Machine Learning Resistant Strong PUF: Possible or a Pipe Dream?

Arunkumar Vijayakumar, Vinay C. Patil, Charles B. Prado*

Prof. Sandip Kundu

University of Massachusetts Amherst

*Inmetro, Brazil

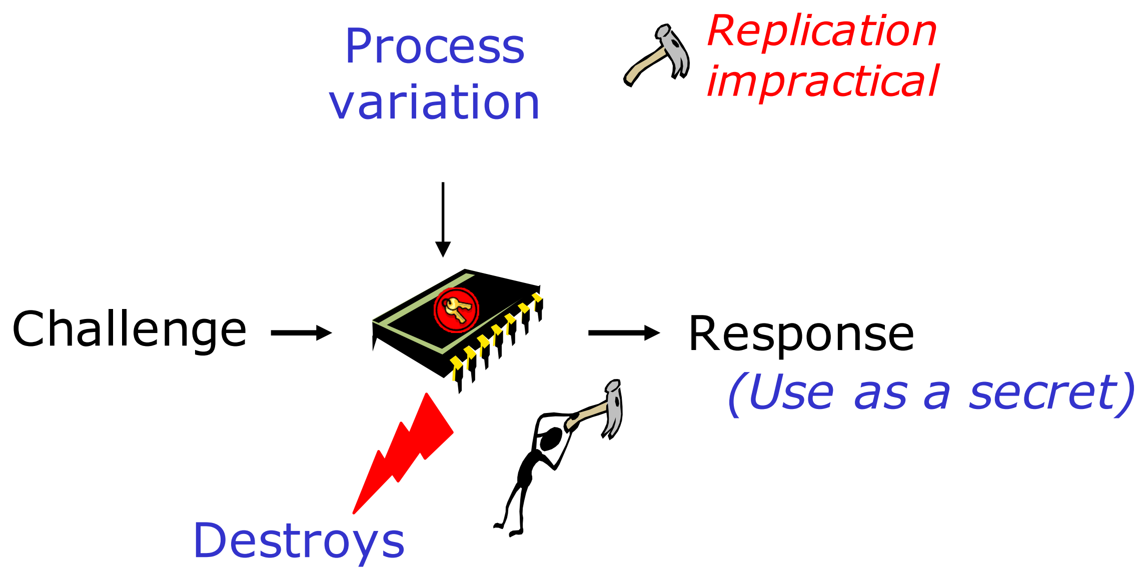
Sponsors: Horizon 2020 European Union Funding for Research & Innovation

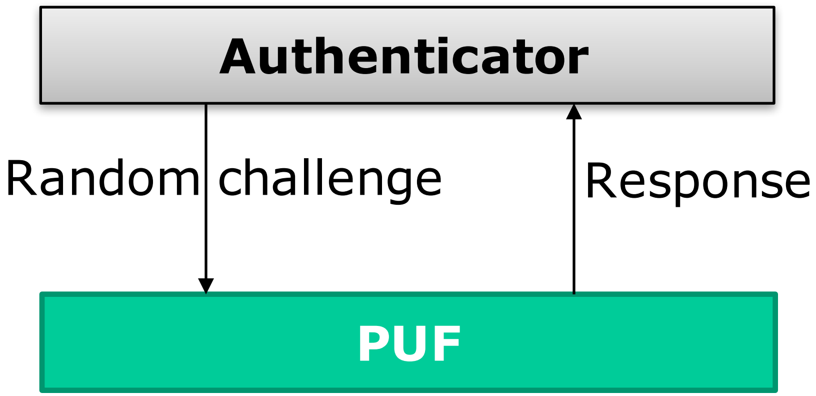
Outline

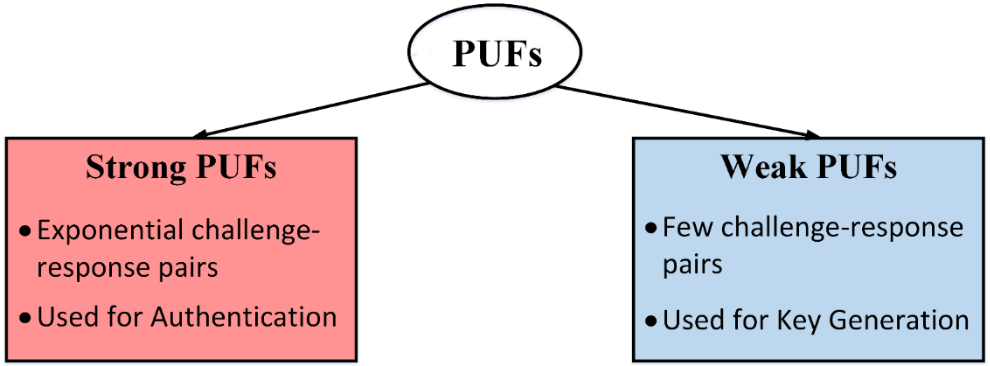
- Motivation
- Problem statement
- Background Work
- Machine Learning Study
- Key Takeaways and Future work

