January 18, 2019

Detection of counterfeit silicon

Miguel Osorio Lozano

Sean Keys

Detection of counterfeit silicon

ABSTRACT

It is important for manufacturers of silicon chips to guard against counterfeit chips, hardware trojans, exploits, royalty-avoiding mechanisms, and other malicious hardware injected into a production stream. The techniques of this disclosure create individualized chip signatures during the manufacturing process. The signatures can be used to detect counterfeit silicon or otherwise suspicious on-board componentry.

KEYWORDS

- counterfeit silicon
- hardware attestation
- silicon attestation
- SRAM fingerprinting
- physically unclonable function
- PUF
- hardware trojan
- hardware counterfeiting

BACKGROUND

It is important for manufacturers of silicon chips to guard against counterfeit chips, hardware trojans, exploits, royalty-avoiding mechanisms, and other malicious hardware injected into a production stream. Secure manufacturing is of special concern for applications such as secure microprocessors, secure storage, etc. When the manufacture of integrated circuits is secured, it also boosts customer confidence.

Secure manufacturing generally seeks to guarantee that no backdoors are left open in the hardware that can be exploited by malicious parties. Further, secure manufacturing includes registration and certification of chip identities, e.g., chip-unique public keys, provision of shared app-layer secrets, e.g., batch keys, etc.

One current technique of unique chip identification is the integration of physically unclonable functions (PUF). PUFs are identifiers on chips that result from manufacturing process variation. An example of a PUF is the SRAM fingerprint. On-chip PUFs have several deficiencies. Such PUFs require expensive normalizing hardware to generate consistent inferences or hashes.

This can be a significant additional expense. Further, it is possible that an attacker can potentially forge a hash. Normalizing algorithms used by on-chip PUFs are lossy. Hardware based PUFs are static. If raw physically unclonable data is collected at the time of attestation, off-line analysis, e.g., using cloud-based machine learning, can spot unexpected differences between silicon dies.

Without creating a per-device unique physical environment (magnetic field, power supply), attackers can intercept genuine silicon, or even buy off-the-shelf parts, and gather sufficient data such that their counterfeit designs can be tuned to produce inferences that match genuine parts.

DESCRIPTION

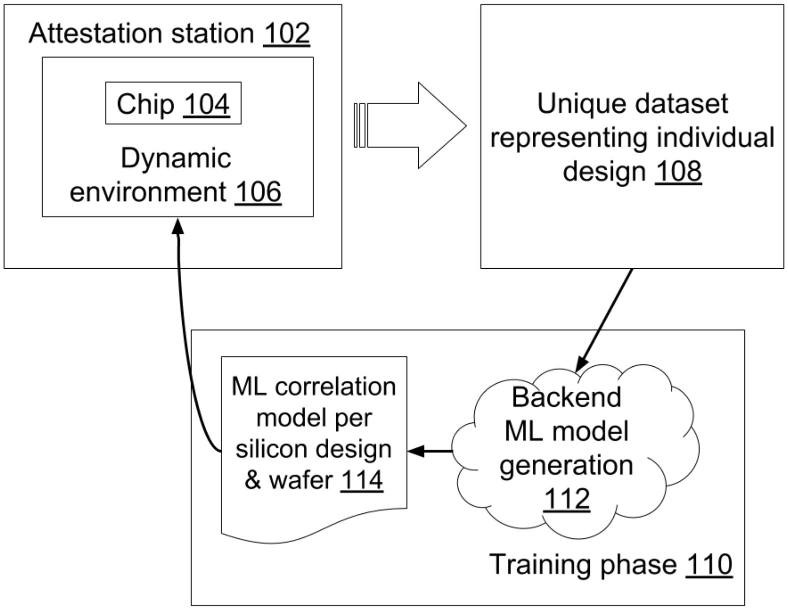
Attestation station 102
Chip 104
Dynamic environment 106
Unique dataset representing individual design 108
Backend ML model generation 112
ML correlation model per silicon design & wafer 114
Training phase 110

Fig. 1: Silicon attestation

Fig. 1 illustrates silicon attestation, per techniques of this disclosure. During the manufacturing process, e.g., near the final system-level test stage and prior to its release to customers, a chip (104) is placed into an attestation station (102). For example, the chip may be hardware that has a small SRAM footprint, e.g., a low power, secure microcontroller that can be used as a security key, or to sign certificates, etc. Further, such hardware may be a trusted platform module in a mobile device, such as a smartphone, tablet or laptop computer. Such hardware can also include secure storage devices. The hardware provides a hardware root of trust, and enables features such as verified boot.

The station subjects the chip to a dynamic environment (106), e.g., collects multiple SRAM snapshots using a different voltage for each snapshot. It can also vary the voltage during an SRAM scan iteration. This enables the collection of multiple datasets from hardware with relatively small SRAM footprints. Each attestation station can be programmed with some entropy, e.g., in a pseudo-random manner. For example, an attestation station may use a unique, randomly-generated voltage step pattern, making it difficult to fabricate data without access to the specific attestation station.

The dataset generated by an attestation station is unique to the chip and its design (108). The dataset is used in a training phase (110) for a backend machine learning model (112) that generates trained ML models (114). The trained ML models are correlation models per silicon design and wafer. Data from chips in the same family, e.g., chips of the same silicon design, wafer, etc., are used in the training phase (110) to generate the machine learning models. Data fingerprints can be used to verify a chip for authenticity. The machine learning model accounts for the association of a dataset with a particular attestation station, enabling the detection of a mismatch during the verification process.

The techniques of this disclosure collect multiple raw datasets in non-lossy form. The relative abundance of data enables continuous improvement in characterizing silicon and detection algorithms. Further, the per-chip cost is lower than off the shelf PUF IP that can cost several dollars per chip.

CONCLUSION

The techniques of this disclosure create individualized chip signatures during the manufacturing process. The signatures can be used to detect counterfeit silicon or otherwise suspicious on-board componentry.