Noise Reduction on Memory-based PUFs

Mafalda Cortez  Said Hamdioui  
Delft University of Technology  
Microelectronics & Computer Engineering Lab  
Mekelweg 4, 2628 CD Delft  
The Netherlands

Vincent van der Leest  Roel Maes  Geert-Jan Schrijen  
Intrinsic-ID B.V.  
High Tech Campus 9  
5656 AE Eindhoven, The Netherlands

TRUDEVICE 2013  
Avignon, France, May 30-31, 2013

Extended version of this work will appear in HOST 2013  

TRUDEVICE 2013
Motivation - Secure Key Storage

- Security based on cryptographic algorithms
  - **Permanent** key storage is highly prone to physical attacks

- **Solution:**
  1. Do not permanently store a key in Non-Volatile Memories (NVMs)
  2. Generate the key only when needed (extract it from a physical structure of the IC)
  3. Delete the key

PUFs’ characteristics
- **Uniqueness**
- **Reproducibility**

~ 50% difference
~ 10% errors
Motivation

shortcomings and contributions

• Most of work focus on
  • Developing new PUFs types
  • Proving PUF uniqueness
  • Validating with silicon PUF reproducibility
  • Developing peripheral circuits (e.g., error correction)

• Shortcomings
  • Techniques to minimize the noise (increase reproducibility) of PUFs

• Our contributions
  • Efficient scheme to reduce the noise in memory-based PUFs
Layout

• Memory-based PUF Secure Systems
  • Enrolment versus reconstruction
• Stability parameters classification
  • Technology versus non-technology parameters
• Simulation Set-Up & Results
• Industrial/Silicon results
• Block diagram of the new scheme
• Conclusion
Enrollment: Define Key

Reconstruction: Reconstruct Key

PRR and PR must be close enough!

Reproducibility is crucial... but expensive!
Stability Parameters Classification

Can non-technology parameters be used for noise reduction?

Simulation Set-Up & Results

Set-up
• HSPICE: SRAM cell with its periphericals
• 45nm Low Power BSIM4 model

Generation of an SRAM fingerprint of 1k

Monte Carlo

Power-up

\[ \begin{array}{cccccccc}
V_{th1} & V_{th2} & V_{th3} & \cdots & V_{thn} & V_{thn+1} & \cdots & V_{th1000} \\
1 & 1 & 1 & 0 & 0 & 1 & 0 & 0
\end{array} \]

PUF fingerprint

Noise Metric: Fractional Hamming Distance (FHD)

Experiments performed
• Voltage Ramp-up Time : 3 \( t_{\text{ramp}} \) (10\( \mu \text{s} \), 50\( \mu \text{s} \) and 90 \( \mu \text{s} \))
• Temperature : 3 Temp (-40\( ^{\circ} \text{C} \), 25\( ^{\circ} \text{C} \) and 85\( ^{\circ} \text{C} \))
• Measurements : 20 Meas (different noise random seeds)

\[ \text{NOTE: } V_{th} \ \text{distribution according to: W. Zhao, F. Liu, K. Agarwal, D. Acharyya, S.R. Nassif, K.J. Nowka and Y. Cao, “Rigorous extraction of process variations for 65nm CMOS design”, ESSDERC, pp. 89–92, 2007.} \]
Simulation Results... Noise at different enrollments

• For Temp below enrollment, max FHD is lower if \( t_{\text{ramp}} \) is longer
• For Temp above enrollment, max FHD is lower if \( t_{\text{ramp}} \) is shorter

Typical approach is not optimal!