Physical Unclonable Functions (PUFs)

- PUFs are circuits which create secrets from complex physical system



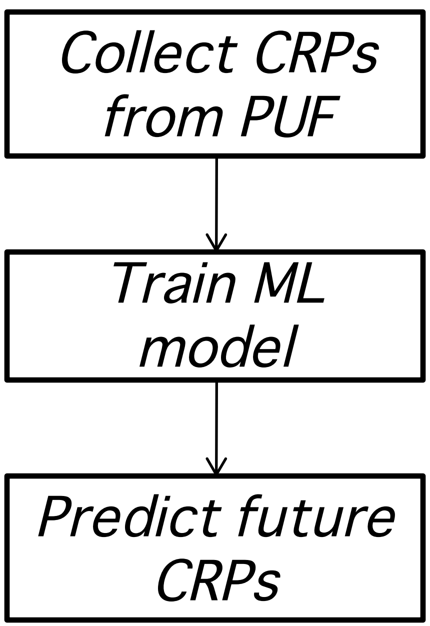


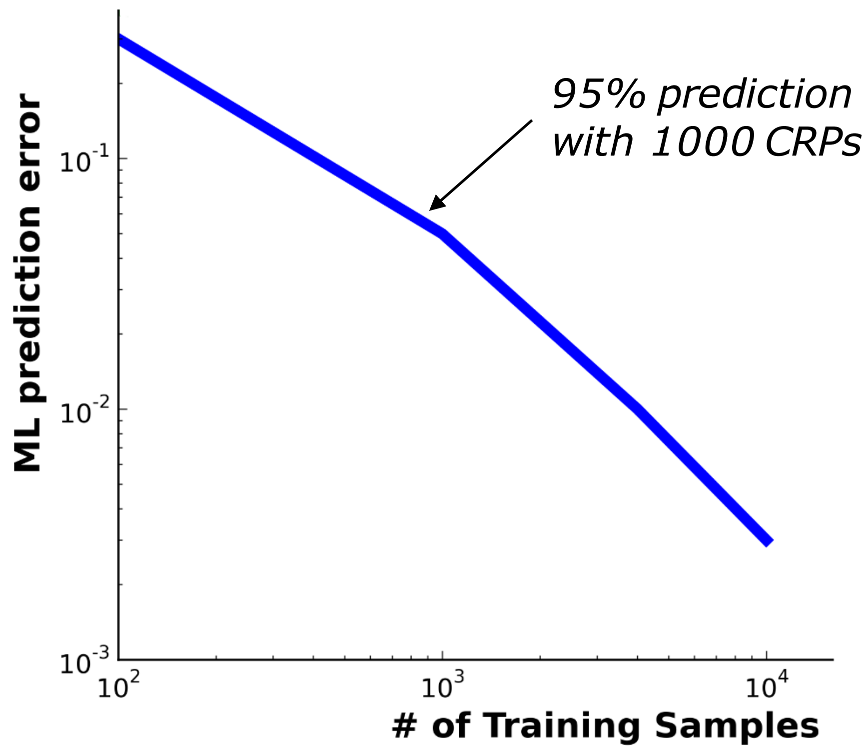


Machine Learning Attack on Strong PUFs

- Attack model *
  - Attacker in temporary possession of PUF → Mine CRPs
  - Attacker can observe CRPs during authentication
- Create software model → PUF cloned !!

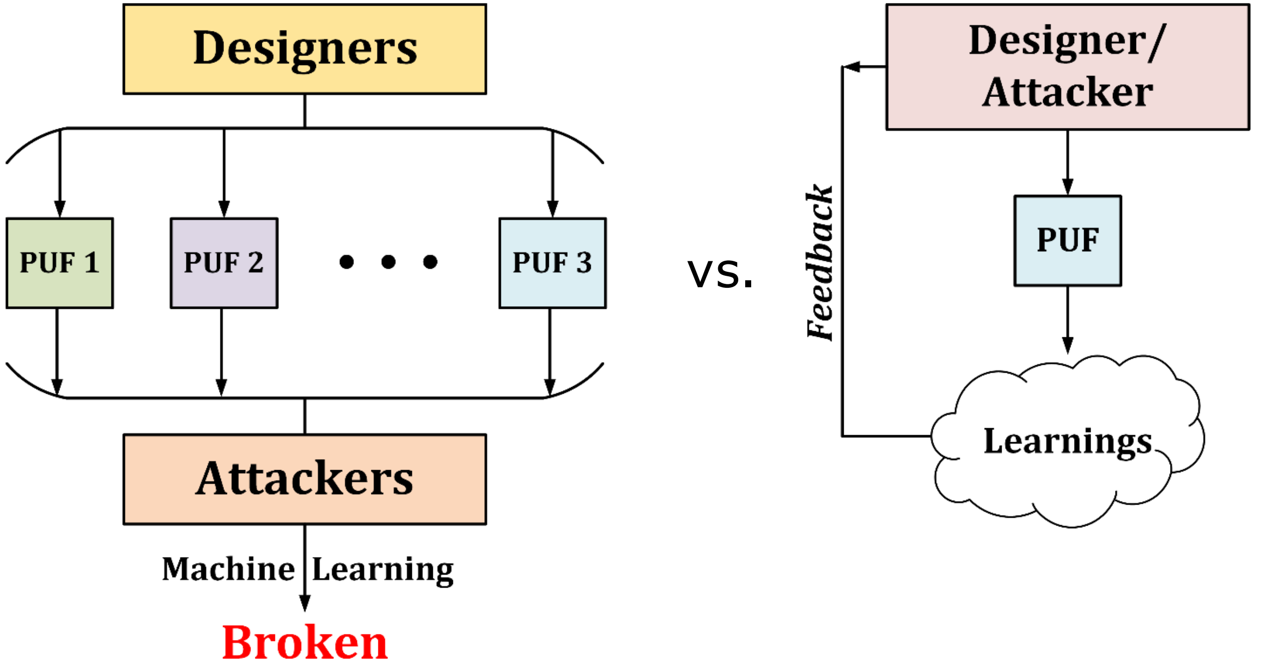
*Arbiter PUF modeled with Support Vector Machine***





