Abstract—Physical unclonable functions (PUFs) take advantage of the effect of process variation on hardware to obtain their unclonability. Traditional PUF design only focuses on the analog signals of circuits. An arbiter PUF, for example, generates responses by racing delay signals. Implementations of such PUFs usually employ large area and power consumption while providing very low throughput.

To address this problem, we propose an energy efficient PUF design in such a way that it races analog signals and computes digital logic simultaneously. More importantly, the analog portion of the circuit (racing) shares a large amount of hardware resources with the digital portion of the circuit (computing) by introducing only small overhead in terms of area and power. Our test results on Spartan-6 field-programmable gate array (FPGA) platforms indicate that by combining the two outputs, our design enables much larger PUF output throughput, better randomness and less power consumption compared to traditional PUFs.

I. INTRODUCTION

The massive deployment of mobile reconfigurable devices has imposed a high-reliability requirement. It has become necessary that these devices should operate reliably irrespective of the change in the external operating environment. In order to monitor the reliability of these mobile reconfigurable devices, measurement of physical operating parameters like on-chip power supply voltages and die temperatures is desired in many situations. An alarm should be raised whenever the reliability is compromised due to abnormal or extreme operating environment. Environmental abnormalities are dangerous because, on the one hand, sudden change in the environment could indicate faulty hardware (circuit short) or logic failure (excessive switching); on the other hand, abrupt variation could also be caused by a variety of physical attack methods, putting the integrity of the operation and the intellectual property embedded in the hardware at the fringe of danger. In both cases, notification of such changes is desired to avoid further damage. Conventionally, detection of abnormal variations requires on-chip temperature and voltage sensors. However, sensor-based monitor system has its drawbacks such as relatively high power consumption and low sample rate. As a matter of fact, many low-end reconfigurable devices like Internet of things (IoT) nodes do not have such sensors embedded. Thus, a low cost, high accuracy abnormal environment detection method is desired to monitor mobile devices without the help of temperature and voltage sensors.

APUF, a unique type of hardware security primitive that has been used in many security applications, has excellent properties including low power and high speed. However, APUFs suffer from the notorious reputation of instability and environmental sensitivity. Challenge-response mapping for a specific APUF is usually not stable: the mapping changes as environmental factors varies. Slight change in temperature and voltage could lead to the unpredictable remapping of many CRP. Several APUF-based security applications intend to minimize the impact introduced by instability using stable CRP selection or error correction code (ECC). While considerable efforts are seeking to eliminate the problem, we see the environmentally sensitive nature of an APUF as an opportunity. The instability due to environmental changes can be exploited to detect abnormalities during the operation cycle of a protected hardware.

We propose to use an APUF as an environmental abnormality detector. The core design idea is based on the observation that changes in an APUF’s challenge-response mapping imply a highly probable variation in operating voltage or temperature. We propose to monitor the remapping activity of only ESCRs instead to save additional energy and latency. Our design is advantageous comparing to current sensor-based system monitors for the following three reasons:

- Implementation is flexible. Detection of environmental variation is not a must for all applications. Our detection framework can be implemented or removed easily on reconfigurable hardware depending on user’s demand.
- High sample frequency. Many abnormalities only affect the device for a few clock cycles, current system monitor sample rates on Xilinx Virtex devices (200kHz maximum) fail to detect these variations. Our APUF-based framework is capable of sampling at a much higher frequency (around 10 Mhz in our implementation).
- Low power and compact size. Our detection framework requires only a single APUF to detect abnormalities, which grants us a huge advantage in terms of area and power. After careful off-line ESC seed generation and alarm threshold calibrations, the system provides comparable detection rates with state-of-the-art detection mechanism.
II. RELATED WORK

A. Physical Unclonable Functions

PUF was first proposed by Pappu et al. using mesoscopic optical systems [1]. Gassend et al. developed the first silicon PUF through the use of intrinsic process variation in deep submicron integrated circuits [2]. More recently, efforts have been made to enhance the security and randomness of PUFs. A robust PUF design is proposed by Xu et al. with a selected PUF challenge-response set [3]. A low power high randomness PUF design was proposed by Gu et al. by combining classic PUFs with hardware random number generators [4]. Xu et al. also proposed to manipulate programmable delay lines to increase the PUF stability [5]. All past PUF studies in the literature intended to reduce environmental impact on PUFs. To the best of our knowledge, we are the first to propose a mechanism that utilizes PUF’s environmental sensitivity as a tool.

