ememory

Embedded Wisely, Embedded Widely

PUF: A Crucial Technology for AI and IoT

WHITEPAPER

Introducing AIoT Trends

Artificial intelligence of things (AIoT) is a new trend that combines artificial intelligence (AI) with the internet of things (IoT) to create networks of digital devices that communicate and process data. While IoT creates vast connections, AI makes these devices come alive.

Here’s one example: an IP camera system can be used for apartment security. Yet, without AI, people need to monitor video from the system in real-time in order to respond to emergencies. With AI, IP cameras can recognize risks automatically and send alerts.

Certainly, AIoT promises to grow rapidly and give rise to new high-valued products in the same way that the internet led to the creation of so many huge businesses. Those who enter the production of AIoT devices will be poised to tap a vast new market.

IoT devices today number in the billions. These small, connected gadgets include electronic devices and appliances networked together and communicating over internet protocols (IPs).

Yet, adding AI to IoT has created new security challenges.

The Challenges of AIoT

One of the key challenges for AIoT is the protection of AI assets. AI functions often need to detect, evaluate and respond in real time. As a result, a critical security concern is the fact that internal databases and interfaces for AI are not suitable for encryption because such an operation would demand too much time and resources. However, big data and interface designs are all proprietary information that need to be securely protected. The data needed by AI systems is typically so large that it usually be stored in an external non-volatile memory (NVM), thereby exposing it to hacking risks that are increasing worldwide.

Meanwhile, in addition to the “internal” security issues of AIoT systems, the external challenges of AIoT security have also increased. Nearly two million cyberattacks in 2018 resulted in more than $45 billion in losses worldwide as governments struggled with ransomware and other malicious incidents.

The Internet Society’s Online Trust Alliance (OTA), which identifies and promotes security and privacy best practices that build consumer confidence in the internet, said in its Cyber Incident & Breach Trends Report that the financial impact of ransomware rose by 60%, losses from business email compromise (BEC) doubled, and cryptojacking incidents more than tripled in 2018.

It’s clear that while security concerns remain unresolved, the deployment of AIoT devices will increase attack vectors for intrusions. Therefore, we say that for AIoT devices, PUF-based hardware security is the perfect solution. With PUF, the existing tradeoff of security for performance is eliminated.

How PUF Solves AIoT Security Concerns

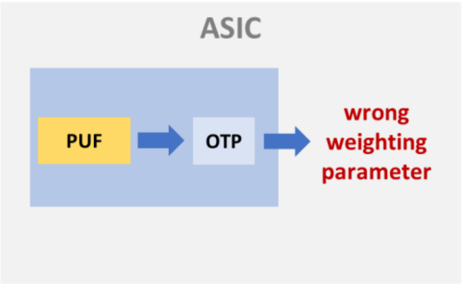
NeoPUF is a hardware security technology based on the physical unclonable variations occurring in the silicon manufacturing process. The underlying benefit of using a PUF (Physical Unclonable Function) in cryptography is its “uniqueness” and “unpredictability”. With eMemory’s NeoPUF, a chip can generate truly random sequences that can be used in applications with high security requirements. Our innovative technology can enable multi-layered security and resolve PUF-related concerns such as the additional costs of complicated ECC (Error Correction Code). The random number extracted via NeoPUF is so unique and unclonable that it can be used as a silicon “fingerprint” for a wide range of security purposes, including encryption, identification, authentication and security key generation.

The dimension of attacks on AIoT include “data and firmware attacks”, “transmission attacks” and “data integrity attacks”. We mentioned that complex encryption and decryption are impractical for the protection of AI assets. PUF has become a relatively simple and fast solution for security. Below are a few application scenarios that could help you have a clearer picture of how PUF solves AIoT security concerns.