Problem Statement

- Many of proposed Strong PUFs have been cloned using ML attacks
  - What learning can circuit designers get from ML studies ?
- Can a stand-alone Strong PUF be built without security enhancing accessories ? E.g. Hash
- Not a new PUF design

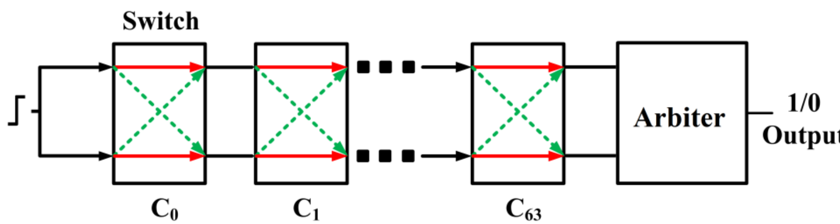


Outline

- Motivation
- Problem statement
- Background Work
- Machine Learning Study
- Key Takeaways and Future work

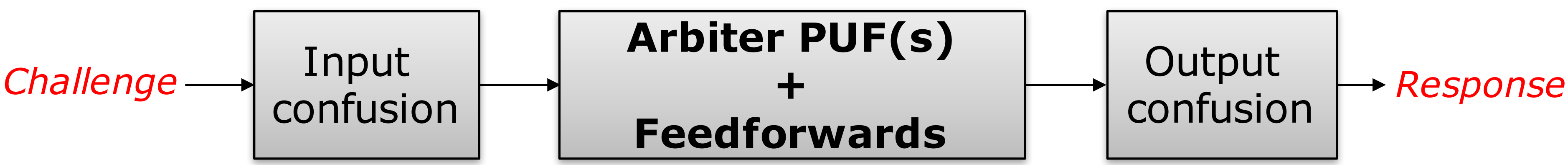
Background – ML Resistant PUFs

- Arbiter PUF



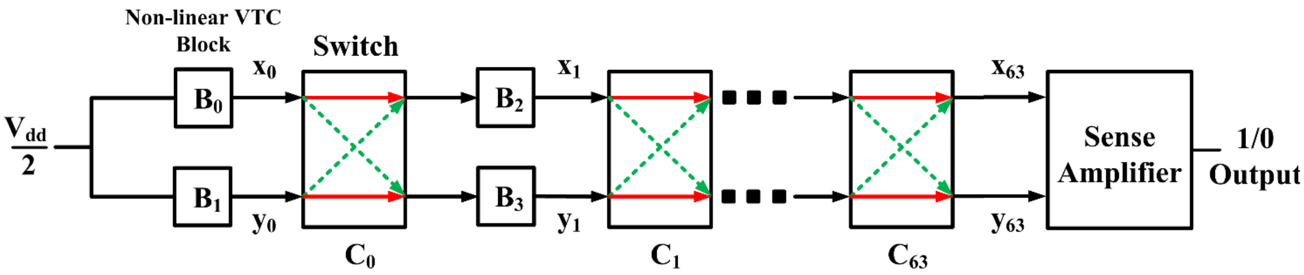
  - Linear additive model → Attacked using Support Vector Machine (SVM)

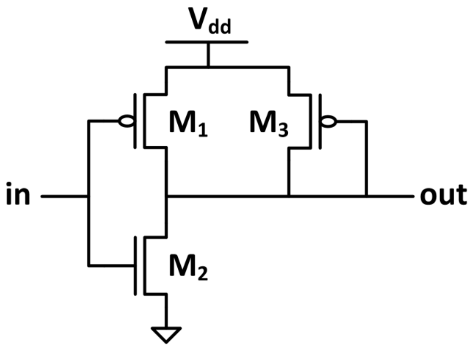
- Increase non-linearity to increase ML resistance
- Digital Modifications to Arbiter PUF
  - XOR PUF, Light-weight PUFs, Feed-forward PUF → All attacked successfully *