Can noise be reduced by manipulating Temp and \( t_{\text{ramp}} \)?
Industrial/Silicon Results: Setup & Experiments performed

Experiments performed
• Temperature : 3 Temp (-40°C, 25°C and 85°C)
• Voltage Ramp-up Time : 10 $t_{\text{ramp}}$ (from 10μs up to 500ms)
• Measurements : 10 Meas (each measurement has different noise)

Measurement flow

- Set temperature (-40°C, +25°C, +85°C)
- Power-up with $t_{\text{ramp}}$ (10μs up to 500ms)
- Read and store fingerprint
- Power-down for 1 second
- X9 (same Temp and Tramp)
- X9
- X2
Industrial/Silicon Results: Original results

Measured noise as FHD (precision 0.5%)

<table>
<thead>
<tr>
<th>Technology</th>
<th>PUF</th>
<th>$t_{\text{ramp}}$</th>
<th>Maximum noise FHD</th>
<th>$\mu$-BCHD</th>
<th>$H_{\infty}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$-40^\circ$C</td>
<td>$+25^\circ$C</td>
<td>$+85^\circ$C</td>
</tr>
<tr>
<td>40nm LP</td>
<td>SRAM</td>
<td>50$\mu$s</td>
<td>23%</td>
<td>6%</td>
<td>20%</td>
</tr>
<tr>
<td></td>
<td>SRAM</td>
<td>10$\mu$s</td>
<td>8%</td>
<td>6%</td>
<td>8%</td>
</tr>
<tr>
<td></td>
<td>DFF</td>
<td>10$\mu$s</td>
<td>28%</td>
<td>8%</td>
<td>25%</td>
</tr>
<tr>
<td></td>
<td>BK</td>
<td>10$\mu$s</td>
<td>10.5%</td>
<td>4.5%</td>
<td>20%</td>
</tr>
<tr>
<td>65nm LP</td>
<td>SRAM</td>
<td>10$\mu$s</td>
<td>13%</td>
<td>6%</td>
<td>12%</td>
</tr>
<tr>
<td></td>
<td>DFF</td>
<td>10$\mu$s</td>
<td>16.5%</td>
<td>5%</td>
<td>28%</td>
</tr>
<tr>
<td>130nm LP</td>
<td>SRAM</td>
<td>10$\mu$s</td>
<td>13%</td>
<td>6%</td>
<td>12%</td>
</tr>
<tr>
<td></td>
<td>DFF</td>
<td>10$\mu$s</td>
<td>16.5%</td>
<td>5%</td>
<td>28%</td>
</tr>
</tbody>
</table>

- Max noise 28%
- Min noise 4.5%
- Worse $\mu$-BCHD = 0.37
- Best $\mu$-BCHD = 0.5
- Better $H_{\infty} = 0.87$
- Worse $H_{\min} = 0.40$
- Best $H_{\min} = 0.40$

Optimization algorithms
- Reproducibility
  - Identify $t_{\text{ramp}}$ for enrollment temp such that the reproducibility is the highest
- Uniqueness
  - Identify $t_{\text{ramp}}$ for enrollment such that Entropy is the highest ($H_{\min}$)
Main observations

- Noise reduction for all devices at all conditions!
- Fastest tramp during enrollment does not lead to the lowest noise
- For Temp below enrollment, max FHD is lower for longer $t_{ramp}$
- For Temp above enrollment, max FHD is lower for shorter $t_{ramp}$
- Algorithm deteriorates $\mu$-BCHD and $H_{\infty}$ for some devices (e.g., 130nm LP SRAM)
Industrial/Silicon Results: Reproducibility optimization algorithm

<table>
<thead>
<tr>
<th>Technology</th>
<th>PUF</th>
<th>$t_{ramp}$</th>
<th>Maximum noise FHD</th>
<th>$\mu$-BCHD</th>
<th>$H_\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td>40nm LP</td>
<td>SRAM</td>
<td>50μs</td>
<td>23% 6% 20%</td>
<td>0.50</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>DFF</td>
<td>10μs</td>
<td>4% 8% 20%</td>
<td>0.50</td>
<td>0.87</td>
</tr>
<tr>
<td>65nm LP</td>
<td>SRAM</td>
<td>10μs</td>
<td>2% 6% 8%</td>
<td>0.37</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>DFF</td>
<td>10μs</td>
<td>28% 8% 25%</td>
<td>0.48</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>BK</td>
<td>10μs</td>
<td>10.5% 4.5% 20%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>130nm LP</td>
<td>SRAM</td>
<td>10μs</td>
<td>13% 6% 12%</td>
<td>0.47</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>DFF</td>
<td>10μs</td>
<td>16.5% 5% 28%</td>
<td>0.43</td>
<td>0.61</td>
</tr>
</tbody>
</table>

