ through a BIAS circuit. These variable resistances regulate peak current available at each inverter stage and thus vary its propagation delay which, in turn, controls the frequency of RO. The frequency of CSVCO, $f_{\text{OSC}}$, is proportional to $(V_{\text{Ctrl}} - V_{\text{th}})^2$ when the current source transistors are in saturation [17]. Fig. 4 (a) illustrates the simulation results when $V_{\text{DD}}$ is 0.9V, showing that the $f_{\text{OSC}}$ increases with increasing $V_{\text{Ctrl}}$ in a non-linear way. To maximize the mismatch, the core inverters are biased in the weak inversion region. In this region, with the increase of temperature, the decrease of threshold $V_{\text{th}}(T)$ dominates the decrease of mobility $\mu(T)$ [14]. Additionally, the current at each stage is proportional to $\mu(T)$ and inversely related to $V_{\text{th}}(T)$. Thus, $f_{\text{OSC}}$ of the CSVCO increases with the increase of temperature. This is confirmed by the simulation results shown in Fig. 4 (b).

Alternatively, we can solve the linearity issue in the CSVCO by controlling the bulk potential of inverters. Fig. 3 (b) shows the schematic of a BCO, which replaces the CS inverters in Fig. 3 (a) with regular inverters. The bulk of all
PMOS, \( M_{P1-P10} \), is connected to a control voltage \( V_{Bulk} \). When PMOS is biased to a potential other than \( V_{DD} \), its threshold voltage is proportional to \( \gamma(\sqrt{2\phi_f-V_{BS}}) \) \[18\], where \( \gamma \) is a process constant and \( \phi_f \) is the surface potential of the PMOS. Therefore, when \( V_{BS} \) increases, \( V_{th} \) reduces, leading to an increased current at each stage and, consequently, an increase of \( f_{OSC} \). This is confirmed by the simulation results illustrated in Fig. 5 (a) (\( V_{DD} \) is 0.9V). It shows a linear relationship between the frequency of BCO and the bulk-to-source potential \( V_{BS} \) although its tuning range is smaller than the CSVCO case. Inverters in the BCO operate in the saturation region, where with the increase of temperature, the decrease of mobility \( \mu(T) \) dominates the decrease of threshold \( V_{th}(T) \) \[12\]. Thus, when temperature is increased, the overall current is reduced leading to a decrease in frequency. This is confirmed by the simulation results shown in Fig. 5 (b).

From the above analysis, we can see that the thermal effect of CSVCO and BCO can be compensated by adjusting their control voltages \( V_{Ctrl} \) and \( V_{Bulk} \) respectively. In Fig. 6, we compare the resulting frequency errors with respect to the reference frequency at 25°C of two temperature compensated ROs\[2\]. Compared to CSVCO, the BCO based approach is much more effective in reducing temperature sensitivity: the frequency error is limited to less than 1%. Fig. 7 plots the distribution of relative frequency deviation due to process variation obtained from 1000 instances of each RO, indicating that the “desired” frequency deviation among different PUF instances is larger than the “undesired” frequency error due to temperature changes. The averaged current consumed by BCO and CSVCO is 14.55\( \mu A \) and 21.34\( \mu A \), respectively\[3\]. Fig. 8 shows the layouts of two ROs, implemented manually in Cadence Virtuoso. Thanks to the elimination of 20 chunky current source transistors used for current starvation, BCO is much more area-efficient than CSVCO. The chip area of the BCO and CSVCO is 39.27\( \mu m^2 \) and 613.74\( \mu m^2 \), respectively.

Fig. 9 illustrates an implementation example of the voltage references\[4,5\] for the proposed BCO and CSVCO, which are derived from a 1.2V external reference \( V_{DD,ext} \). Fig. 10 shows their simulation results. Both control voltages are generated by a simple NMOS voltage divider, whose output \( V_{Bulk} \) or \( V_{Ctrl} \) has a negative temperature coefficient (TC) and thus decreases with increasing temperature.

---

1 We control the bulk of PMOS not NMOS due to the following reasons: (1) Biasing the bulk of NMOS requires a triple well CMOS process while biasing the bulk of PMOS can be applied to any standard CMOS process. (2) Creating a bias voltage around \( V_{DD} \) for PMOS is more straightforward and easier than creating a bias voltage around \( G_{ND} \) for NMOS.

2 For fair comparison, both ROs oscillate at a similar nominal frequency. Thus, BCO is powered by 0.9V and CSVCO uses a higher supply of 1V.

3 Note that essentially the CSVCO can have a much less current consumption with a much lower \( f_{OSC} \) when biased towards deep weak-inversion.

4 In addition to regular transistors (RVT), high threshold voltage (HVT) NMOS transistors are used in the design to reduce static current consumption.

5 Note that these voltage references can be shared by all ROs in the PUF IC in order to minimize hardware overhead.
Fig. 6. Frequency error (compared to the reference at 25°C) after temperature compensation: (a) BCO and (b) CSVCO.

Fig. 7. Histogram of the frequency deviation obtained from 1000 Monte-Carlo simulations: (a) BCO and (b) CSVCO.