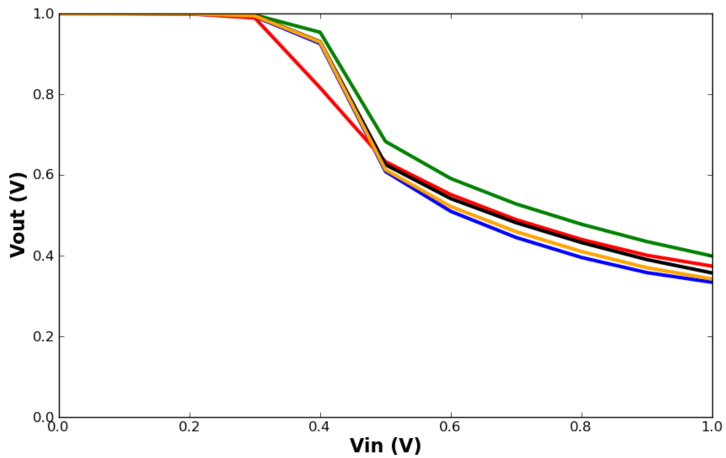


* U. Ruhrmair *et al.*, "Modeling Attacks on Physical Unclonable Functions", ACM CCS, 2010

Analog PUFs – Increase ML resistance







VTC of unit functional block

- Analog PUFs based on
  - non-linear current sources [*]
  - non-linear Voltage Transfer Characteristics (VTC) PUF [**]
- These two works show promise in building ML resistant strong PUFs
  - ~80% SVM ML prediction for 100K CRPs (20% error)

[*] Kumar et.al, HOST 2014

[**] Vijayakumar et.al, DATE 2015

Issues in Analog PUFs

- Verified only against SVM. Many other classes of ML possible
- Checked only an instance of the PUF
  - ML resistance varies in each PUF

Name	Type	Security/ Comments
Arbiter PUF, XOR PUF, Lightweight PUF	Digital	Attacked using Logistic Regression
Feed-forward PUF	Digital	Attacked using Evolutionary Strategies
Non-linear VTC PUF, Non-linear current PUF, SCA PUF	Analog	Resistant against SVM only

- We still don't know how ML-resistant Strong PUFs are !

Outline

- Motivation
- Problem statement
- Background Work
- Machine Learning Study
- Key Takeaways and Future work

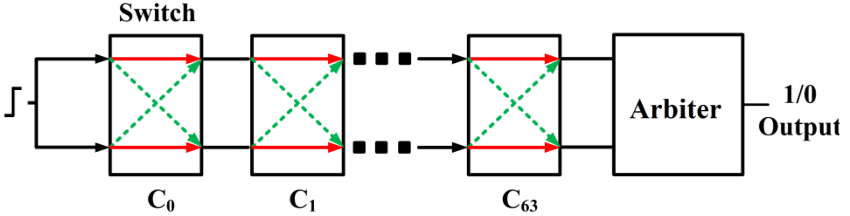
ML Study – Overview of our methodology

1. Build abstract model PUF
   - PUFs are based on delay, voltage, current → can we extract any useful abstraction?

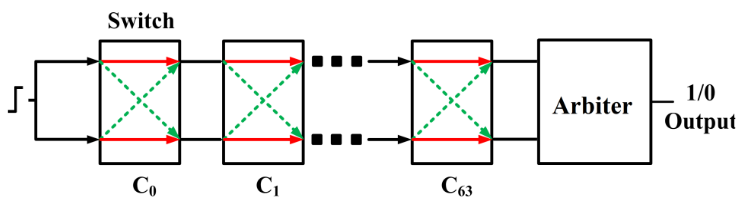
2. Study functions for ML resistance
   - Can we gain general understanding of how to increase the modeling-attack resistance ?

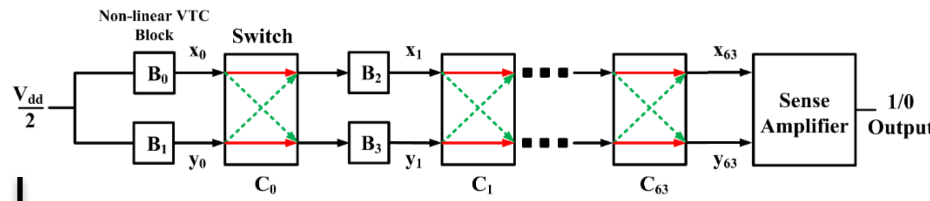
3. Test using meta-ensemble ML techniques
   - Boosting and Bagging ML algorithm