B. System Monitor

Many efforts have been made to monitor the physical parameters of a reconfigurable system. Xilinx has embedded their System Monitor in their recent FPGA products with relatively low sample rate [6]. Le Masle and Luk introduced a ring oscillator-based attack detection system that monitors the core power of a circuit through observing the behavior of embedded ring-oscillators [7]. Even though the ring oscillator-based power monitor framework is capable of detecting abnormalities at a much higher sample rate, it is also much larger in size. Despite the existing studies, our proposed APUF-based abnormality detection framework is capable of achieving high sample rate while remaining simple, compact and low power.

III. PRELIMINARIES

A. APUF Model

The PUF we use in our abnormality detection framework is standard delay-based APUFs. An n-bit APUF takes an n-bit challenge as input and produces a 1-bit response as output. When an n-bit challenge is provided to the APUF, two identically designed paths are generated along a chain of multiplexer pairs. Each challenge bit controls whether the pair of paths should swap positions. A pair of multiplexers is denoted as an APUF segment. To retrieve a response, an impulse signal is fed into the system to excite both paths simultaneously. Because of the uncontrollable process variation, the signal traveling along one of the two paths will

Assuming an APUF maps challenges C to corresponding responses R. The APUF has n segments meaning all challenges C are n-bit long. We assume no delays on the connection wires and all delays are contributed by the APUF segments. Given a specific challenge c ∈ C, the ith APUF segments generates a pair of delays with delay difference of Δd_i. In a stable environment, the corresponding response r ∈ R can be mathematically represented as:

\[
r = \begin{cases} 
0 & \text{if } \sum_{i=1}^{n} \Delta d_i > 0 \\
1 & \text{if } \sum_{i=1}^{n} \Delta d_i < 0 
\end{cases}
\]  

In real life, however, environmental variations change the delay difference in each APUF segment. For simplicity purposes, we assume that a minor change in environmental would change the delay difference in the ith segment by Δd_i ∈ [−e, e], e is the maximum change each delay difference could be altered in a normal operational environment. Δd_i is a function of temperature t and voltage v. Due to on-chip temperature and voltage gradient, t, v are different from segment to segment. When provided with a specific challenge c ∈ C, the response r ∈ R can be mathematically represented by

\[
r = \begin{cases} 
0 & \text{if } \sum_{i=1}^{n} \Delta d_i + \Delta d_i > 0 \\
1 & \text{if } \sum_{i=1}^{n} \Delta d_i + \Delta d_i < 0 
\end{cases}
\]

Noted that our abnormality detection framework is not limited to only standard APUF, many other delay-based PUFs like ring oscillator PUFs serve our need as well. However, APUF is more environmentally sensitive, more compact and low power comparing to ring oscillator-based PUFs, so in this work, our design and implementation are entirely based on the above APUF structure.

IV. CRP ENVIRONMENTAL SENSITIVITY

Physical properties of transistors vary as environment changes, thus a large number of APUF CRPs are sensitive to fluctuations in physical parameters. In this section, we first discuss correlations between CRP consistencies and physical parameters in the operating environment. We then present a method to collect a set of ESCRP with strong correlation with environmental parameters. We intend to show that the CRP inconsistency in APUFs can serve as a good indicator of environment variations. We define CRP inconsistency as the probability for a corresponding response being inverted when the challenge remains unchanged.

A. Environmental Variation

Operating environment is a general concept that includes all on and off chip physical conditions during the process of an operation. Humidity, room temperature, supply power all play an important role during the process of logic operations on a reconfigurable device. In this paper, we mainly focus on studying on-chip voltage and temperature. We carefully examine the core temperature and voltage levels in a reconfigurable platform and their relationship with the on-chip APUF CRP inconsistency.

1) Core Temperature vs. CRP inconsistency: Core temperature can be fluctuated by many factors including but not limited to circuit switching activity, room temperature, heat sink efficiency etc. A sudden, unintended and significant increase or decrease in the core temperature is defined as temperature abnormality. Temperature abnormalities could result in circuit behavior alternation, hardware malfunction or even physical damage. APUF serves as an excellent temperature abnormality detector because transistor delays are sensitive to temperature. Change in transistor delays potentially invert the signal racing results between two delay paths in an APUF.
We first randomly selected 1,000,000 64-bit binary strings as our test challenge set. Each challenge within the test challenge set is repeatedly fed to the 64-bit APUF implemented on a Virtex-5 FPGA. We adopt the majority voting scheme to decide on a standard response. For a challenge \( C \), if a majority of the corresponding responses is \( R \), the tuple \((C, R)\) is recorded as the reference CRP. All reference CRPs are stored in a dictionary \( \text{Dict}_{\text{CRP}} \). Core temperature is being sampled and recorded simultaneously using on-chip sensors. A set of 100 challenges were being evaluated between every two sensor samples. Each set is being evaluated for 1,000 samples until being replaced by a new set of 100 challenges.