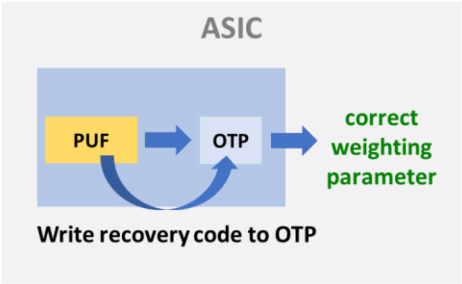
Application Scenario I:

Using a secret derived from PUF as protection, we can mix that secret with parameters. This will prevent encrypted parameter values stored in one-time programmable (OTP) memory from being hacked. When a secure AI module starts processing, we only need to process the mix procedure again with the PUF value so that the encrypted parameters can be simply decrypted to their original value. The encryption concept used in this case is XOR:

A(secret value derived from PUF)⊕B(parameters in OTP) = ciphertext
A(secret value derived from PUF)⊕B(parameters in OTP)⊕ A(recovery code) = plaintext



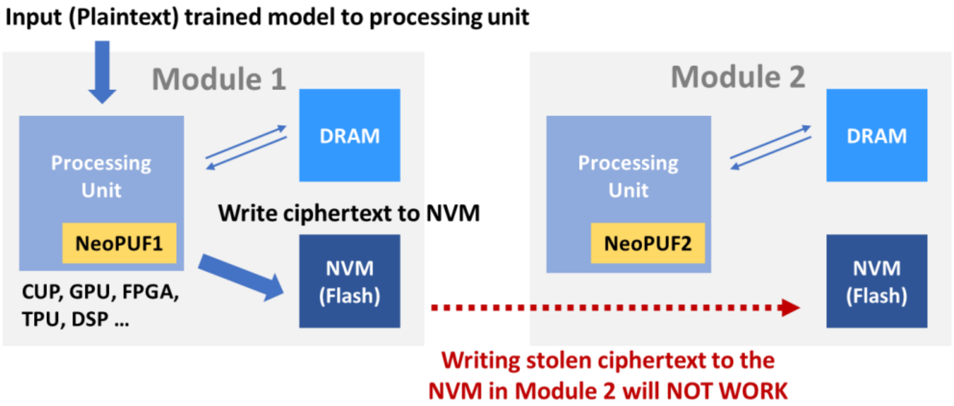
Before Enabling



After Enabling

Application Scenario II:

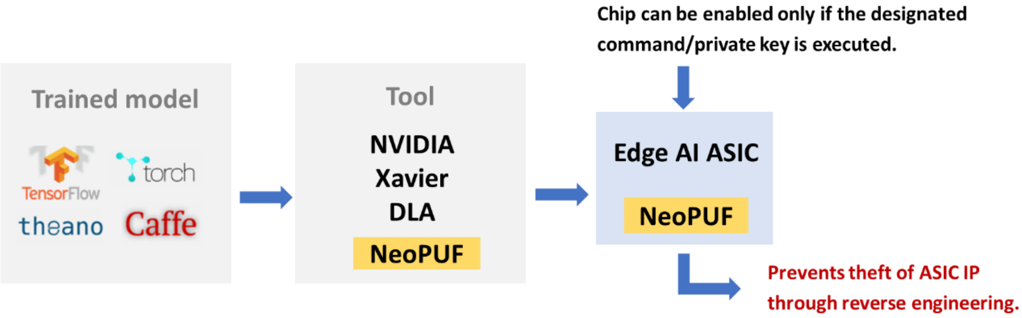
To deal with the transmission attacks, such as data breaches during transmission, we can use this simple idea: tying data to specific models by utilizing unique PUF values in each module. If data and the unique PUF value in a binding module are simply mixed as ciphertext, even if those encrypted data is stolen when it is transmitted to an external NVM, it cannot be used in other modules in anyway. It is because that the data will need the specific PUF value combined with the module to process the decryption procedure.


Input (Plaintext) trained model to processing unit
Module 1
Processing Unit
NeoPUF1
DRAM
Write ciphertext to NVM
NVM (Flash)
CUP, GPU, FPGA, TPU, DSP ...
Module 2
Processing Unit
NeoPUF2
DRAM
NVM (Flash)
Writing stolen ciphertext to the NVM in Module 2 will NOT WORK

Application Scenario III:

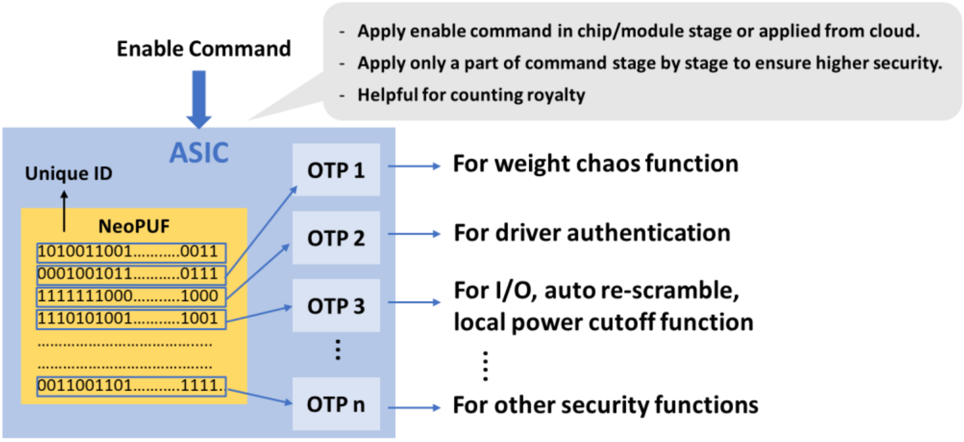
We know that the process of AI machine training is like this: First, collecting a lot of data for training, then extracting and testing the model for performing a prediction or reaction. In the case of edge computing, converting the model into an ASIC could lower the power consumption. Certainly, this model needs to be much leaner, however, this makes the module more vulnerable to reverse engineering.

To counteract reverse engineering, the PUF can be used as a credential. In other words, it can act as a key so that vouchers can only be able to start a chip with its unique PUF value. As a result, malicious reverse engineering hardware analysis and theft of secrets could be prevented.


Chip can be enabled only if the designated command/private key is executed.
Trained model
TensorFlow
torch
theano
Caffe
Tool
NVIDIA Xavier DLA
NeoPUF
Edge AI ASIC
NeoPUF
Prevents theft of ASIC IP through reverse engineering.

Building Safeguard Layers for Your AIoT System with NeoPUF

The same PUF encryption method can be applied to many levels of AI data or system structures as seen through the examples above. In summary, implementing multiple PUF and OTP for layered data protection can greatly strengthen security against hacking. As seen in the diagram below, there are multiple layers of OTP and PUF pairings that can complicate the process of data theft.


Enable Command
- Apply enable command in chip/module stage or applied from cloud.
- Apply only a part of command stage by stage to ensure higher security.
- Helpful for counting royalty
ASIC
Unique ID
NeoPUF
1010011001...........0011
0001001011...........0111
1111111000...........1000
1110101001...........1001
0011001101...........1111.
OTP 1
OTP 2
OTP 3
OTP n
For weight chaos function
For driver authentication
For I/O, auto re-scramble, local power cutoff function
For other security functions

This article is written by PUFSecurity, a PUF-based security IP solutions company and the subsidiary of eMemory. For more information, please visit the website at: http://www.ememory.com.tw and https://www.pufsecurity.com.

About eMemory

eMemory is a global leader in logic process embedded non-volatile memory (eNVM) silicon IP established in 2000. eMemory has devoted itself to research and development of innovative technologies, offering the industry's most comprehensive platforms of patented eNVM IP solutions which are supplied to semiconductor foundries, integrated devices manufacturers (IDMs), and fabless design houses worldwide. eMemory's eNVM silicon IPs support a wide range of applications, including trimming, function selection, code storage, parameter setting, encryption, and identification setting. The company has the world's largest NVM engineering team and prides itself on providing partners with a full-service solution that sees the integration of eMemory eNVM IP from initial design stages through fabrication.

About PUFSecurity

PUFsecurity is a PUF-based security IP solutions company and the subsidiary of eMemory, one of the world's largest SIP holders in Logic non-volatile memory (NVM) technology. We leverage the technical knowledge and industrial expertise of our parent company, eMemory, to bring PUF-based security to the market. With our core IP as the root of trust for security applications and qualified manufacturer platforms, PUFsecurity offers solutions with better performance and cost-efficiency. We believe that embedded hardware security will be widely adopted in this connected world.

eMemory

Embedded Wisely, Embedded Widely

eMemory Technology Inc.
8F, No.5, Tai-Yuan 1st St., Jhubei City, Hsinchu County 30265 Taiwan
T +886-3-5601168 F +886-3-5601169
www.ememory.com.tw