4. Understand limitations of structure if any
   - E.g., Is the cascaded switch architecture itself a limiting factor ?

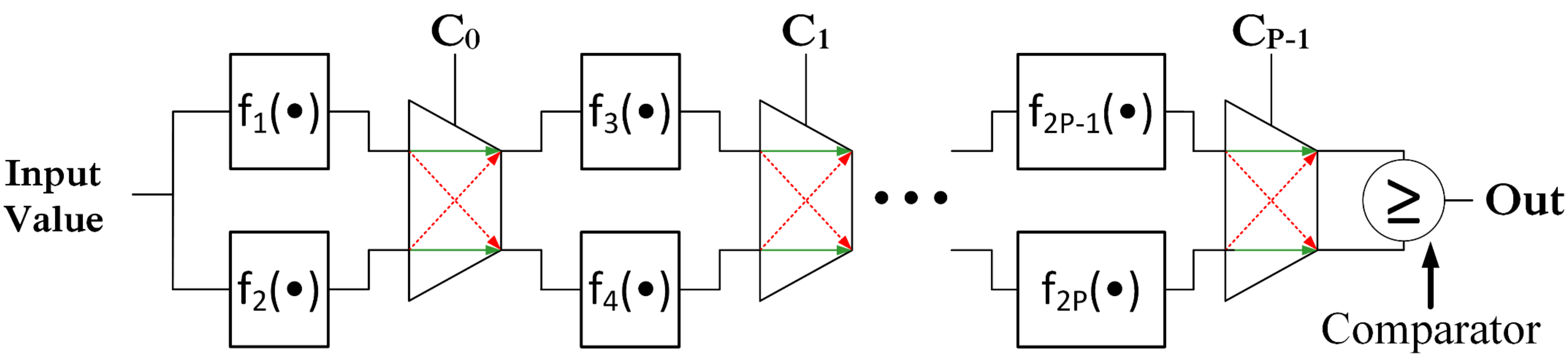


Abstract Model Building



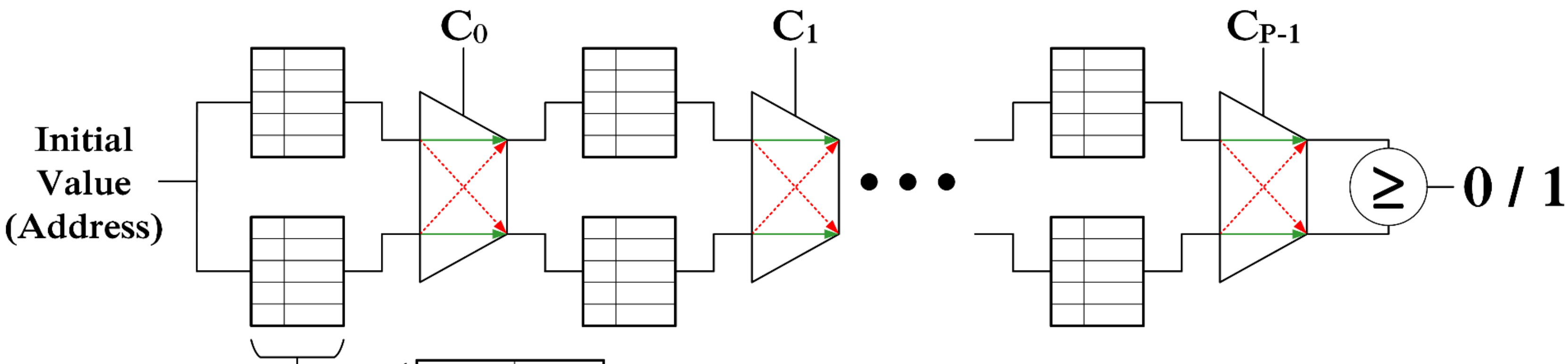


Function composition Model



- E.g. if C_0=0, C_1=0 → $f_3(f_1(\text{Input value}))$

Function of Interest



Address Index	Out
00 *(0)*	10 *(2)*
01 *(1)*	00 *(0)*
10 *(2)*	11 *(3)*
11 *(3)*	01 *(1)*

Table Size: S = 4

- Tables represent abstraction of circuit transfer functions
  - Represented as discrete function
- How ML resistance increases with entropy ?
  - Assume *uniform distribution* for the function
  - *Size of table* –> Amount of *entropy* of PUF unit cell