Fig. 8. Layouts of (a) BCO and (b) CSVCO.
Fig. 9. Voltage reference generation circuits.

Fig. 10. Temperature dependency of voltage references.

temperature. On the other hand, the supply voltage $V_{DD}$ for both ROs should not be sensitive to temperature changes. As shown in Fig. 9, a voltage divider consisting of both PMOS and NMOS is used to generate such a temperature insensitive reference\(^6\). The basic idea is to use the positive TC of PMOS threshold ($\frac{\partial V_{th,P}}{\partial T} > 0$) to compensate the negative TC of NMOS threshold ($\frac{\partial V_{th,N}}{\partial T} < 0$), so as to achieve an almost zero TC\(^7\) at the output.

Next, we briefly discuss the impact of supply variation on both ROs. In both cases, when $V_{DD,ext}$ varies, $V_{DD}$, $V_{Bulk}$ and $V_{Ctrl}$ follow the voltage changes accordingly in the same direction. For instance, when $V_{DD,ext}$ decreases, a drop in both $V_{DD}$ reduces the $f_{OSC}$ of both ROs. For BCO, a drop in $V_{Bulk}$ reduces

\(^6\) In practice, such a temperature-resistance $V_{DD}$ can be generated by a bandgap voltage reference, e.g., a low power bandgap presented in [20].
the $V_{DS}$, and thus increases the $f_{OSC}$ according to Fig. 5 (a). For CSVCO, on the other hand, a drop in $V_{Ctrl}$ reduces the current at each CS inverter stage, and thus further decreases the $f_{OSC}$. Therefore, in comparison with its CSVCO counterpart, the BCO based approach has another advantage in minimizing the PUF’s sensitivity to supply noise.

3 Evaluation Results

3.1 Methodology

The PUF circuits were designed using a standard 65 nm CMOS technology and simulated using Cadence Spectre. Matlab was used to perform post-processing and compute performance metrics. The nominal operating condition was at $25^\circ C$ with $V_{DD} = 0.9V$ for BCO-PUF and $V_{DD} = 1V$ for CSVCO-PUF. To emulate the characterization of 100 different PUF IC chips, 100 runs of Monte-Carlo simulations were performed. Realistic models provided by a commercial foundry were used in the simulations. Both process and mismatch options were enabled to account for intra-die and inter-die variations. Such a transistor-level Monte-Carlo simulation based experimental methodology has been adopted in many PUF literatures, e.g., [11, 3, 7, 12, 13, 9, 19, 4, 15].

3.2 Uniqueness Analysis

We first evaluate the PUF’s uniqueness, $U$, at the nominal condition, using the average inter-chip Hamming Distance ($HD_{int}$) of the responses from $k$ different PUF instances:

$$U = \frac{2}{k(k-1)} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \frac{HD(R_i, R_j)}{n} \times 100\%,$$

where $R_i$ and $R_j$ are the $n$-bit responses of instances $i$ and $j$. Perfect identification of different chips requires $HD_{int} = n/2$ and $U = 50\%$. For both PUFs, we applied 128 random challenges to each of $k = 100$ instances. So, $100 \times 99/2 = 4950$ comparisons were used to compute the $HD_{int}$. As shown in Fig. 11, the average $HD_{int}$ of two PUFs are 63.96 and 64.04, corresponding to the uniqueness of 49.97% and 50.03%.

3.3 Randomness and Entropy

Randomness measures the proportion of ‘0’ or ‘1’ in PUF responses. Ideally, this proportion should be very close to 50%, since any bias towards ‘0’ or ‘1’ make the PUF responses predictable and easier to attack. Fig. 12 (a) and (b) depict output responses from 100 IC instances of BCO-PUF and CSVCO-PUF, characterized at the nominal condition, showing no systematic pattern or noticeable correlation among response bits. Randomness or unpredictability of PUF responses can also be evaluated by Shannon entropy, $H = -p_1 \log_2 p_1 - (1 - p_1) \log_2 (1 - p_1)$, where $p_1$ is the proportion of ‘1’s in $n$ response bits. Entropy $H$ approaches the ideal value of 1 when $p_1$ approaches 50%. Fig. 13 shows that the average entropy of BCO-PUF and CSVCO-PUF are 0.9944 and 0.9923.
Fig. 11. Distribution of inter-chip HD obtained from 100 chip instances of (a) BCO-PUF and (b) CSVCO-PUF.

Fig. 12. Monte-Carlo simulation results of 100×128-bit output responses for (a) BCO-PUF and (b) CSVCO-PUF.

Fig. 13. Entropy of (a) BCO-PUF and (b) CSVCO-PUF.

3.4 Reliability Against Temperature Variation
Reliability is related to the average intra-chip Hamming Distance (HD_{intra}) among responses of the same PUF instance obtained at varying environmental conditions. We characterized 100 instances of each PUF at different temperatures, and for each instance computed how many responses bits are changed out
of the total $n = 128$ responses. This ratio in percentage is defined as the bit error rate (BER): $\text{BER} = \frac{\text{HD}(R_i, R'_i)}{128} \times 100\%$, where $\text{HD}(R_i, R'_i)$ represents the HD$_{\text{intra}}$, $R_i$ is ground truth response of instance $i$, and $R'_i$ is the response taken at a different temperature. As shown in Fig. 14, with (w.i.) or without (w.o.) the aforementioned data screening (DS), BCO-PUF has a better temperature reliability than CSVCO-PUF.