Main observations
- Noise reduction for all devices at all conditions!
- Fastest tramp during enrollment does not lead to the lowest noise
- For Temp below enrollment, max FHD is lower for longer $t_{ramp}$
- For Temp above enrollment, max FHD is lower for shorter $t_{ramp}$
- Algorithm deteriorates $\mu$-BCHD and $H_{\min}$ for some devices (e.g., 130nm LP SRAM)

New best / worse
- Max reduction 6% to 2% (3X)
- New max noise = 17% (previous 28%)
- New min noise = 2% (previous 4.5%)
## Industrial/Silicon Results: Uniqueness optimization algorithm

<table>
<thead>
<tr>
<th>Technology</th>
<th>PUF</th>
<th>$t_{\text{ramp}}$</th>
<th>$\mu$-BCHD $H_{\infty}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>40nm LP</td>
<td>SRAM</td>
<td>50$\mu$s</td>
<td>23% 6% 20% 0.50 0.73</td>
</tr>
<tr>
<td>65nm LP</td>
<td>SRAM</td>
<td>10$\mu$s</td>
<td>8% 6% 8% 0.50 0.87</td>
</tr>
<tr>
<td></td>
<td>DFF</td>
<td>10$\mu$s</td>
<td>28% 8% 25% 0.37 0.40</td>
</tr>
<tr>
<td></td>
<td>BK</td>
<td>10$\mu$s</td>
<td>10.5% 4.5% 20% 0.48 0.75</td>
</tr>
<tr>
<td>130nm LP</td>
<td>SRAM</td>
<td>10$\mu$s</td>
<td>13% 6% 12% 0.47 0.66</td>
</tr>
<tr>
<td></td>
<td>DFF</td>
<td>10$\mu$s</td>
<td>16.5% 5% 28% 0.43 0.61</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Technology</th>
<th>PUF</th>
<th>$t_{\text{ramp}}$</th>
<th>$\mu$-BCHD $H_{\infty}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>40nm LP</td>
<td>SRAM</td>
<td>1ms 100$\mu$s 50$\mu$s</td>
<td>16% 6% 19% 0.50 0.73</td>
</tr>
<tr>
<td>65nm LP</td>
<td>SRAM</td>
<td>500$\mu$s 100ms 50$\mu$s</td>
<td>13% 2% 8% 0.50 0.89</td>
</tr>
<tr>
<td></td>
<td>DFF</td>
<td>500$\mu$s 10ms 250$\mu$s</td>
<td>18.5% 2.5% 8% 0.50 0.90</td>
</tr>
<tr>
<td></td>
<td>BK</td>
<td>1$\mu$s 250$\mu$s 1$\mu$s</td>
<td>7% 5% 9% 0.50 0.58</td>
</tr>
<tr>
<td>130nm LP</td>
<td>SRAM</td>
<td>1ms 10$\mu$s 10$\mu$s</td>
<td>7.5% 6% 12% 0.47 0.66</td>
</tr>
<tr>
<td></td>
<td>DFF</td>
<td>50$\mu$s 500$\mu$s 10$\mu$s</td>
<td>10% 4.5% 9.5% 0.47 0.67</td>
</tr>
</tbody>
</table>

### Main observations
- $\mu$-BCHD and $H_{\text{min}}$ improved (or maintained for SRAM)
- Fastest tramp during enrollment does not lead to the best Entropy (exception 130nm LP SRAM)
- For Temp below enrollment, max FHD is lower for longer $t_{\text{ramp}}$
- For Temp above enrollment, max FHD is lower for shorter $t_{\text{ramp}}$
- Noise increased in some cases (e.g., Buskeeper PUF)
Block diagram of the new scheme

Main observations

• No adaptations of the PUF-circuit required
• Standard components
Conclusion

- Novel idea for reducing noise on memory-based PUF fingerprints
  - Operational conditions
  - Temp versus $t_{\text{ramp}}$

- Validated using both SPICE simulation and silicon measurements

- Easy to implement

- Noise reduction achieved is up to 3x lower
  - Reduce overall cost and improve the robustness