Study I – Increase in entropy

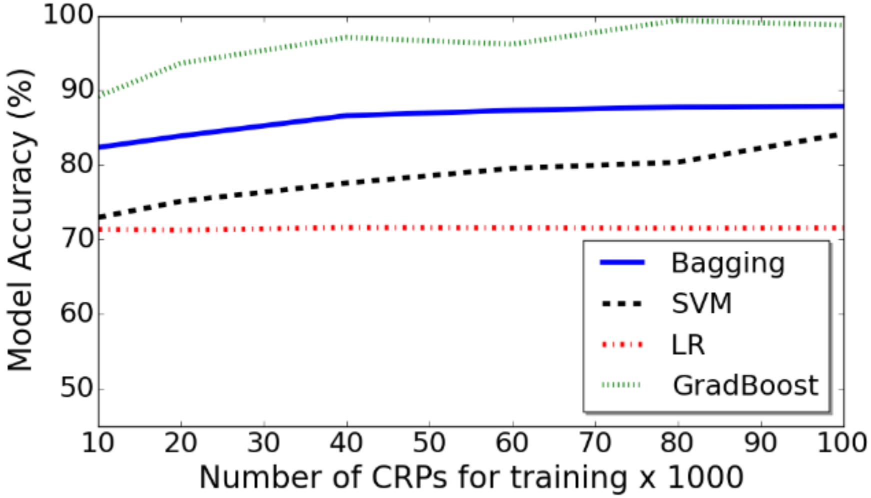


Table Size = 4

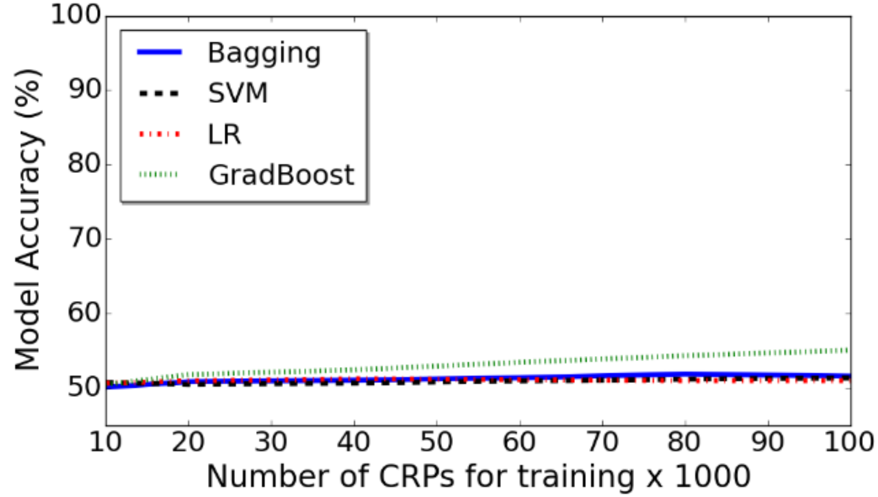


Table Size = 128

- Observation 1: Increasing **size of table** increases ML resistance
  - Higher the (persistent) entropy, higher the ML resistance
- Observation 2: Given sufficient entropy, ML resistance is possible
- Observation 3: Meta-ensemble algorithms are potent
  - Boosting and Bagging perform far better than previous ML algorithms
  - Gradient Boosting technique offer the best known attack

Study II - Impact of bias in function

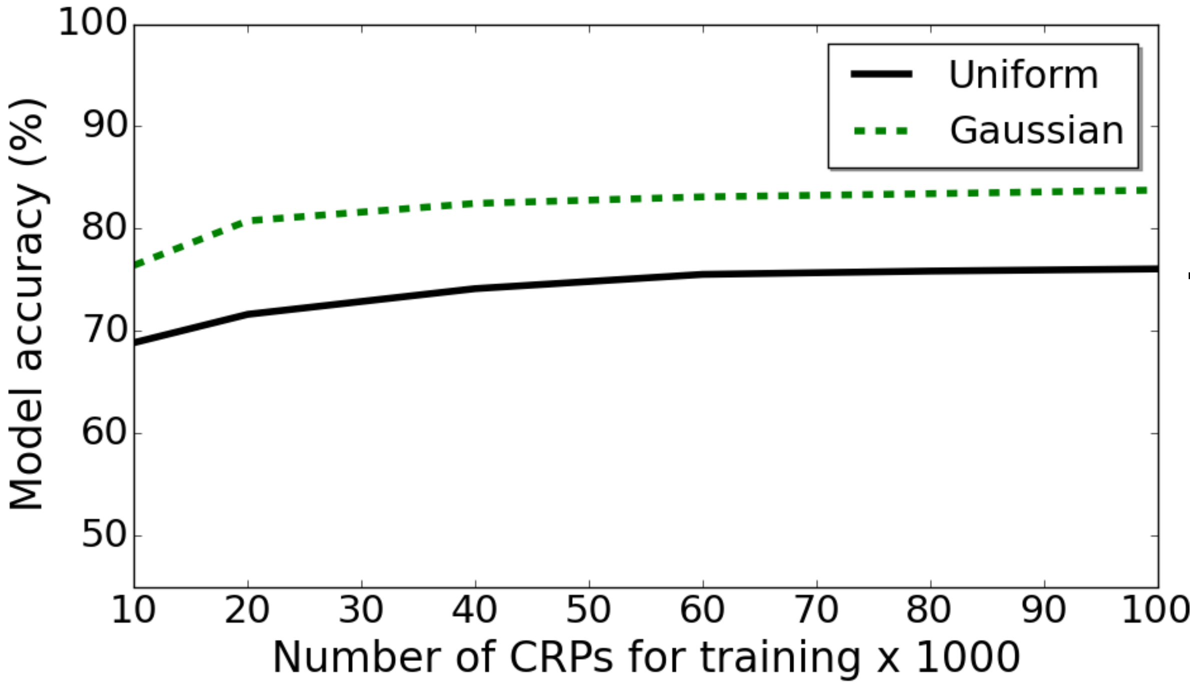
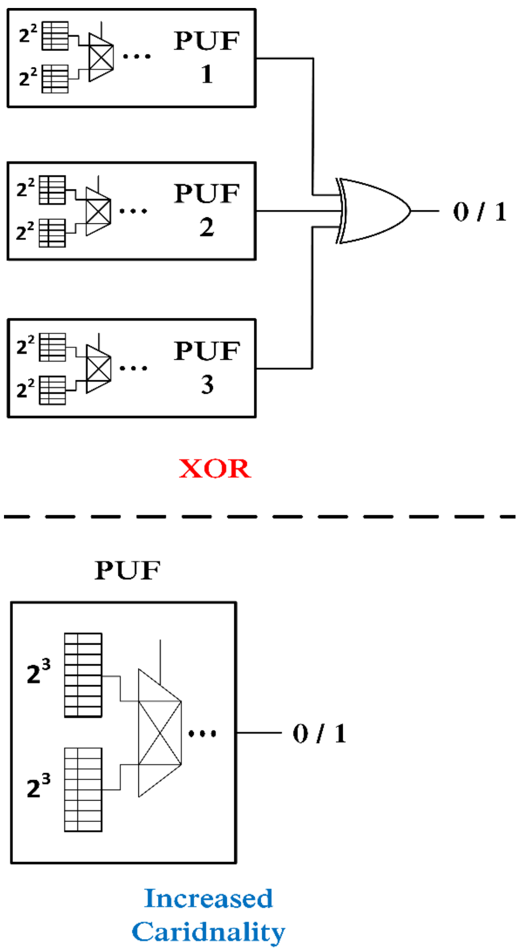