3.5 Resistance to Side-Channel Attacks

RO-PUFs are potentially vulnerable to two types of side-channel attacks [21, 22]: passive power analysis attack and active fault injection attack. Regarding the passive attack, for conventional RO-PUFs with power consuming counters, it is possible to extract the frequencies of ROs by observing power traces and identifying periodic peaks when counters are toggling. The proposed PFD-based RO-PUF dissipates much less power in the DFFs and thus potentially alleviate this issue. With respect to the active attack, we injected a $\pm 10\%$ power supply ripple during PUF operation, following the method described in [21]. Simulation results show that the average BER induced by such a supply ripple attack is only 0.53% and 1.21% for BCO-PUF and CSVCO-PUF, respectively.

3.6 Performance Summary and Future Works

Table 1 summarizes performance metrics of the proposed BCO-PUF in comparison with CSVCO-PUF. Table 2 compares the proposed BCO-PUF to existing RO-PUF ASICs. Note that many RO-PUFs in the literature are not power aware designs and hence do not report their power overhead. Therefore, in this comparison, we only include existing low-power RO-PUF ASICs with available power metrics. We can see that the proposed temperature aware BCO-PUF is one of the best among the state-of-the-arts in terms of power efficiency. Future investigations include evaluating the reliability of PUF responses with aging effects [15]. We also plan to further analyze its security against modeling and power analysis attacks [22] and its susceptibility to physical tempering [23].


Table 1. Performance summary of proposed PUFs.

<table>
<thead>
<tr>
<th></th>
<th>BCO-PUF</th>
<th>CSVCO-PUF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output Throughput</td>
<td>50M bps</td>
<td></td>
</tr>
<tr>
<td>Response Size</td>
<td>128 bits</td>
<td></td>
</tr>
<tr>
<td>Transistor Count</td>
<td>1106</td>
<td>1934</td>
</tr>
<tr>
<td>Power (ROs)</td>
<td>28.7µA x 0.9V</td>
<td>42.2µA x 1V</td>
</tr>
<tr>
<td>Power (PFD+CP)</td>
<td>11.6µA x 0.9V</td>
<td>12.7µA x 1V</td>
</tr>
<tr>
<td>Average Entropy</td>
<td>0.9944</td>
<td>0.9923</td>
</tr>
<tr>
<td>Uniqueness (µ, σ)</td>
<td>(49.97%, 5.57%)</td>
<td>(50.03%, 5.58%)</td>
</tr>
<tr>
<td>Temperature Range</td>
<td>0°C to 100°C</td>
<td></td>
</tr>
<tr>
<td>BER w.i. DS</td>
<td>0.67%</td>
<td>2.81%</td>
</tr>
<tr>
<td>BER w.o. DS</td>
<td>0.82%</td>
<td>3.06%</td>
</tr>
<tr>
<td>Supply Fault Attack</td>
<td>±10% supply ripple</td>
<td></td>
</tr>
<tr>
<td>Average BER</td>
<td>0.53%</td>
<td>1.21%</td>
</tr>
</tbody>
</table>

Table 2. Comparison to low-power RO-PUF ASICs.

<table>
<thead>
<tr>
<th></th>
<th>Technology</th>
<th>RO Freq. per CRP</th>
</tr>
</thead>
<tbody>
<tr>
<td>[13] Optimized supply RO-PUF*</td>
<td>45nm</td>
<td>2.9GHz 82µW</td>
</tr>
<tr>
<td>[13] Negative resistance RO-PUF</td>
<td>45nm</td>
<td>2.6GHz 98µW</td>
</tr>
<tr>
<td>[14] Hybrid RO-PUF</td>
<td>65nm</td>
<td>0.32GHz 32.3µW</td>
</tr>
<tr>
<td>This work: BCO-PUF</td>
<td>65nm</td>
<td>1.13GHz 36.4µW</td>
</tr>
</tbody>
</table>

* Excluding power consumed by multiplexers and counters.
+ Excluding power consumed by sense amplifier.

4 Conclusion

This paper presented effective and efficient approaches for enhancing the reliability of RO-PUFs. We introduced a PFD-based alternative to the slow and power-hungry counter-and-comparator configuration. We also investigated two temperature-compensated ROs, i.e. BCO and CSVCO, and evaluated their usefulness in constructing temperature aware PUFs. Both PUF circuits were implemented in 65nm CMOS circuit and characterized. Results show that they achieve low power, high uniqueness, and satisfying randomness. The BCO-PUF demonstrates superior resistance to environmental variations than its CSVCO-PUF counterpart. In future works, we shall assess the aging effect of the proposed PUFs and perform security analysis.
Acknowledgment

This work was supported in part by the research grant No SM14-0016 from the Swedish Foundation for Strategic Research (SSF).

References