Figure 1a shows a snippet of our experiment. The blue line is the core temperature gathered from the built-in SYSMON hard macro. The red line is the average inconsistency of a set of 100 test CRPs. After the 50th sample, we intentionally increase the core temperature by 5 °C using excessive switching circuits [8]. An intuitive conclusion can be drawn from the figure that the pattern of both core temperature (blue line) and average CRP inconsistency (red line) are correlated. Analysis on the entire CRP test set shows that the CRP inconsistency and collected core temperature are correlated with a correlation coefficient of 0.7461. When the core temperature increases, a rise in CRP inconsistency occurs. The inconsistency falls back as the core temperature recovers. The 5 °C increment in the core temperature leads to approximately 13.5% increment in CRP inconsistency.

However, the CRP test set is chosen at random in our experiment, a large number of stable and ultra unstable CRPs were included. Stable CRPs are always consistent regardless of the increment in the core temperature, thus greatly reduces the environmental sensitivity of the APUF. Ultra unstable CRPs on the other hand always provide near 50% inconsistency, which add additional noise to our evaluation.

2) Core Voltage vs. CRP inconsistency: Core supply voltage inside a reconfigurable device is not static. In most cases, core supply voltage VCCINT varies in a tolerable range. Many events could lead to abnormal variations in VCCINT: simple power analysis(SPA) could lead to an abnormal decrease in the VCCINT since power measurement tool introduces additional resistance to the power rail; Electrostatic discharge(ESD) could lead to incremental increment on VCCINT etc. APUF serves as a good VCCINT detector because a slight variation in supply power results in the remapping of some CRP in APUF, thus by actively monitoring the mapping of APUF CRP inconsistency, variations in internal supply power can be detected through calculation on CRP inconsistency.

We produce a SPA scenario that results in a sudden decrease in VCCINT. The experiment Virtex-5 board uses a power module(TI TPS54620) to provide power to FPGA core with the internal supply voltage(VCCINT). The power module is essentially a buck converter that generates output voltage 0.8V < \( V_{out} < 15V \). Figure 2 shows a simplified diagram of the module modified by us. We produce a SPA attack by adding an additional switch resistor to modify VCCINT directly through the power module. VCCINT can be instantly decreased by opening the switch \( SW \) to add a resistor \( R_{SW} \) into the circuit. The output voltage of the power module before opening the switch \( SW \) can be calculated using equation 3. When mimicking the probe insertion, we open the switch \( SW \) and the output voltage VCCINT can thus be calculated using equation 4.

\[
VCCINT = \frac{R_1 \cdot V_{ref}}{R_2} + V_{ref} \tag{3}
\]

\[
VCCINT = \frac{(R_{sw} + R_1) \cdot V_{ref}}{R_2} + V_{ref} \tag{4}
\]
B. Environmentally Sensitive CRP

The above evaluation above shows that CRP inconsistency in an APUF is correlated with core temperature and supply voltage. However, the correlation is not strong enough to meet our application needs. After review the whole CRP set, we conclude that not all CRPs are equally sensitive to the environment. Based on the notation given in equation 1 and 2, we define stable, ultra unstable and ESCRPs as below.

A stable challenge produces consistent response regardless of operating environment. A stable CRP appears if one or more segments dominate over all remaining segments so that the final delay difference is sufficiently large enough to overcome delay variations. Stable challenges can be defined as:

\[
|\sum_{i=1}^{n} (\Delta d_i^c) + \sum_{i=1}^{n} |\Delta d_i^e| \text{ for all possible } \Delta d_e \approx 0
\]  

so that the APUF generate the same response regardless of environmental variation. Our APUF implementation shows that roughly 30% of CRPs are stable.

An ultra unstable challenge on the other hand creates two paths with near zero delay difference. This can be represented by:

\[
\sum_{i=1}^{n} \Delta d_i^c + \Delta d_i^e \approx 0 \text{ for } \Delta d_e \in [-e, e]
\]

With the given challenge, there is a near 50% probability that the corresponding response would be a 1 or a 0. Our APUF implementation shows that roughly 5-10% of CRPs are ultra unstable.

For the purpose of being an indicator of environmental changes, we are primarily interested in searching for environmentally sensitive challenges(ESCs) that neither produces a stable nor completely unpredictable response. An ideal ESC should at least fulfill the following two conditions:

1) \[
\sum_{i=1}^{n} (\Delta d_i^c) < \sum_{i=1}^{n} |\Delta d_i^e| \text{ for } \Delta d_e \in [-e, e]
\]

when so that no dominating segment group exists in the APUF, and environmental variations could possibly invert the response.