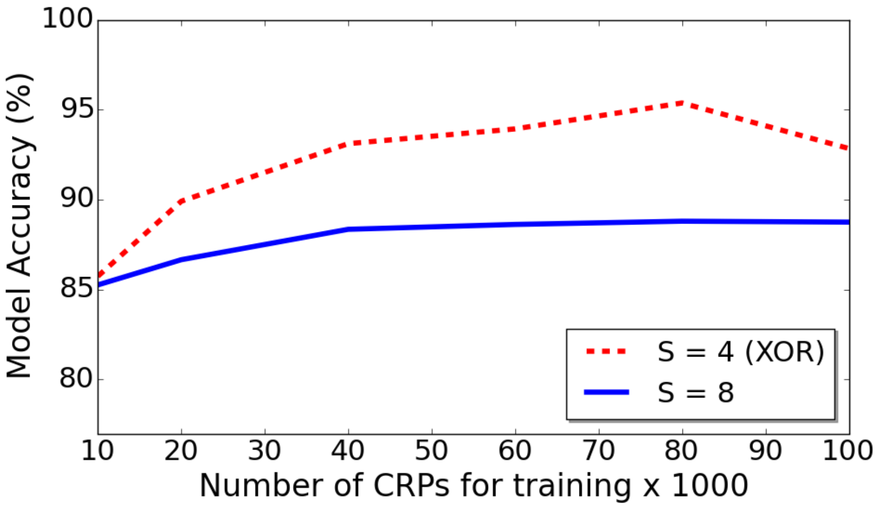
Table size =16

- Gradient Boosting ML attack
  - Uniform vs (Truncated) Normal distribution

- Circuit functions with equiprobable outputs are desirable for ML resistance

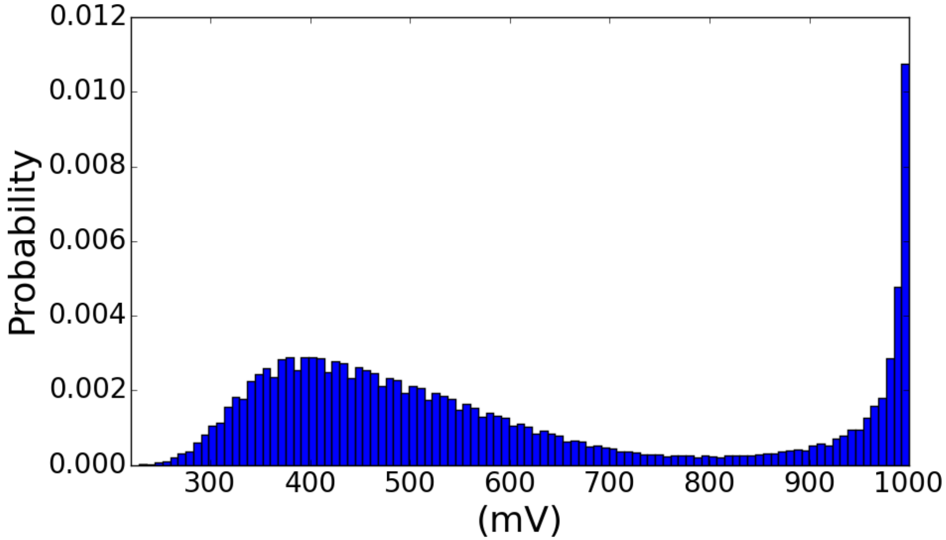
Study III - Impact of Digital Non-linearity



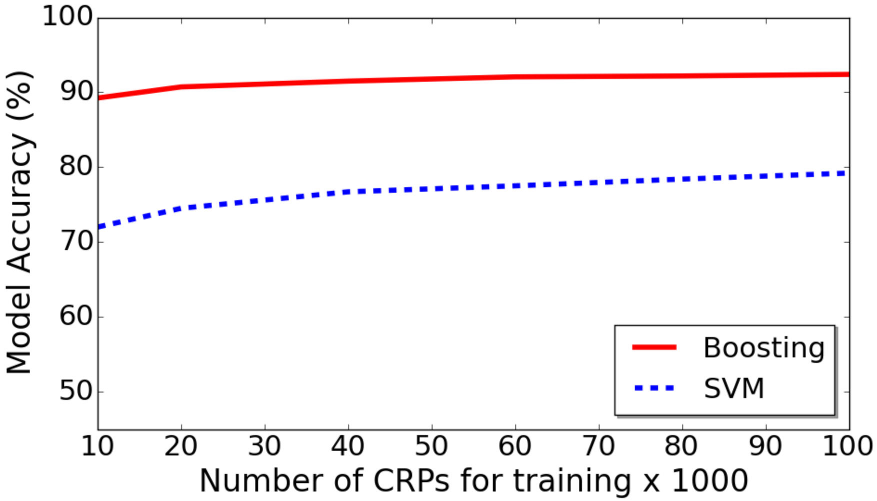


- Single, higher entropy source better than XOR'ing multiple PUFs
  - In context of function composition architecture

Study IV – Boosting vs VTC PUF



PDF of VTC PUF cell output values



Performance of ML attacks

- VTC function output PDF plotted
  - Bias in output value
- Gradient boosting improves prediction accuracy
  - 92% prediction rate in comparison to 80% using SVM*

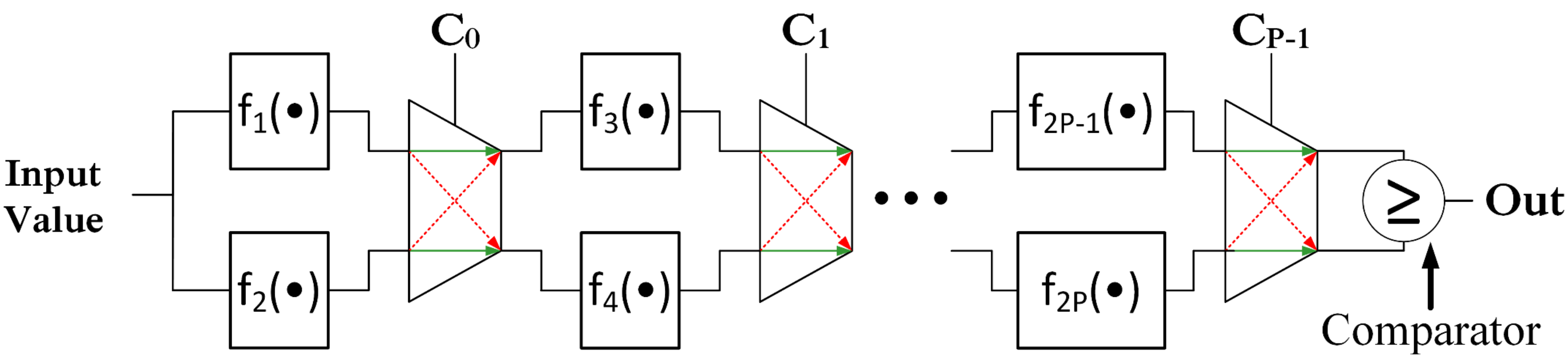
[*] Vijayakumar et.al, DATE 2015

Key Takeaways !

- Non-linear functions increase the machine learning resistance
  - Non-monotonicity needed to prevent saturation in implementation
- Composing non-linear functions using function composition shows promise
  - Can lead to systematic design approaches
- Sufficient entropy from non-linear functions
  - The switch architecture with function composition construction ensures modeling-attack resistance
- Bagging and Boosting algorithms are more potent than traditional ML attacks on PUFs
  - Creates new attack model
- Given function satisfying the properties it is indeed possible to build ML resistant PUF against known attacks

Future PUF design directions

- How it helps **PUF circuit designers** ?
- Properties of the family of functions $f_i()$ identified through study
  - Circuit designers can focus on implanting such function



- **Future work**
  - Circuit implementation of such functions
  - Build silicon and test

Thanks !

Acknowledgement:

Intel Circuit Research Lab members