2) \[
\sum_{i=1}^{n} \Delta d_i^c + \Delta d_i^e \approx 0 \text{ for } \Delta d_e \in [-e, e]
\]

At a given environment the response is stable in a normal operation environment.

In additional to the previous two requirements, an ideal ESC should be able to obtain an inverted response when the environment varies to an abnormal range:

3. 

\[
sign(\sum_{i=1}^{n} \Delta d_i^c + \Delta d_i^e) = -sign(\sum_{i=1}^{n} \Delta d_i^c + \Delta d_i^e)
\]

for \(\Delta d_e \in [-e, e]\) and \(\Delta d_i^e > e\) or \(\Delta d_i^e < -e\).

In our experiment, we defined normal operating environment as: 1. core temperature \(T = 30.5 \pm 5^\circ C\), 2. VCCINT = \(1 \pm 0.5V\). From the 1,000,000 test challenge set, we observe that no challenge fulfills all three requirements for ESCs. Non-uniform physical properties over the chip makes qualifying a single ESC seed using evolution strategies (described in Algorithm 1). Evolution strategies method is inspired by the evolutionary adaptation of a population of individuals to
certain environmental conditions. When applying evolution strategies on ESC generation, one individual in the population is a specific challenge vector and mutation on an individual is defined as randomly flipping multiple bits. The environmental fitness of the individual is determined by the consistency of the corresponding PUF response under a small environmental variation benchmark where environment factors vary beyond the tolerable range using switching circuit/power rail resistor. The tolerable range is defined by the user based on the nature of the application. The pseudocode of ESC exploration is described in Algorithm 1)

**Algorithm 1** Explore ESC seed

Require: (1) Number of candidates in the parent generation \( \mu \).
(2) Number of candidate solutions generated from the parent generation \( \lambda \).
(3) Expected ESC set size \( \kappa \)
(4) A fitness evaluation function \( \text{EvaluatePopulation}() \).
(5) A mutation rate \( \tau \).
(6) Maximum number of generations \( M \).

Ensure: A set of ESCs \( S_{best} \).

Population \( \leftarrow \text{InitializePopulation}(\mu) \).

\[ S_{best} \leftarrow \text{GetBest}(\text{Population}, \kappa) \]

while \( i < M \) do

\[ \text{Children} \leftarrow \emptyset \]

for \( j = 0 \) to \( \lambda \) do

\[ \text{Parent}_j \leftarrow \text{GetParent}(\text{Population}, j) \]

\[ S_j \leftarrow \text{Mutate}(\text{Parent}_j, \tau) \]

\[ \text{Children} \leftarrow S_j \]

end for

\[ \text{EvaluatePopulation}(\text{Children}) \]

\[ S_{best} \leftarrow \text{GetBest}(\text{Population} + S_{best}, \kappa) \]

Population \( \leftarrow \text{SelectBest}(\text{Population}, \text{Children}, \mu) \)

\( i = i + 1 \).

end while

Return(\( S_{best} \))

We observe that when inverting a small number of bits in an ESC seed, the new challenge is most likely to stay environmentally sensitive due to the fact that nearby transistor process variations are somewhat correlated. We thus generate more ESCs by randomly inverting a small number of bits in the ESC seed (described in Algorithm 2).

**Algorithm 2** Generate ESC set

Create empty challenge set \( ESC \), define the number of the size of the ESC set \( m \), and a growth index \( i \). An ESC seed \( esc \) is taken as an input

\[ \text{Count} = 0 \]

while \( \text{Count} < m \) do

\[ \text{for } k \text{ in range}(i) \text{ do} \]

\[ \text{Randomly flips } k \text{ bits in } esc \text{ to obtain } c_k \]

\[ \text{Add } c_k \text{ to } ESC \]

end for

\[ \text{Count} += 1 \]

end while

V. THE DETECTION FRAMEWORK

A. System Design

Figure 4 shows a high level implementation of APUF-based abnormality detection framework. The detection framework consists of three major modules.

The challenge querier is used to generate challenges, and feed them into the APUF at run-time. Before the actual run-time detection, a calibration process generates an ESC seed off-line using algorithm 1. The challenge querier generates a set of ESCs using algorithm 2 at run-time. A random number generator embedded inside of the challenge querier decides at each clock cycle which bits should be inverted.

The APUF process the same challenge multiple times and send all generated responses to the response verifier module. The APUF itself resides immediately next to the logic being monitored to achieve the best accuracy.

The response verifier calculates the CRP inconsistency by counting the frequency of response inversion, and raise an alarm if the inconsistency exceeds a user defined threshold. Since all ESCs are correlated, the 0 and 1 distribution are also very similar, thus a single counter is enough for counting and recording bit inversions. The alarm threshold is recommended to be set to the selected ESC’s fulfillment rate of equation 8 and 9, however it is adjustable as the definition of the normal operational environment may be subject to change based on user requirement.

![Fig. 4: APUF-based abnormality detection framework.](image)

B. Experimental Results

In this section we carefully evaluate our abnormality detection framework.

Our experiments are conducted on two Genesys boards. The Genesys board has a Virtex-5 LX50 FPGA. The system is tested with an implementation of an RSA-1024 and an AES-128 crypto-system [10]. A mixture of 1,000 abnormal core temperature and VCCINT was applied during a 10-hour long operation. In our setting, an abnormality in core temperature is defined as a variation of over 5°C and an abnormality in VCCINT is defined as a variation of over 5%. We generated three distinctive challenge sets, respectively 1,000 random challenges(RC), a single ESC seed, and 1,000 ESCs derived from the ESC seed. The ESC seed is generated based on our
requirement on core temperature and VCCINT, which in our case is our definition of the normal range of temperature and VCCINT defined above. The alarm threshold is set to raise an alarm when CRP inconsistency reaches 40% for 1,000 random challenges, 70% for both single ESC seed and 1,000 ESCs case. A UART core is used to monitor the real-time CRP inconsistency and communicates with RSA/AES cores. The results are displayed in Table II.

We observe that 1,000 ESC challenge set provides the best detection rate while random CRPs provides the worst. ESCs benefit from the elimination of stable and ultra unstable challenges. With a good ESC seed, we are capable of achieving a true positive rate of 100% with 1,000 generated ESCs at runtime, while the false positive rate is as low as 1.1%.

C. Area and Power

We compare our design’s area and power overhead with both Xilinx System Monitor [6] and ring oscillator-based power monitor[7]. Since Xilinx System Monitor only supports Virtex family product and Le Masle’s ring oscillator-based power monitor was originally implemented on Spartan-6 LX45, we implement our design on both platforms for comparison purposes. The sample rate for Virtex-5 implementations is fixed to 200kHz while the implantation on Spartan-6 devices has a sample rate of 8MHz. The FPGA resource and power characteristics for Xilinx System Monitor, ring oscillator-based power monitor and our design are shown in Table I.

Our design uses 63% less FPGA area comparing to Xilinx System Monitor logics while achieving 13% of power savings. To be noted that we do not include area and power of both analog sensor and ADC in the SYSMON hard macro in our calculation due to lack of information. The actual saving is expected to be more. Comparing to Le Masle’s ring oscillator-based power monitor, our design is 98.4% smaller. Since APUF does not impose high switching activities like ring oscillators, our design consumes only 1.5% of the power.

Table I: FPGA resource and power characteristics: Virtex-5 SYSMON vs. our design on Virtex-5 vs. RO-based on-chip power monitor vs. our design on Spartan-6. ADC and analog sensor area and power are not included in calculation.

<table>
<thead>
<tr>
<th>Circuit</th>
<th>1,000 RC</th>
<th>Single ESC</th>
<th>1,000 ESC</th>
</tr>
</thead>
<tbody>
<tr>
<td>AES-128</td>
<td>69/4.0/7</td>
<td>98/1/5.9</td>
<td>100/2.1</td>
</tr>
<tr>
<td>RSA-512</td>
<td>78/4.5/0</td>
<td>100/3.9</td>
<td>100/1.1</td>
</tr>
</tbody>
</table>

Table II: Abnormality detection: true positive rate (%) / false positive rate(%) using 1,000 random challenges vs. single ESC vs. 1,000 ESCs as challenge set per sample. APUFs are placed immediately besides the monitored circuit. A mixture of 1,000 abnormal temperature and VCCINT were applied to the protected circuit.

VI. Conclusion

In conclusion, we discovered that APUF CRP inconsistency is highly correlated with the core temperature and voltage on reconfigurable platforms. Based on the observation we designed a framework to utilize the environmental sensitivity of APUFs to detect suspicious operational environment variations. We propose to apply only ESCs on a given APUF to efficiently and accurately detect environmental abnormalities. When integrated with an AES-128 and an RSA-1024 implementation, our framework is capable of detecting 100% of applied abnormalities with false positive rate as low as 1.1%. Our design provides competitive detection rate with both sensor-based Xilinx System Monitor and Le Masle’s ring oscillator-based power monitor while using much less hardware and power.

VII. Acknowledgement

This work was supported in part by the NSF under award CNS-1059435, and award CNS-1513306.

References